

VAASAN YLIOPISTO
TEKNIIKAN JA INNOVAATIOJOHTAMISEN YKSIKKÖ
TIETOTEKNIikka

Mika Varila

**KONEOPPIMINEN PÄÄTÖKSENTEON TUKIJANA DIABETES
MELLITUKSEN HOIDOSSA**

Tietotekniikan pro
gradu -tutkielma

Tietojärjestelmätieteen maisterin koulutusohjelma

VAASA 2020

SISÄLLYSLUETTELO

1. JOHDANTO.....	6
1.1 Tutkimuksen tavoite ja rajoitukset.	7
1.2 Tutkimuksen kulku ja tulokset.	8
2. KONEOPPIMINEN	12
2.1 Koneoppimismallit ja datan luokittelu.....	17
2.2 Koneoppiminen terveydenhoidossa.....	30
3. TERVEYDENHUOLLON TIETOJÄRJESTELMÄT	40
3.1 Lääketieteellisen hoidon laatu	46
3.2 Krooniset sairaudet.....	47
4. DIABETES MELLITUS	50
4.1 Diabeteksen esiintyvyys ja hoitokäytännöt.....	52
4.2 Diabeteksen syyt ja seuraukset	54
4.3 Kliininen päätöksenteko diabetes mellituksen hoidossa	57
5. TUTKIMUSMETODI JA MALLI.....	61
5.1 Integroiva kirjallisuuskatsaus.....	61
5.2 Kirjallisuuskatsauksen koostaminen	67
5.3 Relevantin tutkimusaineiston kerääminen, analysoiminen ja tulkinta	77
6. KONEOPPIMISMALLIT DIABETES HOIDOSSA.....	86
6.1 Koneoppimisen soveltaminen diabetes hoitoon.....	86
6.2 Koneoppiminen ja diabetes hoito tulosten ennustaminen.....	102
6.3 Koneoppiminen diabeteksen diagnosointityökaluna	114
7. DISKUSSIO	130
8. LÄHTEET	140

LYHENTEET

AIM = Tekoäly lääketieteessä

AUC = Luokittelijan havaintojen erottelukyky

BMTL = Bayesilainen Monitehtävä Oppiminen

CIMT = Kaulavaltimon sisäkalvon paksuuden ultraääni arviointi

CNN = Kierteinen hermoverkko

cVAUC = Ristiinvalidoitu kaariviivan alainen alue

DFU = Diabeettinen jalkamätähaava

DR= Diabeettinen retinopatia

EMR = Elektroninen lääkintärekisteri

FAIMS = Kenttä Assymetria Ioni liikkuvuus

FCN = Täysin kiertein hermoverkko

FMR = Yhteen liittynyt koneoppiminen

FPG = Plasman paastosokeri

GFR = Hiussuoniston suodatus nopeus

Gradient Boost = Kaltevuus suhteen/kaltevuuden/muutosnopeuden vahvistaminen

KNN = K-lähin naapuri

LR= Logistiset Regressiomallit

LDA = Linearinen erottelukyky analyysi

NB = Naiivi Bayesiläinen

NPV = Negatiivinen ennustava arvo

SVM = Tukivektori kone

PDC = Sairaalan sijoittamisen riski ennustus perustuen katettujen päivien osuuteen

PPV = Tarkka positiivinen ennustearvo

PCA = Pääkomponenttien analyysi

OCT = Kerroskuvaus röntgen

QDA = Neliö lineaarinen erottelu analyysi

VAASAN YLIOPISTO**Teknologia ja innovaatiojohtamisen yksikkö****Tekijä(t):**

Mika Varila

Tutkielman nimi:

Koneoppiminen päätöksenteon tukijana diabetes mellituksen hoidossa

Ohjaaja:

Tero Vartiainen

Tutkinto:

Kauppatieteiden maisteri

Oppiaine:

Tietotekniikka

Koulutusohjelma:

Tietojärjestelmätieteen maisteri

Aloitusvuosi:

2016

Valmistumisvuosi:

2020

Sivumäärä: 154

TIIVISTELMÄ

Koneoppiminen on yksi tekoälyn osa-alue, jota voidaan hyödyntää laajasti terveydenhuollossa erilaisiin käyttötarkoituksiin. Diabetes hoidossa koneoppimisteknologioiden käyttöönotto voi merkitä huomattavaa laadullista parannusta ja kustannustehokasta hoitoa. Tutkimuksen tavoitteena on tuottaa käyttökelpoista ja ohjeellistavaa tietoa koneoppimisen soveltamismahdollisuuksista toimivan kliinisen ei-tietämyskantaisen päätöksenteon tukijärjestelmän suunnittelumallin luomiseksi terveydenhuolto-organisaatioihin, terveydenhuollon ammattihenkilökunnan kliinisen päätöksenteon edistämiseksi.

Tutkimuksen teoreettisen viitekehyksen muodostaa koneoppiminen terveydenhoidossa ja kliininen päätöksenteko. Tutkimuksen osioita ovat koneoppimisen sovellettavuus diabeteshoitoon, koneoppimisen soveltaminen diabetes hoitotulosten ennustamiseen ja koneoppiminen diabeteksen diagnosointityökaluna. Tutkimusmenetelmä on kvalitatiivinen, integroiva kirjallisuuskatsaus. Aineisto kerättiin useasta eri tietokannasta, ja se muodostuu pääasiassa tieteellisistä katsaus-, tutkimus- ja konferenssiartikkeleista. Tutkimuksen aineisto analysoitiin ymmärtämään pyrkivällä laadullisella analyysillä. Tämä tehtiin induktiivisella lähestymistavalla aineistolähtöisenä sisällönanalyysinä.

Integroivan kirjallisuuskatsauksen synteessin pohjalta saatu tutkimustulos vastaa esitettyihin tutkimuskysymyksiin ja määrittelee toimivan ei-tietämyskantaisen kptj:n vaatimuksia järjestelmän varsinaista suunnittelua ja teknistä toteutusta varten. Tulokset osoittavat, että koneoppimistekniikoista syväoppiminen, ohjaamaton oppiminen, ohjattu oppiminen, yhteen liittynyt koneoppiminen ja äärimmäinen oppimiskone ovat niitä koneoppimisalgoritmeja, joita pitäisi integroida mukaan ei-tietoon-perustuvaan kliiniseen päätöksenteon tukijärjestelmään, varsinaisen kliinisen päätöksenteko prosessin tukemiseksi diabetes hoidossa. Tutkimuksen tuloksia on selostettu tarkemmin diskussio kappaleessa ja rajoitukset on myös pyritty tuomaan esille.

AVAINSANAT: koneoppiminen terveydenhoidossa, hoidon laatu, diabetes mellitus, kliininen päätöksenteon tukijärjestelmä ja integroiva kirjallisuuskatsaus

UNIVERSITY OF VAASA

The School of Technology and Innovations

Author: Mika Varila
Topic of the Thesis: Koneoppiminen päätöksenteon tukijana diabetes mellituksen hoidossa

Instructor: Tero Vartiainen
Degree: Master of Science in Economics and Business Administration

Degree Programme: Master's Programme in Information Systems
Major: Computer Science

Year of Entering the University: 2016

Year of Completing the Thesis: 2020

Pages: 154

ABSTRACT

Machine Learning is one sub-area of artificial intelligence that can be utilized widely in healthcare to different purposes of use. Adopting machine learning technologies to diabetes care can mean considerable improvement in quality of care and cost-effective treatment of this disease. The aim of this research is to acquire viable knowledge about healthcare systems in diabetes care that utilizes machine learning technologies for a non-knowledge base clinical decision support system planning model designed for diabetes care in order to support health care professionals decision making process.

Machine learning in healthcare and clinical decision making formulate framework in this research. Machine learning as clinical decision making supporter of healthcare management in diabetes care was researched using qualitative research method, in other words integrative literature review. A data is collected from databases that includes a scientific research literature. Articles that have been chosen to this research are review-, survey and conference articles that have been evaluated and accepted by scientific community. The study focuses to machine learning applicability to diabetic care, machine learning and predicting of diabetes care outcome, and machine learning and diagnosing of diabetes mellitus. Research data is analyzed inductively by content analyses that is based on research data.

Based on research methods a non-knowledge base clinical decision support system planning model idea was investigated from integrative literature review for hospitals in order to contribute clinical decision making process of health care professionals in diabetes care. The results show that machine learning techniques that should be integrated into clinical decision support systems are deep learning, non-controlled machine learning, controlled machine learning, federated machine learning and extreme learning machine. These technologies can promote clinical decision making process of healthcare professionals by helping to build a functional non-knowledge base clinical decision making support system designed for diabetes care. Results are reported thoroughly in discussion chapter and limitations are also mentioned.

KEYWORDS: Machine Learning models, Quality of Care, Clinical Decision Making, and Integrative Literature Review

1. Johdanto

Julkinen terveydenhuolto järjestelmämme kohtaa suuria haasteita tulevaisuudessa mm. kroonisten sairauksien lisääntymisen ja väestön ikääntymisen takia. Yksi globaalisti lisääntyvä krooninen sairaus on diabetes mellitus. Se on noussut yhdeksi merkittävimmäksi kansanterveydelliseksi ongelmaksi viime vuosina, ja siitä koituu myös valtavia kustannuksia kansataloudelle terveydenhuoltomenoina. Asiantuntijat uskovat myös diabeetikoiden määrän lisääntyvän yhteiskunnassa lähitulevaisuudessa. Tämä asettaa myös terveydenhuollon tietojärjestelmille paljon uusia vaatimuksia. Niiden on pystyttävä ennen kaikkea tukemaan terveydenhuollon organisaation ammattihenkilöiden diabetes hoitoon liittyviä päätöksentekomekanismeja mahdollisimman täsmällisesti ja asianmukaisesti.

Tekoälyn käyttöönotto terveydenhuollossa voi avata täysin uusia mahdollisuuksia kroonisten sairauksien ennaltaehkäisyssä, sen oikea-aikaisessa diagnosoimisessa ja hoitamisessa. Sen uskotaan voivan kustannustehokkaasti tehostaa terveydenhuollon eri sektorien toimintaa terveydenhuollossa. Koneoppiminen on yksi tekoälyn osa-alue (Vähäkainu & Neittaanmäki 2017). Koneoppimisen soveltamisessa diabeteshoitoon on kuitenkin yleensä ongelmana se, että terveydenhuolto organisaatioilla puuttuu tietopohjana optimaalista kliinistä päätöksentekoa hyvän hoidon laadun näkökulmasta edistävä, hyvin toimiva kliinisen päätöksentekojärjestelmän suunnittelumalli. Tällöin pitäisi selvittää ne osatekijät, mistä onnistunut diabetes hoitoon sovellettu kliinisen päätöksenteko järjestelmän malli rakentuu.

Tutkimus on poikkitieteellinen. Siinä käytetään hyödyksi muiden rinnakkaistieteiden tämän aihepiirin viimeaikaisia löydöksiä ja tutkimustuloksia. Näin voidaan myös edistää tietojärjestelmätieteen kehitystä itsenäisenä tieteenalana. Koneoppimisen soveltamismahdollisuuksien tutkiminen diabetes hoidossa voi antaa täysin uusia mahdollisuuksia, toimivimpien ei-tietopohjaisten kliinisten päätöksenteon tukijärjestelmien suunnittelumallien luomiselle.

Aikaisemmin julkaistun tieteellisen kirjallisuuden perusteella tiedetään koneoppimisen soveltamisen diabetes hoitoon pystyvän huomattavasti helpottamaan diagnosointia, karsimaan kalliita hoitokustannuksia sekä edistämään sairauden kliinistä tutkimusta. Tähänastiset tutkimustulokset ja löydökset ovat kiistatta osoittaneet, että koneoppimista on käytetty sairauksien diagnosoinnissa tehokkaasti. Se toimiikin erinomaisesti erityyppisten sairauksien diagnosointitilanteiden työkaluna, ja antaa historian tähänastisesti tarkimman diagnosointimenetelmän. Sitä pidetään myös erittäin kustannustehokkaana ratkaisuna lääkärille, sairaalalle ja potilaille. (ks. Vähäkainu & Neittaanmäki 2017). Koneoppimista voidaan käyttää myös tautien kulun ymmärtämisessä sekä tehokkaiden hoitojen suunnittelussa esim. 2 tyypin diabetekselle. Koneoppimisella voidaan myös kehittää monia älykkäitä sovelluksia edistämään kliinistä tutkimusta. (ks. Fagella 2018).

1.1 Tutkimuksen tavoite ja rajaukset

Tutkimuksen tavoite on tutkia koneoppimisen soveltamista diabetes hoitoon. Terveydenhuollon elektronisoituminen vaatii terveydenhuolto-organisaatioilta toimivan ICT-teknologian ja korkealuokkaisen tekoälysovellusten onnistunutta soveltamista järjestelmäsuunnittelussa. Tämän tutkimuksen tuloksena syntyy ei-tietämyskantainen kliinisen päätöksentekojärjestelmän suunnittelumallin idea terveydenhuolto organisaatiolle. Tutkimus on toteutettu laadullisella, kvalitatiivisella tutkimusmenetelmällä, eli tutkimusmetodina käytetään tässä tutkimuksessa integroivaa kirjallisuuskatsausta. Laadullinen tutkimus on valittu tutkimuksessa lähestymistavaksi, koska tutkimuksen tarkoituksena on ymmärtää tutkittavaa ilmiötä, ja saada siitä syvälinen, kattava kokonaiskuva ja ymmärtää tarkemmin koneoppimisen käyttötarkoituksia tässä vaikeassa kroonisessa sairaustilassa.

Tämän tutkimuksen ongelma määritellään deskriptiivisesti, eli minkälaisia ongelmia terveydenhuolto järjestelmässä tällä hetkellä vallitsee diabetes mellituksen hoidossa. Onko koneoppimista mahdollista hyödyntää tehokkaasti diabeteshoidossa klinikoiden

kliinisen päätöksenteon tukemisessa? Tutkimusongelmaa pyritään ratkaisemaan seuraavilla tutkimuskysymyksillä: mitä preskriptiivistä tietoa tieteellisessä kirjallisuudessa on koneoppimisen mahdollisuuksista diabeteksen hoidossa? Kuinka sen avulla voidaan edistää ja kehittää terveydenhoidon johdon kliinistä päätöksentekoa tämän yleistyvän kroonisen sairauden hoidossa? Miten koneoppimismallit voivat auttaa diabetespotilaiden hoitotulosten ennustamisessa? Tutkimus toteutettiin laadullisella, kvalitatiivisella tutkimusmenetelmällä, eli tutkimusmetodina käytetään tässä tutkimuksessa integroivaa kirjallisuuskatsausta. Laadullinen tutkimus on valittu tutkimuksessa lähestymistavaksi, koska tutkimuksen tarkoituksena on ymmärtää tutkittavaa ilmiötä, ja saada siitä syvälinen, kattava kokonaiskuva ja ymmärtää tarkemmin tutkimuskohteen käyttötarkoituksia tässä vaikeassa kroonisessa sairaustilassa. Tutkimusaineisto hankittiin pääasiassa tietokannoista, joissa on tallennettuna arvostetuista tiedejulkaisuisuista julkaistuja, tieteellisen arvion läpikäyneitä tutkimus-, katsaus- ja konferenssi artikkeleita. Näistä muodostuu varsinainen työn teoreettinen viitekehys.

1.2 Tutkimuksen kulku ja tulokset

Koneoppimisen onnistunutta soveltamista diabeteshoitoon tutkitaan tässä tutkimuksessa kvalitatiivisen tutkimusmenetelmän avulla, eli ilmiötä tutkitaan integroidun kirjallisuuskatsauksen perusteella. Sen avulla kehitetään toimivaa ei-tietämyskantaisen kliinisen päätöksenteon tukijärjestelmän suunnittelumallia terveydenhuoltoorganisaatioon. Integroiva kirjallisuuskatsaus on valittu tutkimusmenetelmäksi, koska tavoitteena on luoda ei-tietämyskantaisen kliinisen päätöksenteko järjestelmän suunnittelumallin idea terveydenhuolto organisaatioille ja tutkimuksen teorian luomiseen voidaan käyttää uusinta olemassa olevaa tutkimusaineistoa.

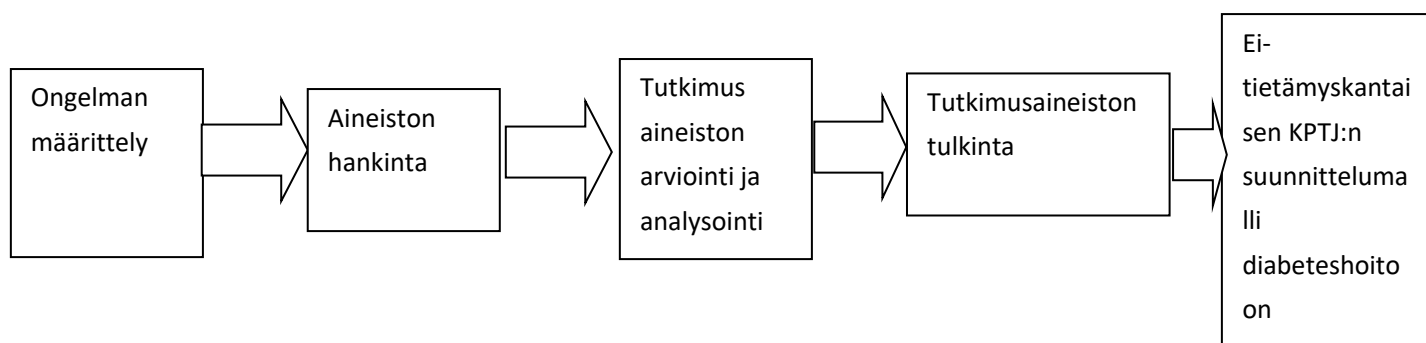
Tutkimusmetodiksi tutkimukseen on valittu laadullinen, integroiva kirjallisuuskatsaus. Integroivassa kirjallisuuskatsauksessa tutkimusongelmat pyritään ratkaisemaan olemassa olevan tutkimuskirjallisuuden avulla. Tässä tutkimusmetodissa tehtyä tutkimusta tutkitaan uudelleen. Sen avulla voidaan koota tutkimuksista tuloksia, nämä puolestaan

ovat perusteena täysin uusille tutkimustuloksille. (Salminen 2011:1). Integroidun kirjallisuuskatsauksen avulla voidaan myös mm. tunnistaa nykyisiä aukkoja tutkimuksessa ja täydentää aukkoja. Sen avulla voidaan myös suorittaa tutkimuksen alaisen kohteen aihealueen tieteellisen näytön vahvuuden arviointi ja teoreettisen tai käsitteellisen viitekehyksen tunnistaminen. (Russel 2005). Tutkimusmenetelmänä integroivan kirjallisuuskatsauksen hyvänä puolena on se, että siinä on mahdollista yhdistää teoreettista ja empiiristä tietoa sekä myös eri metodein tehtyä tieteellistä tutkimusta samaan katsaukseen. Sen päällimmäiset tehtävät ovat tieteen nykytilan kuvaaminen, teorian kehittäminen ja tutkimuksesta saadun tiedon soveltaminen. (Whittermore & Knaff 2005).

Integroiva kirjallisuuskatsaus voidaan jakaa eri vaiheisiin. Integroivassa kirjallisuuskatsauksessa ensimmäisen vaiheen tarkoitus on tunnistaa ja määrittellä selvästi tutkimusongelma. Sen jälkeen tutkimusaineisto hankitaan tutkimuskysymysten kautta määriteltyjen valintakriteerien perusteella. Seuraavassa vaiheessa kerätty tutkimusaineisto arvioidaan, analysoidaan ja analyysivaiheessa esille nousseiden avainkäsitteiden ja merkitysten perusteella voidaan tehdä tutkimuksesta varsinainen synteesi. (Salminen 2011:8).

Tutkimusaineisto koostuu lähinnä tieteellisen arvioinnin läpikäynneistä review-, tutkimus-, ja konferenssi artikkeleista, jotka ovat tutkimuksen ensisijaisia lähteitä. Näin ollen voidaan varmistaa, että kirjallisuuskatsauksessa käytetään tieteellisesti luotettavia lähteitä aineistona. Tutkimuksessa on lähteinä myös kirjoja ja Pro-gradu-tutkielmia, joista on ammennettu tietoa tutkimukseen liittyvän peruskäsitteistön ymmärtämiseksi. Koneoppiminen diabeteshoidossa, ja sen mahdollisuus edistää klinikoiden kliinistä päätöksentekoa, muodostavat tutkimuksen viitekehyksen.

Katsauksen tutkimusaineisto kerätään käyttäen useita tietokantoja. Ne ovat mm. pubmed, Science Direct, IEEE, AIS digital libary ja INFORMS. Niillä saadaan kattava kuvaus aihepiiriin nykytilasta. Seuraavalla sivulla kuvassa 1 on kuvattu kvalitatiivisen integroivan kirjallisuuskatsauksen eri vaiheet.



Kuva 1. Integroivan kirjallisuuskatsauksen vaiheet.

Tutkimuksessa osoitettiin, kuinka koneoppimismallit, algoritmit ja koneoppimisella toteutetut sovellukset voivat huomattavasti edistää terveydenhoidon ammattihenkilöiden kliinisen päätöksenteon prosessia diabetes hoidossa. Koneoppimistekniikoista syväoppiminen, ohjaamaton oppiminen, ohjattu oppiminen, yhteen liittynyt koneoppiminen ja äärimmäinen oppimiskone ovat erittäin toimivia teknologisia ratkaisuja hyvän hoidon laadun takaajina diabetes hoidossa. Koneoppimisteknologian käyttöönotto terveydenhuolto-organisaatioissa mahdollistaa toimivan ei-tietämyskantaisen kliinisen päätöksentekojärjestelmän toteuttamisen hyvän hoidon laadun takaamiseksi. Koneoppiminen tuo tähän merkittävään kansanterveydelliseen ongelmaan erinomaisen, kustannustehokkaan, teknologisen tekoälyyn pohjautuvan ratkaisun. Koneoppimistekniikoiden avulla voidaan paremmin ymmärtää tätä sairautta, ja koneoppimisalgoritmit mahdollistavat ennen kaikkea nopean ja tehokkaan potilasaineiston luokittelun ja käsittelyn.

Neuroverkkoihin perustuvat syväoppimista hyödyntävät päätöksenteontuen järjestelmät antavat nopeasti täsmällistä tietoa diabeteksen liitännäissairauksien tilasta, identifioivat ne, auttavat toimittamaan oikea-aikaista hoitoa niihin täsmällisesti, luokittelevat liitännäissairauksien vakavuusasteen ja pystyvät ennakoimaan diabetekseen liittyvien silmänsairauksien kehityskulkua. Syväoppimisen avulla voidaan myös toteuttaa tehokas diabeteksen liitännäissairauksien seulontajärjestelmä. Syväoppimisjärjestelmiä ollaan nykyisin diabeteshoidossa liittämässä yhä enemmän yhteen ja niihin on integroitu myös automaatiota, jonka avulla voidaan tehdä solutason havainnointia diabetes potilaan halvausriskin määrittämiseksi. Koneoppiminen ja datalouhinnan yhdistäminen voi

edistää myös varsinaista diabetestutkimuksen tekoa merkittävästi. Tärkeää sairauteen liittyvää tietoaaineistoa voidaan lisätä sen avulla, ja tärkeää kliinistä dataa pystytään toimittamaan oikea-aikaisesti terveydenhuollon ammattilaisten käyttöön. Sairauden vakavuusasteen taso voidaan myös määritellä nopeasti. Tämä tukee myös reaaliaikaista kliinistä päätöksentekoa liitännäissairauksien ilmaantuessa. Sairauden puhkeamista on mahdollista ennakoida samoin sen kehityskulkua. Diagnooseista päättäminen onnistuu nopeammin ja hoidon priorisointi tehostuu.

Ohjaamattoman oppimisen avulla voidaan antaa mm. oikeanlaisia ravintosuosituksia potilaille, ja voidaan myös säädellä diabetes potilaan veren sokeri piikkejä tarkoin. Ohjattu oppiminen puolestaan mahdollistaa tehokkaan seulontajärjestelmän rakentamisen ja sairauden komplikaatioiden hallinnan. Hoitoa voidaan myös kehittää yksilöllisempään suuntaan, ja interventioita voidaan myös kohdentaa niitä tarvitseviin hoitoon sitoutumattomiin riskipotilasryhmiin tehokkaasti. Uudet koneoppimistekniikat, kuten äärimmäiset oppimiskoneet voivat tehokkaasti ennustaa potilaan ylipainotilan kehittymistä ja ennakoida siihen liittyviä terveysriskejä. Diabeteksen tapahtumia voidaan myös nykyisin havaita täsmällisesti ns. korkea-ulottuvuus koneoppimisen avulla. Tiivistetysti voidaan sanoa, että koneoppimisen käyttöönotto terveydenhuolto järjestelmissä voi tehostaa diabetes hoidon diagnostiikkaa, oikein ajoitettua hoidon toimittamista laajalle ihmispopulaatiolle kustannustehokkaasti ja auttaa yksilöllisten hoito-ohjelmien toteuttamisessa.

2. Koneoppiminen

Koneoppimista on määritelty lukuisissa lähteissä. Roos(2016)analysoi ja määrittelee koneoppimista käsitteenä Helsingin Yliopiston luentomateriaalissa Johdatus tekoälyyn varsin laajasti ja mielenkiintoisesti. Hänen mukaansa koneoppiminen on käsitteenä määriteltävissä siten, että käsite puretaan kahteen erilliseen sanaan. Tällöin käsitteellä kone viitataan tietokoneeseen ja tietokoneohjelmaan. Käsitteellä oppiminen puolestaan tarkoitetaan hänen mukaansa ongelman ratkaisukyvyyn parantumista kokemuksen avulla. Näin ollen hän päätyy määritelmässään siihen johtopäätökseen, että ohjelmoija ei pelkästään kirjoita tarkkoja sääntöjä jonkun ongelman ratkaisuun vaan ohjelmoija ohjeistaa tietokonetta oppimaan esimerkeistä. Hänen mukaansa tietokone voi oppia ratkaisemaan jotain tehtävää jopa paremmin kuin ohjelmoija itse kykenee. (Roos 2016:31-44).

Roosin (2016) mukaan nykyisin on helppoa ja halpaa kerätä digitaalista dataa halvoilla sensoreilla. Datan tallentaminen on halvan kovalevytilan ansiosta myös halpaa ja helppoa, samoin kuten datan lähettäminen. Se on lähes ilmaista netissä. Tällöin kaikki pystyvät keräämään suuria määriä dataa. Näin tekevätkin mm. yritykset, kaupat (ostotapahtumat, hakukoneet, hakulausekkeet, valinnat), finanssisektori (osakkeet, valuuttakurssit), tehtaot (erilaiset sensorit), sosiaalinen media (facebook, twitter), kaikki palvelimet, tiede, kuten geenitekniikka, geeniespresso, ja hiukkaskokeet fysiikassa. Ongelmaksi muodostuu se, miten näitä voidaan hyödyntää. (Roos 2016:31-44).

Roos (2016) ottaa luentomateriaalissaan hyvin esille havainnollistavia esimerkkejä siitä, miten voidaan ohjelmoida tietokonetta esim. pelaamaan ristinollaa. Näiden esimerkkien kautta päädytään vähitellen vaiheittain siihen, mitä koneoppiminen itseasiassa tarkoittaa. Ensimmäisessä vaihtoehdossa katsotaan koko pelipuu läpi ja valitaan optimaalisesti vain yksinkertaiset pelit. Toisessa vaihtoehdossa ohjelmoija kirjoittaa säännöt sellaisella tyyllillä, että jos vastustajalla on kahden suora ja kolmas on vapaa, sinun tulee estää suora laittamalla oma merkki siihen ym. Tämä on kuitenkin erittäin työlästä ja vaikeaa.

Kolmannessa vaihtoehdossa tullaan varsinaiseen koneoppimiseen. Tällöin tietokoneohjelmoija kokeilee erilaisia sääntöjä, ja pelaa itseään (tai muita) vastaan, kokeilee millä säännöillä voittaa ja millä säännöillä hävitään. Tämä on varsinaista koneoppimista. (Roos 2016:31-44).

Toisena esimerkkinä Roos(2016) kirjoittaa Arthur Samuelin 1950-60 luvulla laatimasta koneoppimista kuvaavasta määritelmästä. Samuelin mukaan alkuvaiheessa tietokoneohjelma oppii pelaamaan tammaa. Silloin ohjelma pelaa itseään vastaan tuhansia kertoja, ja se oppii mitkä positiot ovat hyviä ja mitkä huonoja. Tämä tapahtuu sen perusteella, kuinka usein ne johtavat voittoon tai häviöön. Näin ollen ohjelma tulee lopulta paremmaksi kuin sitä opettava ohjelmoija. Roosin (2016) mukaan eräs koneoppimisen määritelmä on se, että tietokone parantaa suorituskäytännön suorittaessaan jotain tiettyä tehtävää. Tämä tapahtuu aina silloin kun se näkee esimerkkejä. Näin ollen on erotettava tehtävä, eli mitä koneoppiminen yrittää tehdä? Toisin sanoen, minkä ohjelman se ratkaisee? Niin sanottu hyvyysmitta on myös aina erotettava, ja tällöin on mitattava sitä, kuinka hyvin koen ohjelman ratkaisevan tehtävää? Viimeiseksi on vielä erotettava esimerkki/data, eli mikä on se kokemuspohja/esimerkkidata, jonka perusteella kone/ohjelma oppii?(Roos 2016:31-44).

Roos(2016) esittelee mielenkiintoisen modernin sovelluksen, jolla hän pyrkii havainnollistamaan sitä, missä arkipäiväämme kuuluvissa sovelluksissa koneoppimista käytetään. Yksi tällainen on automaattinen kasvojen tunnistus. Tämä ei toimi perinteisesti normaalin ohjelmoinnin periaatteiden mukaisesti, eli että ohjelmoija vain kirjoittaa sääntöjä. Säännöt voisivat esimerkiksi mennä niin, että algoritmilla olisi ehdot ”Jos tumma lyhyt tukka, joskus silmälasit, iso nenä, niin se on Mikko”. Tämä ei mitenkään käytännössä voisi toimia, miten on esim. mahdollista arvioida ”nenän” koko? jne. Sen takia varsinaisessa koneoppimisessa tietokoneelle näytetään (kuva, nimi), ne ovat esimerkkipareja. Tällöin tietokone oppii itse, mitkä piirteet ovat asiaan kuuluvia ja olennaisia. Tämä on yleensä melko vaikea ongelma koneelle, mutta se on mahdollista ratkaista, jos on riittävän paljon esimerkkejä. Seuraavana esimerkkinä on roskapostin suodatus. Sekään ei toimi siten, että ohjelmoija vain kirjoittaa sääntöjä, kuten esim. jos

sisältää sanan ”Viagra”, niin se on roskaa. Tämä on vaikeaa, eikä se ole käyttäjälle räättelöity. (Roos 2016:31-44).

Siksi onkin parempi, että käyttäjä ohjeistaa tietokonetta esimerkiksi siitä, mitkä postit ovat roskaa. Tällöin tietokone kykenee itse oppimaan luokittelusäännöt. Hakutulosten järjestäminen on myös yksi esimerkki, jossa koneoppimista soveltamalla voidaan pystyä ennustamaan sitä, mitä loppukäyttäjä haluaa. Hakutulosten järjestäminen ei toimi optimaalisesti silloin, kun sen toimintaperiaate perustuu siihen, että ohjelmoija kirjoittaa tarkan kaavan, jolla voidaan pisteyttää sanan esiintymismäärä, linkkien määrä, ym. Sen jälkeen ne järjestetään pisteiden mukaan. Tässä ongelmaksi muodostuu se, että ohjelmoija ei voi tietää, miten eri sivujen ominaisuuksia pitäisi painottaa. Tällöin koneoppimista apuna käyttäen tämä tarkoittaisi, että tallennettaisiin ne linkit, mitkä valitaan minkäkin hakulausekkeen jälkeen. Sitten laitettaisiin suosituimmat sivut kärkeen. Siitä voidaan yrittää ennustaa mitä loppukäyttäjä haluaa. Hakulausekkeiden ennustamisessa ohjelmoija antaa valmiin sanakirjan. Tässä ongelmaksi muodostuu sanojen jatkuva muuttuminen, ja edellisiä hakulausekkeitä pitäisikin käyttää esimerkkeinä, joista kone voi oppia. Muita konkreettisia esimerkkejä koneoppimisen soveltamisesta reaali maailman on mm. käsikirjoitettujen merkkien tunnistaminen. Tätä voidaan käyttää hyväksi esimerkiksi postin lajittelussa ja vanhojen kirjojen digitoinnissa. Autonomiset ajoneuvot ovat myös yksi meidän kaikkien tuntema koneoppimisen sovelluskohde. (Roos 2016:31-44).

Muita koneoppimisen sovellusalueita on mm. suoritusjärjestelmät, konekääntäminen ja ns.”kavereiden ehdottaminen”. Suositusjärjestelmien toimintaperiaatteena on se, että mikäli asiakas ostaa kirjan, niin järjestelmä tämän pohjalta olettaa kirjoittajan olevan kiinnostunut myös tästä toisesta kirjasta. Se toimii samalla periaatteella myös silloin kun käyttäjät arvioivat elokuvia. Tällöin järjestelmä ehdottaa sen pohjalta käyttäjälle uusia elokuvia katseltavaksi. Konekääntäminen toimii perinteisellä tavalla. Se perustuu sanakirjaan ja kielioppiin, kuitenkin nykyään kaikki perustuu yhä enemmän ja enemmän tilastolliseen konekääntämiseen, ja se puolestaan perustuu esimerkkeihin ja dataan. Ns. kavereiden ehdottamisessa pyritään koneoppimisen avulla ehdottamaan kavereita, eli voiko Facebook kaverigraafin perusteella arvata, tuntevatko kaksi henkilöä toisensa vaiko ei, ja osaako Twitter arvata ketä olisit kiinnostunut seuraamaan?(Roos 2016:31-44).

Koneoppiminen on tekoälyn osa-alue. Sillä on varsin monia määritelmiä. Fagella (2018) on kirjoittamassaan artikkelissaan määritellyt koneoppimisen tieteenksi, jossa tietokoneita opetetaan toimimaan, kuten ihmiset ja parantamaan heidän oppimistaan ajan kuluessa automaattisella itseohjautuvalla tavalla. Se tapahtuu syöttämällä tietokoneisiin dataa ja informaatiota havaintojen muodossa ja reaali-maailman interaktioilla. Tämä määritelmä on Fagellan (2018) mukaan sellainen, jolla koneoppimisen ideaali, objektiivinen tai äärimmäinen päämäärä kiteytetään. Tämän määritelmän ovat hyväksyneet ja ilmaisseet myös monet muut tämän kyseisen alan tutkijat. (Fagella 2018).

Pyle & San Jose (2015) puolestaan ovat sitä mieltä, että koneoppiminen perustuu algoritmeihin, jotka oppivat datasta. Tämä tapahtuu ilman, että ne saisivat toimintansa tueksi mitään viestiä sääntöihin pohjautuvalla ohjelmalla. Siitä tuli oma tieteellinen oppiaine jo myöhään 1990-luvulla, kun pienet kehitysaskleet digitalisaatiossa ja halvassa laskennallisessa tehossa mahdollistivat data tieteilijöille mahdollisuuden lopettaa valmiiden mallien rakentamisen. (Daniel & Pyle 2015).

Vähäkainu ja Neittaanmäki (2017) ovat raportissaan tekoäly terveydenhoidossa sitä mieltä, että koneoppimisella on yhteisiä piirteitä tilastotieteen kanssa. Näissä molemmissa tehdään päätelmiä aineistoihin perustuen, mutta koneoppimisessa on kuitenkin tarkasteltavana kohteena ohjelmallisten toteutusten laskennallinen vaativuus. Vähäkainun ja Neittaanmäen (2017) mukaan useat eri päättelyongelmat ovat ns. NP-kovia tai jopa vaikeampia, koneoppimisen tutkimiseen kuuluu siis lisäksi likimääräisten päättelyalgoritmien kehitystyö. Koneoppimisen algoritmit luokitellaan niille annettavaan opetusdatan perusteella. Yleisimpiä algoritmien tyyppejä ovat:

1. Ohjaamaton oppiminen (opetusdatasta ei tiedetä mitään aiemmin)
2. Ohjattu oppiminen (opetusdatasta tiedetään haluttu ulostulo)
3. Vahvistusoppiminen (oppiminen tapahtuu mallin ja ympäristön jatkuvan vuorovaikutuksen seurauksena)

(Vähäkainu ja Neittaanmäki 2017).

Vähäkainu ja Neittaanmäki(2017) toteavat, että koneoppimisen ehkä tunnetuin ohjatun oppimisen malli on päätöspuut(decision trees). Ne ovat yksinkertaisia binääripuita, ja niiden avulla järjestelmä kykenee tekemään päätöksiä. Yksittäiset puut eivät pelkästään

ole ns. oppivia järjestelmiä, koska niillä on staattinen luonne. Useiden puiden yhteiskäytöllä ja uusien puiden luomisen avulla on kuitenkin mahdollista saavuttaa oppiva järjestelmä. Päätöspuiden toimivuus yksinkertaisissa järjestelmissä on yleensä havaittu hyväksi. Isommissa järjestelmissä ne eivät kuitenkaan enää ole paras mahdollinen menetelmä. (Vähäkainu & Neittaanmäki 2017).

Vähäkainu ja Neittaanmäki (2017) mukaan koneoppimisessa on mahdollista käyttää myös graafiteorian ja todennäköisyyslaskentaan perustuvia malleja. Esimerkkeinä näistä on ns. Bayers-verkot. Bayer-verkot ovat varsin yksinkertaisia, suunnattomia syklittömiä verkkoja. Niiden toimintaperiaate on sellainen, että mikäli henkilöllä on esimerkiksi koomaan johtava päänsärky, se olettaa hänellä olevan todennäköisesti myös aivokasvain. Bayers-verkkoja käytetään koneoppimisen yhteydessä aina silloin, kun halutaan kerätä tietoja tuntemattomasta systeemistä. Tällöin on mahdollista aloittaa pienellä verkolla ja lähteä vähitellen laajentamaan verkkoa. Koneoppimista käytetään nykyisin useilla erilaisilla ammatillisilla sektoreilla. SAS on ilmoittanut näiksi mm. finanssipalvelut, hallitusten toiminnot, terveydenhuolto, markkinointi ja myynti, öljy- ja kaasuteollisuus, kuljetusala ym. Tunnetuin koneoppimisen käyttötavoista nykypäivänä on kuitenkin ehkä mallien(hahmojen)tunnistaminen(pattern recognition), sen avulla on mahdollista havaita useita eri tyyppisiä kuvia. (Vähäkainu & Neittaanmäki 2017).

Vähäkainu ja Neittaanmäki (2017) määrittelevät kirjoittamassaan raportissa tekoäly terveydenhuollossa koneoppimisen tekoälyn osa-alueeksi ja data-analyysimetodiksi, jonka tarkoitus on automatisoida analyttistä mallin rakentamista. Koneoppimisessa onkin keskeistä iteratiivisesti dataa käyttämällä luoda oppivia algoritmeja. Tällöin koneoppiminen antaa tietokoneelle erinomaisen mahdollisuuden löytää piileviä oivalluksia ja ideoita. Näiden kohdetta ei ole välttämättä oltu edes alun perin algoritmeihin ohjelmoitu, ja on myös huomioitava se tosiasia, että koneoppimisessa ei ole välttämättä aina kirjoitettuna valmista algoritmia kaikkia tilanteita varten. Siinä kone voi oppia täysin itsenäisesti ja päättää haluttuun lopputulokseen. Toiminta on näin ollen lähes samankaltaista kuin hakukoneilla. Ne pyrkivät siis tarjoamaan niin osuvia- ja oikeita hakutuloksia omille käyttäjilleen kuin vain on mahdollista. Lisättäessä tietoa tietokantaan tapahtuu aina koneen oppimiskyvyn kehittymistä. (Vähäkainu & Neittaanmäki 2017).

Koneoppimisen yksi osa-alue on syväoppiminen (Deep Learning), joka on vielä varsin uusi asia. Syväoppimisen katsotaan alkaneen vuodesta 2006, ja se on noussut yhä enemmän esille vuoden 2012 jälkeen. Tällä koneoppimisen osa-alueella sovelletaan Tjoan (2013) mukaan hierarkkisia arkkitehtuureita ja epälineaarisia informaation prosessoinnin tasoja. Syväoppimisessa pyritään luomaan sopivalla algoritmilla ns. neuroverkko, joka tavoittelee sille soveltuvan ongelman ratkaisemista. Syväoppimista käytetään yleensä sellaisten ongelmien ratkaisuun, joiden ratkaisua on vaikea toteuttaa perinteisillä menetelmillä, koska ne vaativat monimutkaisten sääntöjen käyttöä. Syväoppimista hyödynnetään paljon nykyisin sellaisilla alueilla, kuten lääketieteen diagnostiikka, kuva, puhe, tekstien käsittely ja tunnistaminen. Kaikkien tunnetuimpia hyödyntämisen alueita ovat monet arkipäivässä käytetyt sovellukset, kuten Applen Siri, Googlen Street View-karttapalvelu ja puheentunnistuksen palvelu. (Tjoa 2013.)

Syväoppiminen on kokoelma tilastollisia koneoppimisen teknologioita, niitä käytetään oppimaan toimintahierarkkioita, ja ne perustuvat keinotekoisiiin neuroverkkoihin. Syväoppimisen algoritmien suorituskykyä on mahdollista parantaa mm. lisäämällä dataa, tuottamalla/generoimalla lisää dataa, dataa uudelleen skaalaamalla ja dataa muuntamalla. (Akasin 2014.)

2.1 Koneoppimismallit ja datan luokittelu

Koneoppimiseen liittyy paljon käsitteistöä, jota on syytä selventää sekaannuksen välttämiseksi. Koneoppimismallilla tarkoitetaan matemaattista kuvausta reaalimaailman prosessista. Koneoppimismallin muodostamiseksi täytyy hankkia opetusdataa, josta koneoppimisalgoritmi voi oppia. Koneoppimisalgoritmi puolestaan on hypoteesijoukko, joka on omaksuttu alussa ennen opettamisen aloittamista reaalimaailman datan avulla. Esim. lineaarinen regressio algoritmi tarkoittaa joukkoa tehtäväkokonaisuuksia, jotka määrittelevät vastaavia lineaarisella regressiolla rajattuja kuvaajia, ja noista toimintojoukoista voidaan valita toiminto, joka sovitetaan parhaiten opetus datalla. Opettaminen koneoppimista varten tarkoittaa, että algoritmi läpikäydään opetusdatalla. Opettelu algoritmi löytää tällöin polkuja opetusdatalla, ja ne ovat vastaavia

vastaavanlaisten kohteiden syöttömuuttujien kanssa. Opettamisprosessien tuloste taas on koneoppimismalli, jota voidaan käyttää ennusteiden tekemiseen. Tätä prosessia kutsutaan myös oppimiseksi. Regressiotekniikkaa käytetään tulosteiden ollessa todellisia arvoja perustuen jatkuviin muuttujiin esim. milloin tahansa data sarjassa, ja tähän tekniikkaan sisältyy mitan selvittäminen. Luokittelussa taas tietoaaineisto jaetaan ennalta määrättyihin luokkiin. Tällöin esim. sähköposti voi olla roskapostia tai ei roskapostia. Kohde puolestaan on mikä tahansa syöttömuuttujan tuloste. Piirteet taas ovat yksilöllisiä, itsenäisiä muuttujia, jotka toimivat järjestelmän syötteenä. Ennustemallit käyttävät piirteitä tehdäkseen ennusteita. Uudet piirteet voidaan myös saavuttaa vanhoista piirteistä käyttäen metodia, joka tunnetaan piirre koneoppina. Yksi kolumni datajoukossa voi olla esim. yksi piirre, ja niitä kutsutaan myös attribuuteiksi, ja numeroiden piirteitä kutsutaan puolestaan ulottuvuuksiksi. (Bhattacharjee 2017.)

Koneoppimisessa käytettävä tietoaaineisto voidaan jakaa kolmeen osaan eli opetus-, testaus- ja validointiosaan. Validointiosaa käytetään mallin ylisovittamisen välttämisen ehkäisyyn ja optimaalisten hyperparametrien etsimiseen. Hyperparametrit ovat koneoppimisalgoritmien parametrejä, joiden avulla voidaan vaikuttaa siihen, kuinka kompleksinen malli on. Hyperparametrit eivät kuitenkaan ole koneoppimisen malleja vaan algoritmien parametreja. Eri koneoppimisalgoritmeilla on erilaisia hyperparametreja ja monissa algoritmeissa saattaa olla useitakin muutettavissa olevia hyperparametreja. Päätöspuiden hyperparametreilla on mahdollista rajoittaa päätöspuiden maksimietäisyyttä tai lehtisolmuissa olevien havaintojen määrää. Hyvän koneoppimismallin luomiseksi riittää yleensä muutaman hyperparametrien muuttaminen. Ne löydetään yleensä kokeilemalla useita eri arvoja valituille hyperparametreille, ja laskemalla eri parametrien arvoilla luotujen mallien ennustevirheet validointi tai ristiin validointi aineistoa käyttäen. Parhaimmat hyperparametrien arvot johtavat pienempään ennustevirheeseen. Etsittäessä hyperparametrien arvoja on syytä käyttää aina validointi aineistoa tai ristiinvalidointi aineistoa ennustevirheiden laskemiseen, eikä testausaineistoa. Testausaineistoa käytetäänkin vain testaamiseen, ja lopullisen mallin luomiseen, hyperparametrien arvojen valinnan jälkeen. Tällöin mallin ennustevirheestä saadaan tarkka arvo. (Athey 2018.)

Koneoppimismallin ja koneoppimisalgoritmin varsinaiset erot ovat siinä, että lineaarinen regressio algoritmi on tekniikka, jossa sovitaan pisteitä linjalle, kuten esimerkiksi $Y = M X + C$. Sovittamisen jälkeen saadaan tuloste $Y = 10x + 4$. Tämä on koneoppimismalli. Malli on sellainen, jossa annetaan syöttö, ja joka antaa tulosteen. Koneoppimisessa opettamisen jälkeen mikä tahansa koneoppimisalgoritmilla luotu objekti on malli. Tällaisia ovat esimerkiksi tukivektorikone malli (SVM model) ja Satunnaisen metsän malli (Random Forest malli). (Ravindran 2016.)

Koneoppiminen tarkoittaa siis sellaisen mallin luomista, joka ennustaa hyvin uusia havaintoja. Toisin sanoen havaintoja, joita ei ole käytetty mallin sovittamiseen tai opettamiseen. Koneoppimisessa mallin tehokkuutta voidaan arvioida ainoastaan sen ennustamistarkkuudella. Mallin ennustamistarkkuudella pystytään saamaan arvio käyttämällä mallia uusien havaintojen ennustamiseen. Ennustamistarkkuuden testaamiseksi koneoppimisprojektissa aineisto jaetaan opetus- ja testausosaan. Opetusosaa käytetään mallin opettamiseen ja testausosaa puolestaan mallin testaamiseen. Tällöin voidaan saavuttaa realistinen arvio mallin ennustetarkkuudesta, toisin kuin jos mallin ennustetarkkuutta testattaisiin vain aineistolla, jota on käytetty mallin opettamiseen, jolloin saataisiin turhan optimistinen arvio. Koneoppimismenetelmät ovat joustavia ja niiden avulla on mahdollista luoda epälineaarisia, kompleksisia malleja. Näin ollen koneoppimismenetelmillä voidaan sovittaa vaivattomasti malleja, jotka ennustavat opettamiseen käytettyjä havaintoja hyvin, mutta niillä on huono ennustetarkkuus uusia havaintoja ennustettaessa. Tästä seuraa ylisovittamisen ongelma. (Athey 2018.)

Ylisovittaminen (overfitting) on käsite, joka tarkoittaa toisin sanoen koneoppimisalgoritmin luoman mallin kykyä ennustaa opetusaineistoon kuuluvat havainnot liian hyvin. Malli pystytään sovittamaan niinkin hyvin, että se ennustaa koko opetusaineiston havainnot täydellisesti. Kyseessä ei ole kuitenkaan hyvä ratkaisu, koska aineistossa esiintyy aina ajonkin verran satunnaisvaihtelua. Mikäli mallista luodaan liian kompleksinen, se alkaa ennustamaan myös tätä aineiston havaintojen satunnaisvaihtelua. Tällöin malli ei voi yleistää kovin hyvin, ja käytettäessä mallia ennustamiseen uudella aineistolla mallin ennusteet ovat varsin huonoja. (James, Witten, Hastie & Tibshirani 2013: 30-36.)

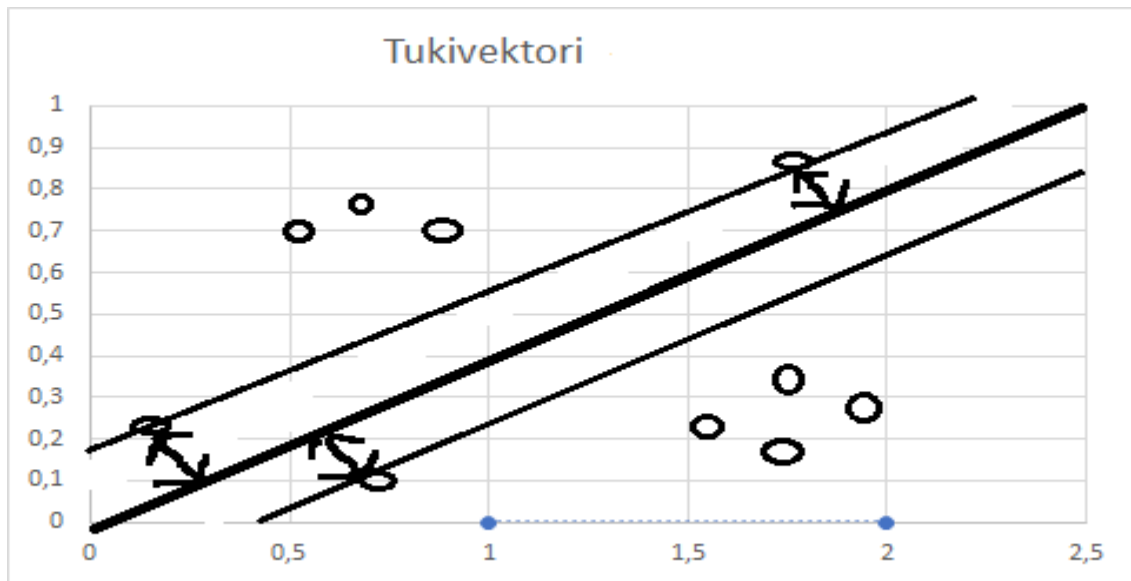
Koneoppimisen avulla voidaan käsitellä suuria tietoaaineistoja, ja tätä aineistoa voidaan käsitellä ja luokitella monella tavalla eri käyttötarkoituksia varten. Käyn tässä läpi muutamia aineiston luokittelutapoja. K:n lähimmän naapurin menetelmää pidetään tehokkaana ja yksinkertaisena luokittelumenetelmänä. K:n lähimmän naapurin luokittelija kykenee ennustamaan uuden havainnon tai instanssin luokan laskemalla kaikkien opetusaineiston näytteiden etäisyyden uusimpaan havaintoon. Tämän jälkeen valitaan k kappaletta näytteitä, joiden etäisyys on suhteessa uuteen instanssiin kaikkein pienin. Uuden instanssin luokka määräytyy aina k:n lähimmän naapurin yleisimmän luokan mukaan. Etäisyys uuteen havaintoon voidaan mitata monella tavalla, yleisimmin käytetään mittana Euklideen etäisyyttä. Kaksiulotteisessa avaruudessa kahden pisteen euklidisen etäisyyden laskemiseen voidaan käyttää Pythagoran lausetta eli $a^2 + b^2 = c^2$. (Wu, Kumar, Quinlan, Ghosh, Yang, Motoda, & Zhou 2008.)

K:n lähimmän naapurin luokittelijaa pidetään niin sanottuna laiskana oppijana. Laiskojen oppijoiden ja ahkerien oppijoiden erot perustuvat siihen, että laiskojen malleja ei rakenneta ekspliittisesti opetusvaiheessa. Ahkeria oppijoita ovat mm. päätöspuut ja tukivektorikoneet. Laiskojen oppijoiden ns. opetusvaihe koostuu käytännössä aineiston tallentamisesta muistiin. Laskenta tapahtuukin vasta uusia instansseja luokiteltaessa. (Wu ym.2008.)

K:n lähimmän naapurin luokittelijaa muodostettaessa, tärkeimmäksi valinnaksi muodostuu sopivan k:n arvon valinta. Mikäli k:n arvo on liian pieni saattaa aiheutua ylisovitus, ja toisaalta sen ollessa liian suuri voi aiheutua alisovitus. Näin ollen k:n arvoksi kannattaa valita luku, jolla vältetään tasatulokseen päätyminen. K:n arvoksi kannattakin valita aina pariton luku, esimerkiksi binäärisen kohdemuuttujan tapauksessa. Wu ym. (2008) huomauttavat myös, että muuttujat kannattaa normalisoida. Se on tarpeen erityisesti muuttujien arvoalueiden vaihteluiden ollessa suuria. Muuttujan arvoalueiden ollessa merkittävästi laajempi kuin muiden muuttujien arvoalue, tämänkaltainen muuttuja kykenee dominoimaan luokitusta. (Wu ym. 2008.)

Tukivektorikoneetta pidetään yhtenä suosituimmista ja parhaimmista luokittelumetodeista, niiden tiedetään olevan erinomaisia ”out of the box”-luokittelijoista. Tukivektorikoneen toimintaperiaate perustuu yksinkertaiseen ja

intuitiiviseen luokittelijaan. Tätä kutsutaan maksimaalisen marginaalin luokittelijaksi (Maximal margin classifier). Maksimaalisen marginaalin luokittelija asettaa aineiston näytteiden väliin päätöstason, se puolestaan erottelee näytteet niiden vastemuuttujan arvon mukaan. Tällöin vastemuuttujan ollessa binäärinen muuttuja, tulisi näytteet erotella aina siten, että päätöstason toisella puolella olevien näytteiden vastemuuttujan arvo on 1 ja toisella puolella olevien näytteiden vasteen arvo on 0. Tällainen päätöstason asettaminen edellyttää, että aineisto on lineaarisesti eroteltavissa. Tällöin eri luokkien yksikäsitteinen erottaminen tasolla pitää olla mahdollista. Mikäli aineisto pystytään lineaarisesti erottelemaan, niin silloin on olemassa loputon määrä tasoja, jotka mahdollistavat aineiston erottelemisen. Maksimaalisen marginaalin luokittelijan periaate perustuu siihen, että näytteiden välille sovitetaan päätöstaso ja näytteiden väliin jäävien marginaalitasojen välimatka tulee mahdollisimman suureksi. (James, Witten, Hastie & Tibshirani 2013: 337-354.)



Kuva 5. Tukivektorikone malli. (James, Witten, Hastie & Tibshiraniin mukaan 2013: 342).

James ym.(2013) kirjoittavat myös päätöspuista. Ne ovat koneoppimisalgoritmeja, joita on mahdollista käyttää sekä regressio- että luokittelutehtäviin. Päätöspuut ovat osoittautuneet hyvin tehokkaiksi ja ne mahdollistavat mallien sovittamisen monimutkaisimpiinkin aineistoihin. Päätöspuut ovat toimintaperiaatteiltaan varsin yksinkertaisia. Ne jakavat aineiston useita kertoja kahteen osaan. Jakamisesta syntyneitä aineiston osia jaetaan toistamiseen siihen asti, kunnes aineiston puhtaita havaintoja on mahdotonta enää parantaa, tai vaihtoehtoisesti saavutetaan jokin etukäteen määritelty pysäytyskriteeri. Aineiston jakaminen suoritetaan aina selittäviä muuttujia käyttämällä. Varsinainen päätöspuu muodostuu solmuista, oksista ja lehdistä. Juurisolmu on puun ensimmäinen solmu, ja siinä on kaikki aineiston havainnot. Jaettaessa aineistoa syntyy oksia, ja niiden tehtävänä on yhdistää juurisolmut, sisäsolmut ja lehtisolmut, näihin on jaettu aineiston osat. Aineisto jaetaan päätöspuun jokaisessa solmussa jonkin muuttujan perusteella. Tätä tehdään siihen asti kunnes aineiston puhtautta ei enää voida parantaa. Jakokriteeri määrittää sen, minkä muuttujan perusteella aineisto jaetaan solmussa. Jakokriteereitä ovat luokitteluasteikon muuttujille giniepäpuhtaus ja Entropia. Regressiopuiden jakokriteereinä puolestaan käytetään yleensä keskimääräistä ennustevirheen neliötä. (James, Witten, Hastie & Tibshirani, 2013: 303-315.)

Hastle, Tibshirani & Friedman (2009) mainitsevat ns. Bagging menetelmästä. Bagging on englanninkielinen termi, joka tulee sanoista bootstrap aggregating. Bootstrap menetelmä tarkoittaa satunnaisotoksen ottamista takaisinpanolla. Suomennettuna termi tarkoittaa sitä, että opetusaineistosta otetaan satunnaisesti näytteitä, ja samoja näytteitä on mahdollista ottaa useampaan kertaan. Näytteitä otetaan yleensä yhtä monta kuin alkuperäisessä aineistossa on havaintoja. Näihin opetusaineistosta otettuihin otoksiin on mahdollista sovittaa koneoppimismalli, kuten esim. päätöspuu. Mallin ennustearvot voidaan saada näiden useiden eri mallien ennusteiden keskiarvosta. On huomioitava se, että luokitteluasteen muuttujaa ennustettaessa ennusteen luokka on aina se, mikä saa eniten ääniä eri luokittelijoilta. Hastle ym.(2009) painottavat, että bagging menetelmän tehokkuus perustuu sen kykyyn vähentää varianssia, ja toisaalta se ei myöskään lisää harhaa. Bagging malli tuottaakin tämän takia parempia ennusteita kuin yksittäinen malli. Se toimii hyvin silloin, kun perusoppijana on epävakaa malli, jolla on suuri varianssi ja pieni harha. (Hastle, Tibshirani & Friedman 2009: 282- 289.)

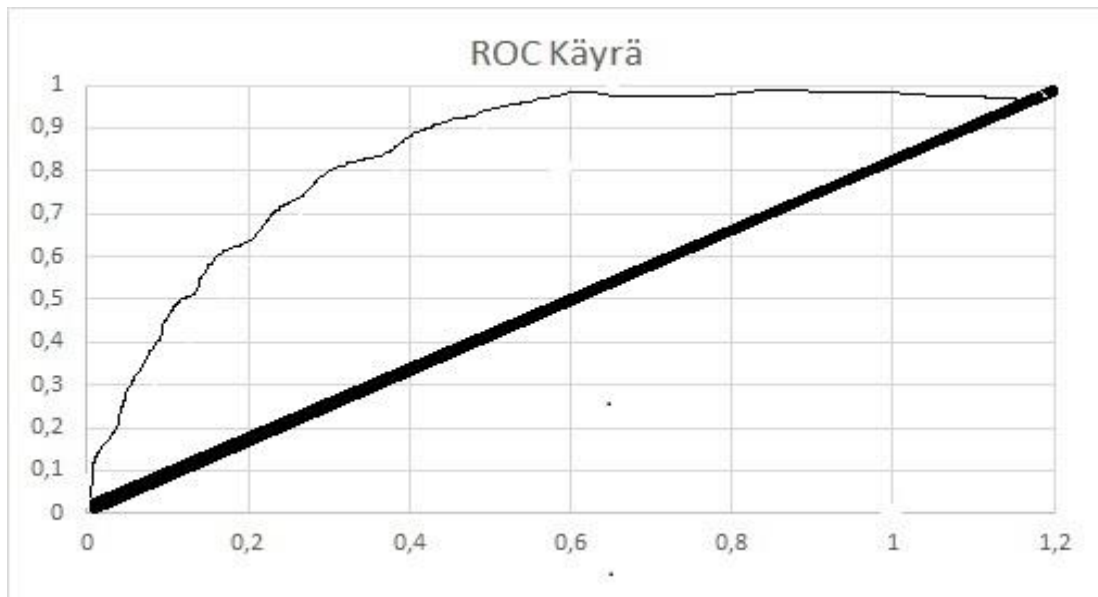
Satunnainen metsä(Random Forest, RF) on bagging-menetelmään perustuva malli, jonka on kehittänyt Breiman (2009). Sen tarkoitus on kasvattaa suuri määrä päätöspuita, niiden välillä ei kuitenkaan saa olla korrelaatiota. Näistä puista otetaan sitten niiden keskiarvo. Päätöspuut soveltuvat hyvin bagging-menetelmään, koska bagging toimii hyvin korkean varianssin ja vähäisen harhan malleille. Päätöspuut kykenevät mallintamaan aineiston monimutkaiset funktiot ja vuorovaikutukset. Mikäli päätöspuista kasvatetaan tarpeeksi suuria, niiden harha jää kohtalaisen vähäiseksi. Satunnaisen metsän ja bagging menetelmän välinen ero syntyy kasvattaessa päätöspuita satunnaisessa metsässä, tällöin syötemuuttujista valitaan satunnaisesti osa, jota käytetään. Satunnaisessa metsässä valitaan satunnaisesti m kappaletta muuttujia, joiden perusteella jako suoritetaan, toisin kuin päätöspuun kasvatuksessa, jossa aineisto jaetaan jokaisessa solmussa sen muuttujan perusteella, mikä parhaiten lisää aineiston puhtautta. Mikäli p on kaikkien syötemuuttujien määrä, niin tällöin $m:n$ arvo on $\sqrt{0}$ tai 1. Tämän seurauksena kasvatettujen päätöspuiden välinen korrelaatio vähenee. Satunnainen metsä on hyvä ja suosittu algoritmi. Sen etuna on se, että ns. hyperparametreja ei tarvitse virittää hyvien ennusteiden saamiseksi. (Hastle, Tibshirane & Friedman 2009: 587-595.)

Ristiinvalidointi on menetelmä, jota käytetään mallin ennustevirheen arviointiin tai mallin parametrien valintaan. Ristiinvalidointi tehdään yleensä jakamalla aineisto useaan yhtä suureen osajoukkoon, jonka jälkeen malli sovitetaan ja testataan niin monta kertaa kuin aineistossa on osajoukkoja. 10-kertainen ristiinvalidointi suoritetaan esim. siten, että aineisto jaetaan ensin 10 osaan. Sitten malli sovitettaisiin kaikkia muita osajoukkoja (lukuun ottamatta ensimmäistä) opetukseen. Ensimmäistä osa joukkoa käytetään ennustevirheen laskemiseen. Seuraavassa vaiheessa samaa toistetaan sillä tavoin, että toisena olevaa osajoukkoa käytetään ennustevirheen laskemiseen. Tämä toistuu siihen asti kunnes kaikki kymmenen osajoukkoa on käytetty ennustevirheen laskemiseen. Ennustevirheen lopullinen arvio saadaan näiden kaikkien kymmenen ennustevirheen keskiarvosta. Ristiin validoinnin avulla laskettua ennustevirhettä voidaan pitää luotettavana, koska ennustevirhettä laskettaessa käytetään aina sellaista aineistoa, jota ei ole käytetty mallin sovittamisessa. Ristiin validointi on myös hyvin lähellä todellista ennustevirhettä. Ristiin validointia pidetään hyvänä työkaluna havaita mahdollinen ylisovittaminen, ja se on myös hyvin käytännöllinen aineiston ollessa niin pieni, että sen järjestäminen erillisiin opetus- ja validointiosiin on mahdotonta. (Hastle, Tibsharani & Friedman 2007:241-247.)

Kaikilla koneoppimismalleilla on myös ns. ennustamistarkkuus. Sen arvioimiseksi on kehitelty muutamia tapoja. Osumatarkkuus (accuracy) on yleinen luokittelijan ennustekyvyn arvioinnin tapa. Siinä mitataan, kuinka monta havaintoa luokittelija pystyy ennustamaan oikeaan luokkaan kaikista ennustettavista instansseista. Mikäli ennustettavan muuttujan luokat ovat epätasapainossa, niin voi seurata kuitenkin ongelmia. Mikäli esim. binäärisen muuttujan tapauksessa toiseen luokkaan kuuluu 90 % havainnoista niin tällöin tähän luokkaan kuuluu kaikki asteet (instanssit) ja luokittelevan luokittelijan osumatarkkuus on 90 %. ROC(Receiver operator characteristics) analyysi on otettukin tämän takia käyttöön, että saataisiin eri luokittelijoiden suorituskyvystä luotettavampi arvio. (Jauhiainen 2019: 33).

Oleellisimmat tunnusluvut, joilla tarkkuutta pystytään arvioimaan ovat herkkyys(sensitivity) ja specifisyys (specificity). Herkkyys tarkoittaa oikeiden positiivisten ennusteiden osuutta ja specifisyys puolestaan tarkoittaa oikeiden negatiivisten ennusteiden osuutta. Herkkyys ja specifisyys on määriteltävä eri kynnyksarvoilla, jotta ROC-käyrä voitaisiin muodostaa. (Zoe, Malley & Mauri 2007).

Kynnyksarvolla tarkoitetaan arvoa, joka kykenee määrittämään sen mihin luokkaan instanssit on luokiteltava arviotujen todennäköisyyksien perusteella. Positiiviseen luokkaan luokitellaan esim. kuuluvaksi kaikki instanssit, joille luokittelija ennustaa 0,5 ylittävän todennäköisyyden. Tämä tapahtuu kynnyksarvolla 0,5. Silloin kun herkkyys ja specifisyys on määritelty eri kynnyksarvoilla, niillä pystytään muodostamaan kuvaajia, joissa vaaka-akselilla on specifisyys eli väärin positiivisten osuus ja pystyakselilla herkkyys eli oikeiden positiivisten osuus. ROC-käyrä on havainnollistettu kuvioissa 4. ROC käyrän alle jäävää pinta-alaa voidaan käyttääkin tunnuslukuna, joka havainnollistaa luokittelijan kykyä erotella eri luokkiin kuuluvat havainnot. ROC käyrä kulkee kuvaajan akseleita pitkin. Luokittelija kulkee tässä ROC-käyrän alapuolisena viivana. Se on 45 asteen suora, joka on sama kuin satunnainen arvaus. ROC-käyrän alle jäävästä pinta-alasta käytetään lyhennettä AUC (area under the curve). Yleensä luokittelijoiden AUC-arvo ovat 0,5 ja 1 välillä. Mitä korkeammasta AUC arvosta on kyse, sitä paremmin luokittelija kykenee erottamaan havainnot eri luokkiin.(Zou, Malley & Mauri, 2007.)



Kuva 6. ROC-Käyrä. (Zou, Malley & Mauri 2007).

On hyvä myös mainita tässä yhteydessä kahdesta usein koneoppimisen käsittelyn yhteydessä esiintyvien käsitteiden merkitysten olennaisimmista eroista. Luokittelu ja regressio ovat molemmat ohjattua oppimista, ja niiden välillä on kuitenkin pieni tärkeä ero. Luokittelumenetelmät esim. kNN valitsevat aina vastauksen tietystä rajoitetusta joukosta vaihtoehtoja (esim. roskapostia\asiallista postia tai 0,1..9). Regressiomenetelmät, kuten esim. lineaariregressio tuottavat vastaukseksi lukuarvon. Tämän ei tarvitse välttämättä olla kokonaisluku. Lineaariregression kaltainen menetelmä on parempi tilanteissa, joissa vastauksena on etäisyys kohteeseen tai tuotteen hinta. (Elements of AI :3 Regressio.)

Luokittelu on tehtävä, jossa pyritään selventämään, mihin luokkaan asia tai olio ominaisuuksiensa puolesta kuuluu. Kasvokuva voidaan esim. ilmeensä perusteella luokitella neutraaleihin, surullisiin, vihaisiin ja yllättyneisiin luokittelussa kannattaa hyödyntää koneoppimista silloin, kun luokiteltavia asioita on enemmän kuin ihminen kykenee laskemaan. Luokittelussa on arvioitava aina mallin luotettavuus ja tarpeen tullen tarkentaa sitä. Ensimmäinen olennainen mittari tähän on ulkoinen tarkkuus (Accuracy), joka kertoo millä prosentilla malli luokittelee oikein. Reaalimaailman epätasaisesti jakautuneessa datajoukossa tarkkuus soveltuu huonosti binääriluokittelumallin (kahden luokan) minimointiin. Binääriluokittelussa datapisteet voi jakaa neljään ryhmän: oikeat

positiiviset eli ne positiiviset, jotka malli onnistui luokittelemaan positiiviksi, väärät positiiviset eli ne negatiiviset, jotka malli onnistui luokittelemaan positiivisiksi, oikeat negatiiviset eli ne negatiiviset, jotka malli onnistui luokittelemaan negatiiviseksi ja väärät negatiiviset eli ne negatiiviset, jotka malli onnistui luokittelemaan negatiiviseksi. (Schroderus 2017.)

Ulkoisen tarkkuuden sijasta binääriluokittelussa mallin luotettavuudesta kertoo paremmin sisäinen tarkkuus(precision) tai herkkyys(sensitivity, true positive rate). Sisäinen tarkkuus = oikeat positiiviset/ oikeat positiiviset + väärät positiiviset. Sisäisellä tarkkuudella ilmaistaan siis, kuinka paljon kaikista positiivisiksi luokitteluista datapisteistä oli oikeasti positiivisia. Binääriluokittelussa sisäisen tarkkuuden tulisi olla mahdollisimman suuri. Herkkyys puolestaan on oikeasti positiiviset/oikeiden positiivisten ja väärin negatiivisten summalla. Herkkyys kuvaa mallin kykyä luokitella positiivisiksi kaikki oikeasti positiiviset tapaukset. Esim. vaarallisten tarttuvien tautien diagnooseissa on tärkeää, että väärin negatiivisten osuus on mahdollisimman pieni, joten mallille halutaan aina suuri herkkyys. (Schroderus 2017.)

Lineaariregressiota voisi verrata kauppalaskuun eli kauppaan mentäessä ostetaan osia eli 2,5 kiloa perunoita ja 1 kilo porkkanoita ja kaksi purkkia maitoa. Perunoiden kilohinta on 2 euroa, porkkanat maksavat 4 euroa ja maitopurkki maksaa 3 euroa, tällöin kassa laskee kokonaishinnaksi $2,5 \times 2 \text{ €} + 1 \times 4 \text{ €} + 2 \times 3 \text{ €} = 15 \text{ €}$. Lineaariregressiossa perunoiden, porkkanoiden ja maidon määrä ovat datasta saatavat syötteet. Tuloste on tällöin kokonaishinta, joka on luonnollisesti riippuvainen kunkin tuotteen ostomäärästä ja hinnasta. Näin ollen sana lineaarinen tarkoittaa sitä, että kun esim. yhtä syötteistä kasvatetaan tietyllä määrällä, tulosteena saatava arvo kasvaa myös aina saman verran. Tämä tarkoittaa käytännössä sitä, että lisättäessä ostoskoriin kaksi kiloa porkkanoita, lasku lisääntyy 8 €. (Elements of AI :3 Regressio.)

Kuitenkin, kun lisätään vain kilo laskuun kasvaa vain 4 €. Lineaariregression termein tuotteiden hintoja kutsutaan kertoimiksi, painoiksi tai painokerroimiksi. Lineaariregression hyvä puoli on sen helppossa tulkittavuudessa, silloin mallin kertoimet voivat olla se kaikkein kiinnostavin asia. Lineaariregressiota on mahdollista käyttää vaikka elinajan odotteen ennustamiseen, esim. tupakoinnin(savukkeita päivässä) kerroin on noin miinus puoli vuotta eli -0,5 vuotta. Käytännössä siis yhden savukkeen lisäys keskimääräiseen kulutukseen vie tupakoitsijan puoli vuotta lähemmäs kuolemaa. Vihannesten kulutus puolestaan (kourallinen vihanneksia päivässä) on painokerroin plus yksi vuosi. Tämä tarkoittaa, että esimerkiksi syömällä kourallinen vihanneksia päivässä saadaan keskimäärin yhden vuoden lisäys elinikään. (Elements of AI :3 Regressio.)

Koneoppimistekniikoiden yksi suurimmista hyödyistä on siinä, että ne mahdollistavat nopean ja tehokkaan tietoaineiston luokittelun silloin, kun sen havainnot eivät ole toisistaan riippuvaisia. Bayesin luokitin on yksi koneoppimistekniikka, joka mahdollistaa asioiden luokittelun. Luokiteltavia asioita voivat olla tekstidokumentit, kahteen tai useampaan eri luokkaan. Luokitin pystytään opettamaan analysoimaan opetusdataa, jossa on merkittynä oikeat vastaukset. (Elements of AI: 4 Naiivi Bayes luokitin).

Naiivi Bayesin luokitin voi laskea eri luokkien todennäköisyydet, kun luokitin päätöksen tueksi on käytettävissä monia eri havaintoja. Idean taustalla on malli, jossa havainnot ovat toistaan ehdollisesti riippumattomia. Yksi esim. naiivin Bayesin luokittimesta on roskapostisuodatin. Siinä sähköpostiviestien luokka kertoo, onko kyseessä roskaposti vai asiallinen viesti. Viesti sisältää sanoja, ja ne ovat havaintoja. Viestin sanojen määrää vastaa saatuja havaintojen määrää. Naiivin Bayesin luokitin on naiivi luokitin, ja sen naiivius liittyy oletukseen havaintojen toisistaan ehdollisesta riippumattomuudesta. Täten roskapostisuodatinta tarkasteltaessa, voidaan olettaa viestien sisältämien sanojen olevan toisistaan riippumattomia. Ne riippuvat vain siitä, onko kyseessä roskaposti vai asiallinen viesti. Tällöin sanat on voitu valita yksi kerrallaan tietysti jakaumasta ja esim. sanojen järjestyksellä ei ole merkitystä. Tämä on naiivi oletus. Siitä tulee myös nimitys naiivi Bayesilainen luokitin. Viestin luokka on ainoa sanoihin vaikuttava tekijä. (Elements of AI: 4 Naiivi Bayes luokitin).

2.2 Koneoppiminen terveydenhoidossa.

Tekoälyn hyödyntäminen lääketieteessä (AIM, Artificial intelligence in medicine) voi tuoda huomattavia parannuksia terveydenhuollon laatuun. Siitä tulee yksilökeskeisempää, kustannustehokkaampaa ja se kykenee myös vähentämään terveydenhoidon henkilökunnan kognitiivista kuormitusta. Näin ollen terveydenhoito henkilökunnan työ voi paremmin keskittyä hoitotyön osa-alueisiin, jossa tarvitaan paljon inhimillistä vuorovaikutusta. Idea tekoälyn hyödyntämisestä terveydenhoidossa (Artificial Intelligence in Medicine, AIM:än) voidaan katsoa syntyneen käsitteenä jo 1970-luvulla. Silloin Schwartz(1970) hahmotteli sitä, kuinka tietokone ja lääketiede tulisivat tukevaisuudessa tekemään yhteenliittymän. (Schwartz 1970).

Schwartz (1970) toteaa nopean, edistymisen tietojen käsittelytieteessä sekä poliittisen sitoutumisen laajentavan terveydenhoitoa, ja lupaavan tuoda perustavanlaatuisia muutoksia lääketieteen harjoittamisen käytäntöjen rakenteisiin. Hänen painotti ohjelmistotieteen todennäköisesti lisäävän vaikutustaan, ja jossain tapauksissa jopa korvaavan laajasti lääkärin älyllisiä toimintoja. Hän ennusti sen väistämättömästi myös vaativan tärkeitä psykologisia, organisatorisia, taloudellisia, oikeudellisia ja teknisiä kustannuksia. Tämän tavoitteen saavuttaminen vaatii uusia interaktioita lääketieteen, tietojenkäsittelyopin ja johtamistieteen kesken, ja uusien taitojen kehittämistä ja terveydenhuoltojärjestelmien poliittisen päätöksentekijöiden osalta myös uusia asenteita. (Schwartz 1970).

Vähäkainu ja Neittaanmäki (2017) ovat myös käsitelleet tekoälyä hyödyntäviä ratkaisuja terveydenhuollossa. 13 D Research Tutkimusprojektin tulokset kertovat sen tosiasian, että arviolta 88 %:ssa diagnosointi tilanteissa joudutaan pyytämään arvioita myös toiselta lääketieteen ammattilaiselta. Tämä saattaa olla esimerkiksi lääkäri. Tehdyistä diagnosointivirheistä jopa 10 % johti potilaiden kuolemaan ja peräti 17 % komplikaatioihin. Sairauksia ymmärretään nykyisin paljon paremmin ja testit ovat parantuneet selvästi. Siitä huolimatta diagnosointi on vieläkin epätarkkaa, se on ikään kuin sekä tiedettä ja taidetta ja se myös altistaa aina jopa hengenvaarallisille virheille. Vähäkainu ja Neittaanmäki (2017) painottavat sitä, että teknologiayritykset ja yliopistojen tutkimuslaboratoriot sekä urauurtavien sairaaloiden koneoppimisen algoritmeja hyödyntävät diagnostiikkasovellukset ovat nopeasti tehneet hyppäyksen teoriasta varsinaiseen käytäntöön. Tässä muutoksessa on valtavasti potentiaalia, ja koneoppiminen antaakin välittömästi toisen mielipiteen diagnoosista. Näin ollen se voi toimia erilaisten sairastapausten varhaisen vaiheen tunnistuksen työkaluna. Koneoppiminen voi myös tarjota tähän mennessä historian kaikkein tarkimman diagnosointimenetelmän. Se on myös kustannustehokas lääkärille, potilaille, sairaalalle ja kykenee parhaimmillaan pelastamaan elämiä. (Vähäkainu & Neittaanmäki 2017).

Koneoppimisen eri menetelmiä on jo käytetty Vähäkainun ja Neittaanmäen (2017) mukaan tehokkaasti erityisen paljon etenkin syövän tunnistamisessa. Stanfordin yliopiston tutkijaryhmä opettikin neuroverkoille, miten on mahdollista löytää 150 000 kuvan tietokannasta syöpätapauksia ja tehdä niistä diagnooseja. Kyseinen tutkija ryhmä testasi neuroverkkoja myös 14 000 iholeesioista koostuvan tietokannan kanssa. Koepaloja ei tarvinnut ottaa ollenkaan, koska siinä neuroverkkoja opetettiin koneoppimisella. Neuroverkko oppiikin opettamisen jälkeen tekemään 72 % varmuudella täysin päteviä diagnooseja. Hyvänä vertailukohteena käytettiin tässä tutkimuksessa ihotautilääkäreitä, ja ne kykenivät tekemään diagnoosin vain 66 % tarkkuudella. Kyseistä testiä laajennettiin, ja se kattoi lopulta 25 lääkäriä. Heillä oli käytössään 2000 iholeesikuvaa, ja niissä oli kaikissa tapauksissa otettu koepalat. Tällöin havaittiin, että neuroverkko pystyi päihittämään asiantuntijalääkärit kaikissa tilanteissa. (Vähäkainu & Neittaanmäki 2017).

Vähäkainu & Neittaanmäki (2017) analysoivat tekoälyä terveydenhuollossa (AIM, Artificial Intelligence in Medicine) käsitteenä. He ottavat esille Rodriguezin (2016) tekemän määritelmän AIM:stä, Siinä hän toteaa sen olevan tietojenkäsittelyn hyödyntämistä älykkäänä työkaluna, se voi myös auttaa kliinisessä arvioinnissa ja päätöksenteossa. Tutkijat tuottavat vielä lisää dataa algoritmeille, ja työkaluista tulee tällöin vielä nykyistä tarkempia. Näin ollen niistä tulee olemaan vielä enemmän hyötyä sairaaloille ja lääkäreille. Tekoälyn uskotaan vaikuttaneen useaan alueeseen sairaalamaailmassa. Se ulottuu aina yksilökeskeisestä lääketieteestä operationaalisten ja ennustettavien kustannusten hallinnan paranemiseen asti. Tästä hyötyvät myös laitteistovalmistajat, sillä sairaalaympäristöt vaativat tehokkaita informaatioteknologian laitteistoja ja laskentakykyä. Tällöin tekoälysovellusten käyttö diagnooseissa onnistuu hyvin. (Vähäkainu & Neittaanmäki 2017).

Deo (2015) käsittelee artikkelissaan koneoppimista lääketieteessä. Hän määrittelee koneoppimisen tieteelliseksi oppiaineeksi, joka keskittyy siihen, kuinka tietokoneet oppivat datasta. Se on eräänlainen tilastollinen polku, joka hakee oppia datan suhteesta ja tietojenkäsittelytieteestä. Koneoppimisessa painotetaan erityisesti tehokkaiden laskennallisten algoritmien käyttöä. Se nousee tilastollisista luvuista, joiden tarkoitus on oppia dataa ja tietojenkäsittelytieteestä, jossa sen painotus on tehokkaissa laskennallisissa algoritmeissa. Tätä matematiikan ja tietojenkäsittelyn ”avioliittoa” ajaa eteenpäin Deon (2015) mukaan eteenpäin ainutlaatuiset laskennalliset haasteet tilastollisten mallien rakentamiselle suuresta datan joukosta. Niihin voi sisältyä biljoonia ja triljoonia datan pisteitä. Näin ollen tämän tyyppinen tietokoneiden oppiminen voidaan perinteisesti luokitella eri kategorioihin, kuten ohjattuun oppimiseen ja ei-ohjattuun oppimiseen. Deo(2015) toteaa, että tätä luokittelua voidaan lisätä vielä yhdellä jaolla, se on hänen mukaansa hyödyllistä silloin kun huomioidaan, kuinka koneoppiminen saattaa ilmoittaa lääketieteen käytännöstä. Tällöin voidaan puhua erottuvasta oppimisesta, siitä erotetaan tehtävät, jotka lääkärit jo osaavat tehdä hyvin ja pyritään oppimaan tehtävistä, joita lääkäri oppii tekemään vain rajoitetusti. (Deo 2015).

Koneoppimisen määrittelyä käsitteenä voidaan lähestyä myös näkökulmasta, jossa tarkastellaan sitä, mitä koneoppiminen ei ole. Tähän päätyvät Obermeyer, Phil, Emanuel (2016). He painottavat artikkelissaan sitä, että useimmat tietokone-pohjaiset algoritmit

lääketieteessä ovat ”asiantuntija järjestelmiä”, niiden avulla voidaan koodata annetusta aiheesta tietoa. Niihin on sovellettu tehtyjä yhteenvedoja erityisistä kliinisistä skenaarioroista. Tällaisia ovat lääkkeiden vuorovaikutuksen havaitseminen ja asianmukaisen päättelyn saaminen radiologisista kuvaantamisista. Asiantuntijajärjestelmät toimivat samalla tavoin kuin ihanteellinen lääketieteen opiskelija toimisi, ne ottavat yleiset periaatteet lääketieteestä ja soveltavat niitä uusiin potilaisiin. (Obermeyer, Phil, Emanuel 2016).

Obermeyer ym.(2016) ovat sitä mieltä, että varsinaisen koneoppimisen soveltaminen lääketieteeseen tarkoittaa sitä, että koneoppiminen lähestyy ongelmia, kuten niitä prosessoiva lääkäri saattaisi tehdä. Se oppii sääntöjä datasta. Aloittamalla potilaan tasolla oleva havainnointi, algoritmit seulovat useiden laajojen numeroitujen muuttujien läpi. Ne etsivät yhdistelmiä, jotka luotettavasti ennustavat tuloksia. Obermeyer ym. (2016) mukaan tämä on tietyssä mielessä samanlainen prosessi kuin perinteiset regressiomallit. Niissä on lopputulos, ja vaihtelut ja tilastollinen funktio on näiden kahden linkittäjänä. Koneoppiminen kykenee käsittelemään suuria määriä ennustavia muuttujia. Se käsittelee joskus enemmän ennustavia muuttujia kuin havaittuja, ja se yhdistää ne epälineaarilla hyvin interaktiivisilla tavoilla. Tämä kapasiteetti mahdollistaa meidän käyttäen uudenlaista dataa, jonka määrä tai monimutkaisuus olisi aikaisemmin tehnyt sen analysoimisen mahdottomaksi. (Obermeyer, Phil & Emanuel 2016).

Fagella (2018) kirjoittaa artikkelissaan erilaisten koneoppimissovellusten käyttötarkoituksista lääketieteessä. Hänen mukaan koneoppimisen kustannustehokkuus ja isompi data takaavat lähes aina paremmat hoitotulokset. Tulevaisuudessa voidaan kehittää koneoppimisen potentiaalisia sovelluksia, joissa datan synkronointi, analysointi ja innovointi ovat jokapäiväistä todellisuutta. Fagella (2018) arvioi, että tulevaisuudessa pystytään tekemään läpimurto useilla sovelluksilla. Ne ovat eräänlaisia ”pioneereja” alueilla, joissa on jatkuvia innovaatioita. Fagella(2018) on luetellut kaikkiaan seitsemän eri sovelluskohdetta, joihin koneoppimissovelluksia voidaan lääketieteen alueella käyttää. Ne ovat tautien identifiointi ja diagnostiikka, personalisoitu hoito, lääkkeiden löytäminen/valmistus, kliininen koetutkimus, radiologia ja radioterapia, älykkäät elektroniset terveys rekisterit ja epidemioiden ja kulkutautien puhkeaminen. Sairauksien identifiointi ja sairastamisen diagnostiikka on lääketieteellisen koneoppimistutkimuksen

tärkeimpiä osa-alueita. Fagella (2018) viittaa yhdysvaltalaisen farmaseuttisen tutkimuksen ja valmistajien julkaisemaan tutkimusraporttiin. Siitä selviää, että tällä hetkellä on yli 800 lääkettä ja rokotetta kehitteillä syövän hoidolle. Knight instituutin tutkija Jeff Tyner huomauttaa, että tämä tuottaa myös haasteita löytää tapoja työskennellä kaiken lopullisen datan kanssa. Sen takia on hyvä idea, että esim. biologi työskentelee tietojärjestelmätieteilijöiden ja laskijoiden kanssa, toteaa Tyner. (Fagella 2018).

Fagella (2018) mainitsee artikkelissaan myös siitä, kuinka koneoppimista voidaan nykyisin käyttää onnistuneesti alkuvaiheen lääkelöydösten kehittämiseen. Niissä voi olla useita potentiaalisia käyttökohteita alkuvaiheen lääkeseoesten seulomisesta onnistumisen arvioimiseen pohjautuen biologisiin tekijöihin. Tähän sisältyy tuotekehitys löydökset teknologiassa, kuten seuraavan sukupolven sekvensointi. Fagella(2018) uskoo myös, että tähän alueeseen on etusijalle vahvasti nousemassa täsmä lääketiede. Siihen sisältyy mekanismien identifioiminen monen tekijän sairauksista, ja vuorostaan vaihtoehtoisten polkujen löytäminen hoidolle ja terapioille. Tähän tutkimukseen sisältyy paljon valvomatonta oppimista, mikä on suuremmaksi osaksi rajattu identifioimaan polkuja datassa ilman ennustuksia. (Fagella 2018).

Fagella (2018) viittaa kohdealueen tutkimuksen merkittävänä avaintekijänä MIT:n kliinisen koneoppimisen ryhmään. Sen täsmälääketieteellinen tutkimus on keskittynyt algoritmien kehittämiseen, jotta voitaisiin paremmin ymmärtää tautien kulkua ja suunnitella tehokkaita hoitoja sellaisille sairauksille, kuten 2 tyypin diabetes. Fagella (2018) viittaa myös Royal Societen huomautukseen, jonka mukaan koneoppimista voidaan soveltaa biovalmistamisessa. Sen avulla optimoidaan farmaseuttista valmistamista. Siinä saadaan dataa kokeellisesta valmistamisprosessista, ja näin ollen prosessilla voi potentiaalisesti auttaa farmaseuttista valmistajaa vähentämään lääkkeiden tuottamiseen vaadittua aikaa. Tästä seuraa lääkkeiden alentuneet valmistuskustannukset ja paremmat vastineet. Kliininen kokeellinen tutkimus on Fagellan (2018) mukaan nousemassa yhdeksi koneoppimisen merkittäväksi käyttökohteeksi. Koneoppimisella voidaan saada aikaiseksi useita hyödyllisiä potentiaalisia sovelluksia, ja niiden avulla voidaan auttaa muodostamaan ja ohjaamaan kliinisiä kokeellisia tutkimuksia. Soveltamalla edistynyttä analytiikkaa parhaiden kliinisten kokeiden identifioimiseen, voidaan saada esille paljon laajempia data-alueita kuin nykyisin. Siihen sisältyy

sosiaalinen media ja lääkärissä käynti. Dataan sisältyy myös geneettinen informaatio, silloin kun erityisesti etsitään kohderyhmää kohdepopulaatiosta. Tämän seurauksena tulisi kokonaisvaltaisesti pienempiä, nopeampia ja vähemmän kalliita kokeita. (Fagella 2018).

Fagella (2018) arvioi, että koneoppimista voidaan myös käyttää etä-monitorointiin. Sen avulla voidaan myös päästä käsiksi reaali-aikaiseen dataan. Se lisää turvallisuutta, yksi esimerkki tästä on biologinen monitorointi ja muut signaalit, jotka toimivat merkkinä haitoista tai kuolemasta osaan ottajille. Fagella (2018) viittaa myös Mckinsey yhtiön (2013) julkaisemaan raporttiin, jonka mukaan on paljon muita koneoppimisen sovelluksia, joiden avulla voidaan auttaa lisäämään kliinisten kokeiden tehokkuutta. Se tapahtuu siten, että löydetään parhaat otoskoot lisääntyntä tehokkuutta varten, on myös mahdollista osoittaa ja käyttöönottaa eri paikkoja potilaiden rekrytointia varten. Käyttämällä elektronisia lääketieteellisiä rekistereitä voidaan myös vähentää data virheitä, esimerkiksi jäljennöksiä. (Fagella 2018).

Älykkäät elektroniset terveystiedot ovat merkittävässä osassa tulevaisuuden lääketieteessä. esim. dokumenttien luokittelussa. Tästä on esimerkkinä potilas tiedustelun lajittelu sähköpostin kautta käyttäen tuki vektori koneita (Support Vector Machine, SVM). Optisia kirjoitusmerkkien tunnistimia muutetaan kursivoiduksi tai muu piirretty käsiala yksilöidyksi merkiksi. Nämä ovat molemmat tärkeitä koneoppimiseen pohjautuvia teknologioita, joiden tarkoitus on auttaa elektronisen terveystiedon kokoamiseen ja numeroinnin edistämiseksi. Tästä on olemassa esimerkkinä kaksi innovaatiota. Ne ovat Matlabin koneoppimisen käsikirjoitusten tunnistamisen teknologiat ja Googlen Cloud Vision API, joka on optinen kirjaimen tunnistaja. MIT:in kliininen koneoppimisen työryhmä on myös viemässä huippuunsa seuraavan sukupolven älykkäitä elektronisia terveystiedot. Ne yhdistävät kokonaisuudeksi sisäänrakennetun koneoppimisen/tekoälyn auttaakseen sellaisissa asioissa, kuten diagnostiikka, kliiniset päätökset ja personoidut henkilökohtaiset hoitosuositukset. Nykyisin on kuitenkin tarve toimiville koneoppimisen algoritmeille. Niiden on voitava oppia vähän tunnuksellisesta datasta, ymmärrettävä luonnollista kieltä ja yleistää hyvin poikittaisia lääketieteellisiä asetuksia ja instituutioita. (Fagella 2018).

Yksi mielenkiintoinen koneoppimisen lääketieteellinen sovellusalue on myös epidemioiden puhkeamisen ennustaminen. Fagellan (2018) mukaan koneoppimista ja tekoäly teknologiaa on mahdollista käyttää monitorointiin ja epidemioiden puhkeamisen ennustamiseen kaikkialla maailmassa. Se perustuu datan keräämiseen satelliiteista, Webin historiallisen informaation reaaliaikaiseen sosiaalisen median päivittämiseen, ja muihin lähteisiin. Tukivektori koneita ja keinotekoisia hermoverkkoja käytetään usein, esimerkiksi malarian puhkeamisen ennustamisessa, ottaen huomioon datan, kuten lämpötilan, keskimääräiset kuukausittaiset vesisateet, positiivisten tapausten määrä, ja muut datan kohdat. (Fagella 2018).

Starr (2018) kirjoittanut koneoppimisen nykyistä käyttökohteista terveydenhoidossa. Starrin (2018) mukaan koneoppimisella on merkittävä vaikutus terveydenhoitoon, maksajat, hankkijat ja farmaseuttiset yhtiöt ovat myös laittaneet merkille sen käyttökelpoisuuden tiloihinsa ja alkaneet käyttää hyväksi koneoppimisen hyötyjä. Starr (2018) huomauttaa, että koneoppimisen malli on luotu syöttämällä dataa oppivaan algoritmiin. Algoritmi on aina se missä kaikki tapahtuu, ja on myös olemassa algoritmeja, jotka havainnoivat potilaan sairaalassa olon pituutta perustuen diagnooseihin. Jonkun täytyy kirjoittaa tuo algoritmi ja sitten testata sitä oikeaan ja luotettavaan dataan. Ajan kuluessa malli voidaan opettaa uudelleen uudella datalla lisäten mallin tehokkuutta. (Starr 2018).

Starr (2018) toteaa koneoppimisen olevan tekoälyn alijoukko. Tekoäly voidaan opettaa käyttämällä tietokonesysteemiä suorittamaan normaalia ihmisälykkyyttä vaativia tehtäviä, kuten visuaalista huomiointia, puheentunnistamista, päätöksentekoa ja kääntämistä kielten välissä. Koneoppimisesta on nykyisin useita esimerkkejä terveydenhoidossa. Nykyisiä esimerkkejä sovelluksista, joiden alkuperä pohjautuu tekoälyyn ovat: Project Inner Eye, se on tutkimuksiin perustuva mekaaninen tekoälyn ohjelmisto radioterapian suunnitteluun. Sitä valvoo Microsoft tutkimus, se on kehittänyt terveydenhoidon tekoälyn ja se on myös kehittänyt terveydenhoitoon useita suunniteltuja algoritmeja. Ne tarjoavat myös laajan näkökulman. Microsoft on myös jatkuvasti sitoutunut resursseihin, joilla terveydenhuolto tehdään tehokkaaksi koneoppimisella. Nykyisin on useita ohjelmia ja julkisesti vapaita resursseja. Näitä ovat esim. Microsoft Medical, Health ja Genomic Microsoftin tutkimuksessa, älykäs terveys tekoälyn ja

pilvien kautta sekä Microsoftin adaptiiviset bioteknologiat, jossa tekoälyä käyttämällä voidaan määrittää immuunijärjestelmä. Kaikki nämä ovat alun perin tehty algoritmeilla, joita ovat kehittäneet tutkijat, datatieteilijät, kehittäjät ja muut. Ennustamisen nopeus tai tunnistaminen riippuvat datasta ja piirteistä, joita käytetään mallissa. Algoritmia käytetään oppimaan datasta. (Starr 2018).

Starr (2018) arvioi, että koneoppimisen edut lääketieteessä ovat suuria, kun tarkastellaan populaationtrendejä. Esimerkkinä tästä, hän mainitsee kysymyksen siitä, missä osassa Yhdysvaltoja suurin osa multipeli skleroosia sairastavista ihmisistä asuu. Tämä voi puolestaan johtaa jatkokysymyksiin, kuten miksi näin on. Koneoppimisen avulla voidaan löytää potentiaalisia huomioimattomia näkökohtia älykkäiden liiketaloudellisten lähestymistapojen kautta. Koneoppimisen käyttö lääketieteessä saattaa tuoda mukanaan, kuitenkin myös ongelmia henkilön intiimien terveystietojen salaamisen kanssa. Starr (2018) mainitsee henkilökohtaisesti identifioidusta tiedosta (Pii). Terveystieteiden tuotteista ja palveluista tulee runsaasti dataa (elektroniset terveysrekisterit ym), tämä voikin sisältää henkilökohtaista identifioitua tietoa (personal identifiable information). Erityisesti on kiinnitettävä huomiota siihen, kuinka organisaatio tulee käyttämään ja kohtelemaan ns. Pii dataa koneoppimisen ratkaisussa. (Starr 2018).

Starr (2018) mainitsee myös kaksi esimerkkiä koneoppimisen soveltamisesta lääketieteeseen. Ne ovat diagnostinen radiologia, ja koneoppimisen soveltaminen lopputulosten arvioimisessa. Starr (2018) mielestä diagnostisten radiologien työ on nykyään paljon kuvien analysointia, jotta he kykenisivät identifioimaan epänormaaleja löydöksiä potilaista ja paljon muuta. He ovat usein kriittisiä diagnoosien teossa, ja heidän diagnoosinsa perustuu siihen, mitä he löytävät. Yksi esimerkki tästä on kasvaimen tunnistaminen. Tekoälyä voidaankin käyttää diagnostisen radiologian avustamisessa. Tästä on esimerkkinä Project InnerEye. Project InnerEye kehittää koneoppimistekniikoita automaattiseen kasvainten luonnosteluun, kuten myös terveeseen anatomiaan 3D radiologista kuvista. Tämä mahdollistaa kohdennettujen mittojen hyödyntämisen kvantitatiivista radiologiaa varten, nopean radioterapian suunnittelun ja navigaation. Project Inner Eye käytännössä kääntää moniulotteiset radiologiset kuvat mittaaviksi laitteiksi. Ohjelmiston tarkoitus on avustaa radiologeja jäljittelemällä kasvainten luonnoksia. Radiologia tuottaa nykyisin suuren määrän skannauksia alueesta (huipulta-

pohjalle). Radiologi käy yleensä läpi jokaisen skannauksen ja jäljittää ääri viivoja kasvaimesta. Sen jälkeen 3D liitetty kasvain voidaan tuottaa, ja tämä tehtävä kestää vain tunnin. Kuitenkin käytettäessä koneoppimisen Project InnerEye:tä tämä vie vain minuutteja. (Starr 2018.)

Toisena esimerkkinä koneoppimisen soveltamisesta lääketieteeseen on ns. lopputulosten ennustus (Predict Outcomes). Starr (2018) kirjoittaa, että muuta hyvin käytännöllistä tietoa voidaan ennustaa käyttäen koneoppimista ja tekoälyä. Voidaan esimerkiksi ennustaa potilaan sairaalassaolon oleskelun pituutta. Tämä on eräs ennustavan analyysin muoto, ja ennustava analyysi voi esimerkiksi auttaa hoitajia pitämään parempaa huolta potilaistaan ja itsestään. Diagnostiikka ja ennustava analyysi ovat Starrin (2018) mukaan suuria hyötyjä. Koneoppimisen ja tekoälyn käytöstä on myös hyötyä muihin lääketieteellisten alan ongelmiin. Monet lääkeyhtiöt ovat innokkaita juuri nyt, koska se auttaa heitä käsittelemään dataa. Nämä teknologiat tulevat muuttamaan olennaisesti terveydenhoitoa. Terveydenhoidossa on kuitenkin vielä paljon säädöksiä, näin ollen näemme koneoppimisen/tekoälyn enneaikaista käyttöönottoa silloin, kun koneoppimisen/tekoälyn hyödyt avautuvat ja uusia algoritmeja kehitetään. (Starr 2018.)

Magoulas ja Prentza (2015) ovat myös kirjoittaneet koneoppimisen lääketieteellisistä sovelluksista. Koneoppiminen tarjoaa heidän mukaansa metodeita, tekniikoita ja työkaluja, jotka voivat auttaa diagnostisten ja esidiagnostisten ongelmien ratkaisuisissa monilla lääketieteen alueilla. Koneoppimista käytetään nykyisin kliinisten parametrien tärkeyden analysoinnissa, niiden kombinaatioiden prognoosissa, tähän sisältyy tautien prognoosien ennustaminen, lääketieteellisen tiedon poiminto tutkimustuloksia varten, taudin hoito ja tuki ja yleinen kokonaisvaltainen potilaiden hoito. Koneoppimista käytetään myös datan analysointiin, kuten säännönmukaisuuksien havainnointiin datassa, käsittelemällä kunnolla käytettävää vaillinaista dataa. Sitä käytetään tehohoitoyksikössä ja älykkäissä hälytyksissä, niiden lopputuloksena saadaan tehokas ja tehostettu monitorointi. (Magoulas & Prentza 2015).

Magoulas ja Prentza (2015) toteavat, että koneoppimismetodien onnistunut toteuttaminen voi helpottaa tietokonepohjaisten järjestelmien integraatiota terveydenhoidon ympäristössä. Tämä voi tarjota uusia mahdollisuuksia lääketieteen ammattilaisille, ja se

voi lopulta parantaa myös lääketieteellisen hoidon laatua. Prentza & Magoulas (2015) pitävät lääketieteellisen diagnostiikan todistamista hyvin tärkeänä lääketieteen sovellusalueena. Asiantuntijajärjestelmät ja malliperusteiset tietokantakaaviot tarjoavat mekanismeja hypoteesien muodostamiselle potilaan datasta. Esimerkkinä tästä on johtosäännöt, jotka ovat hyödynnetty asiantuntijoiden tiedosta asiantuntijajärjestelmien rakentamiseksi. Asiantuntijat eivät kuitenkaan voi valitettavasti tietää, mitä tietoa niillä oikeastaan on heidän ongelmistaan. Symbolista oppimistekniikkaa (sisällyttää induktiivisen oppiminen esimerkeillä) käytetään lisäämään oppimista ja tiedon hallinnan mahdollisuuksia asiantuntijajärjestelmiin. Silloin kun annettu joukko kliinisiä tapauksia toimii esimerkkeinä, älykkäissä järjestelmissä oppiminen voidaan saavuttaa käyttäen koneoppimisen metodeita, jotka kykenevät tuottamaan systemaattista määritelmää niistä kliinisistä piirteistä, jotka epätasaisesti (uniquely) luonnehtivat kliinisiä tiloja. Tällöin tieto voidaan ilmaista yksityisten sääntöjen muodossa tai usein päätöksentekopuuna. (Margoulas & Prentza 2015).

Tätä lähestymistapaa voidaan laajentaa käsittelemään tapauksia, jossa ei ole aikaisempaa kokemusta lääketieteellisen datan tulkitsemisesta ja ymmärtämisestä. Magoulas & Prentza (2015) viittaavat Haeun ja Coficieran määritelmään älykkästä järjestelmästä, joka ottaa reaali-aikaista potilas dataa säilyttäen sen sydämen ohitusleikkauksessa, ja luo malleja normaalista ja epänormaalista sydämen fysiologiasta havaitakseen muutoksia potilaan tilassa. Lisäksi tutkimusasetelmassa, nämä mallit voivat toimia alkuperäisenä hypoteesina, jotka voivat johtaa uudempiin tutkimuksiin. (Magoulas & Prentza 2015).

Potilaan datasta oppiminen tuo Magoulasin ja Prentzan (2018) mukaan monia ongelmia, koska datajoukkoja luonnehtii hajanaisuus/muuttuvia parametrien arvoja, paikkaansa pitämättömyys, (systemaattinen tai satunnainen häiriö datassa, potilasrekisterien saatavuuden puute, epätarkkuus ja epäsopiva valikoima parametrejä annettuun tehtävään). Koneoppiminen tarjoaa työkaluja käsitellä näitä lääketieteellisen datajoukon kirjoitusmerkkejä. Alisymmetriset oppimismetodit, erityisesti neuraaleissa hermoverkoissa, pystyvät käsittelemään näitä datajoukkoja, sitä käytetään usein hahmojen arvoituksen valmiudessa ja heidän ihmismäisissä ominaispiirteissään. Tällöin voidaan kehittää lääketieteellistä päätöksentekoa. (Magoula & Prentza 2018).

3. Terveydenhuollon tietojärjestelmät

Onnistunut ja laadukas terveydenhuolto vaatii toimivia tietojärjestelmiä. Terveydenhuollon tietojärjestelmän on määritellyt mm. sosiaali- ja terveystieteiden tutkimuskeskus ja valvontavirasto Valvira. Se on pyrkinyt määrittelemään, mitä tietojärjestelmät ovat ja mitä ne sisältävät terveydenhuollon näkökulmasta. Valvira pyrkii tuomaan määrittelyssään esille myös sen, mitä käyttäjien asettamia vaatimuksia toimivien tietojärjestelmien tulisi täyttää. Tietojärjestelmä on sosiaali- tai terveydenhuollon asiakastietojen sähköistä käsittelyä varten toteutettu ohjelmisto tai järjestelmä. Sen avulla tallennetaan ja ylläpidetään kaikkia asiakas- tai potilasasiakirjoja sekä niihin sisältyviä tietoja. Tietojärjestelmän on ennen sen käyttöönottoa täytettävä olennaiset vaatimukset liittyen yhteen toimivuuteen, tietoturvaan ja tietosuojaan, kuten myös toiminnallisuuteen. Varsinainen vaatimuksenmukaisuuden osoittaminen kuuluu kuitenkin aina tietojärjestelmien valmistajalle. (Valvira 2015.)

Toinen tapa määritellä tietojärjestelmä, on määritellä se välityspalveluna. Tällöin tietojärjestelmällä tarkoitetaan siis myös varsinaista välityspalvelua. Sillä välitetään sosiaali- tai terveydenhuollon asiakastietoja Kansaneläkelaitoksen ylläpitämiin valtakunnallisiin tietojärjestelmäpalveluihin. Ne on määritelty lain 159/2007 14 pykälässä. Sosiaali- ja terveydenhuollon tietojärjestelmät on myös jaettu ja luokiteltu niiden käyttötarkoitusten ja olennaisten ominaisuuksien mukaisesti joko luokkiin A tai B. Luokka A:han kuuluu kaikki Kansaneläkelaitoksen ylläpitämät kanta-palvelut ja tietojärjestelmät. Ne on tarkoitettu liitettäväksi kantapalveluihin, ja se tapahtuu joko suoraan tai teknisen välityspalvelun kautta. On huomioitavaa, että välityspalvelu kuuluu aina luokkaan A. Luokkaan B puolestaan kuuluu muut tietojärjestelmät. Palvelun antajan on myös aina ilmoitettava tietojärjestelmiin tulevista merkittävistä poikkeamista tietojärjestelmien valmistajalle. Tämä on tehtävä aina silloin, mikäli poikkeama aiheuttaa suuren riskin potilasturvallisuudelle, tietoturvalle tai tietosuojalle. Tästä poikkeamasta on myös aina ilmoitettava Valviralle. Valmistajan puolestaan on aina seurattava ja tehtävä arviointia ajantasaisella järjestelmällisellä menettelyllä tietojärjestelmästä, ja siinä on arvioitava tietojärjestelmän tuotantokäytön aikana saatuja kokemuksia. Tietojärjestelmän

valmistajan on aina myös ilmoitettava kaikista tietojärjestelmiin liittyvistä merkittävistä poikkeamista suoraan Valviralle. (Valvira 2015.)

Tietojärjestelmä on järjestelmä, jonka tarkoitus on aina edistää erityisen tiedon tukea päätöksenteossa kaikilla organisaation tasoilla. Winter, Haux, Ammenwerth, Brigl, Hellrung & Jahn (2011) määrittelevätkin tietojärjestelmän ihmisten ja teknisten komponenttien muodostamaksi kokonaisuudeksi, jossa tapahtuu varsinaista tiedon käsittelyä. Vane (2016) puolestaan huomauttaa, että tietojärjestelmään tallennetaan dataa, ja tietojärjestelmä tallennetaan ihmisten ja teknisten komponenttien muodostamaksi kokonaisuudeksi, jossa tapahtuu tietojen käsittelyä. Usein käytetään myös termiä sosio-tekniinen järjestelmä. Se on tiedonhallintajärjestelmä, joka koostuu ihmisistä ja teknisistä komponenteista. Käsitteellä sosio viitataan ihmisen olevan osa tiedonkäsittelyä, tekninen puolestaan viittaa informaation käsittelyn työkaluihin, joita ihmisellä on käytössä. Näitä ovat tietokoneet, puhelimet ja potilasarkistot. (Vane 2016:2 & Winter ym.2011.)

Sairaalan tietojärjestelmä(Hospital Information System) koostuu joukosta komponentteja ja organisoituja toimintatapoja, joiden tarkoitus on muodostaa objekteista informaatiota. Tämän on tarkoitus parantaa päätöksentekoa kaikilla terveydenhuolto järjestelmän tasoilla. Vane(2016) määrittelee tietojärjestelmän tehtäväksi käsitellä informaatioita, dataa ja tietoa. Terveydenhuollossa tietojärjestelmien pitäisi myös aina pystyä antamaan laaja kokonaiskuva yksittäisen potilaan lisäksi, koko sairaalasta ja sen toiminnasta. Tieto on keskeinen sairaalaan tietojärjestelmiin liittyvä komponentti ja avainkäsite, ja se tarkoittaa mm. tietämystä lääkkeiden toimivuudesta ja hoitotoimenpiteistä. Informaatio puolestaan määritellään Winterin, Hauxin, Ammentwerthin, Briglin, Hellrungin & Jahnin (2011) mukaan käsitteeksi, joka on johdettavissa tiedon avulla tapahtumista, konsepteista ja faktoista. Data puolestaan on heidän mukaan ihmiselle tai koneelle käyttökelpoisessa muodossa olevaa dataa. Tämä voi olla vaikka paperiselle taulukolle kerättyjä tuloksia verenpainemittauksista tai vaihtoehtoisesti samoja arvoja syötettynä digitaalisesti elektroniseen käyttöjärjestelmään. Paperi ja tietojärjestelmä ovat molemmat informaation käsittelyvälineitä, näin ollen kyseistä informaatiota tarvitseva voi tarkistaa mittaustulokset molemmista. (Vane 2016:2 & Winter ym.2011.)

Winter ym. (2011) toteavat sairaalan tietojärjestelmien olevan tietojärjestelmien muodostama järjestelmä. Sen varsinainen tehtävä on aina kattaa kaikki informaation käsittely sairaalassa. Sairaalan tietojärjestelmän osat koostuvat fyysisestä informaation käsittelystä, sairaalan toiminnoista, liiketoiminta prosesseista ja sovelluksista. Sairaalan tietojärjestelmä käsittää kuitenkin kaiken sairaalassa käsitellyn informaation, ja siinä ei välttämättä tarvitse olla osallisena ihminen tai jonkinlainen tekninen väline. Sairaalan tietojärjestelmä on olemassa heti sairaalaan toiminnan aloittamisesta lähtien, sitä ei näin ollen vain hankita jossain vaiheessa. Tällöin on kysymys aina siitä, kuinka sen toimintaa kyetään parantamaan ja tehostamaan. Tehokkaalla tietojen käsittelyllä pystytään myös vähentämään toiminnasta johtuvia kuluja. (Vane 2016:3.)

Sairaalan tietojärjestelmän tehtävä on pyrkiä takaamaan riittävän sujuva sairaalan toimintojen suoriutuminen, ja toimintoihin sisältyy esim. potilaan hoito, potilaiden hallinta ja sairaalan hallinta. Lain asettamat vaatimukset esim. potilaiden tietojen katselun suhteen pyritään toteuttamaan sairaalan tietojärjestelmissä. Potilaiden hoitoa tukevan toimintojen tarjoaminen on sairaalan tietojärjestelmien keskeisin tehtävä. Potilaiden tietojen saattaminen saataville, oikea-aikaisesti, oikeaan paikkaan, oikeassa muodossa ja oikeutetuille henkilöille sisältyy tähän olennaisesti. Vane (2016) huomauttaa, että lääkärin tai sairaanhoitajan on saadaksesen tietoja mainittujen ehtojen mukaisesti kerättävä, käsiteltävä ja tallennettava ne. (Vane 2016:6.)

Tämä systemaattisen dokumentoinnin mukanaan tuoma tietomäärä vaati tietojärjestelmiltä aina suurta tehokkuutta, ja tämän takia ollaankin siirretty käyttämään sähköisiä, elektronisia tietojärjestelmiä. Aktorin pitäisi saada järjestelmästä ajantasaista, kliinistä tietoa, näitä voivat olla esim. uusimmat hoitotoimenpiteet ja tiedot lääkkeiden sivuvaikutuksista. Tämän lisäksi sairaalan hallinnolle pitäisi antaa tietoa sairaalan sen hetkisisistä kustannuksista ja toiminnoista. Sairaalan tietojärjestelmiä käyttää aina suuri joukko eri työtehtävien hoitajia, ja samalle tiedolle onkin yleensä tarvetta eri tyyppisessä muodossa. Tärkeää osa sairaalan tietojärjestelmää on myös esim. tutkimuksen tukeminen yliopistosairaaloissa. Tietojärjestelmän tehtäviin sisältyy myös aina erilaisen tiedon tarpeiden vaatimukseen sopeutuminen. (Vane 2016:6.)

Terveystietojärjestelmät on uusi vakiintunut termi, josta on alettu puhumaan viime aikoina. Tämä on täysin uusi käsite, jota on alettu käyttämään sairaalatietojärjestelmä termin ohessa. Näiden kahden käsitteen suurimmat erot ovat niiden toimijoiden laajuudessa, jotka käyttävät tietojärjestelmiä. Käsitettä terveystietojärjestelmä (Health Information Systems, HIS) kutsutaan transinstitutionaaliseksi järjestelmäksi. Se tarkoittaa, että sama järjestelmä on monen terveydenhuollon laitoksen yhteiskäytössä. On olemassa myös kansallisia tietojärjestelmiä. Yksi tällainen esimerkki on kansallinen terveysarkisto(kanta), se on myös transinstitutionaalinen tietojärjestelmä, toiminnoiltaan se on kuitenkin hyvin suppea sairaalatietojärjestelmiin verrattuna. Sairaaloiden kustannustehokkuuteen ja toimintokuluihin kohdistuu nykyisin myös suuria paineita, ja myös potilaskeskeisempään hoitoon siirtyminen aiheuttaa niitä. Tämän seurauksena tunnetaan erityistä mielenkiintoa yhteiskäyttöisten tietojärjestelmien kehittämiseen. Yhteiskäyttöisen tietojärjestelmän käyttö johtaa väistämättä siihen, että sairaalan toiminnot siirtyvät yksittäisistä kokonaisuuksista potilaskeskeisempää hoitoa kohti. Potilaskeskeisempi hoito laajenee myös useille erikoistuneille laitoksille. (Vane 2016:7.)

Suuret sairaala tietojärjestelmät voidaan jakaa kahteen luokkaan: hallinnollisiin ja kliinisiin. Yksinkertaisiin tapa erottaa ne toisistaan on tarkastella niiden käyttötarkoitusta ja niiden sisältämää data tyyppiä. Hallinnollinen tietojärjestelmä(tai hallinnollinen sovellus) sisältää ensisijaisesti hallinnollista tai taloudellista dataa, ja sitä käytetään yleisesti tukemaan johdon toimintoja ja terveydenhuolto organisaation yleisiä toimenpiteitä. Hallinnollinen tietojärjestelmä saattaa esim. sisältää sellaista informaatiota, jota käytetään henkilöstön johtamisessa, raha-asioissa, materiaaleissa, hankkimisessa tai välineissä. Se voi olla järjestelmä henkilöstön resurssin hallintaan, materiaalien hallintaan, potilas kirjanpitoon tai laskutukseen, tai myös potilaslistaan. Vastakohtana, kliininen tietojärjestelmä tai (kliininen sovellus), puolestaan sisältää kliinistä tai terveyteen liittyvää informaatiota, joka käyttää potilaan monitorointia hyväkseen tarjotakseen hoitoa ja diagnosointia. Kliininen tietojärjestelmä saattaa olla osastokohtainen järjestelmä, kuten esim. radiologia-, farmasia- tai laboratoriojärjestelmä, tai vaihtoehtoisesti kliininen päätöksenteontuki, lääkitsemisen hallinta, tietokoneilla järjestetty luokiteltu syöttö tai elektroniset terveysrekisteri(Electronic health care record, EHR) järjestelmät, nämä ovat niistä muutamia esimerkkejä. Niiden laajuus saattaa olla

rajoitettu yhteen kliiniseen informaation alueeseen (esim. radiologiaan, farmasiaan tai laboratorioon) tai ne saattavat olla ymmärrettäviä ja kattaa käytännöllisesti katsoen kaikki potilas hoidon näkökulmat, kuten esim. elektroniset terveystietorekisterit tekevät. (Emont 2011.)

Tyypillisiä kliinisiä tietojärjestelmiä ovat muun muassa kliinisiin sovelluksiin sisältyvät avustavat tietojärjestelmät, kuten laboratorio tietojärjestelmä, radiologi tietojärjestelmä ja farmasia tietojärjestelmä. Muita kliinisiä tietojärjestelmiä ovat sairaanhoito dokumentaatio ja elektroninen terveystietorekisteri (Electronic health record(EHR)). Hoito dokumentaatio hyödyntää hoitodokumentointia arvioinnista arvosteluun, potilas hoidon päätöksen tekoa(hoito suunnittelu, arviointia, lohkosuunnittelu taulukko, potilas tarkkailu ja hyödyntäminen). Elektroninen terveystietorekisteri (EHR) hyödyntää elektronista kaappausta ja potilaan terveyshistorian raportointia, ongelma listausta, hoitoa ja tuloksia. Se sallii kliinikoiden dokumentoida kliinisiä löydöksiä, progression muistiinpanoja ja muuta potilas informaatiota. Sen tehtävä on antaa myös päätöksentekotuki työkaluja, huomautuksia sekä varoituksia. (Wager, Lee, Glaser 2017:68.)

Terveydenhuollossa tarvitaan ennen kaikkea tietojärjestelmiä, jotka tuottavat aktiivisesti ajankohtaista tietoaineistoa hoidettavista potilaista kliinikoiden käyttöön. Tällöin on kyseessä kliininen päätöksenteon tukijärjestelmä. Berner (2007) määrittelee kliinisen päätöksenteon tukijärjestelmän (Clinical decision support systems, CDSS) kliinisen päätöksen teon vaikuttamista varten suunnitelluiksi tietokonejärjestelmäksi. Sen on tarkoitus vaikuttaa yksittäistä potilasta koskevaan kliiniseen päätöksentekoon juuri sillä hetkellä, kun päätöksiä tehdään. Kliinisten päätöksenteon tukijärjestelmien uskotaan olevan avainasemassa silloin, kun yritetään parantaa potilasturvallisuutta. Oikealla tavalla käytettynä kliininen päätöksenteon tukijärjestelmä kykenee muuttamaan tapaa, jolla lääketiedettä opetetaan ja harjoitetaan. Kliiniset päätöksenteko tukijärjestelmät voidaan luokitella tietämuskantaisiin ja ei-tietämuskantaisiin tietojärjestelmiin. Tietämuskantaan perustuvat kliiniset päätöksenteon tukijärjestelmät koostuvat tietokannasta, päättelymoottorista ja kommunikaatiomekanismista. Ei-tietämuskantaan perustuvat kliiniset päätöksenteko tukijärjestelmät puolestaan käyttävät koneoppimista ja muita tilastollisia hahmontunnistus lähestymistapoja. (Berner 2007:2-4.)

EBMEDS on lyhenne termistä (Evidence Based Medicine Electronic Decision Support), eli suomennettuna se tarkoittaa lääketieteelliseen näyttöön perustuvaa sähköistä päätöksenteon tukea. Päätöksenteontuki itsessään antaa terveydenhuollon organisaatioille välineitä johtaa ja ohjata toimintaansa, ja erityisesti se tarjoaa mahdollisuuden prosessien ohjaukseen. Tämä voi tapahtua esim. alueellisen hoitoketjun sisältämien sopimusten kautta. Päätöksentuen avulla on myös mahdollista seurata hoidon laatua ja mm. käytettyjä voimavaroja. Päätöksentuki terveydenhuollossa antaa mahdollisuuden hoitokäytäntöjen ja potilasturvallisuuden parantamiselle, ja sen avulla voidaan myös tuoda tutkittua tietoa osaksi kliinistä päätöksentekoa. Käsitteellä tietokoneavusteinen päätöksentukijärjestelmä (electric decision support systems) (EDSS) tarkoitetaan sellaisia terveydenhuollon järjestelmiä, joista hoitava ammattihenkilö saa neuvoja potilaskohtaisesti. Sitä voidaan pitää yksityisen terveydenhuollon ammattilaisen näkökulmasta ennen kaikkea työvälineenä, joka avustaa muistamisessa ja auttaa hallitsemaan suurta tietotulvaa. (Varonen, Kaila, Kunnamo & Mäntyranta 2005.)

Päätöksenteon järjestelmän toimivuutta, vaikuttavuutta ja kustannustehokkuutta terveydenhuollossa on myös tutkittu viime vuosina. Lääkäriseura Duodecim on ollut mukana hallinnoimassa ja johtamassa päätöksentukihanketta, jossa tapahtuu varsinainen päätöksentukijärjestelmän kehitys. Hankkeella pyritään luomaan laaja kansallinen kliininen päätöksentukijärjestelmä. Tämä kliininen päätöksenteon tukijärjestelmä on sitten mahdollista liittää nykyisiin sähköisiin potilaskertomuksiin, jotka on määritelty alun perin kansallisessa terveyshankkeessa. Tutkimuksen tausta on lähtenyt tiedon keräämisestä, aiemmista päätöksenteon tutkimuksista, ja siinä on pyritty myös huomiomaan terveydenhuollon eri ammattiryhmien näkemyksiä ja asenteita päätöksentukea kohtaan. Päätöksentuki tutkimus on oleellinen osa Päätöksentukihanketta (EBMcDS), jolla pystytään kehittämään ns. EBMEDS-palvelua. EBMEDS on lyhenne termistä (Evidence Based Medicine Electronic Decision Support), eli suomennettuna se tarkoittaa lääketieteelliseen näyttöön perustuvaa. Tämä jaetaan vielä neljään osahankkeeseen, joita ovat päätöksentukihankkeen valmistelu ja tuottaminen, päätöksentuen pilotointi ja päätöksentuen toimivuuden ja vaikuttavuuden tutkimus. Hankkeella ohjataan sähköistä päätöksentukea. (Duodecim 2019).

3.1 Lääketieteellisen hoidon laatu

Lääketieteellisen hoidon laatua on määritelty puhtaasti lääketieteellisestä näkökulmasta kirjallisuudessa varsin monilla tavoilla. Yksi määritelmä lähtee siitä näkökulmasta, että se on selvästi kliinisillä metodeilla mitattavissa oleva ominaisuus. Tähän tulokseen ovat päätyneet Rutstein, Berenberg, Chalmers & G.Child (1976). Heidän mukaan hoidon laatu on mitattavissa metodilla, joka laskee tarpeettomien sairauksien ja vammojen sekä ennen aikaisten tarpeettomien kuolemien tapaukset. Näiden ilmentyminen on listattu terveydentilana, ja siihen oikeuttaa aina kysymys siitä, miksi tämä tapahtui. Toiseksi määritelmän ominaispiirteeksi, he ovat valinneet sellaiset terveydentilat, joissa kriittisesti lisääntyneet sairauksien ja vammojen asteet, tai ennen aikaiset kuolemat, toimivat hoidon laadun merkkeinä. He painottavat vielä sitä, että sairauksien laajat kategoriat on pantu merkille, ja intensiivinen tutkimus voi paljastaa sellaisia erityispiirteitä, jotka toimivat varsinaisina terveyden merkkeinä. (Crutstein, Berenberg, Chamcers, G.Child 1976.)

Maailman terveysjärjestö eli WHO (World Health Organisation) puolestaan määrittelee terveydenhoidon laadun ulottuvuutena, jota terveydenhoidon palvelut tarjoavat yksilöille ja potilaiden populaatioille parantaakseen toivottuja terveystuloksia. Tämän saavuttamiseksi, terveydenhoidon tulee olla turvallista, tehokasta, oikea-aikaista, taloudellista, oikeudenmukaista ja ihmiskeskeistä. WHO myös määrittelee terveydenhoidon laatua olennaisesti kuvaavat käsitteet. Turvallisuus on WHO:n määritelmän mukaan terveydenhoidon toimittamista tavalla, joka minimoi riskit ja vahingot palvellakseen käyttäjiä, tähän sisältyy myös ehkäistävässä olevien vammojen ja hoitovirheiden vähentäminen. Tehokkuus määritellään palveluksi, jotka perustuvat palvelujen tarjoamiseen tieteellisestä tiedosta ja myös todistus-pohjaisiin suosituksiin. Oikea-aikaisuus puolestaan tarkoittaa viivästysten vähentämistä terveydenhoidon tarjoamisessa ja vastaanottamisessa. Taloudellisuus määritellään terveydenhoidon toimittamiseksi tavalla, joka maksimoi resurssien käytön ja välttää hukkaamista. Oikeudenmukaisuus tarkoittaa sellaisen terveydenhoidon toimittamista, joka ei eroa laadussa sen perusteella, mikä on henkilökohtaisesti tunnusomaista. Tällaisia piirteitä ovat sukupuoli, rotu, etnisyys, maantieteellinen sijainti tai sosioekonominen status. Tärkeää on myös ihmiskeskeisyys, joka ottaa huomioon mieltymykset ja pyrkimykset

yksittäisissä käyttäjien palveluissa ja heidän yhteisön kulttuureissa. (WHO:Quality of Care 2018.)

3.2 Krooniset sairaudet

WHO on myös julkaissut maailmanlaajuisen raportin ikääntymisestä ja terveydestä. Tässä raportissa nousee esille erityinen huoli maailmanlaajuisesta populaation ikääntymisestä. 60 vuotta ylittävien määrän odotetaan lisääntyvän vuoden 2015 12 %:tista 22 %:tiin vuoteen 2050 mennessä. Ennustetaankin, että 60 vuotta täyttäneiden ja niiden ylittäneiden määrä tulee saavuttamaan 2 biljoonaa seuraavan 35 vuoden kuluessa. Aasian maissa tilanne tulee olemaan kaikkein huonoimmillaan, ja esimerkiksi Japanissa tulee yli 30 %:tia väestöstä olemaan vanhusväestöä seuraavan 10 vuoden kuluttua. (Hwang & Cheng 2017:296).

Hwang ja Chen (2017) painottavat sitä, kuinka tämä kaikki kuormittaa terveydenhuollon järjestelmiä, samalla kun terveyden hoidon henkilöstöä ja laitteistoa on liian vähän. Yhtenä mahdollisena ratkaisuna on langattomien laskenta ja IoT teknologioiden yhdistäminen terveyden seuranta palveluihin. Hwang & Chen (2017) huomauttavat myös, että kroonisista sairauksista on tullut terveydenhoitoon yhä keskeisempiä ja vaikeampia ongelmia verraten niitä pelkästään väestön ikääntymisestä koituviin kustannuksiin terveydenhoidossa. He arvioivatkin, että yhteiskunnan talouden ja ympäristön muuttuessa kroonisista sairauksista on tullut päällimmäinen uhka ihmisten terveydelle. Sairastavuuden uskotaan myös lisääntyvän entisestään. Ihmisten ei ole kuitenkaan helppo käyttää julkisen terveyden hoidon palveluita ja laitteistoja, koska niistä on pulaa. (Hwang & Cheng 2017:296).

Hwang ja Chenin (2017) mukaan nämä lääketieteelliset järjestelmät ja systeemit keskittyvät myös yhä enemmän kaupunkeihin, joissa on paljon väestöä ja pääomaa. Tähän asti lääketieteelliset ja terveydenhoito järjestelmät esim. kehitysmaissa, ovat enimmäkseen tarkoitettu kestämään akuuttien sairauksien ja tarttuvien tautien kanssa. Kroonisten sairauksien ennalta ehkäisyä ja hoitoa ei siellä juuri painoteta. Hwang ja Chen (2017) huomauttavat myös, että elintason nousu lisää myös kroonisten sairauksien

määrää. Yhdysvalloissakin jopa 50 %:ttia ihmisistä kärsii yhdestä tai useammasta eri tasoisesta kroonisesta sairaudesta, ja jopa kahdeksankymmentä prosenttia lääketieteeseen rahoituksesta kuluu kroonisten sairauksien hoitoon. Tämä summa on lähes 18 % :ttia Yhdysvaltain bruttokansantuotteesta. Kalliit terveydenhoidon kustannukset asettavat valtavan taloudellisen taakan hallitukselle ja yhteiskunnalle. Hwang ja Chen (2017) ovat jakaneet kroonisten sairauksien syntymisen syyksi kolme pääasiasia. Nämä ovat muuttumattomat asiat, muuttuvat asiat ja ne asiat, joita on vaikea muuttaa. Ikä ja perinnöllisyys kuuluvat muuttumattomiin asioihin, ja ne selittävätkin 20 %:ttia kroonisten sairauksien syistä. Elinolosuhteet ovat aina kriittisiä fyysisille tiloille, ja niitä on myös vaikea muuttaa sattumanvaraisesti. (Hwang & Chen 2017:298).

WHO on vuonna 2015 tuottanut raportin kroonisista sairauksista. Se on listannut neljä suurinta kroonista sairaustyyppeä, ja ne ovat kardiovaskulaariset sairaudet, syövät, krooniset hengitystiesairaudet ja diabetes mellitus. Kardiovaskulaariset sairaudet ovat joukko sydän ja verisuonisairauksia, ja niihin sisältyy mm. sepelvaltimo sydänsairauksia. Ne ovat sairauksia, joita välittyy verisuonten kautta sydänlihakseen. Näihin sisältyy esim. aivosokeri sairaudet, ja verisuonisairaudet, jotka välittyvät käsiin ja jalkoihin. Niitä ovat reumaattinen sydänsairaus, sydänlihakseen ja läppään tulleet vahingot reumaattisesta kuumesta, jonka on aiheuttanut streptokokkinen bakteeri. On myös synnynnäisiä sydänsairauksia, sydämen rakenteen epämuodostumia, jotka ilmenevät syntymästä. Yleisiä ovat myös syvä laskimon verisuonitukos ja keuhkoembolia. Keuhkoemboliassa esiintyy verikokkareita jalkalaskimossa, joita voi irrottautua ja siirtyä sydämeen. Raportti osoittaa, että vuonna 2012 enemmistö tarttuvista kuolemaan johtavista taudeista alle 70 vuotiaiden joukossa, johtui näistä neljästä sairaudesta. (Hwang & Chen 2017:296).

Kardiovaskulaariset sairaudet aiheuttivat suurimman osuuden kroonisista kuolemista alle 70 vuotiaissa. Osuus oli 37 %:ttia, seuraavaksi suurin kuolin syy oli syöpä 27 %:ttia, ja sitten tulivat krooniset hengitystiesairaudet 8%:ttia. Diabeteksen osuus oli 4 %:ttia kuolemista, ja muiden tekijöiden osuus oli noin 24 %:ttia kaikista kuolemista. Ympäristötekijöihin lisättyinä maailmanlaajuiset sosiaaliset ja taloudelliset trendit, kuten väestön ikääntyminen, on suurin yksittäinen syy kroonisista sairauksista kärsivien potilaiden määrän kasvuun. Kaupungistuminen aiheuttaa myös ympäristön saasteiden runsasta lisääntymistä. Keuhkosairaudet ovatkin lisääntyneet, koska raskas PM 2.5 ja utu

ovat stimuloineet niiden syntymistä. Toinen vaikuttava puoli asiassa on globalisaatio, ja sen seurauksena ihmiset ovat tottuneetkin kommunikoidaan ystävien kanssa mobiili laitteilla. Uudet teknologiat yhteiskunnassa, matkapuhelimissa ja verkoissa tekevät kaupunkimaisen urbaanin elämän sopivammaksi. Hwang ja Chen (2017) huomauttavat kuitenkin myös, että tämä voi myös luoda epäterveellisiä elämäntapoja. Tiedetään, että esimerkiksi tietokoneen edessä istuminen liian pitkään ja jatkuva liikunnan puute lisäävät ylipaino-ongelmia ym. (Hwang & Chen 2017:296).

Viimeisimmän WHO:n raportin mukaan terveyttä määrää viisi eri tekijää. Itseasiassa terveyspalvelut ja muut vastaavanlaiset voivat ratkaista vain 10 % lääketieteellisistä ongelmista. Viisikymmentä prosenttia on riippuvaisia elämäntavoista, kuten yleisistä elintavoista, ruokailutottumuksista ja liikunnasta. Kaksikymmentä prosenttia riippuu ympäristöstä ja jäljellä oleva 20 % johtuu biologisista tekijöistä, kuten perimästä. Tämä kaikki osoittaa, että suurimmat syyt terveysongelmiin ovat elämäntapoihin liittyviä. Hwang ja Chen (2017) painottavatkin, että kaikkien pitäisi keskittyä terveyden tarkkailuun jälkihoidon sijaista. Johtuen pitkäaikaista ominaispiirteistä, krooniset sairaudet eivät ole tarpeeksi akuutteja hoidettavaksi sairaalassa. Siksi kansalliset hallitukset käyttävätkin paljon rahaa tämän ongelman ratkaisemiseksi, ja jatkuva terveyden monitorointi on ratkaisevaa tämän ongelman ratkaisemisessa. (Hwang & Chen 2017:296).

4.Diabetes mellitus

Diabetes mellitus on vakava krooninen sairaus. Sairauden aiheuttaa perinnöllinen ja hankittu puute haiman insuliinin tuotannossa. Aiheuttaja voi myös olla pelkästään hankittu puute insuliinin tuotannossa, ja haiman tuottaman insuliinin tehottomuus voi myös aiheuttaa sairastumista diabetes mellitukseen. Insuliinin puute johtaa aina suureneviin glukoosipitoisuuksiin veressä. Tällöin monia kehon järjestelmiä vahingoittuu, näitä ovat etenkin hermot ja verisuonet. Diabeteksessa esiintyy kaksi muotoa. Diabeteksen tyyppissä yksi haima ei tuota lainkaan insuliinia, joka on eloonjäämiselle välttämätöntä. Tällainen diabeteksen muoto kehittyy usein lapsilla ja nuorilla. Tyypin 1 diabetes tunnistettiin aiemmin insuliiniriippuvaisena. Tyypin 2 diabetes puolestaan johtuu kehon kyvyttömyydestä reagoida oikealla tavalla haiman tuottamaan insuliinin toimintaan. Tämä diabetes tyyppi on paljon yleisempää, ja se muodostaakin lähes 10 % kaikista maailmanlaajuisista diabeteksen tapauksista. Useimmiten tämä esiintyy aikuisilla, ja sitä on kuitenkin havaittu aiempaa enemmän myös nuorilla. On myös todistettu, että tietyntyypiset geneettiset markkerit lisäävät tyypin 1 diabeteksen kehittymisen riskiä. Tyypin 2 diabetesta pidetään myös hyvin perinnöllisenä. Eräissä populaatioissa esiintyy myös geenejä, joihin liittyy jatkuvaa kasvanutta riskiä tyypin 2 diabetekselle. Voidaan todeta, että molemmat diabeteksen tyypit ovat hyvin monimutkaisia sairauksia. Niitä aiheuttavat mutaatiot ja ympäristötekijät, nämä mutaatiot ilmenevät useassa geenissä. (WHO: diabetes mellitus 2016.)

Lääkäri voi diagnosoida diabeteksen tiettyjen biomarkkereiden mukaan. Glukoosin sietokyky (IGT) on vähentynyt ja paasto glykemia (IFG) on heikentynyt. Mikäli ne alittavat veren glukoosipitoisuuden normaalialueen, voidaan diabetes diagnosoida. Henkilöillä, jotka sairastavat IGT:ä ja IFG:tä, tai pelkästään IFG:tä, on huomattavasti suurempi riski sairastua diabetekseen ja sydän- ja verisuonitautien sairauksiin kuin normaalin glukoosintoleranssin omaavilla henkilöillä. Kliinisestä interventtiosta koituu eniten hyötyä potilaille, joiden glukoosi intoleranssi on kohtuullisella tasolla. Diabeteksen oireet ilmenevät voimakkaina, heikkoina, tai niitä ei ole välttämättä ollenkaan. Tyypin 1 diabeteksessa esiintyy oireita, kuten virtsan liiallista

erittymistä(polyuria), janoa(polydipsia) väsymistä ja laihtumista. Oireet ovat vähemmän merkittäviä tyypin 2 diabeteksessa. Tämän muodon diabeteksessa ei välttämättä esiinny varhaisia oireita lainkaan, ja tautikin diagnosoidaan yleensä vain useita vuosia sen alkamisen jälkeen, silloin kun komplikaatioita on jo ilmentynyt. (WHO: diabetes mellitus 2016.)

Diabetes on varsin yleinen sairaus. Hiljattain kootut tiedot osoittavat, että lähes 150 miljoonalla ihmisellä on diabetes mellitus maailmanlaajuisesti tarkasteltuna. Tämän määrän oletetaan jopa kaksinkertaistuvan vuoteen 2025 mennessä. Useimpien diabeetikoiden ikä tulee olemaan myös varsin korkea. Diabeetikoiden oletetaan olevan kehittyvissä maissa 65-vuotiaita tai jopa vanhempia. Kehitysmaissa suurin osa diabeetikoista kuuluu 45-64-vuotiaiden ikäryhmään, ja vaikutukset tulevat olemaan suurimpia tuottavina vuosina. Diabetes mellituksen diagnosoimisen helpottamiseksi WHO on julkaissut suositukset veren glukoosipitoisuuden diagnostisesta tasosta, ja tätä tasoa on muutettu viimeksi vuonna 1999. Diabeteksen hoito voi tapahtua farmakologisesti tai ei-farmakologisesti. Ei farmakologisen diabeteksen hoidon perusta on oikeanlaisessa ruokavaliossa ja liikunnassa. Kuitenkin noin 40 % diabetesta sairastavista tarvitsee suun kautta annettavia aineita, jotta verensokeria voitaisiin tyydyttävästi kontrolloida ja lähes 40 %: tta tarvitsee insuliinipistoksia. Tämän Frederic Bautingin ja Charles Bestin vuonna 1921 eristämä hormoni mullisti diabeteksen hoidon ja komplikaatioiden ehkäisyn. Se muutti aiemmin varmaan kuolemaan johtavan sairauden sellaiseksi, että potilaalla oli mahdollisuus pitkäaikaiseen selviytymiseen. (WHO: diabetes mellitus 2016)

Tyypin 1 diabetes tekee ihmisen yleensä täysin riippuvaiseksi insuliini pistoksista, niillä turvataankin heidän elossa selviäminen. Nämä ihmiset tarvitsevat päivittäin insuliinia. Diabetesta sairastavista suurin osa kuitenkin sairastaa tyypin 2 diabetesta. Tällöin heidän eloonjäämisensä ei riipu insuliinista, kuitenkin noin kolmasosa heistä tarvitsee insuliinia alentaakseen verensokeritasoaan. Insuliinia ei kuitenkaan ole kaikille sairastaville tarjolla, ja se on myös varsin kallis lääke monissa köyhissä maissa siitä huolimatta, että WHO on luetellut sen olennaiseksi lääkkeeksi. Tästä onkin muodostunut kansainvälisten

terveysvirastojen ja kansallisten terveystieteiden huoli. (WHO: diabetes mellitus 2016)

4.1 Diabeteksen esiintyvyys ja hoitokäytännöt

Diabetes mellitus on varsin vakava sairaus, johon liittyy paljon ikäviä komplikaatioita. Yksi niistä on ns. diabeettinen retinopatia, ja siitä on muodostunut yksi johtavista sokeuden ja näkövammaisuuden syistä. Diabetes mellituksen ilmenemiseen on liittynyt yhtenä osana verkkokalvon pienten verisuonten vaurioituminen, ja näiden mekanismien kautta voi tapahtua varsinainen näön menetys. Tutkimusnäytön perusteella tehdyt havainnot viittaavat siihen, että kärsittyään 15 vuotta diabeteksestä noin 2 % ihmisistä on tullut sokeaksi ja 10 %:lla on vakava näkövamma. Diabeetikoilla on myös yleisemmin glukooman ja kaihin aiheuttamaa näön menetystä kuin diabetesta sairastamattomalla verrokiryhmällä. (WHO: diabetes mellitus 2016)

Diabeettisen retinopatian eteneminen ja alkaminen voi hidastua käyttämällä hyväksi hyvää aineenvaihdunnan hallintaa. Diabetesta sairastavilla henkilöillä voidaan ehkäistä näön tai näkövammaisuuden menetystä havaitsemalla varhain näköä uhkaava retinopatia ja myös hyvällä hoidolla. Mikäli retinopatia on edistynyt, tärkeäksi nousee säännölliset silmäkokeet ja myöskin oikea-aikainen laserhoito tai leikkaus. Tutkimusnäyttö osoittaa, että kehittyneissäkin maissa suuri osa hoitoa tarvitsevista ei saa tällaista hoitoa yleisen ammatillisen tietoisuuden ja hoitolaitoksen puuttumisen vuoksi. Monissa kehitysmaissa, joissa diabetes esiintyy yleisesti, ei tällainen hoito pääse väestön keskuuteen lainkaan. Diabetes aiheuttaa myös munuaisten vajaatoimintaa, sen esiintymistiheys tosin vaihtelee paljon populaatioiden välillä, ja se liittyy myös keston ja vakavuuteen. Munuaisten vaurioitumisen hidastumiseksi on monia toimenpiteitä. Näitä ovat korkean veren glukoosin hallinta, korkea verenpaineen hallinta, interventio lääkkeiden kanssa munuaisvaurion varhaisessa vaiheessa, ja myös ruokavalion proteiinia pitäisi rajata. Diabeettisen munuaissairauden seulonta ja varhainen havaitseminen ovat tärkeitä keinoja ennaltaehkäisyssä. (WHO: diabetes mellitus 2016.)

Diabetesta sairastavien kuolemista teollisuusmaissa 50 %:ttia aiheutuu sydänsairauksista. Sydänsairauksien riskitekijät diabeetikoilla ovat tupakointi, lihavuus, korkea verenpaine ja korkea seerumin kolesteroli. Diabeettista neuropatiaa voidaan pitää todennäköisesti yleisimpänä komplikaationa diabeteksessa. Tutkimusnäyttö osoittaa, että diabeetikoista jopa 50 %:ttia kärsii jonkin verran näistä komplikaatioista. Tässä tilassa tärkeimmät riskitekijät ovat kohonnut verensokerin taso ja kesto. Neuropatia voi pahimmillaan johtaa aistinhäviöön ja raajojen vaurioitumiseen. Sitä pidetään myös merkittävimpänä syynä diabetesta sairastavien miesten impotenssiin. Diabetes voi aiheuttaa myös muutoksia verisuoniin ja hermoihin. Tämä aiheuttaa diabeettista jalkatautiä, ja se voi johtaa haavaumiin ja usein myöhemmin myös raajan amputointiin. Se onkin yksi diabeteksen kalleimmista komplikaatioista, tämä esiintyy erityisesti sellaisissa yhteisöissä, joissa käytetään puutteellisia jalkineita. Verisuonten ja neurologisten sairauksien prosessi on aiheuttamassa tätä. Diabetesta voidaankin pitää yleisimpänä syynä alaraajojen ns. ei-traumaattiseen amputoimiseen, tämä voidaan kuitenkin estää tavanomaisella jalkahoidolla ja tarkastuksella. (WHO: diabetes mellitus 2016.)

Diabetekseen ennaltaehkäisystä tehdyt tutkimukset osittavat, että taudin puhkeamista on mahdollista ennaltaehkäistä ja viivästyttää lievässä glukoosi-intorelansissa (IGT). Tämän vahvistavat suurilla väestöpopulaatioilla tehdyt tutkimukset Suomessa, USA:ssa ja Kiinassa. Niissä todettiin, että vain puoli tuntia kävelyä päivässä ja painon lasku pienentävät diabeteksen esiintyvyyttä yli puolella. Diabetesta pidetään vakavana ja kalliina sairautena. Sen myös odotetaan yleistyvän erityisesti kehitysmaissa ja heikossa asemassa olevissa vähemmistöissä. On kuitenkin olemassa keinoja estää sen puhkeaminen, seurata diabeteksen edistymistä, ja on myös mahdollista lisätä julkista ja ammatillista tietoisuutta sen oireista ja riskeistä. Tämän uskotaan olevan ehkä tärkein askel ennalta ehkäisyssä ja valvonnassa. (WHO: diabetes mellitus 2016.)

Yhdysvalloissa on The American Diabetes Association (ADA) antanut omat hoitokäytäntösuositukset diabeteksen hoitoon. ADA on diabeteksen hyvää ja laadukasta hoitoa kehittävä järjestö. Siinä on monitieteinen asiantuntijakomitea, joka tarjoaa komponentteja mm. diabeteksen hoidolle ja se myös päivittää hoitostandardeja vuosittain tai useamminkin. ADA on lisäksi antanut hoitosuosituksia diabeteksen hoitoa ja

kansanterveyden edistämistä varten. Hoitosuosituksiin kuuluu varmistaa, että hoito päätökset tehdään ajoissa ja että ne pohjautuvat todistusperusteisiin suosituksiin. Ne myös tehdään yhteistyössä potilaiden kanssa perustuen yksilöllisiin mieltymyksiin, prognooseihin ja sairauksiin. Suositusten tarkoitus on suunnata diabeteksen hallinnoinnin lähestymistapa kroonisen hoitomallin mukaiseksi. Siinä painotetaan tuottavaa interaktiota valmistetun ennakoivan hoitotyöryhmän asioihin perehtyneiden aktivoitujen potilaiden kanssa. Hoitojärjestelmien pitäisi helpottaa työryhmä perusteista hoitoa, potilasrekistereitä, päätöksentekoa työkaluja ja yhteisön mukaantuloa huomioitaessa potilastarpeita. Pitää myös olla sujuvaa arviointia diabetes hoidon laadusta ja luoda laadun parantamisen strategioita, joiden pitäisi yhdistää luotettavaa dataa edistääkseen parempia hoitoprosesseja ja terveyttä painottaen samalla kustannustehokkuutta. (Diabetes Care 2019:57.)

Kliinisten suositusten soveltaminen terveydenhoitoon antaa työkaluja, joilla voidaan hallita populaation terveyttä kokonaisuudessaan, ja diabeteksen hoito täytyy myös olla yksilöllisesti jokaiselle potilaalle suunniteltu. Näin ollen kansanterveyden parantaminen vaatii sekä järjestelmätason että potilas-tason lähestymistapoja. Yhdysvaltojen diabetesyhdistys korostaa potilaskeskeisen hoidon tekevän yksilöä vastuunottovaksi ja hänen yksilöllisiä tarpeita, mieltymyksiä ja arvoja kunnioittavaksi. Potilaskeskeinen hoito pyrkii myös takaamaan sen, että potilaan arvot ohjaavat kaikkea kliinistä päätöksentekoa. (Institute of Medicine Committee on Quality of Health Care in America 2001.)

4.2 Diabeteksen syyt ja seuraukset

Tärkeimmäksi yksittäiseksi syyksi diabeteksen yleistymisessä on noussut länsimaalaiskehitys, joka tarkoittaa Helsingin Yliopiston dosentin ja Itä-Suomen yliopiston epidemiologian professori emeritus Jukka T.Salonen mukaan ennen kaikkea ruokavalion muuttumista epäedulliseksi. Diabeteksen lisääntymisen uskotaan olevan jopa nopeampaa kuin muistisairauksien ja Alzheimerin. Salonen huomauttaa myös, että ylipainon lisääntymien on perinteisillä lääketieteen keinoilla täysin mahdoton hoitaa. Tämä onkin hänen mukaansa enemmänkin yhteiskuntasidonnainen ja kulttuurinen

ongelma, ja siihen pitäisikin Salosen mukaan löytyä poliittinen lähestymistapa. Nykyään myös uskotaan perusterveydenhuollon pystyvän vaikuttamaan ennaltaehkäisevästi diabetekseen. Siihen ei Salonen mukaan kuitenkaan kiinnitetä julkisella puolella juuri huomiota. Diabeteksen maailmanlaajuisesti ehkäisemiseksi on esimerkiksi YK:n toimesta asetettu konkreettisia tavoitteita. YK on asettanut tavoitteeksi pysäyttää aikuisten diabeteksen yleistymisen vuoden 2010 tasolle viimeistään jo vuonna 2025. Tilanteen kuitenkin kehittyessä ja jatkuessa 2000-luvun trendien suuntaisesti, tähän tavoitteeseen on erittäin epätodennäköitä päästä. Tavoitteeseen pääsyn todennäköisyys on maailmanlaajuisesti tarkasteltuna miehillä alle yhden prosentin, ja naisilla puolestaan tasan yksi prosenti. (Torvinen 2016.)

Diabetes aiheuttaa maailmanlaajuisesti merkittäviä kustannuksia terveydenhuoltoon. On arvioitu, että yksinomaan Suomessa diabeteksen hoitoon käytetään 15 prosenttia terveydenhuollon menoista. Diabeteksestä aiheutuvat hoitokulut ovat joissain maissa niin korkeat, että terveydenhuolto on melkein romahtamassa diabeteskuluihin. Maailmanlaajuiset kustannukset terveydenhuollon menoihin olivat 720 miljardia euroa vuodessa. Diabetes on joissain maissa niin yleinen sairaus, että se saattaa tuhota niiden koko terveydenhuoltojärjestelmän. Laajassa väestöpohjaisessa tutkimuksessa on todettu, että etenkin miesten diabetes on yleistymässä merkittävästi. The Lancet lehden 2016 julkaisemassa vertailtiin diabeteksen yleisyyttä miehissä ja naisissa vuosien 1980 ja 2014 välillä. Tutkimuksessa oli käytössä paljon dataa, siinä oli dataa kerättyä kaikkiaan 751 eri väestötutkimuksesta. Näin ollen mukana oli 4,4 miljoonaa aikuista lähes kaikista maailman maista. On huomioitava, että tutkimuksen aineistossa ei oltu mitenkään erotettu tyyppien 1 ja 2 diabetesta toisistaan. Tutkimuksen tuloksen perusteella voidaan todeta, että miesten diabetes on maailmanlaajuisesti tarkasteltuna yli kaksinkertaistunut viimeisten 34 vuoden aikana. Tarkasteltaessa vuotta 1980 niin voidaan todeta, että diabetesta sairasti 4,3 prosenttia miehistä, ja 2014 diabetesta sairasti jo yhdeksän prosenttia miehistä. Tarkasteltaessa naisten osuutta diabeteksessa voidaan havaita, että sitä sairasti vuonna 1980 viisi prosenttia, ja vuoteen 2014 mennessä se oli lisääntynyt 7,9 prosenttia. Maailmanlaajuisesti tarkasteltuna diabetes ei ole vähentynyt yhdessäkään maailman maassa suhteessa vuoteen 1980. (Torvinen 2016.)

Vuonna 1980 diabetesta sairasti 108 miljoonaa aikuista, ja nämä luvut kasvoivat vuonna 2014 peräti 420 miljoonaa. On arvioitu, että määrän kasvusta lähes 40 prosenttia voidaan selittää johtuvaksi väestökasvusta ja ikääntymisestä, ja 29 % selittyisi pelkästään taudin yleistymisellä. Jäljelle jäävä osuus selittyisi näiden kahden syyn vuorovaikutuksella. Tässä tutkimuksessa oli mukana Maailman terveysjärjestö WHO:n lisäksi noin 500 eri tutkijaa eripuolilta maailmaa. Kuopiosta valmistunut sepelvaltimotaudin vaaratekijätutkimus SVVT oli tutkimusmateriaalina mukana. Suurimmat vaikeudet diabeteksen yleistymisessä koskettavat erityisesti kehittyviä maita. Suomessakin diabetes on yleistynyt erittäin nopeasti, ja Suomessa diabetesta sairastaa lähes puoli miljoonaa suomalaista. On myös huomioitava, että heistä lähes 150 000 ei tiedä sairastavansa diabetesta. Diabetesta sairastavista tyyppi 1 diabetesta sairastaa Suomessa 50000 ihmistä. Isoin ongelma Suomessa on tyyppi 1 diabetes, ja sen osuus onkin maailman suurin. The Lancetin julkaisemassa tutkimuksessa mukana ollut Helsingin Yliopiston dosentti ja Itä-Suomen yliopiston epidemiologian professori emeritus Jukka T.Salonen toteaa, että varsinaista syytä ei vielääkään tiedetä vaikka aiheesta on tehty jo runsaasti tutkimusta. Suomessakin diabetes on nopeasti yleistynyt ja sen hoito vastaa 15 %: ttiä kaikista terveydenhuollon menoista. (Torvinen 2016.)

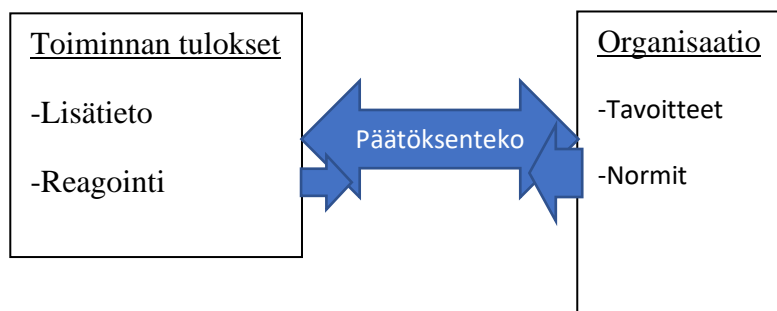
Salonen huomauttaa, että suurimmat ongelmat tauti aiheuttaa kuitenkin kehittyvissä maissa. Suurimmat ongelmat esiintyvät erityisesti Lähi-Idässä ja Aasiassa. Kiinan ja Intian tilanne on esimerkiksi todella huono. Tutkimukset osoittavat, että Intiassa diabetes maksaa jopa 63 miljardia euroa vuodessa. Kiinassa puolestaan kustannukset nousevat 148 miljardiin ja Yhdysvalloissa 92 miljardiin euroon. Näissä laskelmissa ei ole edes huomioitu diabeteksen takia menetettyjä työpäiviä, näin ollen todelliset kustannukset saattavat kohota todennäköisesti paljon suuremmiksi. Salosen mukaan useissa maissa terveysarvoja ei juuri mietitä varsinaisessa yhteiskunnallisessa suunnittelussa ollenkaan. Täten pelkästään diabeteksen hoito voi aiheuttaa suuria vahinkoja maan taloudelle. Suomessa asiantuntijoiden esille tuomia mielipiteitä ei juuri kuitenkaan huomioida terveyspolitiikassa mitenkään. (Torvinen 2016.)

4.3 Kliininen päätöksenteko diabetes mellituksen hoidossa

Higgs, Smith ja Ellis (2008) ovat määritelleet kliinisen päätöksenteko käsitteen sisältöä ja merkitystä tieteellisenä terminä. He painottavat sitä, että tutkimus kliinisessä päättelyssä on keskittynyt vahvasti siihen sisältyvien prosessien näkökulmiin. Heidän mukaan kliininen päätöksenteko on kliinisen päättelyn tulos ja komponentti. Sille on myös annettu keskeinen asema terveydenhoito alan ammattilaisten ammatin harjoittamisessa. Se on myös elintärkeä niiden tekijöiden tunnistamisessa ja ymmärtämisessä, jotka vaikuttavat positiivisesti ja negatiivisesti päätöksentekoon terveydenhoidon laadun näkökulmasta. Tämä tulee esille tilanteissa, joissa päätöksentekoon vaikuttavat tekijät edesauttavat sellaisten virheiden syntymistä, joilla on mahdollisia epäsuotuisia vaikutuksia terveydenhoidon asiakkaille. Päätöksentekoon vaikuttavat tekijät voivat myös edistää terveydenhoito kokemuksia, tuloksia ja seurauksia. (Higgs, Jones, Loftus & Christensen 2008.)

Thomas (1991) ym. toteavat kliinisen päätöksenteon olevan laaja käsite. Se sovelletaan prosessiin, jossa tehdään valintoja vaihtoehtojen välillä. Tämä tapahtuu toiminnan kuluessa. Terveydenhoidon ammattilaisten tekemä kliininen päätöksenteko on puolestaan paljon monimutkaisempi, kompleksisempi prosessi, ja se vaatii yksilöltä paljon enemmän kuin tarkoin määriteltujen valintojen tekemistä rajoitettujen vaihtoehtojen välillä. Smith ym.(2008) huomauttavat, että terveydenhoito ammattilaisilta vaaditaan monenlaisia päätöksiä, jotka tarkentavat mm. diagnooseja, interventioita, interaktioita ja arviointeja dynaamisessa yhteydessä käyttäen monipuolista tietämystä (sisältyen lisääntynyt todistusperusteinen kirjallisuuden pääsisällys). Tähän sisältyy useita muuttujia ja yksiköitä. Tämän lisäksi kliinisiä päätöksiä luonnehtii epävarmuuden tilanteet, missä kaikki tarvittava tieto niiden tekemiseksi on tai voi olla. Tässä yhteydessä kliinisen päätöksenteon tekeminen on harvoin yksittäinen päätös, joka on tehty yksittäisistä vaihtoehtoista, joissa yksi päätös voidaan eristää muista. Päätökset ovat pikemminkin sulautettuna päätös toiminta sykleihin, joissa tilanteet purkautuu, ja joissa päätökset sekä toiminta vaikuttavat toinen toisiinsa. (Higgs, Jones, Loftus & Christensen 2008.)

Orasanu ja Connolly (1993) kuvasivat päätöksenteon luonnetta erityisessä paikassa, kuten terveydenhoito paikassa. Heidän mukaan ongelmat ovat vahingollisesti jäsennettyjä ja epäselväksi tehtyjä, epätäydellisen dynaamisen tiedon ja monien vuorovaikutus tavoitteiden läsnä ollessa. Päätöksenteko ympäristö on epävarma, ja saattaa muuttua silloin kun päätökset tehdään. Tavoitteet saattavat myös muuttua tai olla huonosti määriteltyjä tai kilpailevia. Päätöksenteko ilmenee toimita-takaisinkytkentä silmukassa, missä toiminnan tulokset vaikuttavat ja muodostavat lisää tietoa, joihin päätöksentekijöiden täytyy reagoida, jotta voidaan tehdä lisää päätöksiä. Päätökset sisältävät aikapaine tekijöitä, henkilökohtaista stressiä ja hyvin merkittäviä seurauksia osanottajille. Useat toimijat ovat yhteydessä eri roolien kanssa. Organisatoriset tavoitteet ja normit vaikuttavat päätöksen tekoon. Alapuolella on kuva kliinisestä päätöksenteosta, joka tehdään takaisinkytkentä silmukassa, epävarmassa ympäristössä organisaatioiden tavoitteet ja normit huomioon ottaen. (Higgs, Jones, Loftus & Christensen 2008.)



Kuva 4. Kliinisen päätöksenteon malli terveydenhoito paikassa. (Orsanu & Connolly 1993).

Viime aikoina painotus on ollut kliiniseen päätöksenteon tekeminen yhteistyöprosessina käsittäen jaetun ja rinnakkaisen päätöksenteon potilaiden ja terveydenhoito ammattilaistyöryhmien välillä. Päätöksenteon yhteistyön laatu tarkoittaa, että mikä tahansa tarkastelu ammatinharjoittajien kliiniseen päätöksentekoon vaikuttaviin asioihin voi myös tarkastella työryhmien ja potilaiden päätöksentekoon vaikuttavia asioita. (Higgs, Jones, Loftus & Christensen 2008.)

Kliiniseen päätöksentekoprosessiin vaikuttaa nykyään yhä enemmän koneoppiminen ja datalouhinta. Tämä mahdollistaa yhä yksilöllisemmän hoidon ja parempien lääkehoitojen kehittämisen. Lääkehoitojen tullessa yhä yksilöllisemmäksi tulee jatkuvasti siirtymistä pois ns. ”kaikille yhtä ja samaa” periaatteesta kohti ajattelutapaa, jossa mm. lääkehoitoa annetaan siitä hyötyvälle. (Ruokoniemi & Rannanheimo 2018).

Tekoäly ei kuitenkaan kykene toimimaan yksin vaan se tarvitsee aina oppaakseen syväoppineen ihmisen. Tekoälyn nykyisilläkin sovelluksilla pystytään jo kertomaan mainostajille, mihin kellonaikaan olemme erityisen alttiina nettiostoksille, missä kaupassa käymme ja myös sen paljonko haluamme ostoksistamme maksaa. Niitä voidaan hyödyntää myös esim. syöpähoidossa siten, että ne hälyttävät hoitavalle lääkärille syöpähoitojen aikana ilmentyneistä hoidon vaarantavista haittavaikutuksista. Ne suosittelevat tällöin myös sopivia lääkehoitoratkaisuja ja kykenevät seulomaan massoittain tieteellisiä julkaisuja, ja tällöin myös uusien innovatiivisten lääkkeiden kehityksen tarpeisiin voidaan esittää hypoteeseja. Monimutkaisen datan tehokas ja nopea järjesteleminen on mahdollista syväoppivien neuroverkkojen avulla, ne ovatkin keskeisiä mm. kuvien ja tekstien louhinnassa. Tekoäly hyödyntää asiantuntijoiden tuottamia tietoarkistoja ja kykenee niistä saadun aineiston avulla kehittämään algoritminsa positiivista ennustearvoa pitkälle. Ihmisen kognitiivisilla kyvyillä on täysin mahdotonta vähentää sellaisella nopeudella väärin positiivisten ja väärin negatiivisten tulkintojen lukumäärää. On kuitenkin otettava huomioon, että tekoäly kykenee tekemään päätöksiä vain ja ainoastaan sen mukaan, minkä laatuista dataa sille on syötetty ja kuinka toimiva algoritmi sitä ohjaa. Tekoäly pystyy tuottamaan suuresta terveystiedonmassasta eri muuttujien välisiä assosiaatioita ihmisten käyttöön. Ongelmaksi tässä muodostuu se, että tiedonlouhinta ei vielä pysty takaamaan laadukasta tietoon perustuvaa päätöksentekoa. (Ruokoniemi & Rannanheimo 2018).

Terveysteknologinen kehitys ja tutkittu tieto saattavat kohdata oikeissa olosuhteissa. Tämä tapahtuu aina potilaan parhaaksi. Diabeteksessa käytettävien verensokerimittareiden teknologisen kehityksen edistyessä, on yhä enemmän siirrytty pois ”kaikille yhtä samaa” periaatteen mukaisista ajoittain tehdyistä sormenpäänmittauksista, henkilökohtaiseen ja yksilölliseen verensokeriseurantaan ja yhä

yksilöllisempään päätöksentekoon. Nykyiset älymittareiden erilaiset lisäominaisuudet tekevät mahdolliseksi erilaisia toimintoja aina tulosten mobiilianalyysistä alkaen jatkuvaan reaaliaikaiseen glukoosiseurantaan saakka. Samanlaisia teknologisia edistysaskeleita on tehty mm. Parkinsonin taudin tilavaihteluissa. Nykyisin on myös alettu yhä enemmän puhumaan siitä, kuinka tärkeää olisi ottaa data haltuun. Tiedetään, että parhaimmillaan big data ja sitä louhiva tekoäly mahdollistavat järkevän, yksilöllisen ja laadukkaan lääkehoidon. (Ruokoniemi & Rannanheimo 2018.)

Tamhane, Rodriguez-Gutierrez, Hardraves ja Monitori (2015) ovat kirjoittaneet jaetusta päätöksenteosta diabeteksen hoidossa. Jaettu päätöksenteko (Shared decision-making (SDM), tarkoittaa yhteistyöprosessia, jossa potilaat ja klinikot työskentelevät yhdessä harkitussa keskustelussa. Tämän keskustelun tarkoitus on tunnistaa järkeviä johtamisen vaihtoehtoja, jotka sopivat parhaiten ja osoittavat potilaan ainutlaatuisen tilanteen. Jaettu päätöksenteko tukee potilaskeskeistä tutkimuksen kääntämistä käytäntöön. Se myös auttaa toteuttamaan todistusperustaisen lääketieteen periaatetta; siinä todistus on välttämätön, mutta ei koskaan riittävää tekemään kliinistä päätöstä, koska siihen tarvitaan potilaan arvoja ja vaaditaan myös asiayhteys. Jaetun päätöksenteon keskustelut perustuvat klinikon ja potilaan yhteistyölle. Se perustuu todisteisiin, jotka ottavat huomioon erilaiset hoitovaihtoehdot, ja tarkastelevat vaihtoehtoja, ja tarkastelevat vaihtoehtoja erilaisten arvojen ja mieltymysten valossa sekä potilaan taustasta. Nykyisin uskotaan, että jaettu päätöksenteko on sopiva diabeteksen hoidolle, koska diabeteshoito vaatii usein johtamisen vaihtoehtojen harkintaa. Ne eroavat siitä tavasta, jolla on merkitystä potilaille, ja ne myös asettavat merkittäviä vaatimuksia potilaan elämälle ja elämäntavalle. Viime vuosikymmeninä jaettu päätöksenteko on osoittautunut onnistuneeksi vaihtoehdoksi toteuttamiskelpoisuudessa ja hyödyllisyydessä todisteiden jakamisessa potilaiden kanssa, ja potilaat ovat myös saaneet osallistua päätöksentekoon lääkäreiden kanssa. (Tamhane 2015.)

5. Tutkimusmetodi ja malli

Tässä kappaleessa esittelen tarkemmin tutkimuksessa käytettävän metodin ja mallin. Tutkimusmallina käytetään arkistotutkimusta. Tutkimuksen tarkoitus on analysoida aihepiiristä julkaistua dataa, jota on tallennettu eri tietokantoihin. Arkistotutkimuksen metodeihin sisältyy suuri alue toimia, joita on sovellettu hyödyntämään dokumenttien tutkimuksia ja tekstuaalisia materiaaleja. Näitä on yleensä tuotettu organisaatioiden toimesta, ja ne voivat myös koskea organisaatioita. Perinteisessä mielessä arkistometodit ovat niitä, jotka koskevat historia tutkimukseen liittyviä dokumentteja. Tällöin dokumentit on luotu jossain suhteellisen kauas menneisyyteen ajoittuvassa ajankohdassa antaen mahdollisuuden päästä käsiksi organisaatioihin, yksilöihin ja tapahtumiin, joihin ei muuten olisi mahdollista perehtyä syvällisemmin. Arkistotutkimus metodeita voivat käyttää myös tutkijat, jotka tekevät ensisijaisesti ei-historiallisten dokumenttien tutkimusta, tällöin he tutkivat tekstejä, joita ovat tuottaneet nykyaikaiset organisaatiot. Julkaisut tekstit voivat myös käsitellä näitä aikalaisorganisaatioita. Arkistotutkimusmetodit voivat usein myös toimia työkaluina täydentämään muita tutkimusstrategioita. Arkistotutkimus metodeita voidaan soveltaa analysoimaan digitaalisia tekstejä, kuten elektronisia tietokantoja, sähköposteja ja web sivuja. (Ventresca & Mohr 2017).

5.1 Integroiva kirjallisuuskatsaus

Tutkimusmenetelmänä tutkimuksessa on integroiva kirjallisuuskatsaus. Sen avulla pyritään siis vastaamaan määriteltyihin tutkimuskysymyksiin. Tarkoituksena on pyrkiä ratkaisemaan tutkimuksen tutkimusongelmat olemassa olevan tutkimuskirjallisuuden avulla. Kirjallisuuskatsausta pidetään tutkimusmetodin ja tekniikkana, jossa tehtyä tutkimusta tutkitaan uudelleen. Kirjallisuuskatsauksen avulla kootaan tutkimuksista tuloksia, nämä voivat olla sitten puolestaan pohjana täysin uusille tutkimustuloksille. (Salminen 2011: 1).

Haaparanta ja Niiniluoto (1986: 12-169) puolestaan määrittelevät kirjallisuuskatsauksen siten, että sen on ehdottomasti täytettävä kaikki tieteen metodille asetetut yleisen hyväksynnän saaneet vaatimukset. Tämänhän on toteuduttava myös kaikissa muissa käytettävissä tieteellisissä tutkimusmetodeissa. Kaikkien tieteellisten tulosten on myös oltava aina julkisia, ja tiedeyhteisön on myös kyettävä arvioimaan niitä kriittisesti. Niiden pitää olla myös ns. itsekorjattavia, eli tutkimuksen virheet ja puutteellisuudet on kyettävä poistamaan uusilla tutkimustuloksilla. On myös erittäin tärkeää tietää, että tutkimuskohteen ominaisuudet ovat aina kaikista auktoriteeteista riippumattomia tai ainoastaan tutkijan omia mielipiteitä. (Haaparanta ja Niiniluoto 1986: 12-169).

On olemassa monia hyviä syitä sille, miksi kirjallisuuskatsaus valitaan tutkimuksen metodiksi. Baumstein ja Learyn (1997) ovat tuoneet esille muutamia esimerkkejä syistä, joiden vuoksi kirjallisuuskatsaus on syytä tehdä. Heidän mukaansa kirjallisuuskatsauksen avulla voidaan kehittää edelleen vallitsevaa teoriaa sekä sen avulla voidaan myös kehittää täysin uutta teoriaa. Sen avulla voidaan myös arvioida teoriaa. Kirjallisuuskatsauksen tiedetään myös pystyvän luomaan eheän kokonaiskuvan erityisistä asiantiloista. Sillä voidaan myös tunnistaa ongelmia ja syntyy myös erinomainen mahdollisuus kuvailla teoriaa kehityshistoriallisesta näkökulmasta. (Baumstein ja Learyn 1997:312.)

Kirjallisuuskatsausta voidaan kuvata myös kahden lähestymistavan kautta. Intellektuaalisella historialla voidaan antaa katsaukselle tietynlaiset puitteet. Tällöin kirjallisuuskatsauksessa pysytään vain oman tieteenalan sisällä ja yritetään antaa kyseisen alan tutkijoiden luomasta aineistosta selvä kokonaiskuva. Van Wantin (2001) mukaan se voi toimia hyvänä havainnollistavana esimerkkinä tutkimus- ja julkisjohtamisen teorioista. Hän kuvaakin pitkäaikaistutkimuksessaan johtajuuden teorioiden ja sen tutkimuksen kehittymistä 61 vuoden ajalta, voidaan näin ollen sanoa kyseessä olevan varsin kattava kuvaus ilmiöistä. (Van Wantin 2001.)

On olemassa myös poikkitieteellisempi tapa jäsentää kirjallisuuskatsausta. Siinä jäsennellään kirjallisuuskatsausta ideoiden historian avulla. Tutkimusongelman idea yritetään kuvata tässä tutkimustavassa sillä tavoin, että tieteen alojen rajoista huolimatta

sen kehitys esitetään. Tämän lähestymistavan hyvänä puolena on se, että yhtenäisyydet ja poikkeavuudet saadaan esille. (Bearfield & Ellen 2008:64.)

Kirjallisuuskatsauksia voidaan myös jaotella eri tyyppeihin. Metsämuuronen (2009) mukaan systemaattisessa kirjallisuuskatsauksessa kiinnitetään erityistä huomiota käytettyjen lähteiden keskinäiseen tekniikkaan ja yhteyteen, jolla kaikki siteeratut tulokset on hankittu. Tämän tyyppinen kirjallisuuskatsaus edellyttää muun tieteellisen tutkimustyön tavoin samanlaista tarkkuutta tutkijalta. (Salminen 2011: 4.)

Fink (2005) on myös antanut määritelmiä kirjallisuuskatsauksesta. Hän painottaa sitä, että tutkimuskirjallisuuteen perustuva kirjallisuuskatsaus on aina systemaattinen, täsmällinen ja aina toistettavissa oleva menetelmä. Sen avulla arvioidaan, tiivistetään ja tunnistetaan tiedemiesten, tutkijoiden ja käytännön asiantuntijoiden valmiiksi julkaistu tutkimusaineisto. Fink (2005) huomauttaa myös, että kirjallisuuskatsaus perustuu aina alkuperäisestä laadukkaasta tutkimustyöstä tehtyihin johtopäätöksiin. (Fink 2005:3.)

On yritetty myös paljon määritellä sitä, mitä kirjallisuuskatsaus ei ole. Salmisen (2015) mukaan sitä ei voi koskaan pitää pelkästään selityksiin puettuna lähdeluettelona tai vain kirja-arvosteluna. Syynä tähän on se, että kirjallisuuskatsauksen vaatimuksiin sisältyy aina kriittistä tarkastelua. Luetteloihin, joihin sisältyy lyhyitä tiivistelmiä tiettyjen aihealueiden kirjoista ja artikkeleista, ei juuri sisälly mitään kriittistä tarkastelua. Toisinaan mm. kirja arvostelut saattavat sisältää kriittisiäkin arvioita, ne kuitenkin analysoivat kerrallaan vain yhden teoksen. (Bearfield & Eller 2008:62.)

Salmisen (2011) painottaa myös sitä, että kirjallisuuskatsaus termin määrittely vaatii täsmennyksen. Termiä kirjallisuuskatsaus (review, literature review, research literature review) voidaan pitää epätäsmällisenä. Tällöinhän ei ole kyse katsauksesta, joka arkipäivän kielenkäytössä tarkoittaa lyhyttä tai tiivistä katsantoa, tai ilman analyttistä yhteenvedoa tai tarkasteltavaa, perusteellista otetta katsauksessa. Review on termi, jolla viitataan uudelleen tai kriittisesti tehtävään arviontiin. Termi voi tarkoittaa, myös arvostelua, selontekoa, tarkistusta tai mahdollisesti jopa historiikkia. Kirjallisuuskatsaus ymmärretään yleensä suhteellisen kapeasti. On kuitenkin hyvä tietää, että se pitää sisällään varsin monia tyyppisiä. (Salminen 2011:6.)

Green, Claire, Johnson & Adams (2006), Evans (2008) ja Cronin, Frances & Michael (2008) ovat jaotelleet kirjallisuuskatsauksen kolmeen perustyyppiin. Ne ovat kuvaileva kirjallisuuskatsaus, systemaattinen katsaus ja lopuksi meta-analyysi sekä kvalitatiivisena että kvantitatiivisena tyyppinä. (Green ym.2006: 101–117.)

Kuvaileva kirjallisuuskatsaus on yksi kaikkien käytetyimmistä kirjallisuuskatsauksen tyypeistä. Kuvailevan kirjallisuuskatsauksen tiedetään olevan ilman tiukkoja ja tarkkoja sääntöjä oleva yleiskatsaus. Siinä käytetään laajoja aineistoja, ja niiden valintaa eivät rajaa mitkään metodiset säännöt. Se kuitenkin mahdollistaa tutkittavan ilmiön kuvaamisen hyvinkin laaja-alaisesti ja sen ominaisuuksia voidaan tarvittaessa myös luokitella. Kuvailevan kirjallisuuskatsauksen tutkimuskysymyksistä muodostuu väljempiä kuin systemaattisessa katsauksessa tai meta-analyysissä. Kuvailevasta kirjallisuuskatsauksesta käytetään joskus myös nimitystä traditionaalinen kirjallisuuskatsaus. Se toimii aina itsenäisenä metodina, mutta se voi toisaalta tarjota täysin uusia tutkittavia ilmiöitä systemaattista kirjallisuuskatsausta varten. Kuvailevasta kirjallisuuskatsauksesta voidaan vielä erottaa kaksi eri tyyppistä kirjallisuuskatsausta. Ne ovat narratiivinen ja integroiva kirjallisuuskatsaus. Etenkin integroivalla kirjallisuuskatsauksella tiedetään olevan paljon yhtymäkohtia systemaattiseen katsaukseen. (Salmisen 2011:6.)

Painotetaan sitä, että kun tutkittavasta ilmiöstä halutaan monipuolisesti tietoa on kirjallisuuskatsauksista valittava integroiva kirjallisuuskatsaus. Tällöin voidaan tutkitusta aiheesta tuottaa uutta tutkimustietoa. (Salminen 2011:6). Tämä auttaa lisäksi tutkimuskirjallisuuden kriittisessä arvioinnissa, tarkastelussa ja syntetisoinnissa. (Torraco 2005: 356.) Systemaattisen kirjallisuuskatsaukseen verrattuna integroiva kirjallisuuskatsaus tarjoaa selvästi laajemman kuvan aihepiirin kirjallisuudesta. (Evans 2008:137.)

Integroiva katsaus ei valikoi tai seulo tutkimusaineistoa alkuunkaan yhtä tarkasti kuin systemaattinen katsaus. Tällöin tutkittavasta aiheesta voidaan kerätä huomattavasti suurempi otos. Integroiva ote voi sallia erilaisin metodisin lähtökohdin tehdyt tutkimukset analyysin pohjaksi. (Salminen 2011:8).

Kirjallisuuden tyypit ja niissä olevat näkökulmat voivat olla tutkimusaineistossa myös huomattavasti laajemmat ja vaihtelevammat verrattuna systemaattiseen katsaukseen. (Whittermore 2008:149.)

Painotetaan sitä, että integroiva kirjallisuuskatsaus voidaan sijoittaa osaksi systemaattista kokonaisuutta narratiivisilla keinoilla. Integroiva kirjallisuuskatsaus sisältää kuitenkin yleensä kriittistä tarkastelua toisin kuin narratiivinen katsaus. Tässä on kyseessä metodinen vaatimus, sillä kriittinen arviointi mahdollistaa tärkeimmän tutkimusmateriaalin tiivistämisen katsauksen perustaksi. (Birmingham 2000: 33-34.) Tämän tutkimuksen integroiva kirjallisuuskatsaus on tarkoitus laatia tutkimusongelman pohjalta. Tarkoituksena olisi siis määriteltyjen tutkimuskysymysten mukaan etsiä oikeilla avainsanoilla, hakutermeillä ja -ehdoilla tietoa aihepiiristä. Hakujen tuloksena saatuja arvostetuissa tieteellisissä julkaisuissa julkaistuja asiantuntijoiden analyysoimia artikkeleita käydään läpi, ja niistä valitaan tutkimuksen kannalta mukaan vain kaikkein olennaisimmat. Sen jälkeen niiden keskeisimmistä löydöksistä tehdään päätelmiä ja yhteenvetoja.

Kirjallisuuskatsausta voidaan pitää metodina, joka mahdollistaa tiedon yhteen keräämisen tietystä aihepiiristä tai esim. mahdollisesta tutkimusongelmasta. Kyseinen tieto voi tulla varsin monista lähteistä, näitä saattavat olla tieteelliset artikkelit, lehdet, kirjat, pro-gradu-tutkielmat tai esim. arkistomateriaalit. Kirjallisuutta tutkimalla voidaan pro-gradu-tutkielmalle antaa vahva tieteellinen perusta, koska olemassa olevan tutkimuskirjallisuuden tutkiminen tuo siihen ajantasaisen uusimman tutkimustiedon ja tieteelliset teoriat aihepiiristä. Mikäli kirjallisuuskatsaus on tehty oikealla asianmukaisella tavalla se ei ole vain listaus tiedosta. Kirjallisuuskatsauksen laatijan tavoitteena pitäisikin olla tärkeimpien kriittisten ideoiden ja informaation pohtiminen, joita hän on löytänyt osana tutkielman teoreettista viitekehystä. (Iiron 2017).

Kirjallisuuskatsaus antaa tukijalle syvemmän ymmärryksen nykyisestä tiedosta ja tieteellisistä teorioista. Se kykenee periaatteessa luomaan aina vahvan tieteellisen lähtöasetelman hänen tutkimukselleen. Tutkimusongelman ja tutkimuskysymysten selvän hahmottamisen jälkeen tutkijan täytyy mennä aihepiiriin syvemmälle ja perehtyä tutkimukseen liittyvään olennaiseen kirjallisuuteen. Tämä voikin muodostua suureksi

haasteeksi, koska saatavilla olevan kirjallisuuden määrä sekä aikaraja on yleensä varsin rajallinen tutkielman kirjoittamiselle. Tutkimuksen prosessiin onkin siksi käytävä käsiksi mahdollisimman tehokkaalla ja systemaattisella tavalla. (Iiro 2017).

Kirjallisuus pitäisikin Iiron (2017) mukaan luokitella ensin sen relevanssin mukaan ja vasta sitten tieteellisen laadun mukaan. Relevantti julkaisu sopii tutkimusaiheeseen ja ongelmaan aina hyvin. Kirjan ja artikkelin relevanssin määrittely ilman sen lukemista alusta loppuun kokonaan edellyttää ainoastaan sen johdannon ja yhteenvedon lukemista. Tämä antaa tutkijalle usein jo riittävästi tietoa arvioitaessa, onko julkaisu oman tutkimuksen kannalta tärkeää. Julkaisun laatua määrittelevät useat tekijät, ja yleinen tutkijoille suunnattu ohje on, että käytetään ainoastaan artikkeleita, jotka on julkaistu johtavissa tieteellisissä julkaisuissa. Julkaisuihin osallistuvien kirjoittajien asiantuntijuuden arviointi saattaa myös olla hyödyksi. Asiantuntevat kirjoittajat linkittyvät monesti akateemiseen instituutioon, koska ovat julkaisseet laajasti ja muut viittaavat kirjoituksissaan juuri heihin. Tutkijan tulisi myös muistaa, että verkkosivuilla oleva informaatio ei usein ole luetettavaa. Ainoastaan verkkosivut, joita tiedeyhteisön hallitus pitää tai hallitustenvälejä ylläpitävät instituutiot ovat poikkeus tähän sääntöön. Tutkijan on käytettävä myös kaikkein uusinta tutkimuskirjallisuutta, jotta tutkimusta ei rakenneta vanhentuneen tiedon pohjalle. Neljäs vaihe on kirjallisuuden käsittely, tähän vaiheeseen kuuluu sen kirjallisuuden läpi käyminen, johon tutkija aikoo keskittyä. Sen jälkeen löydettyä tietoa käsitellään, ja tämä voi tapahtua esim. ongelma-analyysin tai teoreettisen viitekehyksen kautta. Tässä vaiheessa korostuu huolellisesti valittujen julkaisujen läpi käyminen. (Iiro 2017).

Kuten integroivassa kirjallisuuskatsauksessa tässä tutkimuksessa yritetään saada laaja katsaus tutkittavasta ilmiöstä. Integroivalla kirjallisuuskatsauksella integroidaan aiemmin tehtyjä tutkimuksia ja julkaisuja koneoppimisen soveltamisesta terveydenhoitoon, sairaalaan tietojärjestelmistä, kliinisistä päätöksen tekojärjestelmistä, kliinisestä päätöksenteosta ja diabetes mellituksesta, uudempaan tutkimusaineistoon koneoppimisen soveltamisesta diabetes hoitoon, josta on kuitenkin vielä melko vähän täysin relevanttia julkaistua tutkimustietoa varsinkaan tietojärjestelmätieteellisestä näkökulmasta. Tutkimustuloksia ei ole myöskään aikaisemmin juurikaan pyritty

integroimaan kattavasti puhtaasti suunnittelutieteelliseen kontekstiin, siksi integroivaa kirjallisuuskatsausta on käytetty tutkimuksessa tutkimusmenetelmänä.

Tutkimuksen tavoitteena on aineistolähtöisellä tutkimuksella tarkastella ja analysoida olemassa olevia diabetes hoidossa käytettyjä koneoppimismalleja ja metodeja ja tuottaa ohjeellistavaa teoriaa koneoppimisen soveltamisesta diabetes hoidon kliinisen päätöksen tekojärjestelmän suunnittelumallin luomiseksi. Tähän katsaukseen on kelpuutettu mukaan ainoastaan tieteelliset kriteerit täyttävää uusinta ja relevanttia julkaistua tutkimuskirjallisuutta. Se on julkaistu viimeisen viiden vuoden aikana. Katsauksen synteesi tehdään siis kaikkein uusimman tiedeyhteisössä arvioidun, tunnustetun ja yleisesti hyväksytyt tutkimustiedon pohjalta.

5.2 Kirjallisuuskatsauksen koostaminen

Kirjallisuuskatsauksella on filosofisesti ajateltuna oma olemuksensa. Kirjallisuuskatsauksen ontologia koostuu neljästä vaiheesta, eli kohteesta, lopputuloksesta, kehuksesta ja vaiheesta. Kohde sisältää otsikon, alueen ja tiedonhaaran. Lopputulos sisältää synteessin, tulkinnan ja ohjauksen. Kehys sisältää asiankohtaisen, marginaalisen, sisällöllisen ja käsitteellisen. Vaihe puolestaan sisältää etsimisen ja arvioinnin, synteessin, tulkinnan ja ohjauksen sekä johtopäätöksen. Schyren (2013) painottaakin sitä, että kirjoitettaessa kirjallisuuskatsausta kirjoittajan tulee tehdä huolellisia päätöksiä sen etenemisessä tarkasteltaessa sen kohdetta, lopputulosten tyyppiä, rajauksia ja vaiheita. (Schyren 2015:291).

Kirjallisuuskatsaus voi sisältää valitun aiheen, valitun alueen tai valitun tiedonhaaran nousevassa järjestyksessä. Ponnistusten määrä, joka tarvitaan kirjallisuuskatsaukseen, eroaa toisistaan riippuen katsauksen kohteesta. Kirjoittajan tulisi myös olla Schyrenin (2013) mukaan selvillä sen lopputuloksesta. Kirjallisuus ei koskaan myöskään tarjoa ainutlaatuista, uniikkia määritelmää tai ymmärrystä siitä, mitä kirjallisuuskatsauksen pitäisi olla. Useimmat kirjallisuuskatsauksen tekijät ovat sitä mieltä, että kirjallisuuskatsauksen pitäisi syntetisoida ja tulkita kirjallisuutta. Silloin kun on tarkoitus identifioida tutkimus määrätynlaisena tulkintana, voidaan mahdollinen

kirjallisuuskatsauksen lopputulos tiivistää kolmeen tyyppiin eli kirjallisuuden syntetisoimiseen, kirjallisuuden tulkintaan ja tulevan tutkimuksen ohjaamiseen. Schyren (2013) toteaa, että on vaikea lähteä analysoimaan ristiriitoja perspektiivistä, joka painottaa katsauksen lopputuloksen olevan välttämätön kokonaisuuden osa kirjallisuuskatsauksessa. Jää kokonaan tulevan kirjallisuuskatsauksen laatijan itsensä päätettäväksi, minkä lopputuloksen he haluavat osoittaa katsaukseen. Kirjallisuuskatsauksen rajaus puolestaan viittaa sen mittaavuuteen, laajuuteen, katkelmallisuuteen ja löytyvyyteen. Rajaaminen voi olla ajankohtaista, marginaalista ja käsitteellistä abstraktion järjestämistä. (Schyren 2015:291.)

Yhtenä esimerkkinä on kirjallisuuskatsaus, johon sisältyy yksinkertainen Google haku jostain aiheesta, ja joka toimittaa tekijä-keskeisen esityksen tuloksista. Schyren (2013) huomauttaa, että marginaalisessa kirjallisuuskatsauksessa jokainen askel määrää seuraavan. Näin ollen esim. kirjallisuus haun tulokset määrittää sen tavan, jolla esityksen löydökset on rakennettu. Käsitteellisessä kirjallisuuskatsauksessa yksi tai useampi käsitteistö, kuten mallit, kehykset tai teoriat perustellaan, ja niitä käytetään esityksen rakentamiseen ja löydösten tulkintaan. Kirjallisuuskatsauksen rajauksella on olennainen vaikutus kirjallisuuskatsauksen useisiin vaiheisiin kirjallisuuskatsauksen johtamisessa. (Schyren 2015:291.)

Schyren (2013) erottaa viisi vaihetta: Ensimmäinen vaihe on etsi ja arvioi, se kertoo asiaankuuluvan, relevantin kirjallisuuden hankkimisen mahdollisuuksista. Vaiheet synteesi, tulkinta ja ohjaus kertovat vastaavan kirjallisuuskatsauksen tuloksen saavuttamisesta. Vaihe johtopäätökset viittaavat siihen, kuinka kirjallisuuskatsaus lopetetaan. (Schyren 2015:291.)

Kirjallisuuskatsauksen avainkysymys on, kuinka jäsentää sen käsittelyä ja lopullista artefaktia, toisin sanoen varsinaista kirjallisuusartikkelia. Menettelytapa ja artefakti eivät ole Schyrenin (2013) mukaan itsenäisiä ja riippumattomia. Kirjallisuuskatsauksen tulisi myös sisältää seuraavat vaiheet. Ne ovat katsauksen rajaaminen, kirjallisuuden etsiminen ja arvioimisen vaihe, synteessin vaihe, tulkinta vaihe, ohjausvaihe ja johtopäätös vaihe, tämän ovat todenneet mm. Fink (2010), Wolfswinkel, Furtmueller & Wilderom (2013)

ja Rowe (2014). Rajaus on prosessi, jolla on olennainen vaikutus kaikkiin vaiheisiin. Kirjallisuuskatsauksessa pitäisikin olla täysin sen kuvaamiseen omistettu osa, ja katsauksen alussa pitäisi myös määritellä, mikä on päämotiivi kirjallisuuskatsauksen kirjoittamiseen valitusta aiheesta. (Schyren 2015:292.)

Tämän jälkeen tulisi myös määritellä, miten kirjallisuuskatsaus eroaa muista julkaistusta katsauksista, eli pitäisi määritellä tarkoin myös katsauksen päämäärä, laajuus, rajaus sekä kirjallisuuskatsauksen rakenne, kuten Webster ja Watson (2002), Wolfswinkwel ym. (2013:47) ja Okoli ja Schrabram (2010:7-14) toteavat. Attribuuttien määrittely ja kuvaileminen auttavat kirjoittajaa keskittymään aiheen keskeisiin osiin työssä. Ne auttavat kirjallisuuskatsauksen tiivistelmissä ja havainnollistavat, mitä voidaan saavuttaa, ja mitä ei voida saavuttaa. Ne auttavat lisäksi selittämään, miten kirjallisuuskatsaus näyttää sen merkityksellisyyden opiskelijoille ja /tai ammatinharjoittajille ja myös katsauksen metodologisen täsmällisyyden. (Schyren 2015:292.)

Etsintä ja arvioimisvaiheen tarkoituksena kirjallisuuskatsauksessa on kirjallisuuden etsintä ja kerättyjen artikkelien arvioiminen. Schyrenin (2013) mukaan tähän vaiheeseen kuuluva kirjallisuuden etsimisen prosessi voidaan kuvailla laajalti itsenäisesti, irrotettuna kirjallisuuskatsauksen aiheesta ja päämääristä. Kerättyjen artikkelien arviointi puolestaan riippuu nimenomaan kirjallisuuskatsauksesta, ja täten se voidaan kuvailla vain yleisellä tasolla. Synteesi vaiheessa taas on tarkoituksena tehdä synteesi muiden tutkijoiden aiheesta tehdyistä tutkimuksista. Schyren (2013) painottaakin synteesin kokonaisvaltaisen tehtävän olevan juuri synteesin tekeminen siitä, mitä muut tutkijat ovat tutkittavasta aiheesta löytäneet ja julkaisseet. Se on Schyrenin (2013) mukaan pakollista huolimatta siitä, minkä tyyppistä kirjallisuus katsausta ollaan laatimassa. Synteesi sisältää sekä niiden käsitteiden kuvaamisen, joita käytetään jäsentämään sitä, kuinka tutkimuksen löydökset esitetään ja myös varsinaisen esityksen. (Schyren 2015:293.)

Sen jälkeen kirjallisuuskatsauksessa seuraa ns. tulkintavaihe. Se on myös tärkeä osa hyvin laadittua kirjallisuuskatsausta, koska se mahdollistaa lähdekritiikin. Schwartz ym.(2006) huomauttavatkin, että kirjallisuuskatsauksen hyödyt pitäisi laajentaa tutkimuslöydösten synteesin yli, kirjallisuuskatsauksen pitäisikin olla ennen kaikkea kriittinen. Boel ja Cecer-Kecmanovic (2014) toteavat, että kriittinen arviointi ei pelkästään paljasta, vaan

vielä tärkeämpänä seikkana myös haastaa mahdollisen merkitysten ja ymmärrysten tason muodostusta tiedon rungosta. Schyren (2013) toteaa, että useista kirjallisuuskatsauksista voidaan ymmärtää se, että niissä on tarkoitus identifioida ns. tutkimusaukot. Kuitenkaan ei ole tarpeen paljastaa, mikä kirjallisuudesta puuttuu, jotta siitä tulisi kriittistä. Jotkut kirjoittajat kuten, mm. Lepine & Wilcox.-King (2010), Webster & Watson (2002) ehdottavat, että kirjallisuuskatsaus asettaa tiedon puun laajentamaan nykyisiä teorioita. Nämä molemmat voivat myös yhdistyä, kuten Wolfswinkel (2013) ym. huomauttavat. Tämä tapahtuu silloin, kun kirjallisuuden analyysi johtaa tutkimusaukkojen löytämiseen tiedossa, joka on tärkeää tarkalle teoriaan suuntautuvalla tutkimukselle. (Schyren 2015:293.)

Kirjallisuuskatsaukseen liittyy myös aina ohjausvaihe. Schyren (2013) huomauttaa, että kirjallisuudessa ei olla samaa mieltä siitä, pitäisikö kirjallisuuskatsauksen johtaa tulevaa jatkotutkimusta ja millä tavoin se tapahtuisi. Tutkimuskatsaukset ovat näkemyksellisiä ja taitavasti laadittuja artikkeleita, jotka käsitteellistävät tutkimus alueet, tekevät synteetin aikaisemmista innovatiivisista löydöksistä, edistävät kentän ymmärrystä, ja identifioivat sekä kehittävät tulevaisuuden tutkimuksen suuntaa. Webster ja Watson (2002) ovat myös sitä mieltä, että katsauksen kirjoittaminen ei ainoastaan vaadi menneen tutkimuksen tutkimusta, vaan se tarkoittaa uuden kartan tekemistä tulevalle tutkimukselle. Rowe (2014) sitä vastoin mainitsee, että saman tutkimusartikkelin ei välttämättä tarvitse selittää, kuinka voidaan varsinaisesti päästä perille siinä. Tämä ei hänen mukaansa ole kirjallisuuskatsauksen olemus. Viimeisenä vaiheena kirjallisuuskatsauksessa on johtopäätös vaihe, jossa tutkielmasta tehdään varsinainen yhteenveto. Schyrenin (2013) mukaan tämä on viimeinen vaihe kirjallisuuskatsauksen avain oivallusten soveltamisessa. Tähän sisältyy Wolfswinkelin ym.(2013) mukaan rajoitukset ja väistämätön suuntaus, joka on saattanut ilmetä yhtenä tai useampana kokonaisprosessin askeleena ja se myös motivoi jatkotutkimukseen erityisellä kentällä. (Schyren 2015:293.)

Kirjallisuuskatsauksen rajauksella voidaan saada tutkimusaiheeseen aivan uudenlaisia näkökulmia ja keksiä ratkaisemattomia aiheeseen liittyviä tutkimuskysymyksiä. Schyren (2013) toteaa, että aikaisemman kirjallisuuskatsauksen puute voi antaa hyvän motiivin kirjallisuuskatsauksen laatimiselle. On myös tärkeää huomioida muiden opiskelijoiden ja tutkijoiden mielipiteet uuden kirjallisuuskatsauksen laatimisen tärkeydestä. Useimmiten

kuitenkin kirjallisuuskatsaus on aiheesta laadittu, ja näin ollen tehtäväksi jääkin vain tarve selittää oman kirjallisuuskatsauksen eroavaisuudet toisista kirjallisuuskatsauksista. Katsauksen ainutlaatuisuus näyttäytyy etenkin silloin, kun se kykenee antamaan täysin uuden näkökulman aiheeseen tai ja/tai se keskittyy uusiin vielä ratkaisemattomiin tutkimuskysymyksiin. (Schyren 2015:295.)

Kirjallisuuskatsauksen vaiheisiin kuuluu myös tutkimuskirjallisuuden etsimisen ja arvioimisen vaihe. Schyren (2013) mukaan tähän haku ja arviointivaiheisiin sisältyy kirjallisuuden etsimisen ja arvioimisen tehtävät. Schyren (2013) painottaa kirjallisuuskatsauksen tutkielmakirjallisuuden etsimisen hyvään lähtökohtaan sisältyvän erityisesti oppikirjat ja muiden opiskelijoiden kirjallisuuskatsaukset. Nämä yleensä sisältävät kattavan viittauksen kappaleisiin ja urauurtaviin töihin tieteenalalla. Toisen tyyppisiä lähdekirjallisuuden varantoja aineiston etsimisessä ovat kirjallisuus tietokannat, yhteisöjen ja organisaatioiden julkaisuluettelot, julkisten ja yliopistojen kirjastojen hakemistot, useiden julkaisijoiden online luettelot ja online kirjakaupat, kuuluisien akateemisten aikakauslehtien sisällykset ja konferenssi toimituskokoukset, standardointiorganisaatioiden tarjoamat standardi hakemistot, ammatti-alan lehtien julkaisemat artikkelit, tutkimukset ja eri sanomalehdet. (Schyren 2015:297-298.)

Oikean kirjallisuuden löytämisessä ja valitsemisessa tarvitaan oikeanlaisia hakusanoja ja termejä etsittäessä sitä kirjallisuustietokannoista. Schyren (2013) painottaakin, että lähdekirjallisuustietokannoista kysely vaatii tarkoituksenmukaisen bibliograafisen tai arkistotietokannan ja myös oikeanlaisen haku termien valitsemisen. (Schyren 2015: 300.)

Oikeiden ja tarkoitustenmukaisten lähdekirjallisuus tietokantojen valitsemisen jälkeen kirjallisuuskatsauksen vaiheittaisessa etenemisessä tullaan vaiheeseen, jossa tulisi löytää niistä haun kautta tutkimukseen integroitavaksi relevanttia kirjallisuutta. Wolfswinkel ym.(2013) toteavat, että sopivien kirjallisuus tietokantojen identifioimisen ja valitsemisen jälkeen, seuraavaksi pitää määrittellä haku merkkijonot, jotka ovat tarkoituksenmukaisia identifioimaan relevanttia kirjallisuutta tutkimukseen. Sopivien haku merkkijonojen määrittäminen on kriittistä, koska se määrittää sen laajuuden, mitä asiaankuuluvaa, relevanttia ja mitä asiaankuulumatonta, irrelevanttia kirjallisuutta löytyy. Hyvän lähtökohdan tälle antaa avainsanojen ottaminen jo aikaisemmin identifioiduista

artikkeleista. Toinen vaihtoehto on tehdä luokittelu siitä aineistosta, joka on tarkoituksenmukaista kirjallisuuskatsauksen aiheelle. Jotkut tietokannat tarjoavatkin avainsanojen luokittelun taxonomian. Monet kirjallisuustietokannat sallivat myös rakentaa loogisia haku merkkijonoja, joihin sisältyy avainsanojen ilmaiseminen, liitettynä loogisilla operaattoreilla, mukaan lukien JA(AND), TAI(OR) ja EI(NOT). Ei ole oikeaa tai väärää luetteloa haku merkkijonoista, ja tutkijoiden tulee todennäköisesti toimia niiden kanssa hieman ennen kuin he löytävät lopullisen listan haku merkkijonoista. (Schyren 2015:300-301.)

Toinen tärkeä asia relevantin kirjallisuuden löytämisessä kirjallisuuskatsaukseen on löytää oikea aikaväli, josta tutkimustietoa haetaan. Schyren (2013) mainitsee, että hakumerkkijonon määrittelemisen lisäksi, pitäisi haussa valita myös tarkasteltava aikaväli. Useinkaan ei ole juuri olemassa mitään vakuuttavaa syytä, miksi pitäisi rajoittaa aikaväliä, mutta erityisissä tapauksissa ajallisen rajoittaminen voi tulla kyseeseen esim. silloin, kun bibliograafiset tutkimusartikkelit julkaistaan tietyssä ajankohtana. Sen jälkeen täytyy vielä valita haun ulottuvuus. Tällöin voidaan soveltaa haku merkkijonoja otsikoihin, avainsanoihin, abstraktioihin ja kokonaisiin julkaistuihin teksteihin. Voidaan myös etsiä muita nimenomaisia kirjoittajia. Se voi olla hyödyllistä sellaisten kirjoittajien julkaisujen kanssa, jotka ovat julkaisseet tärkeitä artikkeleita liittyen laadittavaan kirjallisuuskatsaukseen. Tutkijoiden täytyy luultavasti toimia yhdistelmien kanssa identifioidakseen sopivia hakupolkuja. Haku prosessin aikana usein useimmat saman julkaisijan julkaisemat julkaisut ja niiden organisaatiot nousevat esille. (Schyren 2015:301-302).

Kirjallisuuden etsimisprosessin jälkeen, tutkijan täytyy hankkia ja arvioida kirjallisuus. Hakuprosessin aikana pitäisi säilyttää paljon relevantista kirjallisuudesta, Schyren (2013) huomauttaa. Määrätty osa kirjallisuudesta ei yleensä ole saatavilla. Tähän on syynä se, että yleensä käyttöoikeudet tai kirjat eivät ole joko saatavana onlineina eikä paikallisista kirjastoista. Artikkelien abstraktit ja yhteenvedot, sisällysluettelot tai kirjojen eri painokset ovat yleensä vapaana. Kirjojen suhteen yliopistot ovat yleensä yhteistyössä muiden yliopistojen ja kansallisten kirjastojen kanssa. Kirjallisuuden hankkimisessa täytyy ottaa huomioon se, että etsittyä kirjallisuutta arvioitaessa sen täytyy olla määritelty sekä laadukkaaksi että kirjallisuuskatsaukseen sopivaksi. Schyren (2013) huomauttaa,

että koskien kirjallisuuden laatua tutkijoiden pitää määrittellä laatu kriteerit, joka on usein vaikeaa johtuen laadun määrittelyn vaikeudesta. Hakemiston asetuksen minimivaatimukset ovat kuitenkin hyödyllisiä ja mahdollisia. Vaatimukseen voi kuulua esim. se, että kirjallisuuskatsaukseen integroidut tutkimukset ovat tehty riittävän suurella otannalla ym. Schyren (2013) on sitä mieltä, että on hyödyllistä määrittellä eri laatu kriteerit ja/tai eritasoiset laatu kriteerit riippuen julkaisun tyypistä. Näin ollen aikakauslehdissä julkaistut tutkimukset ja konferenssien toimituskokoukset pitäisi selittää tarkasti metodologisesta ja teorian näkökulmasta, kun taas julkaisujen aikakauslehdissä tulisi keskittyä käyttökelpoisuuteen ja merkityksellisyyteen käytännössä. Schyren (2013) on sitä mieltä, että arvioitaessa löydetyn kirjallisuuden laatua ja sopivuutta tutkijoiden tulisi ensin lukea abstrakti, ja päättää tulisiko artikkeli sulkea pois tai siirtää päätöstä tuonnemmaksi, kun artikkeli on kokonaan analysoitu. (Schyren 2015:304).

Kirjallisuuskatsauksessa seuraa kirjallisuuden etsimisen ja arvioinnin jälkeen ns. synteesivaihe, jonka tarkoituksena on esittää kirjallisuuskatsaukseen integroidusta tieteellisistä julkaisuista koottu tieto lukijalle. Schyren (2013) kuvailee tätä vaihetta kirjallisuuden hakemisen jälkeiseksi vaiheeksi, jonka tarkoituksena on esittää katsauksen lukijalle tiedon runko siihen valituista julkaisuista. Okoli(2012) pitää tätä tehtävää yhtenä avain asiana kirjallisuuskatsauksessa. Hänen mukaan kaikkein tärkein askel kirjallisuuskatsauksessa on aina ollut tutkimuksen synteesi, joka on sisällynyt katsaukseen. Tämän osan johtaminen ei ole suoraviivaista, ja Rowe (2014) toteaa oman kokemuksensa perusteella sen olevan usein tehty epäsopivalla tavalla, etenkin kun kokemattomat opiskelijat kirjoittavat katsausta. Kirjallisuuden synteesissä pitäisi havaita ja luokitella useita tutkimus osia selviin kategorioihin. Hart (1998) huomauttaa myös, että kirjallisuudesta tehdään synteesi aina tietystä näkökulmasta ja täten siihen sisältyy aina tulkinta. Schyren (2013) taas painottaa, että tämä osa kirjallisuuskatsausta pitäisi olla pääasiassa deskriptiivinen. (Schyren 2015:305).

Kirjallisuudessa ollaan laajasti yhtä mieltä siitä, että synteesi tiedon puusta pitäisi esittää käsite-(idea)keskeisesti eikä kronologisella tai tekijä(laatija)-keskeisellä tavalla. Webster & Watson (2002) ovat sitä mieltä, että käytetty käsitteistö myös määrittää organisoitua kehystä. Tekijä-keskeisen lähestymistavan käyttöönoton haittana on ns. ”hän sanoi”

ongelma. Zornin & Campellin (2006) mukaan, siinä tekijä kertoo meille, mitä jokainen lähde sanoo, mutta ei välitä-lähteiden välisistä suhteista mitään tietoa. Rowe (2014) yhtyy tähän näkemykseen todetessaan, että kirjallisuuskatsauksen ei tarvitse integroida kaikkia kirjallisuuden antamia tiedon kokonaisuuden osia ”kokonaisvaltaiseksi logiikaksi”, vaan pikemminkin ”joukoksi yhdenmukaisia yleiskäsitteitä.” Schyren (2013) korostaa vielä, että kirjallisuuden löydökset pitäisi syntetisoida uuden tai vanhan käsitteistön ympärille. Se voi olla yhdenmukaista ja noudattaa kaiken käsittävää logiikkaa tai käsitteistöt ei välttämättä kytkeydy täysin toisiinsa. Kirjallisuudesta voidaan tehdä synteesi eri näkökulmista, jotta saadaan täydentäviä näkemyksiä kirjallisuudesta. Jasperson, Carte, Saunders, Butler, Croes & Zheng (2002) käyttävät teknologia ja energia linssejä tutkiakseen keskinäisiä suhteita energian/tehon ollessa päälle kytkettynä ja informaatioteknologian vaikutuksia, levittämistä, hallinnointia tai käyttöä. (Schyren 2015:305).

Kirjallisuuskatsauksen tulkintavaiheen tarkoituksena on tulkita valitusta tutkimuskirjallisuudesta luotu esitys katsauksen lukiojoille. Schyren (2013) mainitsee, että tiedon puun tulkitseminen kuuluu kaikkein luovimpiin tehtäviin kirjallisuuskatsauksessa. Tämän vaiheen yleisimpiin asioihin sisältyy tutkimusaukkojen tulkitseminen, uuden näkökulman käyttöönotto kirjallisuuden rungossa, ja kirjallisuuden analysoiminen tekemällä ehdotuksia tai myötävaikuttamalla uuteen teoriaan. Tämän tyyppiset myötävaikutukset peittävät usein toisiaan. Tutkimus aukkojen identifiointi on Hartin (1998) mukaan hyödyllistä, koska se auttaa paikantamaan tutkimuksen kartoittamattomat alueet, ja näin ollen se menee askeleen yli tutkimuksen synteisiin. Ensin mainitun viitatessa siihen, mitä on tehtävä, jälkimmäisen puolestaan on suhteessa siihen, mitä on tehty. (Schyren 2015: 309).

Zorn & Campell (2006) ovat sitä mieltä, että näiden aukkojen esittäminen osoittaa mahdollisen tulevaisuuden tutkimuksen suunnan ja motivoi tutkijoita päättämään tutkimuksen, kuten Webster & Watson (2002) toteavat. Tutkimus aukot voivat ilmetä eri muodossa, ja esimerkiksi tietyt näkökulmat/ilmiöt saattavat olla jätetty kokonaan huomioimatta. Tällöin esim. tutkimus tulokset eivät ole vaikuttavia, tai ne ovat erittäin ristiriitaisia, ja tietämys suhteessa kohdennettuun ongelmaan saattaa olla riittämätöntä, kuten Boel & Cerez-Kecmanovic (2014) toteavat. Schyren (2013) huomauttaa myös siitä,

kuinka kaksi haastetta ilmenee kirjallisuuskatsauksen laatijan keskittyessä tutkimusaukkoihin. (Schyren 2015: 309).

Schyren (2013) on sitä mieltä, että kirjallisuus katsauksen laatijan on syytä käyttää mallia, jossa kirjallisuuskatsauksen löydöksistä tehdään synteesi ja identifioidaan tieteellisen tutkimustyön varsinaiset aukot. Tällöin tutkijan täytyy aivan ensiksi identifioida se tutkimus alue, jossa jatkotutkimusta tarvitaan ja se pitäisi yksityiskohtaisesti kuvailla irrallisilla lauseilla erityisistä puutteista tutkimuksessa sekä viittaamalla selvästi kirjallisuuteen. Näihin puutteisiin perustuen, tutkija kehittää varsinaiset tutkimuskysymykset yhdessä tutkittavan alueen mukana. Powell (2014) puolestaan painottaa sitä, kuinka tutkijat voivat piirtämällä samanlaisen kehyksen kuin mitä käytetään kirjallisuus synteessin tekemisessä identifioida tärkeitä tutkimusalueita. Kehys on rakennettu syöttöjen, sosio-emotionaalisten prosessien, tehtävä prosessien ja tulosteiden ympärille. Perustuen neljään vähän tutkittuun tutkimus alueeseen, kirjoittajat voivat ehdottaa monia osittain yhdisteltyjä tutkimuskysymyksiä. (Schyren 2015:311).

Täydentävät kirjallisuuskatsaukset, jotka identifioivat tutkimus aukkoja nousevat Dibbern, Goles, Hirschheim, & Jayatilaka (2004) mukaan tietojärjestelmien ulkoistamisesta ja ratkaisemattomien asioiden aukoista tietämyksessä. Kohl ja Grover (2008) ehdottavat neljää teemaa jatkotutkimuksessa informaatioteknologian liikearvossa. Roberts, Galluch, Dinger, & Grover (2012) identifioivat rajoitukset tietojärjestelmätieteen absorptiokyvyssä. Alavo ja Leidner (2001) ehdottavat myös varsinaisten tutkimuskysymysten tekoa tiedon hallinnasta ja tiedon hallinnan järjestelemisestä. (Schyren 2015:311).

Kirjallisuus katsauksen tehtävänä voidaan nähdä varsinaisen tieteellisen teorian rakentamisen avustamisen. Le Pine ja Wilcox-King (2010) näkevät katsaukset ikään kuin ajoneuvoina teorian kehitykselle, niiden tehtävänä on ehdottaa tai avustaa uuteen teoriaan. Wolfwinkel ym. (2013) yhtyvät tähän ja väittävät teorian että teorian rakentaminen on kasvavassa määrin yksi tärkeistä seurauksista käytettäessä mainetta saavuttanutta kirjallisuus segmenttiä. Vaikka kirjallisuuden piirteet eivät varsinaisesti ole yleinen mielipide siihen, mitä teoria varsinaisesti on, kuten Sutton & Straw(1995) toteavat. Gregor (2006) toteaa, että kirjallisuuden tulkitsemisen yhteydessä se on

kuitenkin sopivaa. Hän myös väittää, että kaikki teorat sisältävät ”keinot esitykseen”(fyysisen esityksen sanoilla, logiikalla, kaavioilla ja taulukoilla), ”rakentamisen” (mielenkiinnon ilmiön), ”suhteiden toteutumisen” ja ”laajuuden”. Tämä laaja ymmärtäminen ei vaadi teorialla olevan selittävää komponenttia. (Schyren 2015: 312).

Gregor (2006) ehdottaa edelleen viittä eri tyyppistä teoriaa: ensimmäinen on teoria analysoinnista (tyyppi 1), teoria selittämisestä (tyyppi 2), teoria ennustamisesta (tyyppi 3), teoria selittämiseen ja ennustamiseen (tyyppi 4), ja teoria suunnitteluun ja toimintaan tyyppi (5). Kirjallisuuskatsaus voi esittää tai vähintään myötävaikuttaa uuteen teoriaan silloin kun se tulkitsee tiedon runkoa näkökulmasta, jota ei ole omaksuttu aikaisemmin. Näin ollen varsinaiseen uuteen teoriaan myötävaikuttamista, voidaan pitää uuden näkökulman omaksumisessa kirjallisuuskatsauksessa tärkeänä. Uusi näkökulma kirjallisuudesta omaksutaan keskustelemalla samankaltaisuuksista ja eroista, jotka esiintyvät käytettäessä erilaisia linssien joukkoja samanaikaisesti. Tähän keskusteluun perustuen kirjoittajat kehittävät teoreemia, jotka voidaan tulkita monista näkökulmista ja näihin voidaan viitata eräänlaisina ”metaotaksumina”, tämän toteavat Jaspersson ym. (2002). (Schyren 2015: 312).

Kirjallisuuskatsauksen ohjausvaihe on vaihe, jossa tutkijoiden on tarkoitus antaa ehdotuksia siitä, minkälaista jatkotutkimusta tieteenalalla pitäisi tehdä. Schyren (2013) painottaa, että jatkotutkimuksen ohjaaminen voi ilmetä eri muodoissa ja yksityiskohtien tasolla. Useat katsauksien laatijat antavatkin lyhyitä ehdotuksia jatkotutkimuksia varten johtopäätöksissään tekemissä huomautuksissa. Muut osoittavat jatkotutkimuksen suunnan yksityiskohtaisemmin ilman heidän suosituksiensa juurtumista yhdenmukaiseksi ideaksi. (Schyren 2015: 314.)

Kirjallisuuskatsauksen viimeinen vaihe on ns.johtopäätös vaihe, jossa tutkija tekee yhteenvedon katsauksen löydöksistä ja johtopäätöksistä. Schyrenin (2013) mukaan kirjallisuus katsauksessa täytyy muiden tutkimus artikkeli lajityyppien mukaan tehdä johtopäätös kirjallisuus katsauksesta. Schyren (2013) suosittelee tutkijoiden tekevän yhteenvedon, siitä mitä kirjallisuus katsaus on löytänyt, mitä johtopäätökset ovat tutkimukselle ja käytännölle, ja myös mitä ovat kirjallisuus katsauksen rajoitukset.

Yhteenvedossa pitäisi olla lyhyt synteesi jokaisesta kirjallisuuskatsauksen myötävaikutuksesta teorian kehityksessä. Erityisesti pitäisi mainita käsitteet, jotka omaksuttiin katsaukseen ja tulkita kirjallisuus ja kehittää toimiva tutkimussuunnitelma. Pitää tietenkin tehdä myös yhteenveto siitä, mitä kirjallisuus löydöksiä, tutkimus aukkoja, tiedon laajentamista ja jatkotutkimus polkuja katsaukseen löydettiin. Kirjallisuuskatsauksen johtopäätökset voivat viitata tutkimukseen ja käytäntöön ja ne pitäisi esittää, kuten Webster & Watson (2002) toteavat. Tutkimus suunnitelman järjestäminen tarkoittaa, että on jo osoittanut olennaiset tutkimuksen johtopäätökset. Tämä ei kuitenkaan välttämättä tarkoita, että ei olisi enää mitään sanottavaa jatkotutkimuksesta. Schyren (2013) antaa erillisen osan nimettynä ”jatkotutkimukselle mahdollinen”, jossa hän lyhyesti luonnostelee jatkotutkimus alueita, joita ei ole käsitelty tutkimus suunnitelmassa. Lopuksi pitäisi vielä ilmoittaa rajoitukset, koska mitään ”täydellistä” kirjallisuus katsausta ei ole olemassa. (Schyren 2015:319.)

5.3 Relevantin tutkimusaineiston aineiston kerääminen, analysoiminen ja tulkinta

Tämän tutkimuksen pääasiallisin aineisto on kerätty PubMed Health tutkimustietokannoista, ja haussa käytetään myös muita tietokantoja esim. (IEEE, Science Direct, Google Scholar, AIS digital library, MIT ja INFORMS) apuna tutkimusaineiston keräämisessä tutkimusaihepiiriin liittyvän tieteellisen julkaisujen (journal, review ym.). Näitä hakujen kautta löydettyjä aihepiiriin kuuluvia artikkeleita luetaan läpi ja niiden löydöksiä analysoidaan tarkoin. Sen jälkeen niistä valitaan parhaimmat ja relevantimmat julkaisut, ja tehdään niistä katsauksen lopussa synteesi.

Ensimmäisenä tietokannoista etsittiin tietoa hakumerkkijonoilla machine learning AND diabetes care, koska tarkoituksena oli kartoittaa, miten koneoppimisteknologiaa on käytöön otettu terveydenhuollossa yleisesti. Tieteellisten julkaisujen tarkasteluajanjaksoksi valittiin viimeiset viisi vuotta, koska se on yleensä tutkimustiedossa melko tyypillinen ajanjakso, jossa tutkimus tiedon tieteellinen arvo vähenee selvästi. Valitsin tietokannaksi Pubmed Healthin ja suoritin em. hakutermeillä haun tästä tietokannasta, ja haun tuloksena sain kaikkiaan 84 julkaistua artikkelia

vuodesta 2014 nykyiseen ajankohtaan 2019. Nämä artikkelit luettiin läpi ja analysoitiin niiden tärkeimmät löydökset. Saadakseni vielä tarkempia ja relevantimpia haku tuloksia tein samoilla hakuehdoilla haun myös Pubmed Health tietokannasta. Se antoi haun tuloksena 20 julkaistua artikkelia. Nämä 20 julkaistua artikkelia sisälsivät varsin vaihtelevasti varsinaista relevanttia tietoa koneoppimisen soveltamisesta diabeteksen lääketieteelliseen hoitoon. Tämän takia olikin vielä syytä valikoida työhön integroitavaksi vain sellaisia artikkeleita, jotka antoivat asianmukaista ja uutta tietoa aihealueesta. Näitä artikkeleita löytyi kaikkiaan 16 kpl. Tein niiden löydöksistä yhteenvedon ja pyrin nostamaan esille tutkimukseen liittyvää avainkäsitteistöä, josta tutkimukseen integroitavaa kirjallisuuskatsausta voidaan johtaa eteenpäin.

Tämän jälkeen käytin Science Direct tietokantaa käyttäen samoja haku sanoja ja hakumerkkijonoja aineiston etsimisessä. Haun tulokset olivat kuitenkin sen verran ylimalkaisia ja epätarkkoja, joten päätin aivan ensiksi käydä kaikki artikkelit läpi ja arvioida, onko niissä sellaista tietoa, jota tarvitsen omassa tutkielmassani. Haun tuloksena saatiin tässä haussa ns. kaksoiskappaleita(duplikaatteja), eli nämä artikkelit olivat jo edellisissä tietokannoissa tehdyissä hauissa mukana. Analysoituani artikkeleita valitsin niistä katsaukseen mukaan vain relevantimmat. Arvioin niitä ja analysoin niiden tärkeimmät löydökset. Toistin saman haun myös mm. IEEE, AIS digital libary ja INFORMS tietokannoista.

Sen jälkeen hain samoista tietokannasta hakumerkkijonolla Machine learning algorithm AND prediction of diabetic patients treatment outcome tietoa siitä, kuinka koneoppimisen avulla voidaan ennustaa diabeettisten potilaiden hoitotuloksia. Haun tuloksena saatiin 8 kpl julkaistuja artikkeleita. Toistin saman haun myös muihin tutkimuksessa käytettäviin tietokantoihin. Arvioin ja analysoin niitä, ja poistin haun tuloksena tulleet duplikaatit. Viimeiseen tutkimuskysymykseen hain tietokannoista tutkimusaineistoa hakumerkkijonoilla Machine learning AND diagnosing diabetes mellitus. Katsaukseen valittiin luettujen artikkelien otsikoiden, abstraktien ja yhteenvetojen perusteella 44 julkaistua tieteellisen arvioinnin läpikäynyttä artikkelia.

Taulukko 1. Lopullinen tutkimusaineiston määrä.

Katsaukseen valittujen tutkimusartikkelien määrä	Tietokanta
25	Pubmed
4	IEEE
2	AIS
10	Science Direct
3	Muut tietokannat

Webster ja Watson (2002) nostavat esille tutkimuksen kannalta relevantin kirjallisuuden identifioinnin tärkeyden tutkijalle. Heidän mukaansa korkea-laatuinen katsaus on valmis ja keskittyy ennen kaikkea sisältöön. Valmiiksi laadittu kirjallisuuskatsaus kattaa relevantin kirjallisuuden aiheesta ja ei ole rajoittunut yhteen tutkimusmetodologiaan, yhteen joukkoon julkaisuja tai yhteen maantieteelliseen alueeseen. Tietojärjestelmätieteen tutkimuksiin integroidut kirjallisuuskatsaukset keskittyvät pelkästään pieneen joukkoon huippujulkaisuja. Webster ja Watson (2002) suosittelvatkin rakenteellista lähestymistapaa katsauksen lähdemateriaalin määrittelyyn. Suurin vaikutus katsaukseen on todennäköisesti johtavilla julkaisuilla, ja täten kirjallisuuskatsauksen laatijan tulisi aloittaa katsauksen laatiminen niistä. Katsauksen laatijan tulisi myös tutkia valittujen konferenssien edistymistä, erityisesti laadultaan korkeatasoisten. Webster ja Watson (2002) korostavat tietojärjestelmätieteen olevan poikkitieteellinen tieteenala, joka ulottuu muihin tieteisiin. Tutkijan tuleekin tämän takia katsauksen ja teorian kehittämisen vaiheessa katsoa tietojärjestelmätieteen tiedonhaaran lisäksi oman tieteenalansa ulkopuolelle. (Webster & Watson 2002: 16.)

Webster ja Watson (2002) toteavat kirjallisuuskatsauksen myös olevan sisältö-keskeinen. Täten käsitteistö määrittää katsauksen rungon. Jotkut kirjoittajat ottavat kuitenkin tekijä keskeisen lähestymistavan ja tutkimukselle välttämätön ja olennainen esitetään

relevanteissa artikkeleissa. Tämä metodi ei kuitenkaan onnistu syntetisoimaan kirjallisuutta. He korostavatkin sitä, kuinka ns.tekijä keskeisestä lähestymistavasta voidaan helposti siirtyä sisältökeskeiseen lähestymistapaan sisältömatriisin avulla. Siinä tutkimukseen liittyvät avainkäsitteet on matriisissa sarakkeisiin jaoteltuna, ja aineisto ,josta ne on löydetty, riveihin jaoteltuna. (Webster & Watson 2002:17).Tämän tutkimuksen keskeiset käsitteet ovat eriteltynä seuraavan sivun taulukossa 2. Ne ovat koneoppiminen terveydenhoidossa, kliininen päätöksenteko, diabetes mellitus, kliininen päätöksenteon tukijärjestelmä ja koneoppiminen diabetes hoidossa.

Taulukko 2. Sisältömatriisi. (Webster ja Watson 2002).

Tutkimusaineisto	Käsite				
	A	B	C	D	E
1.	x				
2.		x		x	
3.			x		x
4.				x	
5.					x

Tutkimusaineisto näihin käsitteisiin on hankittu suomen- ja englanninkielisistä painetuista kirjoista, elektronisista lähteistä, kuten erilaisista tekniikan raporteista, arvostetuissa tieteellisissä journal-julkaisuissa julkaistuista katsaus-, tutkimus- ja konferenssiartikkeleista sekä elektronisesta tietokirjallisuudesta. Täten se on tieteelliset kriteerit täyttävää ja relevanttia aineistoa. Koneoppimisen soveltamista diabeteshoitoon tarkastellaan aivan uuden tutkimustiedon perusteella (viimeisemmän viiden vuoden ajalta julkaistut artikkelit). Taulukon riveillä on aihealuetta käsittelevä ryhmitelty aineisto. Mukana ei ole pelkästään yksittäinen aineisto, vaan se saattaa sisältää useitakin eri lähteitä, jotka määrittelevät tutkimuksen käsitteistöä.

Käsite A matriisitaulukossa kuvaa koneoppimista terveydenhoidossa. Tutkimusaineiston etsimisessä on pyritty huomioimaan erityisesti se, että tiedot ja käsitteiden määrittely tulee alkuperäislähteistä ja tieteellisesti korkea luokkaisista julkaisuista. Tällöin tieto on autenttista ja luotettavaa, ja lähde antaa täsmällistä tietoa aiheesta, ja käsite määritellään universaalisti. Koneoppiminen terveydenhoidossa on vielä varsin laaja aihe alue, ja se sisältää monella tavalla toteutettuja sovelluksia, tutkielmaan integroitiin vain kaikkein keskeisimpiä aihetta kuvailevia lähteitä.

Tutkimusmatriisin toisella rivillä on aineisto 1. Tämän aineiston on tarkoitus antaa selvä määritelmä koneoppimisen soveltamisesta terveydenhoitoon. Aineistossa on määritelty käsitettä eri merkitysyhteydessä, ja pyritty tuomaan esille siihen liittyviä eri konnotaatioita. Koneoppiminen (machine learning) on ennen kaikkea tekoälyn alakäsite, ja aineistossa kuvaillaan koneoppimisen liittyminen eri terveydenhoidon sektoreihin yhä enemmän erilaisilla tekoälyyn pohjautuvilla ratkaisuilla. Siinä tuodaan esille myös vaatimuksia ja odotuksia koneoppimisesta, joita ympäröivä yhteiskunta ja kehittyvä tiede tuo. Aineisto sisältää lähinnä elektronista kirjallisuutta, kuten raportteja ja tieteellisiä artikkeleita. Sitä on etsitty pääasiassa suomenkielisillä sanoilla, kuten koneoppiminen ja terveydenhoito. Englannin kielellä etsin aineistoa hakusanoilla Machine Learning sekä hakumerkkijonoilla Machine learning and Health care.

Käsitteellä B kuvataan tässä tutkimuksessa kliinistä päätöksentekoa. Kliinistä päätöksentekoa tarkastellaan terveydenhoidon ammattilaisten epävarmoissa tilanteissa tehtyjen päätösten näkökulmasta. Tähän on hankittu validia tutkimustietoa sitä julkaisevasta lähteestä.

Matrisiin kolmannella rivillä on aineisto kaksi. Tämä aineisto antaa hyvän yleiskuvan terveydenhoidon ammattihenkilöiden kliinisestä päätöksenteosta rajoitettujen vaihtoehtojen joukosta terveydenhuolto-organisaatiossa. Aineistoa on etsitty hakusanoilla kliininen päätöksenteko ja englanniksi clinical decision making.

Käsite C kuvailee tässä tutkimuksessa diabetes mellitus sairautta. Sillä pyritään kuvaamaan tämän kroonisen sairauden puhkeamisen syitä ja eri tyyppisiä

ilmenemismuotoja. Käsite sisältää myös kuvauksen sen eri liitännäissairauksista, esiintyvyyden yleisyydestä, sen kehityskulusta, ennaltaehkäisystä ja hoitokeinoista. Aineisto on kerätty tuoreista diabetes tutkimusraporteista, jotka on julkaistu elektronisena tutkimusraporttina. Tarkoitus on ollut saada ajankohtaista tietoa sairaudesta.

Kolmas aineisto tutkimusmatriisissa kuvailee diabeteksen esiintymisen yleistymisen ongelmia yhteiskunnallisista ja kulttuurisidonnaisista näkökulmista. Aineistossa kuvaillaan terveydenhuollon menojen kansallisia ja maailmanlaajuisia kustannuksia. Se sisältää elektronisia lähteitä, kuten tunnetun organisaation julkaiseman raportin, jossa on julkaistu tieteelliset kriteerit täyttävää tutkimustietoa sairaudesta, sekä elektronisesti julkaistun kattavan tutkimusraportin. Aineisto on hankittu hakutermillä diabetes mellitus.

Käsite D puolestaan tarkoittaa tutkielmassa kliinistä päätöksentekojärjestelmää. Se kuvailee kliinisen päätöksentekojärjestelmän osana laajaa sairaalatietojärjestelmää. Käsite määrittää kliinisen päätöksenteon tukemista varten suunnitelluksi tietojärjestelmäksi, jossa voidaan hyödyntää myös koneoppimista.

Aineisto numero neljä koostuu lähinnä elektronisesta tietokirjallisuudesta. Aineisto antaa laajan kuvan kliinisen päätöksentekojärjestelmän käyttötarkoituksesta kroonisten sairauksien, kuten diabetes mellituksen hoidossa. Se sisältää tärkeää tietoa mm. kliinisen päätöksentekojärjestelmän mahdollisuudesta tukea terveydenhoidon ammattilaisten päätöksentekoa kompleksisissa tilanteissa. Aineisto on hankittu hakutermeillä kliininen päätöksentekojärjestelmä ja clinical decision support system.

Käsite E kuvaa koneoppimista diabeteshoidossa. Käsitteeseen sisältyy koneoppimisen kuvaaminen terveydenhuollossa diabetes hoidon kontekstissa. Käsitteellä pyritään antamaan kokonaiskuva koneoppimisen sovellettavuudesta tämän vakavan kroonisen sairauden hoidossa. Se kuvailee myös koneoppimisen käyttökohteita diabetes hoidossa, koneoppimisen hyödyntämistä kliinisen päätöksenteon apuna diabetes hoidossa ja koneoppimisen mahdollisuuksista antaa diabetes hoidon ennustemalleja. Käsitteellä E pyritään tuomaan esille ennen kaikkea koneoppimisen mahdollisuudet tehdä diabetes hoidosta kustannustehokkaampaa ja laadukkaampaa. Käsite ulottuu myös kliinisten

kokeiden ja potilaiden terveydentilan seurantaan ja tarkasteluun. Tieto on hankittu ja lainattu tutkimusaineistosta, joka on lähes uutta ja tieteellisesti korkeatasoista.

Aineisto viisi tutkimusmatriisissa kuvaa koneoppimisen mahdollisuuksia diabeteshoidon laadun kehittäjänä. Aineiston pohjalta on tutkittu käsitettä E. Aineisto muodostuu elektronisista lähteistä, kuten tekniikan raporteista ja tieteellisissä julkaisuissa julkaistuissa artikkeleista. Aineisto hankittiin hakumerkkijonoilla Machine learning and diabetic care.

Tutkimuksessa pitäisi aina sen aineiston hankinnan ja arvioinnin jälkeen tehdä analyysi tutkimukseen integroitavasta aineistosta. Salminen (2018:8) painottaa, että tässä tutkimusvaiheessa täytyy analysoida ja tulkita tutkimukseen valikoitu ja hyväksytty tutkimusaineisto. Aineistoa voidaan analysoida monella menetelmällä. (Hirsjärvi, Remes & Sarjavaara 2008) huomauttavat, että analyysitavat on jäsennettävissä selittämiseen ja ymmärtämiseen pyrkiviin analyysimenetelmiin. Selittämiseen pyrkivässä analyysitavassa käytetään tavallisesti tilastollista analyysia ja johtopäätösten tekoa, ja ymmärtämiseen pyrkivässä tavassa laadullista analyysia ja johtopäätösten tekoa. (Hirsjärvi ym.2008:219). Tässä tutkimuksessa käytetään aineiston analyysisissa ymmärtämiseen pyrkivää laadullista analyysia, jonka perusteella saadaan tutkimuksen tulokset ja tehdään johtopäätökset.

Kvalitatiivisessa tutkimuksessa aineiston analyysi tehdään usein osittain samanaikaisesti aineiston keräyksen yhteydessä (Hirsjärvi ym. 2008:218). Tutkimusaineiston analyysia tehdään tässäkin tutkimuksessa samanaikaisesti keräyksen yhteydessä. Sisältöanalyysissä aineistoa tarkastellaan Tuomen & Sarajärven (2002) mukaan eritellen, yhtäläisyyksiä ja eroja etsien ja tiivistäen. Sisältöanalyysi on diskurssianalyysin tapaan tekstianalyysia, siinä tarkastellaan jo valmiiksi tekstimuotoisia tai sellaiseksi muutettuja aineistoja. Tutkittavat tekstit voivat olla esim. kirjoja, päiväkirjoja, haastatteluita, puheita ja keskusteluita. Sisältöanalyysi mahdollistaa tiivistetyn kuvauksen tekemisen tutkittavasta ilmiöstä, ja sillä voidaan myös kytkeä tutkimuksen tulokset laajempaan kontekstiin ja aihealuetta koskeviin muihin tutkimustuloksiin. (Tuomi & Sarajärvi 2002:105.)

Sisällönanalyysillä tarkoitetaan Tuomen & Sarajärven (2002) mukaan enemmänkin laadullista sisällönanalyysia kuin sisällön määrällistä erittelyä, ja näitä kumpiakin

voidaan hyödyntää analysoitaessa samaa tutkimusaineistoa. Sisällönanalyysia on mahdollista jatkaa eteenpäin tuottamalla esim. sanallisesti kuvatusta aineistosta määrällisiä tuloksia. Tutkimusaineiston laadullisessa sisällönanalyysissa aineisto jaotellaan ensin pieniin osiin, käsitteellistetään ja järjestetään lopuksi vielä täysin uudenlaiseksi kokonaisuudeksi. Sisällönanalyysi on mahdollista tehdä aineistolähtöisesti, teoriaohjautuvasti tai teorialähtöisesti. Erona näissä käsitteistössä on se, että analyysi ja luokittelu perustuu aineistoon tai valmiiseen teoreettiseen viitekehykseen. (Tuomi & Sarajärvi 2002: 106-116.)

Laadullisessa tutkimuksessa teoriaa on mahdollista käyttää apuvälineenä. Tällöin voidaan tehdä tulkintoja kerätystä tutkimusaineistosta. Teoria voi olla kuitenkin myös päämäärä, mikäli tavoitteena on uuden teorian luominen tai vanhan teorian kehittäminen paremmaksi. Tutkimusprosessin aloittamisvaiheessa tutkijan on myös aina päätettävä, mikä asema teorialla on hänen tutkimuksessaan. Kolme vaihtoehtoista tutkimustapaa ovat teorialähtöinen, teoriasidonnainen ja aineistolähtöinen tutkimus. (Eskola 2001a:135-140.) Teorialähtöinen tutkimus tarkoittaa sitä, että tutkimusaineiston analyysi perustuu olemassa olevaan teoriaan tai malliin. Aineiston analyysia ohjaa tällöin valmis malli, ja tarkoituksena on yleensä tämän mallin tai teorian testaaminen. Uudessa yhteydessä tätä analyysimallia on käytetty tavallisesti luonnontieteellisessä tutkimuksessa. (Tuomi & Sarajärvi 2002:95-99.) Suoraan teoriaan pohjautuvaa aineistoanalyysia kutsutaan myös teorialähtöiseksi analyysiksi. Teorialähtöisessä analyysissä aineiston keruu on aina etukäteen määritelty. (Tuomi 2007:131).

Aineistolähtöisessä tutkimuksessa tutkimus perustuu aineistoon. Se tarkoittaa sitä, että esim. analyysiyksiköt eivät ole täysin ennalta määrättyjä ja teoria rakennetaan lähtökohtaisesti tutkimusaineistosta. Tutkimuksen lähestymistapa on tällöin induktiivinen, eli siinä esitetään väitteitä, jotka etenevät yksittäisistä havainnoista yleisempiin väitteisiin. (Eskola & Suoranta 1998:83). Induktiivinen lähestymistapa ei siis perustu teorian tai hypoteesin testaamiseen, tutkija ei voi myöskään määritellä sitä, mikä on tutkimuksen kannalta tärkeää. (Hirsjärvi ym. 2004: 155.) Aineistolähtöisessä tutkimusaineiston analyysissä yritetään luoda tutkimusaineistosta teorettinen kokonaisuus tutkimuksen tarkoituksen pohjalta valittujen analysoitavien kohteiden aineistosta. Analysoitavat kohteet eivät myöskään ole millään tavoin etukäteen päätettyjä,

toisin kuin teoriaohjautuvassa analyysissä, vaan ne nousevat esille tutkimusaineistosta ja niillä ei ole varsinaisia kytköksiä aiempiin teorioihin ja havaintoihin. (Tuomi 2007: 130-131.)

Teoria- ja aineistolähtöisen tutkimuksen välimaastoon voidaan puolestaan sijoittaa teoriasidonnainen tutkimus. Siinä aineiston analyysi ei perustu välttämättä suoraan teoriaan, mutta siihen on havaittavissa yhtymäkohtia. Tutkimusaineistolle etsitään siis teoriasta selityksiä tai vahvistuksia tukemaan teoriaa. Tutkija kykenee tekemään huomioita empirian vastaamattomuudesta aiempiin tutkimustuloksiin. (Eskola 2001). Tämä teoria ohjautuva analyysi on puhtaasti teorialähtöisen ja aineistolähtöisen analyysin välissä, joka myös alkaa yleensä niin kuin aineistolähtöinen analyysi. Teoriaohjautuvassa analyysissä on kuitenkin tunnistettava aiempaan teoriaan pohjautuvat vaikutukset ja kytkökset. (Tuomi 2007:131).

Tässä tutkimuksessa on käytetty tutkimusaineiston analysoinnissa pääasiassa aineistopohjaista analysointitapaa, mutta joitakin aineistoja on analysoitu myös teoriaohjaavasti ja teoriaan pohjautuvaa analysointitapaa käyttäen. Aineistoja, jotka liittyvät koneoppimisen soveltamiseen terveydenhoitoon, on analysoitu jo kertaalleen hyväksytyjä teorioita ja muita tutkimusaineistoja hyödyntäen. Kliinisen päätöksenteon ja kliinisen tietojärjestelmän tutkimusaineistojen analysoinnissa on käytetty myös teoriaohjaava aineiston analyysitapaa, koska tutkimusaineiston analysointi perustuu paljon aihealueen valmiiseen vanhan teoriaan ja julkaistuihin teoksiin sekä raportteihin, joihin perustuen on analysoitu aiheeseen liittyvää aineistoa ja integroitu se mukaan tähän tutkimukseen. Tutkimuksessa on käytetty myös puhtaasti teoriaohjaava aineiston analysointitapaa, ja tämä tapahtuu etenkin kliinisen päätöksenteon ja kliinisen päätöksentekojärjestelmän tutkimuksen käsitteistön kohdalla. Kliiniseen päätöksentekoon liittyvää aineistoa on analysoitu paljon Higgisin, Jonesin, Loftuksen ja Christensenin (2009) ja kliinisiä tietojärjestelmiä Bernerin (2007) teosten pohjalta.

6. Koneoppimismallit diabetes hoidossa

Koneoppimismalli on matemaattinen kuvaus reaali maailman prosessista. Koneoppimismallin muodostamiseksi täytyy ensin hankkia opetustietoaineistoa, josta koneoppimisalgoritmi voi oppia. (ks. Bhattarcharjee 2017.)

Diabeteksen lääketieteellisessä hoidossa tarvitaan tietokonejärjestelmä, joka voi vaikuttaa terveydenhuollon ammattihenkilöiden päätöksentekoprosessiin nopeasti ja tehokkaasti. Tarvitaan koneoppimista hyödyntävä päätöksenteon tukijärjestelmä. Tässä kappaleessa pyrin tekemään ideakeskeisen ja kuvailevan synteesin valitusta tutkimusaineistosta saadusta tutkimustuloksesta, diabetes hoidon kliinisen päätöksenteon tukemisen edistämisen näkökulmasta. Tarkastellaan sitä, mitkä ovat ne koneoppimismallit, joilla voidaan tukea kliinikoiden päätöksentekoa diabeteshoidon yhteydessä terveydenhuolto organisaatioissa. Tulosten selvyuden parantamiseksi olen jaotellut tutkimusosioiden tuloksia eri kategorioihin, ja muodostanut käsitteistä yhdenmukaisia, kuvailevia yleiskäsitteiden ryhmittymiä. Lähdeaineisto on myös luokiteltu toisiinsa liittyviksi joukoiksi, jotta tutkimuksen keskeisempien käsitteiden välille voisi muodostua merkityksellisiä suhteita. (ks. Rowe 2014.)

6.1 Koneoppimisen soveltaminen diabetes hoitoon

Tässä tutkimuksessa on etsitty uusimmasta tutkimuskirjallisuudesta tietoa koneoppimisen sovellettavuudesta diabeteshoitoon, jotta voitaisiin antaa ohjeellistavaa tietoa ei-tietämyskantaisen kliinisen päätöksenteko järjestelmän suunnittelumallin toteuttamiseksi terveydenhuolto organisaatioihin. Tutkimuksessa tarkastellaan millä tavoin koneoppimista sovelletaan diabetes hoidossa nykyisin, ja mm. Schwartzin (1970) sekä Vähäkainun ja Neittaanmäen (2017) tekemän määritelmään perustuen on etsitty toimivan ei-tietämyskantaisen kliinisen päätöksenteko tukijärjestelmään sisällytettyjä koneoppimismalleja. Seuraavalla sivulla kuvatussa taulukossa on eritelty diabetes hoidossa yleisesti sovellettavia koneoppimismalleja ja algoritmeja.

Taulukko 3. Koneoppimisen soveltamisen mahdollisuudet diabetes hoidossa.

Koneoppimismallit, algoritmit ja sovellukset	Käyttötarkoitus Diabetes hoidossa	Aineisto
Syväoppimiseen pohjautuva järjestelmä/syväoppimisen teknologia	Silmäsairauksien tekninen ja kliininen tarkastelu, diabeteksen liitännäissairauksien täsmällinen tunnistaminen digitaalisten silmänpohjan verkkokalvokuvien perusteella. Tunnistaa sairaustiloja, kuten diabeettinen retinopatia (DR), glaukooma (silmänpainetauti), ikään liittyvä silmän pilkun makulaarinen rappeutuminen (ADM, Age Related Macular Degeneration) varhaiskypä retinopatia.	Ting & Wong (2018)
Syväoppimiseen pohjautuva järjestelmä/syväoppimisen teknologia	Silmäsairauksien seulonta, lääketieteellinen kuvaantaminen, hahmon tunnistus, kuvien luokitus, lähete optomologille koskien diabeettista retinopatiaa ja diabeettista makulaarista turvotusta, oikea-aikaista hoitoa ja näön menetyksen ehkäisyä varten. Digitaalinen valokuvaus näköyhteys röntgenkerroskuvauksella tapahtuvaa näkökenttien sairaus prosessin arviointia varten. Optomologiseen hoitoon integrointia silmänpainetautiin, silmään liittyvää makulaarista rappeutumista ja diabeettista retinopatiaa varten.	Cheung, Tang, Ting, Tan & Wong (2019) Grewal, P. S., Oloumi, F., Rubin, U., & Tennant, M.T.S. (2018)
Syväoppimisjärjestelmä suunniteltu diabetekseen liittyviin silmän sairauksiin	Diabetes epidemian hoito, diabetes hoidon saatavuus laajassa globaalissa populaatiossa suuresta monietnisestä populaatiosta. Havaitsee viittauksia diabeettiseen retinopatiaan, silmänpaine tautiin, ikään liittyvään makulaariseen rappeutumaan yhteisössä ja kliinisessä hoidossa olevassa moni-etnisessä populaatiossa. Diabeettisen retinopatian havaitseminen keskivaikeassa, lisääntyvässä diabeettisessa retinopatiassa tai sitä vaikeampiasteisessa näköä uhkaavassa diabeettisessa retinopatiassa	Ting, D.S.W, Peng, L., Varadrajana, A. V., Keane, P. A., Burlina, P.M., Chiang, M, F., Schemetter, L., Pasquale, L. R., Bressler, N. M., Webster, D.R. , Abramoff, M.& T.Y. Wong (2018)
Syväoppiminen	Diagnostisen teknologian kehittäminen silmäsairauksien diagnosointi/silmän verkkokalvokuvien oivallukset, digitaaliset kuvat tarjoavat miljoonia morfologisia datajoukkoja. Ne voidaan analysoida ymmärrettävästi tekoälyä käyttäen. Identifoi, paikallistaa ja saa selville patologiset piirteet melkein jokaisessa makulaarisessa ja verkkokalvo sairaudessa.	Schmidt-Erfurth, Sadeghipour, Gerendas, Waldstein & Bogunovic (2018)
Koneoppimisjärjestelmä Gaussilaiseen dataluokitteluun perustuen	Kerää diabetes mellitus sairauden kannalta dataa nopeasti. Tilastollisen datan luokittelun hyödyntäminen tuo merkittäviä taloudellisia säästöjä terveyden hoitomenoissa. Diabeteksen ennakointi helpompaa jos diabetestä pysyttäisiin ennakoimaan tilastollisesti perustuen johonkin sen yhteisvaihteluihin.	Maniruzzaman, Md., Kumar, N., Abedin, Md, M., Islam, S, MD., Suri, M.S., El-Baz, A, S., & J.S. Suri (2017)
Modifioitu Hopfield Hermoverkko (MHNN, Modified Hermenuth Neural Network)	Diabeettisen retinopatian diagnosoiminen verkkokalvokuvista.	Hermanth, D. J., Anitha, J., Son, L.H., & J. Mittal (2018)

Koneoppiminen ja laskennalliset ohjelmointimetodit keinotekoiset hermoverkot		
Koneoppiminen ja datan louhinta menetelmät	Diabetes tutkimuksen teko, bioteknologian ja terveystieteiden edistäminen. merkittävä tietoaaineiston lisääminen. Suuri toimitettu määrä geneettistä dataa ja kliinistä tietoa. Nämä muodostetaan suurista elektronisista terveysrekisteistä(electronic health records, EHR)	Kavaktosis, I., Tsave, O., Salifoglor, A., Maglaveras, N., Vlahavas, I. & J. Chowarde (2017)
Laskennallinen menetelmä Syväkoneoppimiseen perustuva menetelmä	Kliinikoiden hoitotyössä tehokas ja täsmällinen diabetekseen liittyvien silmäsairauksien havaitsemisessa. Havaitsee herkästi ja tarkasti viittauksia diabeettiseen retinopatiaan ja arvio verkkokalvon silmänpohjan valokuvia aikuisikäisille diabeetikoille.	Guishan, V., Peng, L., Coran, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venuagopalan, S., Widner, K., Maradas, T., Cuadros, J., Kim, R., Raman, R., Nelson, P., Mega, L. & Webster, D.R (2016)
Koneoppimisalgoritmi, joka integroi verimuuttujat, ravintotottumukset, antropometrit, fyysisen aktiivisuuden ja suolen mikrobiston	Kykenee vertailemaan diabeetikoita nykyisiin ravintosuosituksiin perustuvassa aterioiden jälkeisessä verensokerin muutoksissa. Näin ollen diabeetikoiden ruokavaliosta ja dieettihoidosta voidaan tehdä suunnitelmallista ja yksilökeskeistä. Koneoppimisalgoritmin avulla voidaan täsmällisesti ennustaa aterian jälkeinen glykeeminen vastine tosielämän aterioihin.	Zeevi, D, Kovem, T., Zmora, N., Israeli, D., Rotschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnii-Sagi, T., Lotan-Pomper, M., Suez, J., Ali Mahdi, J., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Kosower, N., Rein, M., Zillerman-Schapira, G., Dohnalova, L., Pereoner-Fischer, M., Bikovsky, R., Halper, Z., Elinav, E. & E. Segal (2015)
Piirvektorien joukko, puoliohjattu oppiminen ja graafiin pohjautuva luokittelu	Tulehduseritteiden identifiointi, makulan paikallistaminen	Ren, F., Cao, P., Zhao., D & C. Wan (2018).
Ennustava mallinnus MIMIC-3 tietokanta, koneoppimisteknologia ja kaksijäseninen tietokantainen regressiomallinnus	Diabetes potilaiden kuolleisuuden ennustaminen, vakavuusasteen osoitus, diabeteksen suhteellinen vaikutus.	Anand, R.S., Stey, P., Jain, S., Biron, D.R., Bhatt, H., Monteiro, K., Feller, E., Ranney, M.L, Sarkar, J.N. & Chen, S. E.(2018)
Syvä kaareutuva monimutkainen hermoverkko	Diabeetikoiden terveyden hoito, tehohoitoa ja potilaiden sairauksien yhteisvaikutusten tulos. Hoidon aiheen ennakoiminen perustuen keskeiseen verkkokalvon kuvien näköyhteys röntgen kerroskuvaus seulontaan ilman ihmisen interventioita.	Prahs, P., Radek, V., Mayer, C., Cvetkova, Y., Cvetkova, N., Helbig, H. & D. Märker (2018)
Riskikaavan kehittäminen ja arviointi Ristiinvalidointi koneoppimismetodi	Kakkostyyppin diabeteksen komplikaatioiden riskin oikeanlainen arviointi. Kardiovaskulaaristen riskien toimenpiteistä kardiovaskulaaristen riskien torjuntaan. Vahvistavat mikrovaskulaaristen tapahtumien kaavoja käyttäen diabeteksen ehkäisyohjelman tutkimustuloksia ja kardiovaskulaarisia tapahtumia käyttäen tutkimuksen dataa. Valitsee ennustavia muuttujia demograafisista merkeistä, kliinisistä muuttujista, sairauksien yhteisesiintymisestä(monihäiriöisyydestä), lääkkeistä ja biomarkkereista.	Basu, S. Sussman, J. Berkowiz, S.A., Hayward, R. & J.S. Yudkin(2015)

<p>Tapahtumapohjainen todistus viitekehys</p> <p>Tapahtumapohjainen todistus</p>	<p>Lääketieteellisten ongelmien ratkaisu.</p> <p>Ongelman ratkaisu paradigma. Siinä käytetään vanhaa tietoa uusien ongelmien ratkaisussa ja tutkinnassa.</p> <p>Sopii asiantuntija pohjaiseen ja teoriaa vähemmän sisältävään ongelmaan.</p>	<p>El-Sappagh, S., Elmogy, E., Riad, A.M.(2015)</p>
<p>Tutkimuskuvien ja koneoppimisalgoritmien yhdistelmä</p>	<p>Diabeettisen retinopatian seulonta.</p>	<p>John (2016)</p>
<p>Älykäs autonominen järjestelmä(Intelligent Diabetic Assistant(IDA))</p>	<p>Kliininen päätöksenteon tukeminen ja lähestyvän reaali-aikaisen jalka mätähaavan havaitseminen ja rajapinnan seulonta.</p>	<p>Wiejinsigne (2019)</p>
<p>Kuvaperusteinen kehysten hyödyntämisen maksimointi ruokavalion virheiden määrittämisen maksimoimiseksi. (Syväoppimista hyödyntävä koneoppimissovellus diabetes hoidossa)</p>	<p>Mallin ja uuden kehysten puitteiden rakentaminen ruokavalion datan analysoimista varten kuvien ja tekstin muodossa. Ravintotuotteiden arvoa ennustava, ja suosittelee myös terveellisiä vaihtoehtoja diabeetikoille reaali-ajassa.</p> <p>Automaattinen eri ruokakomponenttien tunnistaminen ravinto kuvista. Korkean terveyshyöty ravintosuosittelun tekeminen yksilöiden yhdenmukaisten makumieltymysten kanssa.</p>	<p>Qiu, L., Bhattarya, P.& P. Tuan (2017)</p>

Integroivan kirjallisuuskatsauksen avulla tutkimuksessa on yritetty etsiä vastauksia siihen, miten kliinistä päätöksentekoa Higgisin (ym.2008) määritelmän pohjalta voidaan edistää koneoppimisen käyttöönotolla diabeteshoitoon, jotta hoidon laatu toteutuisi terveydenhuolto-organisaatioissa mahdollisimman hyvin. Terveystenhoito ammattilaisilta vaaditaan nykyisin kliinisissä päätöksissä monien tekijöiden huomioon ottamista. Täytyy pystyä tekemään päätöksiä. Päätökset voivat tarkentaa mm.diagnooseja, interventioita, interaktioita ja arviointeja hyödyntäen monipuolista tietämuskantaa.(Higgis ym.2008.) Kliininen päätöksentekojärjestelmä pitäisi myös aina suunnitella terveydenhoitopaikkaan aina sellaiseksi, että se antaa tukea päätöksentekohetkellä. (ks. Berner 2007.)

Aivan viime aikoina diabeteshoidossa on käyttöönotettu ja sovellettu paljon syväoppimiseen perustuvia järjestelmiä (ks. Ting & Wong 2018; Cheung, Tang, Ting, Tan & Wong 2019; Grewal, Oloumi, Rubin & Tennant 2018; Ting, Peng, Varadrajan, Keane, Burlina, Chiang, Schemetter, Pasquale, Bressler, Webster, Abramoffm & Wong 2018; Schimdt-Erfurth, Sadeghipour, Gerendas, Waldstein & Bogunovic 2018; Guishan, Peng, Coran, Stumpe, Wu, Narayanaswamy, Venuagopalan, Widner, Maradas, Cuadros, Kim, Raman, Nelson, Mega, & Webster 2016; ja Qiu, Bhattarya.& Tuan 2017).

Tutkimuksesta nousi esille se, että syväoppimiseen pohjautuvilla järjestelmillä pystytään teknisesti ja kliinisesti tarkastelemaan tehokkaasti diabetekseen liittyviä silmäsairauksia. Syväoppimiseen pohjautuva järjestelmä mahdollistaa etenkin diabeteksen liitännäissairauksien täsmällisen tunnistamisen silmänpohjan verkkokalvokuvien perusteella. (ks. Ting & Wong 2018.) Järjestelmää sovelletaankin diabeteksen hoidossa paljon diabetes potilaiden sairaudesta aiheutuvien silmäsairauksien kliinisessä ja teknisessä tarkastelussa, koska se tunnistaa sellaisia sairaustiloja, kuten diabeettinen retinopatia(DR), glaukooma(silmäpainetauti), ikään liittyvä silmän pilkun makulaarisen rappeutuminen (ADM, Age Related Macular Degeneration), varhaiskypsän retinopatian ja silmän taittovirheet. Tekoälyä hyödyntävien koneoppimis- ja syväoppimistekniikoiden suorituskyky on viime aikoina kehittynyt johtuen automaattisten kuvankäsittely yksiköiden edistymisestä, matemaattisista malleista ja big datan skaalattavuudesta. (ks. Ting & Wong 2018; Grewal, Oloumi, Rubin & Tennant 2018.)

Tekoäly on tullut vahvasti mukaan diabetekseen liittyvien silmäsairauksien seulontaan. Koneoppimis teknologian uusi haara syväoppiminen (deep learning) on tehnyt huomattavan läpimurron lääketieteellisessä kuvantamisessa erityisesti hahmon tunnistamisessa ja kuvien luokittelussa. Systemaattiseen ja kansalliseen diabeettisen retinopatian tai diabeettisen makulaarisen turvotuksen seulontaan käytetään digitaalista pohja valokuvausta, ja näköyhteys kerroskuvaus röntgensäteillä (OCT) suoritettavaa kuvausta ollaan nykyisin toteuttamassa jo ensisijaisen perushoidon tasolla. Tämän tarkoituksena on antaa oikea-aikainen lähete optamologille koskien diabeettista retinopatiaa, diabeettista makulaarista turvotusta, oikea-aikaista hoitoa ja näön menetyksen ehkäisyä varten. (ks. Cheung, Tang, Ting, Tan & Wong 2019.)

Syväoppiminen on nopeasti kasvava teknologia ja sillä on useita potentiaalisia sovelluksia erityisesti optomologiassa. Syväoppimis työkaluja on käytetty myös erilaisin tavoin mukaan lukien digitaalinen valokuvaus, näköyhteys röntgen kerroskuvaus ja näkökentät. Nämä työkalut ovat osittaneet hyödyllisyytensä useiden sairausprosessien arvioimisessa mukaan lukien silmäpainetauti, silmään liittyvä makulaarinen rappeutuminen ja diabeettinen retinopatia. Syväoppimis teknologiat kehittyvät nopeasti, ja ne tulevat laajasti integroiduksi optomologiseen hoitoon. (ks. Grewal, Oloumi, Rubin & Tennant 2018.)

Koneoppimiseen perustuvista metodeista erityisesti syväoppiminen voi identifioida, paikallistaa ja saada selville patologiset piirteet melkein jokaisessa makulaarisessa ja verkkokalvo sairaudessa. Perinteiset hermoverkot jäljittelevät ihmisen aivojen polkuja objektien tunnistamisessa, oppimalla patologisia piirteitä opetus joukosta, ohjautusta oppimisesta tai jopa hahmojen tunnistamisen yleistämisestä itsenäisesti ja myös ohjaamattomasta oppimisesta. (ks. Schimdt-Erfurth, Sadeghipour, Gerendas, Waldstein & Bogunovic 2018.) Tutkimuksessa nousi myös esille se, että täysin automatisoidut tekoäly pohjaiset järjestelmät on hyväksytty laajasti mukaan diabeettisen retinopatian seulontaan. Koneoppiminen ja syväoppiminen kehittääkin merkittävästi seulontaa, diagnostista luokittelua, kuten myös hoitoa ja terapian ohjaamista, sairaus aktiviteetin uusiutumisen valvontaa, automaattista havaitsemista, parantaa hoitavia vaikutuksia ja asiaankuuluvien kohteiden identifiointia/tunnistamista uusia hoitosuosituksia varten. (ks. Schimdt-Erfurth ym. 2018.)

Diabetekseen liittyvien silmäsairauksien suunniteltua syväoppimisjärjestelmää on kehitetty ja arvioitu. He arvioivat tutkimuksessaan syväoppimisjärjestelmän suorituskykyä havaita viittauksia diabeettiseen retinopatiaan, mahdolliseen silmänpaine tautiin ja ikään liittyvään makulaariseen rappeutumaa yhteisössä ja kliinisessä hoidossa olevassa moni-etnisessä populaatiossa. Koneoppimisteknologioita hyödynnetään parhaiten maailmanlaajuisen diabetes epidemian hoidossa kehittämällä sairauden seulontajärjestelmää. Tällöin on mahdollista kehittää erityisesti diabetes hoidon saatavuutta laajassa globaalissa populaatiossa, koska koneoppimisteknologioilla pystytään nykyisin seulomaan suuresta moni-etnisessä populaatiosta diabetekseen liittyviä silmäsairauksia hyvin tehokkaasti. (ks. Ting, Peng, Varadraján, Keane, Burlina, Chiang, Schemetter, Pasquale, Bressler, Webster, Abramoff & Wong 2018.) Diabeteksen tullessa maailmanlaajuisesti epidemiaksi pitää terveydenhuoltoon käyttöönottaa tehokas sairauden seulontajärjestelmä, joka pystyy tarjoamaan päätöksentekijöille relevanttia tietoa diabeteksen liitännäissairauksien tilasta tarkasteltaessa suurta populaatiota.

Syvään koneoppimiseen perustuvat algoritmit kykenevät tutkimusnäytön mukaan herkästi ja tarkasti havaitsemaan viittauksia diabeettisen retinopatiaan arvioitaessa verkkokalvon silmänpohjan valokuvia aikuisikäisillä diabeetikoilla. Syväoppiminen määritellään laskennalliseksi metodiksi, joka sallii algoritmin ohjelmoida itsensä oppimalla isosta joukosta esimerkkejä. (ks. Guishan, Peng, Coran, Stumpe, Wu, Narayanaswamy, Venuagopalan, Widner, Maradas, Cuadros, Kim, Raman, Nelson, Mega, & Webster 2016.) Syväoppimisesta saadaan diabeteshoitoon erinomainen työkalu havainnoitaessa teknisesti sairauteen olennaisesti liittyvän näköhäiriön ilmenemistä ja automatisoimalla tätä tunnistamisprosessia.

Tutkimuksessa nousi esille myös se, että koneoppimisteknologialla pystytään edistämään klinisen päätöksentekojärjestelmän suorituskykyä diabeetikoiden suunnittelemtomien uusiutuvien sairaalahoitojaksojen riskien identifioinnissa. Tämä voi parantaa hoidon laatua sairaalassa olon aikana ja vähentää sairaalahoitojaksojen esiintyvyyttä. Sairaalalle ominaisesta sairaalaan takaisinsijoittamisesta tulee monesti vaikea ja kriittinen asia maailmanlaajuisesti, koska suunnittelemtomat uusiutuvat sairaalassa olojaksot vievät suuren osan sairaalan kokonaiskustannuksista. Näin ollen tätä diabeteshoidon ongelmaa on pyritty ratkaisemaan tukivektorimallin ja geneettisen algoritmin yhdistelmällä, joka on kokeellinen, uusi metodi riskien ennustemallien rakentamiseksi. Tähän malliin kuuluu piirteiden valinta ja epätasapainoisen datan käsittely, ja se antaa päätöksenteon tukea klinikoille. (ks. Cui, Wang, Wang, Yu & Yaochu 2018.)

Tutkimustuloksista käy ilmi myös se, että keinotekoisista hermoverkoista Modifoitu Hopfield neuraali hermoverkko kykenee täsmällisesti diagnosoimaan diabeettisen retinopatian silmän verkkokalvokuvista. (ks. Hermanth, Anitha, Son & Mittal 2018.) Abnormaaliuden identifiointi verkkokalvokuvista on erittäin haastava tehtävä, ja keinotekoisien hermoverkkojen käyttöä suositaankin paljon diabeettisen retinopatian diagnosoimisessa verkkokalvokuvista. Modifoitu Hopfield Neural Network (MHNN) hermoverkko on tehokas työkalu ihmisen abnormaalien verkkokalvokuvien luokitteluun. MHNN:ään pohjautuvalla Lotus Eye Care Hospital sovelluksella tehdyssä kokeessa kerättiin kuvia, jotka osoittivat MHNN:än herkkyyden ja specificisyyden arvoksi 0,98 ja tarkkuuden 99,25 %. MHNN osoittautui paremmaksi kuin (Hopfield Neural Network, (HNN) ja muut neuraali hermoverkko lähestymistavat diabeettisen retinopatian diagnosoimiseen verkkokalvokuvista. (ks. Hermenanth 2018.)

Koneoppimisen yksi suurimmista hyödyistä diabeteksen hoidossa on sen antama mahdollisuus ennustaa diabetes potilaiden kuolleisuutta. Tämä ilmenee etenkin tehohoidossa olleiden diabetes potilaiden kuolleisuuden ennustamisessa, jossa voidaan käyttää hyväksi koneoppimista ja vakavuusasteen osoitusta. Diabetes muodostaa huomattavan terveystingon, joka johtaa moniin pitkä-aikaisiin terveys ongelmiin, mukaan lukien verkkokalvo, kardiovaskulaariset ja neuropaattiset komplikaatiot. Monet näistä ongelmista voivat johtaa lisääntyneisiin terveyden hoidon kustannuksiin, kuten myös riskiin teho-hoitoon jäämisestä ja kuolleisuudesta. MIMIC-3 tietokanta, koneoppimisteknologia ja kaksijäseninen regressio mallinnus soveltuvat ennuste malliksi kuolin todennäköisyyden riskien ennustamiseen. Koneoppimisella voidaan saavuttaa hyvä käyttökelpoisuus ja vaativa tietoaaineiston luokittelu viiden muuttujan yhdistelmänä, toisin kuin aikaisemmissa koneoppimismalleissa, jotka ovat saattaneet vaatia lähes 35 muuttujaa samanlaisen riskien arviointiin ja ennustamiseen. (ks. Anand, Stey, Jain, Biron, Bhatt, Monteiro, Feller, Ranney, Sarkar & Chen 2018.) Koneoppimisella voidaan ennustaa teho hoidossa olevan diabeetikon kuolleisuutta tarkoin tietyistä piirteistä ja suorittaa täsmällistä riskiarviointia vain muutamaa keskeisimpään muuttuajaan perustuen.

Tutkimuksessa havaittiin myös se, että koneoppimisjärjestelmällä voidaan nykyisin luokitella myös diabeetikon makulaarisen turvotuksen vakavuusaste verkkokalvokuvista. Makulaarinen turvotus on yksi diabeettisen retinopatian vakava-asteinen komplikaatio, joka aiheuttaa vakava-asteista näön menetystä ja se johtaa hoitamattomana sokeuteen. Makula(silmäpilkku) pystyttiin luokittelemaan käyttämällä anatomisia piirteitä, ja tulehduseritteen alueet pystyttiin havaitsemaan uudella metodilla. Taudin vakavuusaste pystyttiin luokittelemaan eri vaiheisiin perustuen tulehduseritteen sijaintiin ja makula täplän koordinaatteihin. Makula pystyttiin paikallistamaan 0,975 ja 0,942 täsmällisyyden keskiarvolla. Tutkimus osoitti myös, että tulehduseritteen luokittelussa voidaan käyttää hyväksi puoliohjattua oppimista ja tulehduseritteen tunnistamisessa vektori kvantisointia. (ks. Ren, Cao, Zhao & Wan 2018.) Koneoppimisella voidaan luokitella ja paikallistaa diabeettisen retinopatian vakava-asteinen komplikaatio täsmällisesti, joka auttaa kliinisiä päätöksentekijöitä mm. sairauden vakavuusasteen seuraamisessa.

Käyttämällä suuria määriä verkkokalvokuvia syväoppimisjärjestelmä voidaan opettaa havaitsemaan keskivaikeaa ja vaikea-asteista (näköä uhkaavaa) diabeettista retinopatiaa, ikään liittyvää makulaarista rappeutumaa ja mahdollista silmäpainetautia. Syväoppimisjärjestelmä kykenee verkkokalvo spesialisteille ja yleisille optomologeille asetettujen standardien mukaisiin suosituksiin, näiden sairaustilojen havaitsemisessa herkkyyden ja tarkkuuden suhteen. (ks. Ren 2018.) Syväoppimisjärjestelmä pystyy havaitsemaan diabetekseen liittyviä näköhäiriöitä tarkoin ja antamaan asiantuntevia suosituksia klinikoille nopeasti.

Samoihin johtopäätöksiin on päädytty myös muissa syväoppimisjärjestelmän suorituskykyä mittaavissa diabeettista retinopatiaa, silmäpainetautia ja ikään liittyvää makulaarista rappeutuman havaitsemista mittaavissa tutkimuksissa. Syväoppimisjärjestelmällä on korkea herkkyys ja tarkkuus diabeettisen retinopatian identifioinnissa ja siihen liittyvissä silmäsairauksissa. Syväoppimisjärjestelmän antama AUC arvo (Area Under Receiver Operating Characteristic Curve) diabeettiselle retinopatialle on 0,936, herkkyys on 90,5% ja tarkkuus on 91,6%. Näköä uhkaava diabeettinen retinopatia AUC arvo on 0,958, herkkyys on 100 % ja tarkkuus on 91,1 %. Mahdolliselle silmänpainetaudille AUC on 0,942, herkkyys on 96,4 % ja tarkkuus on 87,2%. Ikään liittyvään diabeettiseen makulaariseen rappeuman AUC arvo on 0,931 ja herkkyys on 93,2%. (ks. Ting, Peng, Varadrajan, Keane, Burlina, Chiang, Schemetter, Pasquale, Bressler, Webster, Abramoff & Wong 2018.) Suurella moni-etnisellä otannalla tehdyt tutkimukset ovat osoittaneet, että syväoppimisella pystytään havaitsemaan ja luokittelemaan tarkoin diabeteksen aiheuttamia näköhäiriöitä verkkokalvokuvista ja antamaan automaattisesti jopa toimivia hoitosuosituksia potilaille.

Kliinistä päätöksentekoa voidaan edistää nykyisin tehokkaasti käyttäen koneoppimiseen pohjautuvia metodeita. (ks.Prahs, Radek, Mayer, Cvetkova, Cvetkova, Helbig, & Märker 2018). Viime vuosina tapahtunut tekoäly ja konenäkö tutkimuksen edistyminen on antanut mahdollisuuksia kehittää toimiva sovellus, jolla pystytään ennakoimaan hoidon aiheita perustuen keskeiseen verkkokalvonkuvien näköyhteys röntgensädekerroskuvaus (OCT) seulontaan ilman ihmisen interventioita. Tätä tarkoitusta varten on opetettu kompleksista tekoäly hermoverkkoa. Näköyhteys kerroskuvaus röntgensäde (OCT) seulonnat keskiverkkokalvoista antavat yksityiskohtaisen datan ja ovat laajasti kliinikoiden käytettävissä hoito aiheen päätöksenteko prosessissa. (ks. Prahs ym. 2018).

Kerroskuvaus röntgensäde kuvat voidaan luokitella eri kategorioihin ja kuvien etukäteiskäsitellyn jälkeen OCT kuvien joukko voidaan jakaa eri suhteiseksi opetus ja testidata joukoksi. Hoidon aiheeseen pohjautuva syväkaareutuva kompleksinen hermoverkko GoogleNet voidaan onnistuneesti opettaa louhitusta kliinisestä datasta ja laskea ennustamistäsmällisyys herkkyys, tarkkuus ja ROC. Opetettu neuraalinen hermoverkko pystyy luokittelemaan ja saavuttamaan 95,5 % ennuste täsmällisyyden datajoukon kuvissa. Yksittäisen verkkokalvon kuvauksissa järjestelmän antamaksi herkkyysarvoksi saatiin 90,1% ja tarkkuus puolestaan oli 96,2 % ja ROC arvo oli 0,968. Tekoäly hermoverkoilla saadaan hyvä suorituskyky verkkokalvon kerroskuvaus röntgensäde kuvauksissa. Koneoppimismetodilla voidaan taata kliininen tuki päätöksenteko prosessissa, mutta tutkimus on osoittanut, että hoidossa ei saa kuitenkaan virheellisesti ottaa vain hermoverkkojen lähtöarvoja suosituksina vaan arvion tulisi tapahtua hoitavan lääkärin toimesta. (ks. Prahs ym. 2018).

Koneoppimisella ja datan louhinta metodeilla pystytään nykyisin tekemään varsinaista diabetes tutkimusta ja edistämään Bioteknologiaa ja terveystieteitä. Pystytään myös lisäämään merkittävästi relevanttia tietoaaineistoa. Suurista elektronisista terveysrekistereistä (electronic health records, EHR) voidaan muodostaa toimitettavaksi suuri määrä geneettistä dataa ja kliinistä tietoa. (Kavaktosis, Tsave, Salifoglor, Maglaveras, Vlahavas & Chowarde 2017.)

Nykyään diabetes tutkimuksen tarkoitus on systemaattisen katsauksen johtaminen koneoppimisen sovelluksista, data louhinta tekniikasta ja näyttövälineistä diabeteksen tutkimuskentällä, niissä otetaan huomioon ennustaminen ja diagnoosi, diabeetikon komplikaatiot, geneettinen tausta ja ympäristö, terveyden hoito ja johtaminen. Ennustamista ja diagnoosia pidetään nykyisin kaikkein suosituimpana tutkimus aihealueena. Koneoppimisalgoritmeilla toteutettu tutkimus, jossa on testattu laajasti sekä ohjatun oppimisen että ohjaamattoman oppimisen algoritmeja on osoittanut, että Tukivektorikone (Support Vector Machine) SVM on nousemassa yhdeksi käytetyimmistä algoritmeista data tyypissä, joka koostuu pääasiassa kliinisestä datajoukosta. (ks. Kavaktosis ym. 2017.)

Koneoppimiseen perustuvilla luokittelutekniikoita pystytään osoittamaan epälineaarista, epänormaalia ja sisäistä korrelaatorakennetta lääketieteellisessä datassa. Gaussin menetelmään perustuvaa koneoppimisluokittelutekniikkaa on verrattu tutkimuksessa muihin koneoppimis luokittelu tekniikoihin, kuten lineaariseen erotteluanalyysiin (Quadratic discriminant analysis (QDA) ja Naiiviin Bayesilaiseen (Naive Bayes, NB). Edellä mainittujen luokittelutekniikoiden suorituskyvyn arvioinnissa käytettiin täsmällisyyttä (accuracy, AUC), herkkyyttä (sensitivity SE), erityisyyttä (specificity, SP), tarkkaa ennustavaa arvoa (positive predictive value, PPV) ja negatiivista ennustavaa arvoa (Negative predictive value, NPV) ja ROC käyrää (Receiver operating characteristic curves, ROC). (ks. Maniruzzaman, Kumar, Abedin, Islam, Suri, El-Baz & Suri 2017.)

Tutkimus suurella diabetes datajoukolla osoittaa, että koneoppimisjärjestelmä, jossa sovelletaan Gaussilaiseen menetelmään perustuvaa datan luokittelutekniikkaa on tuloksiltaan huomattavasti parempi verrattuna muihin metodeihin. Koneoppimisjärjestelmä osoitti diabetes potilaiden data luokittelun täsmällisyyden arvoksi 81,97 %, herkkyydeksi 91,71 %, tarkkuudeksi 63,33 %, tarkaksi positiiviseksi ennustearvoksi 84,91 % ja negatiiviseksi ennustavaksi arvoksi 62,5 %. (ks. Maniruzzaman, Kumar, Abedin, Islam, Suri, El-Baz & Suri 2017.) Gaussilaista luokittelutekniikkaa soveltava koneoppimisjärjestelmä on erittäin tehokas diabetespotilaiden datan luokittelija, jolla pystytään louhimaan suuresta potilasdatasta keskeisimpiä sairauden hoidon kannalta tärkeitä muuttujia kliiniseen päätöksentekoon.

Kakkostyyppin diabetekseen liittyvien komplikaatioiden riskejä voidaan nykyisin arvioida käyttäen ristiinvaldointi koneoppimismetodia, ja sillä pystytään vaikuttamaan hoitotulokseen myönteisesti. (ks. Basu, Sussman, Berkowiz, Hayward & Yudkin 2017.) Ristiinvaldointi koneoppimismetodia on käytetty valitsemaan ennustavaa muuttujaa demograafisista merkeistä, kliinisistä muuttujista, sairauksien yhteisesiintymisestä (monihäiriöisyydestä), lääkkeistä, biomarkkereista ja ne kehittävät riski kaavan tyyppin 2 diabeteksen komplikaatioille. Tämän uskotaan parantavan tyyppin 2 diabeteksen arviointia. (ks. Basu ym. 2017.) Ristiinvaldointi koneoppimismetodien avulla pystytään diabeteshoitoon kehittämään tarkkoja komplikaatioiden ennustamiskaavoja, joiden avulla klinikot voivat ennakoita kakkostyyppin diabeetikoiden komplikaatioiden ilmentymistä. Tämä mahdollistaa diabeteshoidon onnistuneen suunnittelun ja seurannan.

Tutkimus osoitti, että syväoppimiseen perustuvien sovellusten, kuten EyePACS-1 ja Messidor-2 pystyvän tehokkaasti optimoimaan kuvien luokittelua. Syvä kierteinen hermoverkko kyetään opettamaan vanhoista tapahtumista kehittyistä suuresta määrästä verkkokalvokuvia koostuvasta datajoukosta. Suuresta määrästä verkkokalvokuvia koostuva datajoukko on mahdollista lajitella optomologiaan erikoistuneiden asiantuntija lääkärin toimesta, lajitella useaan otteeseen ja luokitella diabeettista retinopatiaa ja diabeettista makulaarista turvotusta kuvan luokittelua varten. Näin pystyttiin luomaan algoritmi, jonka herkkyys ja tarkkuus sekä viittaamisen diabeettiseen retinopatiaan on muodostettu perustuen optomologiseen asiantuntijaryhmään ja vertailuihin standardoituin enemmistöpäätöksiin. Herkkyys ja tarkkuus arvot olivat molemmilla syväoppimissovelluksilla korkeat (yli 87,5 %-98,5 %). Algoritmin käyttökelpoinen toteutettavuus kliinisiin puitteisiin vaati vielä jatkotutkimuksia, pitäisikin ennen kaikkea arvioida johtaako algoritmin käyttö parempaan hoitoon ja tuloksiin verraten sitä nykyisiin optomologisiin arviointeihin. (ks. Guishan, Peng, Coran, Stumpe, Wu, Narayanaswamy, Venuagopalan, Widner, Maradas, Cuadros, Kim, Raman, Nelson, Mega & Webster 2016.) Koneoppimisjärjestelmä oppii tunnistamaan diabetekseen liittyviä silmäsairauksia kokeneiden asiantuntijoiden tavoin, mikä auttaa klinikoita päätöksen tekoprosessissa merkittävästi.

Tietokoneavustettu seulontateknologia diabeettisen retinopatian luokittelua varten verkkokalvopohja kuvista, joukolla koneoppimisalgoritmeja on noussut merkittäväksi keinoksi ennalta ehkäistä laajassa diabeetikoiden joukossa sokeuden puhkeamista. Julkisista datajoukoista ja vanhoista verkkokalvokuvista rakennettu koneoppimisjärjestelmä on osoittanut hyvää herkkyyttä ja tarkkuutta alustavassa pilottitutkimuksessa. (ks. John 2016.)

Koneoppimisalgoritmeilla kyetään myös vertailemaan diabeteksestä kärsivien yksilöiden, diabeetikoiden nykyisiin ravintosuosituksiin perustuvien aterioiden jälkeisiä verensokerin muutoksia tarkoin. Diabeetikoiden ruokavaliosta ja dieettihoidosta on näin ollen mahdollista tehdä suunnitelmallista ja yksilökeskeistä. (ks. Zeevi, Kovem, Zmora, Israeli, Rotschild, Weinberger, Ben-Yacov, Lador, Avnil-Sagi, Lotan-Pomper, Suez, Ali Mahdi, Matot, Malka, Kosower, Rein, Zilberman- Schapira, Kosower, Rein, Zillerman- Schapira, Dohnalova, Pereoner-Fischer, Bikovsky, Halper, Elinav & Segal 2015.)

Tutkimus osoitti, että on mahdollista suunnitella sellainen koneoppimisalgoritmi, johon voidaan integroida verimuuttujat, ravintotottumukset, antropometrit, fyysinen aktiivisuus ja suolen mikrobisto. Mittaamalla näitä muuttujia voidaan täsmällisesti ennustaa henkilökohtaista aterianjälkeistä glykeemistä reagointia aterioihin. Nämä ennusteet pystyttiin vahvistamaan itsenäisessä 100 henkilön ihmisryhmässä. Kyseiseen algoritmiin perustuva satunnaistettu sokko ruokavalio interventio, johti huomattavasti alempiin aterian jälkeisiin ja yhdenmukaisiin vaihteluihin suolen mikrobiston asetelmassa. Tutkimustulokset osoittivat kiistatta henkilökohtaisesti räätälöityjen ruokavalioiden pystyvän onnistuneesti muuntelevan aterian jälkeistä kohonnutta veren sokeria ja sen aineenvaihdunnallisia tuloksia. (ks. Zeevi ym. 2015.)

Tapahtumapohjainen todistus viitekehys koneoppimismetodia on alettu soveltamaan diabetes hoidossa viime aikoina. Tapahtuma-pohjainen todistus on ongelman ratkaisu paradigma, joka hyödyntää vanhaa tietoa tulkitakseen ja ratkaistakseen uusia ongelmia ja se sopii asiantuntijapohjaiseen ja teoriaa vähän sisältävään ongelmaan. Monia diabeteshoidon lääketieteellisiä ongelmia voidaan ratkaista tehokkaasti rakentamalla semanttisesti älykäs tapahtuma todistus-pohjainen järjestelmä. Metodi voidaan integroida terveydenhuolto järjestelmiin, ja tämä pystyy vastaamaan kompleksisten lääketieteellisten käsitteiden sisällön ymmärtämiseen ja epäselvän termin käsittelyyn. Järjestelmä on saavuttanut jopa 17,67 % tarkkuuden. (ks. El-Sappagh, S., Elmogy, E & A.M Riad 2015).

Uusin kehitys datan digitalisoinnissa, konenäössä, koneoppimisessa ja tekoälyratkaisuissa on laajentunut myös aikaisemmin hyvin koulutettujen lääkärin tutkimille alueille. Tutkimuksessa nostettiin esille älykäs autonominen järjestelmä(Intelligent Diabetic Assistant(IDA) prototyyppi, joka kykenee päättämään diagnooseista ja hoidon priorisoinnista riippuen näyttöruudulla ilmenevistä havainnoista. Älykäs autonominen järjestelmä koostuu tietopohjaisista moduuleista, joilla voidaan suorittaa vakavuusastetasoon perustuvaa luokittelua, kliinistä päätöksenteon tukemista ja lähestyvän reaali-aikaisen jalkamätähaavan havaitsemista ja rajapinnan seulontaa. Diabeettiseen hoitoon osallistuvien kliinikoiden kanssa suoritettu järjestelmän käytettävyysskoe todistaa järjestelmän hyvää, mutta ei kuitenkaan mitään poikkeuksellisen hyvää käytettävyyttä. (ks. Wiejishne ym. 2016). Diabetespotilaista voidaan koneoppimisella nykyisin kerätä ja analysoida paljon sairauteen liittyvää dataa, joka voidaan saattaa reaaliaikaisesti kliinikoiden saataville päätöksentekotilanteissa. Tällöin on mahdollista saada nopeasti kokonaiskuva henkilön sairaudesta, ja antaa esim. kliinisen päätöksenteon tukea sairauden kriittisissä vaiheissa.

Tutkimuksessa nousi esille myös kuvaperusteinen kehysten hyödyntämisen maksimointi ruokavalion virheiden määrittämisen maksimoimiseksi, joka on uusi syväoppimista hyödyntävä koneoppimismalli diabetes hoidossa. Siinä voidaan tunnistaa automaattisesti eri ruoka komponentteja ravinto kuvista eli voidaan analysoida ruokavalion dataa ja suositella terveellisiä vaihtoehtoja diabeetikoille reaali-ajassa. Näillä ravintosuosituksilla ei ole pelkästään korkea terveyshyöty, sillä ne ovat yhdenmukaisia yksilöiden makumieltymysten kanssa. (ks. Qiu, Bhattarya & Tuan 2017). Internet ja mobiilisovellukset ovat vahvasti tulossa diabeteshoitoon, ja niiden avulla voidaan toteuttaa onnistuneita sovelluksia potilaille ja vähentää kliinikoiden ja terveydenhuolto järjestelmien kuormitusta.

6.2 Koneoppiminen ja diabetes hoitotulosten ennustaminen

Taulukko 4. Koneoppimismallit diabetes hoitotulosten ennustamisessa..

Koneoppimismallit/Algoritmit	Käyttötarkoitus diabeteksen ennustemalleissa	Lähdeaineisto
<p>Pääkomponenttien analyysi(Principal Component Analysis, PCA)</p> <p>Datan louhinta tekniikat(mm. luokittelu ja esikäsitteily)</p> <p>Datalouhinnan useiden algoritmien lähestymistapa</p>	<p>Useiden ohjattujen koneoppimisalgoritmien täsmällisyyden edistäminen käyttäen vaadittua minimi lukumäärä attribuutteja</p> <p>Diabetes potilaiden taudin kehityskulun ennustaminen ja diabeteksen luokittelu</p> <p>Diabetes sairauden kehityskulun ennustaminen ja diabeteksen luokittelu</p>	Kanchan, D, B. & M. M. Kishor. (2016)
<p>Datan louhinta tekniikat(mm. luokittelu ja esikäsitteily)</p>	<p>Päättää diagnooseista ja hoidon priorisoinnista ottaen huomioon näyttöruudulla ilmenevät havainnot</p> <p>Sairauden vakavuus asteen tasoon perustuva luokittelu, kliinisen päätöksenteon tukeminen ja lähestyvän reaali-aikaisen jalkamätähaavan havaitseminen ja rajapinnan seulonta.</p>	Wiejsignhe, I., Gamage, L., Perera, J.& Chitaranjan, C. (2019)
<p>Ohjattu koneoppiminen</p> <p>Ohjatun koneoppimisen luokittelun menetelmät</p>	<p>Ennustaa henkilön diabeteksen ja liitännäissairauksista johtuvaa sairaalaan joutumista täsmällisesti. Voidaan luoda ennustavia malleja niiden potilaiden identifioimiseen, jotka ovat vaarassa joutua sairaalaan seuraavan vuoden kuluessa kakkostyyppin diabeteksen aiheuttamien komplikaatioiden takia.</p> <p>Löysivät piileviä potilas ryhmiä positiivisessa luokassa</p>	Brismi, Xu, Wang, Dai & Paschalidis (2018)

Hajanainen, lineaarinen vektorikone luokitin	<p>Erottelee positiiviset näytteet negatiivisista(sairaalan sijoittamista kuvaavista)</p> <p>Metodi auttaa tuottamaan tietoa potilasryhmän sisällöstä.</p>	
Koneoppimista soveltava ennustava mHealth sovellus	<p>Toimitetaan potilaan itse-hoitoa varten, potilaiden diagnosoimiseksi ja se määrittää todennäköisyyden joutua sairauden vaikutuksen alaiseksi. Analysoi diabeettisena potilaana olemisen riskejä seikkaperäisesti.</p> <p>Koneoppimisen tekemä arviointi diabeettiseen, prediabeettiseen tai esidiabeettiseen potilaan mahdollisuuksiin selvittää ilman lääkäreitä tai lääketieteellisiä testejä.</p>	Khan, Muaz, Kabir & Islam (2017)
<p>Koneoppimistekniikat ja data louhiminen kliinisestä datasta</p> <p>Satunnainen metsä (Random Forest)</p> <p>Yksinkertaiset logistiset regressiomethodit (Simple Logistic Regression Methods)</p>	Mahdollistaa diabeettisen nephropatian ja ennustavien biomarkkereiden identifioimisen kakkostyyppin diabetes potilaiden joukossa.	Rodeiquez-Romero.,V, Bergstrom.,R.,F, Decker, B.,S, Lahu, G., Vakilynejad, M. & R.R. Bies.(2019)
Koneoppimisen ja elektronisen lääkintä rekisterien yhdistäminen	Kehittää kakkostyyppin diabeteksen ennustava malli sellaisen populaation datajoukosta, jolla ei ole diabetes historiaa, koneoppimista ja elektronista lääkintärekisteri tietokantaa hyödyntäen	Choi, B, G., RHA.,S,W, Kim,S,W, Kang, J,H, Park,J.Y& Y.K. Noh. (2019)
Datakäyttöinen mallinnus ja veren sokeri dynamiikan ennustaminen ykköstyyppin diabeteksessä koneoppimista käyttäen	Biomarkkereiden hyödyntäminen	Ashenafi Zebene Woldargay (2019)
Lineaarista regressioista kehitetty ennuste malli	Diabeettinen seulonta ja diabeteksen lyhyt- ja pitkäaikaisten komplikaatioiden hallinta.	Lebech Cichosz, Denker Johansen & Hejlesen (2015).
Koneoppimismalli Naivi Bayesiläinen	Ns. esidiabetes vaiheeseen olennaisesti liittyvän kaulavaltimon rasvoittumistaudin kehityskulun ennustaminen	Hu, Reaven.,; Saremi, Liu,Abbasi, Liu & Migrino (2016)
Koneoppimismallit hoidon lopputulosten ennustamisessa	<p>Diabetes hoidon kehittäminen potilaskeskeisempään ja yksilöllisempään suuntaan. Vastikään diagnosoitujen ensisijaisessa hoidossa hoidettujen kakkostyyppin diabeetikoiden glykeemisen, kehityskaaren yhdenmukainen tunnistaminen, ennustaminen ja selittäminen.</p> <p>Koneoppimismalleja rakennettiin ennustamaan glykeemista kehityskaarta käyttäen helposti hankittavia potilas ominaispiirteitä päivittäisessä kliinisessä ammatin harjoittamisessa.</p> <p>Kolmen glykeemisen kehityskaaren tunnistaminen. Ne olivat vakaa, riittävä glykeeminen kontrolli, parantunut glykeeminen hallinta ja heikentynyt glykeeminen hallinta. Kehon massa indeksi ja pitkäveren sokeritasot olivat kaikkein tärkeimmät kehityskaaren jäsen ennustajat. Kehitetty malli voi tehokkaasti selittää heterogeenisyyden kakkostyyppin diabeetikoiden tulevaisuuden glykeemisessä reagoinnissa. Sitä voidaan käyttää kliinisessä soveltamisessa nopeana ja helppona työkaluna yksilöllisesti</p>	Hertroijs, D.F., Elisen, A.M., Brouwers, M.C.G.J., Scharper, N.C, Köhler, S., Popa, M.C.,Asterodis, S., Heendriks, S.H, Bilo, H.J.& Ruwald.D. (2018).

	toteutetussa diabetes hoidossa, yleisempien sairaalaan joutumisen syiden tunnistamisessa. Pystyttiin ennustamaan myös kaikkein erottelukykyisin hoitoon sitoutumisen raja-arvo yleisten kaikkien sairaalan joutumis riskien takia potilaan leikkauksien ja terveyden mukaan. Koneoppiminen saattaa olla arvokas tarkoituksenmukaisen potilas-keskeisen hoitoon sitoutumisen raja-arvon tunnistamisessa ja kohdennettaessa hoitoon sitoutumattomia potilaista varten interventioita	
Satunainen metsä Ohjattu oppiminen	Sairaalaan sijoittamisen riskin ja lääkityksen selvitys. Erotelee sairaalaan sijoittamisen riskin käyttäen katettujen päivien osuutta(PDC).	LO-Ciganic, W.-H., Donohue, J. M., Thorope, J.M, Perera, S., Thorope, C., Marcum, Z, A. & W.F Gellad . (2015)
Tilastolliset mallit	Ennustavat aikaista sairaalan sijoittamista. Koneoppiminen tukee hoitotyötä ja hoitajat voivat kehittää sen merkityksellisyyttä ja suorituskykyä. Hoitajat voivat myötävaikuttaa koneoppimisalgoritmeihin täyttämällä data aukkoja hoitoon asiankuuluvalla datalla. He voivat myös parantaa datan esikäsittelyn tekniikoita, ja arvioida potentiaalista arvoa käytännössä. Hoitajien täytyy käsitellä koneoppimisen antamaa informaatiota ja soveltaa näkemyksiään työssään tehdäkseen kunnollisia klinisiä päätöksiä. Yhdistämällä koneoppiminen ainutlaatuisen hoitotiedon kanssa voi antaa hoitajille mahdollisuuden kiinnittää huomiota varsinaiseen hoitotyöhön, kehittää hoitotiedettä ja parantaa yksilöllistettyä potilas hoitoa.	Kwon., Karim, Topaz & Currie (2019)
Principal Component Analysis (PCA) Six Sigma lähestymistapa kynnys arvo marginaaliryöstys klusterointi algoritmi (TBC)	Anti diabeettisten ravinto-tuotteiden sokeria vapauttavan säätö potentiaalnin analysoiminen.	Mulay, Joshi & Chaudhari (2018)
Ennustava analyysi metodi (integroidaan useita datan louhimis tekniikoita, koneoppimisalgoritmeja, tilastotietoja, jotka käyttävät nykyistä ja mennyttä dataa saadakseen lisää oivalluksia ja ennustaakseen tulevia riskejä) Hadoop MapReduce (Koneoppimis algoritmi)	Etsitään virheellisiä arvoja diabetesjoukosta ja yritetään löytää polkuja siitä. Antaa mahdollisuuden ennustaa tulevia riskejä, ja potilaan riskitasojen mukainen hoitotyyppi voidaan tarjota.	Kalyankar, Shivanande &. Dharwakar (2017).
Satunainen Metsä (Random Forest, RF)	Mahdollistaa datajoukon syöttöpiirteiden analysoimisen. Niiden tärkeys luokittelussa voidaan huomioida myös diabeettiseen retinopatia sairauteen altistavat tekijät. Näin aikaansaadut sairaus diagnoosit perustuvat nopeasti saatavilla olevaan elektroniseen dataan. Ne tulevat olemaan olennainen osa älykästä terveydenhoitoa ja mobiili terveydenhoitoa.	Sun, Zhang (2019)

Tukivektorikoneet Kielioppiilinen kehityskulku (Grammatical Evolution)	Hypoglykeemisten tapahtumien ennustamiseen ja datan louhintaan diabeteksen hallinnan ennusteissa.	Vehi Contrearas, Ovido, Bialgi & Bertachi (2019)
--	---	---

Tutkimuksessa havaittiin ohjatun koneoppimisen pystyvän tarkoin ennustamaan diabetes sairautta. Tässä käytetään ns. pääkomponenttien analyysia (Principal Component Analysis, PCA). Tiedetään, että maailmanlaajuisesta sydänsairauksista oireyhtymistä johtuvien kuolemien syyt ovat suurin kuolemiin johtanut syy ja 23,6 miljoonaa ihmisen oletetaan kuolevan sydänsairauksiin tämänkaltaisen kehityssuunnan jatkuessa. Terveiden hoito toimiala kerää suuren määrän sydänsairaus dataa, jota ei ole valitettavasti ole louhittu löytämään piilevää informaatiota tehokkaaseen päätöksentekoon. (ks. Kanchan & Machale 2016.)

Koneoppimisen data louhinta teknikoiden käyttöönoton avulla klinikot voivat päättää diagnooseista ja hoidon priorisoinnista perustuen näyttöruudulla ilmeneviin havaintoihin. Tämä mahdollistaa sairauden vakavuusasteen tasoon perustuvan luokittelun, kliinisen päätöksenteon tukemiseksi ja reaali-aikaisen jalkamätähaavan seulonnan. (ks. Wiejissignhe, Gamage, Perera & Chitaranjan 2019.) Koneoppiminen mahdollistaa terveyden hoidon ammattihenkilöille nopeasti diabetekseen vakavuusasteen kartoittamisen.

Ohjatulla koneoppimisella voidaan nykyisin ennustaa henkilön diabeteksen liitännäissairauksista johtuvaa sairaalan joutumista täsmällisesti. Tämän takia niiden potilaiden identifioimiseen, jotka ovat joutumassa sairaalaan seuraavan viiden vuoden kuluessa kakkostyyppin diabeteksen aiheuttamien komplikaatioiden takia, on pyritty luomaan ennustava malli. Sairaalaan sijoitettavien potilaiden selvittämiseksi kehitettiin metodi, jossa vektoritukikone luokitin tuotettiin erottelemaan sairaalan sijoitetut potilaat sairaalan sijoittamattomista. Uusi yhdentynyt metodi pystyttiin vakiintumaan ja todistettua, kuinka sen luokittimen tuottama tieto on yleistettävissä testijoukkoon, jota ei ole aikaisemmin nähty opetusvaiheen lainkaan. Isossa potilasjoukossa metodi saavutti 89 % täsmällisyyden ROC alueen termillä, ja metodi kykenee myös tuottamaan tietoa sisältävät potilasryhmät. Lääkäri voi näin ollen tuottaa algoritmin tuotoksen ja antaa ohjausta kohti preventiivisiä arvioita. Tutkimus osoitti, että kohtuullisetkin toimenpiteet ovat riittäviä luomaan merkittäviä säästöjä sairaalahoitoon, ja sairaalaan sijoittamista voidaan ehkäistä, terveydentilaa voidaan parantaa ja pystytään myös radikaalisti vähentämään sairaala kustannuksia. (ks. Brismi, Xu, Wang, Dai & Paschalidis 2018.)

Diabeteksen kontrolloimiseen on lopputulosten ennustamiseksi luotu ennustemalli, ns. prosessimittareita käyttäen. Diabeteksen kontrolloimiseksi on pyritty identifioimaan tehokas metodi ennustavan metodin rakentamiseksi, ja on käyttöön otettu prosessimittareita, jotka ovat olennaisen tärkeitä hoidon laadun kehittämisstrategioiden tehokkuuden ohjaamisessa. Ne voivat ennustaa toivottuja hoitojen lopputuloksia, ja identifioida tehokkaan metodin ennustavan mallin rakentamisessa. (ks. Khanji, Lalonde, & Schnitzer 2019.)

Tutkimus on osoittanut, että ilman kontrollia olleiden diabetespotilaiden kohdalla parhaaksi tekniikaksi osoittautui perinteistä metodia ja hajautettua koneoppimis tekniikkaa sisältävä hybridimetodi. Metodien erottelukyky kapasiteettia kuvaava paras cVauc arvo, ja se on diabeteksen kohdalla 0,79. Tutkimuksessa havaittiin, että kompleksisessa ja moniaineisessa potilaspopulaatiossa kardiovaskulaaristen sairauksien ehkäisyyn viiden prosessimuuttujan joukolla hybridi metodeilla päästään hyvään ennustekelpoisuuteen. (ks. Khanji ym. 2019.) Suunniteltaessa kliinistä ei-tietämyskantaista päätöksentekojärjestelmää terveydenhuolto-organisaatioon on otettava huomioon se, minkälaisista koneoppimismetodeista sovelletaan järjestelmän toteuttamisessa, jotta päästään parhaimpaan mahdolliseen tulokseen onnistuneiden hoito tulosten ennustamisessa ja hoitostrategioiden ohjaamisessa.

Koneoppimis tekniikoita on mahdollista soveltaa nykyisin erinomaisesti diabetekseen olennaisesti liittyvien biomarkkereiden tunnistamisessa. (ks. Rodriquez-Romeo & Bergström 2019.) Soveltamalla koneoppimis tekniikoita ja datan louhintaa kliinisen dataan, saadaan mahdollisesti identifioitua diabeettisen nefropatian (munuaissairauden) ennustavia biomarkkereita. Diabeettinen nefropatia on hyvin yleinen komplikaatio kakkostyypin diabetes mellituksessa. Aikaisemmat ja myöhäisemmät ennustajat on jaoteltu pitkäaikaistutkimuksen datassa. Satunainen metsä (Random Forest) ja yksinkertaiset logistiset regressiomethodit (Simple Logistic Regression Methods) ovat osoittaneet parasta tehokkuutta arvioitujen algoritmien joukossa. Diabeettisen nefropatian varhaiset ennusteet olivat lähtökohtaisesti hiussuoniston suodatusnopeus (GFR), systolinen verenpaine, plasman paastosokeri (fasting plasma glucose) ja kalium muutokset hiussuoniston suodatusnopeudessa. Triglycereideissä tapahtuvat muutokset puolestaan tunnistettiin myöhäisen kehityksen ennustajina. Tutkimus osoitti, että koneoppimiseen pohjautuvat metodit pystyvät tunnistamaan onnistuneesti ennustavat tekijät koskien diabeettista nefropatiaa kakkostyypin diabetes potilaiden joukossa. (ks. Rodriquez-Romeo & Bergström 2019.)

Koneoppimista voidaan myös käyttää diabetes mellituksen puhkeamisen ennustamisessa myös potilasjoukossa, jossa ei ole diabeetikoita vaan ainoastaan kardiovaskulaaristen riskien omaavia potilaita. (ks. Choi, RHA, Kim, Kang Park & Noh. 2019.) Kakkostyypin diabetekseen kehitettiin ennustava malli käyttäen elektronisia lääkintärekistereitä (Electronic Medical Records, EMRs) ja koneoppimista sekä vertailtiin tämän mallin suorituskykyä muiden perinteisten tilastollisten metodien kanssa. Ennustemallin mukaisen ristiinvalidointi testin todistamista varten on muodostettu logistinen regressiomalli(LR), lineaarinen erottelukyky analyysi(Linear discriminant analysis, QDA) ja K-lähin naapuri (K-nearest neighbour, KNN.) (ks. Choi ym. 2019.)

Kaikkien ennustavien mallien joukossa, lineaarinen oppimismalli osoitti korkeinta ennuste tehoa kaarenalaisen arvon alueen arvolla 0,78 lineaariseen malliin verrattuna. Lineaarinen erottelukyky analyysi(LDA), neliöllinen erottelukyky (Quadratic discriminant analysis, QDA) ja K-lähin Naapuri (K-Nearest Neighbour) mallit eivät näyttäneet tilastollisesti merkittävän erilaisilta. Kakkostyyppin diabetekseen on näin ollen kehitetty ja vahvistettu kakkostyyppin diabeteksen ennustejärjestelmä käyttäen koneoppimista ja elektronista lääkintärekisteri tietokantaa. (ks. Choi ym.2019.) Koneoppimista hyödyntävän kliininen päätöksentekojärjestelmän hyvänä puolena on se, että voidaan antaa diabetekseen sairastumisriski arvioita potilas populaatiosta, joka kärsii jostain muusta vakavasta kroonisesta sairaudesta. Koneoppimisen ja elektronisten lääkintärekisteri tietokantojen hyödyntäminen mahdollistaa sellaisen kakkostyyppin diabeteksen ennustejärjestelmän toteutuksen, joka antaa täsmällistä tietoa klinikoille diabeteksen puhkeamisuhasta.

Ykköstyyppin diabetekseen liittyvien biomarkkereiden kehittymisen ennustamiseen ja sairauteen olennaisesti kuuluvien hoitotoimenpiteiden seuraamiseen on kehitetty erillinen koneoppimissovellus, jonka avulla terveydenhoidossa voidaan tehdä paranneltu ja älykkäämpi järjestelmän käyttöönotto ykköstyyppisen diabeteksen hoidossa. (ks. Woldargay, Årsand, Batsis, Albey, Mamykina & Hartgivsen 2019.)

Data mallinnus ja veren sokeri dynamiikan ennustaminen ykköstyypin diabeteksessa käyttäen koneoppimista on nousemassa tärkeäksi teknologiaksi terveyden huollossa. Tällä hetkellä diabetekseen ei ole vielä parannuskeinoa, mutta verensokeritasojen pitäminen suositellulla tasolla on hoidossa keskeistä. Viimeisimmät edistysaskeleet diabetes teknologiassa ja itsehoito sovelluksissa ovat tehneet potilaille relevantin datan saannin helpommaksi. Keinotekoisien haiman kehitys(suljettu järjestelmä), personoitu päätöksenteko järjestelmä ja verensokeri tapahtumien hälytykset ovat myös nousseet aikaisempaa näkyvimmin esille. Keskeiseksi asioiksi näiden diabetes hoito teknologioiden kehittämisessä on tullut verensokeri dynamiikan mallintaminen. Lisääntyneen potilas historiallisen datan katsotaan myös luoneen otolliset olosuhteet koneoppimisen käyttöönotolle, ja sen sovellukselle älykkäisiin ja paranneltuihin diabeteksen hoitoa varten toteutettuihin järjestelmiin. Koneoppimisen kyky ratkaista kompleksisia tehtäviä dynaamisessa ympäristössä on merkittävästi edistänyt sen menestystä diabetes tutkimuksessa. (ks. Woldargay 2019.)

Diabeteksen hoidon ja siihen liittyvien komplikaatioiden ennuste mallit ovat kehittyneet ja ne on käyttöön otettu klinikoiden avuksi päätöksen tekoon. Diabetesta pidetään nykyisin yhtenä pääasiana lääketieteessä ja terveydenhoidon johtamisessa, siihen liittyvä runsas data ja tieto on aina myös saatavilla potilaille. Diabetekseen liittyvä data saattaa olla peräisin tilastollisesta malleista tai monimutkaisista hahmon tunnistusmalleista, ja ne voidaan sulauttaa ennuste malleiksi, jotka yhdistävät potilastiedot ja enteelliset seurausten lopputulokset. Tällaista tietoa käytetään kliinisessä päätöksenteon tukemisessa sairauksien tarkkailussa ja julkisen terveydenhoidon johtamisessa potilaan hoidon parantamiseksi. Tutkimus osoitti, että voidaan kehittää ennustemalleja diabeteksen hallintaa ja sen komplikaatioita varten, ja julkaisut näistä malleista ovat myös kasvaneet viimeisten vuosikymmenten aikana. Monia lukuisia tai vastaavia lineaarisia regressioita käytetään ennustemallin kehittämistä varten todennäköisesti niiden koodiriippumattoman toimintatavan vuoksi. (ks. Lebech Cichosz, Denker Johansen & Hejlesen 2015.)

Koneoppimisen avulla pystytään nykyisin myös ennustamaan ns. esidiabetes vaiheeseen olennaisesti liittyvää kaulavaltimon rasvoittumistaudin kehityskulku. Koneoppiminen pystyykin nopeasti ennakoimaan ateroskleroosin (valtimon rasvoittumistaudin) etenemisen heikentyneestä glukoosin sietokyvystä kärsivässä potilas ryhmässä. (ks. Hu, Reaven, Saremi, Liu, Abbasi, Liu & Migrino 2016). Esidiabeteksen tiedetään olevan suuri epidemia ja sen tiedetään olevan yhteydessä haitallisiin sydän ja aivoverisuoni peräisiin sairauksiin. Riskien vähentämisessä varhainen nopea ateroskleroosin kehittyvän potilasryhmän tunnistaminen voi olla hyödyllistä. Tutkimus on osoittanut, että koneoppimis metodeita voidaan käyttää nopeassa kaulavaltimon sisäkalvon verisuonten paksuuden kehittymisen ennakoimisessa heikentyneestä glukoosin sietokyvystä kärsivillä potilailla. (ks. Hu 2016).

Esille nousi vahvasti myös koneoppimismallien mahdollisuus kehittää diabetes hoitoa potilaskeskeisempään ja yksilöllisempään suuntaan. (ks. Hertroijs, Elisen, Brouwers, Scharper, Köhler, Popa, Asterodis, Heendriks, Biló & Ruwald 2018). Tutkimuksessa havaittiin pistemäärä niiden muuttujien keskuudessa, johon sisältyy kehon massaindeksi, pitkäaikainen verensokeri, triglyceridi ja niillä tiedetään olevan kyky ennustaa glykeeminen hallinta kakkostyyppin diabeetikoiden keskuudessa. Tutkimus on auttanut tunnistamaan, ennustamaan ja selittämään yhdenmukaisen glykeemisen kehityskaaren, vastikään diagnosoitujen, ensisijaisessa hoidossa hoidettujen, kakkostyyppin diabeetikoiden keskuudessa. Tämän tarkoituksena on ollut tehdä päätös siirtyä kohti potilaskeskeisempää ja tehokkaampaa hoitoa. Koneoppimismallit ennustavat glykeemista kehityskaarta käyttäen helposti hankittavia potilasominaispiirteitä päivittäisestä kliinisestä ammatin harjoittamisesta. (ks. Hertroijs, Elisen, Brouwers, Scharper, Köhler, Popa, Asterodis, Heendriks, Biló & Ruwald 2018).

Tutkimus osoitti, että kehon massa indeksi ja pitkäveren sokeri tasot olivat ne kaikkein tärkeimmät kehityskaaren jäsenten ennustajat. Kehittyvästä kohortista opetetulla ennustavalla mallilla oli ROC arvo alle kaaren 0,96 vahvistetussa kohortissa viitaten erinomaiseen täsmällisyyteen. Tutkimus osoitti, että kehitetty malli voi tehokkaasti selittää heterogeenisyyden kakkostyyppin diabeetikoiden tulevaisuuden glykeemisessä reagoinnissa. Sitä voidaan täten käyttää kliinisessä soveltamisessa nopeana ja helppona työkaluna yksilöllisesti toteutetussa diabetes hoidossa. (ks. Hertroijs ym. 2018).

Koneoppimista on mahdollista soveltaa mainiosti diabeetikoiden hoitoon sitoutumisen täsmällisessä mittaamisessa ja heidän sairaalahoidon tarpeiden kartoittamisessa. (ks. LO-Ciganic, Donohue, Thorope, Perera, Thorope, Marcum. & Gellad 2015). Tutkimuksessa havaittiin, että koneoppimista on mahdollista käyttää kakkostyyppin diabeetikoiden lääkehoitoon sitoutumisen raja-arvon ja sairaalan sitoutumisen riskin tutkimisessa. Siinä on myös osoitettu hoidon laadun pyrkimysten olevan yhteydessä potilaisiin, jotka ovat onnistuneesti saavuttaneet vähintään 80 % tai enemmän lääkehoitoon sitoutumisen tavoitteista. Kuitenkaan ei ole juuri vielä empiirisiä todisteita siitä, että tämä raja-arvo optimaalisesti ennustaisi tärkeitä terveys tuloksia. (ks. LO-Ciganic ym.2015.)

Tutkimuksessa havaittiin, kuinka sitoutuminen oraalisiin hypoglykemisiin (glykeemisiin) lääkkeisiin on yhteydessä sairaalaan sijoittamisen välttämiseen ja tunnistamaan hoitoon sitoutumisen raja-arvoa optimaaliseen sairaalassa olon riskin erotuskykyyn. Satunnaisia olemassaolon metsiä sovellettiin tunnistamaan sairaalaan sijoittumisen ennustajat ja olemassaolon puut sovitettiin empiirisesti polveutumaan hoitoon sitoutumisen raja-arvosta, joka parhaiten erottelee sairaalaan sijoittamisen riskin käyttäen katettujen päivien osuutta (PDC) ja tutkimuksessa huomioitiin ajankohta, jolloin henkilö ensimmäisen kerran joutuu sairaalaan johtuen yleisesti jostain tai diabeteksen liittyvästä syystä. Koneoppimisen uskotaan olevan lähestymistapana arvokas tarkoituksenmukaisessa potilaskeskeisessä hoitoon sitoutumisen raja-arvon tunnustamisessa ja kohdennettaessa hoitoon sitoutumattomia potilaita varten toimenpiteitä suunnitellussa interventiossa. (ks. LO-Ciganic ym.2015.)

Diabeteksen hoidon ja siihen liittyvien komplikaatioiden ennuste mallit ovat kehittyneet ja ne on käyttöön otettu klinikoiden avuksi päätöksen tekoon. Diabetesta pidetään nykyisin yhtenä pääasiana lääketieteessä ja terveydenhoidon johtamisessa, ja siihen liittyvä runsas data ja tieto on aina myös saatavilla potilaille. Diabetekseen liittyvä data saattaa olla peräisin tilastollisesta malleista tai monimutkaisista hahmon tunnistumalleista, ja ne voidaan sulauttaa ennuste malleiksi, jotka yhdistävät potilastiedot ja enteelliset seurausten lopputulokset. Tällaista tietoa käytetään kliinisessä päätöksenteon tukemisessa sairauksien tarkkailussa ja julkisen terveydenhoidon johtamisessa potilaan hoidon parantamiseksi. Tutkimus osoitti, että voidaan kehittää ennustemalleja diabeteksen hallintaa ja sen komplikaatioita varten, ja julkaisut näistä malleista ovat myös kasvaneet viimeisten vuosikymmenten aikana. Monia lukuisia tai vastaavia lineaarisia regressioita käytetään ennustemallin kehittämistä varten todennäköisesti niiden koodiriippumattoman toimintatavan vuoksi. (ks. Lebech Cichosz, Denker Johansen & Hejlesen 2015.)

Tutkimuksessa selvisi, että koneoppimisella voidaan myös ennustaa ja ehkäistä hypoglykeemisiä tapahtumia ykköstyypin diabetesta sairastavissa. Koneoppimisella voidaan tiukasti kontrolloida veren sokeria, joka vähentää mikro- ja makrovaskulaarisia komplikaatioita ykköstyypin diabetes potilailla. Tämä on vielä kuitenkin varsin vaikeaa johtuen laajasta yksilöllisestä, vaihtelevuudesta ja muista tekijöistä, jotka vaikuttavat glykeemiseen kontrolliin. (ks. Vehi, Contrearras, Ovido, Bialgi. & Bertachi 2019.)

Vakava hypoglykemia on suurin rajoittava tekijä intensiivisessä diabetes hoidossa olevilla potilailla, ja sen tiedetään vaikuttavan negatiivisesti tästä sairaudesta kärsivien potilaiden elämän laatuun. Koneoppimismetodeihin perustuvien työkalujen tiedetään olevan toteutettavissa oleva tapa edistää potilasturvallisuutta ennakoiden glykeemisiä tapahtumia. Järjestelmä, jossa on toteutettu ennustava yhdistelmä ja onnistunut luokittelu, voi vähentää merkittävästi hypoglykeemisten jaksojen määrää, parantaa turvallisuutta ja antaa potilaille enemmän itseluottamusta päätöksen teossa. (ks. Vehi ym. 2019.) Onnistuneen kliinisen päätöksenteko tukijärjestelmän suunnittelu diabetes hoitoa varten, edellyttää potilaiden yksilöllisten tarpeiden vaatimusten huomioimista. Koneoppimisen käyttöönotto järjestelmiin mahdollistaa tarkan sairauten liittyvien tapahtumien selvittämisen ja kliinisen päätöksenteon edistämisen tiedon saatavuuden lisääntyessä.

6.3 Koneoppiminen diabeteksen diagnosointityökaluna

Taulukko 5. Koneoppiminen diabeteksen diagnosoinnin edistämässä.

Koneoppimismetodi\algoritmi	Käyttötarkoitus diabeteksen diagnosointi työkaluna	Lähdeaineisto
Syväoppiminen ja kohdealue-tieto Kierteisten hermoverkkojen oppimat piirteet ja kokonais vektorikuvaajat DeepDR-Net (Kompakti syväoppimis algoritmi)	Koneoppiminen diabeteksen diagnosointi työkaluna Käytetään havaitsemaan punaisia viottumia, eli identifioimaan todellisia viottumilta vaikuttavia objekteja. Avustaa lääkäreitä diabeettisen retinopatian vakavuusasteen luokittelussa verkkokalvokuvista. Diabeettisen retinopatian oireiden vakavuusasteen luokittelu.	Orlando, Prokofyeva, Del Fresno & Blaschko (2018)
Tukivektorikoneet (SVMs), päätöksentekopuut, tekoäly hermoverkot ja logistiset regressiot (Ohjattu oppiminen ja syväoppiminen)	Diabeettisen retinopatiaan liittyvien riskitekijöiden täsmällinen havaitseminen, jotta voitaisiin tehdä toimenpiteitä suurimpien riskiryhmien diabeettisen retinopatian ennaltaehkäisy strategioiden luomiseksi. Tarkoituksenmukainen koneoppimis algoritmi yhdistettynä kliinisiin piirteisiin voi tehokkaasti havaita diabeettisen retinopatian. Tukivektorikoneilla aikaansaatu ennuste suorituskyky oli parempi kuin muiden koneoppimisalgoritmien. Tämä metodi myös identifioi insuliinin käytön ja diabeteksen keston uutena tulkinnallisena piirteenä, ja avustaa kliinisissä päätöksissä niiden riskiryhmien tunnistamisessa, jotka altistuvat diabeettiselle retinopatialle. Metodi	Hsin-Yi., Chan & Chia-Yu Su.(2018)

	osoitti, että todennäköisyys saada diabeettinen retinopatia lisääntyy 3,561 kertaiseksi potilaille, jotka käyttivät insuliinia verrattuna insuliinia käyttämättömiin potilaisiin.	
Korkea-ulottuvuus Koneoppiminen Satunnainen metsä (Random Forest, (RF))	Diabetes tapahtumien täsmällinen ennustaminen Potentiaalia diabetes tapahtumien ennustamisessa käsiteltäessä korkea-ulotteista dataa	Casanova, Saldana, Simpson, Lacy, Subauste, Blackshear, Wagenknecht, & Bertoni. (2016)
Satunnaisen metsän kaltevuuskulman vahvistaminen ja suoraviivainen erottelukyky analyysi kiintopisteinä	Kakkostyyppin diabeteksen seulontajärjestelmän kehittäminen. Metodi on suorituskyvyltään hyvä, se ei tarvitse veren talteenottoa ja tuottaa tuloksia alle viidessä minuutissa.	Moreno, Lujan, Fernandez, Manrique, Trivino, Miquel, Rodriquez & Burquiclos (2017)
Yhteensovitettu data-analyysi	Uusi automatisoitu metodi diabeettisen retinopatian seulontaan	Zsolt, Peto, Csoz, Tukacs, Molnar, Berta, Tozser, Hajdu, Nagy, Donokos. & Csutak..(2015)
Kenttä Assymetria Ioni Liikkuvuus Spektometria (FAIMS)	Tehokas lääketieteellinen diagnoosien tekeminen virtsan häilyvyyden ja vaihtelevuuden perusteella. Sallii orgaanisten vaihtelevien seosten mittaamisen biologisista näytteistä. Nämä orgaaniset vaihtelevat seokset sisältävät asiaankuuluvan tiedon henkilön aineenvaihdunnasta. Niitä voidaan käyttää sairauden diagnostiseen tarkoitukseen.	Martinez-Vernon, Covington, Arasarndnam, Esfahani, O'Connell & Savage.(2018)
Äärimmäiset oppimiskoneet (Extreme learning machine)	Ylipainon täsmällisen tilan havaitseminen. Veren indeksejä voidaan käyttää ylipainoisuus tilanteen ennustamisessa.	Chen, Yang, Liu, Liu, Liu, Zhang, Xy. & Lu. (2015)
Automaattinen syväoppimis perusteinen järjestelmä Kahden syväoppimisjärjestelmän yhteensovittaminen Täysin kierteiset verkot (Fully convolution network, (FCN))	Soluontelon havaitseminen. Diabetes potilaiden halvaus riskin luonnehtiminen käyttäen kaulavaltimo ultraääntä. Koodittamisesta ja koodinavaamisohjelmasta solukalvon lohkoihin jakamiselle. Käyttää täysin kierteisten verkkojen kolmea näyttöönotto kerros osaa solukalvon havaitsemiseen.	Biswas, Kuppli, Laba, Edla, Suri, Shauma, Chadrado-Godia, Laird, Nicolaides & Suri (2019)
Automaattinen syväoppimis perusteinen järjestelmä Kahden syväoppimisjärjestelmän yhteensovittaminen Täysin kierteiset verkot (Fully convolution network, (FCN))	Jalkamatähaavan piirteiden louhiminen ja terveen ihon alueiden erojen ymmärtäminen.	Goyal, Reeves, Davisson, Rajbhandari, Spragg & Yap. (2018)
Kierteiset hermoverkot (Syväoppiminen)	Automatisoitu ja nopea kuivan ja märän ikaan liittyvän rappeutumisen luokittelu. Kuiva ja märkätyypisten rappeutumien täsmällinen löytäminen oikea-aikaista hoitoa varten.	Serner & Serte (2018)
Reaali-aikaisen datan louhimiseen perustuva koneoppimisjärjestelmä	Järjestelmä voi havainnoida aikaisin diabeettista retinopatiaa.	Fitriati & Muratko (2016)
Kevyet laskennalliset tekniikat. Luokittelijat, kuten Tukivektori-kone(SVM), Random Forests (RF), Kaltevuussuhteen/kaltevuuden/muutosnopeuden vahvistaminen Gradient boost, AdaBoost, Gaussian Naive Bayes	Tietokannoista louhittujen verkkokalvokuvista muodostetun epätasapainoisen datajoukon luokittelu ja täsmällisyyden vahvistaminen diabeettisen retinopatian havaitseminen.	Dnyanshawar Labhade (2016)
Kierteiset hermoverkot (syväoppiminen) DFUNet	Jalkamatähaavan piirteiden erotteleminen louhimalla dataa.	Goyal & Royce (2018)

<p>Kierteiset hermoverkot</p> <p>Bayesialainen Monitehtävä Oppimis Lähestymistapa (Bayesian Multitask Learning), (BMTL)</p>	<p>Terveydenhoidon ennustavan analyysi tutkimuksen edistäminen. Kroonista sairauksista kärsivien potilaiden riskiprofiloinnin haasteellisten tehtävien ratkaisu. Epäsuotuisten terveysvaikutus tapahtumien vähentäminen. BMTL lähestymistapa voi vaikuttaa kliiniseen käytäntöön vähentämällä virheitä ja viiveitä interventioissa.</p>	<p>Normahani, Agrawal, Bravis, Falinska., Bloomfield, Mehar, Gaulton, Sangster, Ackle., Gamm, Aslam., Stanfield & Usman. (2019)</p>
<p>Yhteen liittynyt koneoppiminen</p>	<p>Edistää koneoppimisen ja data analyysin käyttöä, käytetään etenevässä tutkimuksessa. Siinä pyritään valjastamaan tietoa perustana olevista tieteistä ja kliinisestä tutkimuksesta edistämään terveydenhoitoa.</p>	<p>Thomas, Abraham & Liu (2018)</p>

Syväoppimiseen ja kohdealue tietoon perustuvia uusia metodeja on kehitetty ns. punaisten viottumien havaitsemiseen. Punaisten viottumien liittyvät olennaisesti diabeettiseen retinopatiaan (DR), ja termillä pyritään ryhmittelemään sekä Microaneuryrsmat (MAS) ja veren vuodet (HEs). Nämä viottumat havaitaan päivittäisessä kliinisessä työssä manuaalisesti lääkärin toimesta käyttäen apuna silmän pohjan valokuvia. Se on kuitenkin osoittanut vaikeaksi tehtäväksi johtuen punaisten viottumien koosta ja niiden eroavaisuuksien puutteista. Diabeettisen retinopatian tietokoneavusteista diagnosointia havainnoimalla punertavia viottumia, onkin aktiivisesti tutkittu johtuen sen kehittävästä vaikutuksista sekä kliinikoiden yhtenäisyyteen että täsmällisyyteen. Tämä voi lisäksi antaa ymmärrettävää palautetta, jota lääkäri voi helposti arvioida. (ks. Orlando, Prokofyeva, del Fresno & Blaschko 2018).

Ykköstyypin diabetekseen liittyvien biomarkkereiden kehittymisen ennustamiseen ja sairautteen olennaisesti kuuluvien hoitotoimenpiteiden seuraamiseen on kehitetty erillinen koneoppimissovellus, jonka avulla terveydenhoidossa voidaan tehdä paranneltu ja älykkäämpi järjestelmän käyttöönotto ykköstyypin diabeteksen hoidossa. (ks. Woldargay, Årsand, Batsis, Albey, Mamykina & Hartgivsen 2019.)

Data mallinnus ja veren sokeri dynamiikan ennustaminen ykköstyypin diabeteksessä käyttäen koneoppimista on nousemassa tärkeäksi teknologiaksi terveyden huollossa. Tällä hetkellä diabetekseen ei ole vielä parannuskeinoja, mutta verensokeritasojen pitäminen suositellulla tasolla on hoidossa keskeistä. Viimeisimmät edistysaskeleet diabetes teknologiassa ja itsehoito sovelluksissa ovat tehneet potilaille relevantin datan saannin helpommaksi. Keinotekoisien haiman kehitys(suljettu järjestelmä), personoitu päätöksentekojärjestelmä ja verensokeri tapahtumien hälytykset ovat myös nousseet aikaisempaa näkyvimmin esille. Keskeiseksi asioiksi näiden diabetes hoito teknologioiden kehittämisessä on tullut verensokeri dynamiikan mallintaminen. Lisääntyneen potilas historiallisen datan katsotaan myös luoneen otolliset olosuhteet koneoppimisen käyttöönotolle, ja sen sovellukselle älykkäisiin ja paranneltuihin diabeteksen hoitoa varten toteutettuihin järjestelmiin. Koneoppimisen kyky ratkaista kompleksisia tehtäviä dynaamisessa ympäristössä on merkittävästi edistänyt sen menestystä diabetes tutkimuksessa. (ks. Woldargay 2019.)

Huolimatta siitä, että koneoppimista käytetään yhä enemmän tukemaan kliinistä päätöksentekoa hoitotyössä koneoppimisalgoritmeja ja tilastollisia malleja on asetettu ennustamaan aikaista sairaalaan sijoittamista. Parhaiten suoriutuva koneoppimis algoritmi osoitti kohtuullisen ennustekyvyn lisäksi parantamismahdollisuuksia. Hoitajilla on mahdollisuus myötävaikuttaa koneoppimis algoritmeihin täyttämällä data aukkoja asiaankuuluvalla datalla. Hoitajat voivat myös parantaa datan esikäsittelyn tekniikoita ja arvioida potentiaalista arvoa käytännössä. Tutkimuksessa nousi esille myös se, että hoitajien täytyy käsitellä koneoppimisen antamaa informaatiota ja soveltaa näkemyksiään työssään tehdäkseen kunnollisia päätöksiä. Hoitajat varmistavat, että koneoppimisalgoritmit ovat muotoiltu henkilöidyllä yhteydellä jokaisen potilaan kanssa. Koneoppimisen yhdistäminen ainutlaatuisen hoitotiedon kanssa antaa heille mahdollisuuden keskittää enemmän huomiota varsinaiseen hoitotyöhön, kehittää hoitotiedettä ja parantaa yksilöllistettyä potilas hoitoa. Osallistumalla koneoppimisen täysipainoiseen kehittämiseen, toteuttamiseen ja arvioimiseen hoitajat voivat integroida ja maksimoida koneoppimisen hyötyjä. (ks. Kwon, Karim, Topaz. & Currie 2019). Koneoppimis sovellus voi tukea hoitotyötä tehokkaasti ja hoitohenkilökunta voi myös kehittää sen merkityksellisyyttä ja suorituskykyä.

Diabeteksen ollessa kaikissa ikäryhmissä yksi suurimmista terveyshuolista, on elintarviketeollisuus toimialat luoneet markkinoille erilaisia Anti-diabetes ravintotuotteita (vehnä jauhoja, ruokaöljyjä, maito tetra pakkauksia ym.) brändeja, joiden väitetään pystyvän hoitamaan diabeetikoiden korkea verenpainetta ym. Tämän todistamiseksi ohjaamattomaan koneoppimiseen liittyvä marginaaliryöstös(klusterointi) algoritmi (TBC) on toteutettu ongelman ratkaisuun. Kyseinen koneoppimisalgoritmi pystyy ottamaan huomioon Anti Diabeettisten ruoka tuotteiden ravinto koostumuksen ja analysoimaan niiden sokeria-vapauttavaa-säätö potentiaalia. Tarkasteltaessa ilmiötä, on vahvistettu Diabetes mellitusta ja Anti-diabeettisia tuotedatajoukkoja toisella ohjaamattomaan koneoppimiseen liittyvällä koneoppimisalgoritmillä eli Pääkomponentti Analyysillä(Principal Component Analysis)(PCA) ja BCA-integroidun Dmaic kuuden sigman askeleiden kautta. Tutkimuksessa havaittiin, että tuotteilla pystytään säätämään diabeettisen potilaan veren sokeripiikkejä tehokkaasti. (ks. Mulay, Joshi & Chaudari 2018).

Diabeettista retinopatiaa voidaan nykyisin ennustaa ja identifioida sujuvasti tulkinnallista biomedikaalisista piirteistä koneoppimisalgoritmeja käyttäen. Nykyisin tiedetään, mitkä riskitekijät liittyvät enemmän diabeettisen retinopatiaan. Niiden täsmällinen havaitseminen, antaa mahdollisuuden tehdä toimenpiteitä suurimpien riskiryhmien diabeettisen retinopatian ennaltaehkäisy strategioiden luomiseksi. Tutkimus osoitti, että on mahdollista rakentaa toimiva ennustemalli diabeettiselle retinopatialle kakkostyyppin diabeteksessa käyttäen datan louhinta tekniikoita, kuten tukivektorikoneita, päätöksentekopuita, tekoäly hermoverkkoja ja logistisia regressioita. Tutkimuksessa havaittiin, että ennuste suorituskyky suoriutui paremmin kuin muut koneoppimisalgoritmit ja ne saavuttivat 0,839 arvon täsmällisyydessä ja AUC alueessa. (ks. Tsao, Chan & Su 2018.)

Tutkimuksessa havaittiin myös, että diabeteksen keston lisääntyessä todennäköisyyden aste saada diabeettinen retinopatia puolestaan lisääntyy 3,5 kertaiseksi potilaille, jotka käyttivät insuliinia verrattuna insuliinia käyttämättömiin potilaisiin. Nämä tutkimustulokset ovat hyödynnettävissä klinisen päätöksen tekojärjestelmän kehittämiseksi tulevaisuuden kliinisessä työssä. (ks. Tsao ym. 2018.)

Esille nousi myös ns.korkea ulottuvuus koneoppimisen rooli diabeteksen tapahtumien täsmällisessä ennustamisessa. Tämä mahdollistaa nykyisin täsmällisen diabeteksen tapahtumien ennustamisen korkea-ulottuvaisissa asetuksissa ei rajoittuneista muuttujista määrittelystä laajasta havainto datajoukosta. Satunnaisella metsällä (RF) on paljon potentiaalia diabeteksen tapahtumien ennustamisessa käsiteltäessä korkea ulotteista dataa. Tilastollisten mallien kyky ennustaa diabetekseen liittyviä tapahtumia perustuu yleensä vain rajoittuneisiin muuttujiin, ja satunnaisen metsän (Random Forest) suorituskyvyn tutkimus diabeteksen tapahtumien havaitsemiseksi laajasta havainto datajoukosta, on osoittanut, että satunnaisen metsän (RF) mallilla pystytään arvioimaan jopa 93 muuttujaa ja luomaan muuttujien arvoasteikkoa sen mukaisesti, miten ne myötävaikuttavat diabeteksen ennustamiseen. Tutkimuksessa havaittiin, että toteutettaessa malleja, jotka perustuvat logistiseen regressioon ja satunnaiseen metsään, jossa piirteet on valikoitu ennalta, satunnaisen metsän (RF) mallin suorituskyvyksi saatiin AUC arvo 0,82. (ks. Casanova, Saldana, Simpson, Lacy, Subauste Blackshear, Wagenknecht & Bertoni 2016.)

Ohjatun oppimisen avulla voidaan nykyisin rakentaa toimiva seulontajärjestelmä kakkostyyppin diabetekseen. Analyysi pulssi oximetristä ja koneoppimis moduulista koostuvassa järjestelmässä, joukko piirteitä oli koneoppimisalgoritmien syötteenä, ne määrittävät syöttöotoksen luokan, toisin sanoen, oliko koehenkilöllä diabetes tai ei. Koneoppimisalgoritmia järjestelmässä oli satunnaisen metsän kaltevuuskulman vahvistaminen ja suoraviivainen erottelukyky analyysi kiintopisteenä. Järjestelmän testaamisen jälkeen tuloksista on havaittu, että tämä seulonta metodi diabeteksen havaitsemiseen on suorituskyvyltään verrannollinen ns. hemoglobiini A1cHb A1c(glycated haemoglobin) testiin, ja se tuottaa tuloksia nopeasti ilman veren talteenottoa (n.5 min.) (ks. Moreno 2017.)

Diabeettisen retinopatian seulomista varten on myös kehitetty uusi automatisoitu metodi. Kyseessä on kyynelnesteen proteomiikka biomarkkereita hyödyntävä yhteen soviteltu metodi. (ks. Torok, Peto, Tukacs, Molnac, Berta, Tozeen, Haydu, Navy, Domokos & Csutak. 2016). Diabeettisesta retinopatiasta johtuen lähes 5 miljoonaa ihmistä on sokeutunut, aikainen diagnosoiminen ja hoito voivat kuitenkin hidastaa diabeettisen retinopatian etenemistä. Diabeettisen retinopatian merkkejä osoittaneesta diabetespotilaiden digitaalisia verkkokalvokuvia ja kyynel neste näytteitä otettiin jokaisesta silmästä erikseen, ja tuloksia niiden antamasta proteomiikka analyysistä ja digitaalisista microaneyrysmasta (MA) tehtyjä havaintoja käytettiin tutkimuksessa koneoppimisjärjestelmän syötteenä. Tutkimus osoitti (MA) havainto metodien yksistään antavan tulokseksi 0,84 herkkyudessa ja 0,89 tarkkuudessa. Yhteensovitettu data analyysi muodostuu proteomiikka datan piirteistä yhdessä havaittujen Micro Aneurysma numeroiden kanssa liitettynä kuvaan, ja ne saavuttivat herkkyys/tarkkuus arvot 0,95|0,78. Havaittiin, että yhteensovitettu malli johti luotettavaan seulontaa metodiin. (ks. Torok ym. 2015.)

Koneoppimismetodin avulla voidaan tehdä tehokkaita lääketieteellisiä diagnooseja perustuen virtsan häilyvyyteen ja vaihtelevuuteen. Tämän on tehnyt mahdolliseksi se, että tauteja voidaan mitata helposti saatavilla olevista ruumin nesteistä, joka on avannut ovia uuden tyyppiselle ulkopuolelta vaikuttaville lääketieteellisille diagnooseille. Uusia teknologioita on kehitetty ja hienosäädetty, jotta tämä olisi mahdollista. Kenttä Assymetria Ioni Liikkuvuus Spektometria (FAIMS) on yksi näistä teknologioista, ja sen avulla sallitaan orgaanisten vaihtelevien seosten mittaaminen biologisista näytteistä (esim. virtsasta). Nämä orgaaniset vaihtelevat seokset sisältävät asiaankuuluvan tiedon henkilön aineenvaihdunnasta, ja niitä on mahdollista käyttää sairauden diagnostiseen tarkoitukseen. Hyvin kehittyneet data prosessoinnin kanavat mahdollistavat tämän datan tehokkaan käytön, ja ne ovat välttämättömiä kaikkein hyödyllisimmän datan louhimisessa kompleksisista taustalla olevista biologisista rakenteista. (ks. Martinez-Vernon, Covington, Arasarndnam, Esfahani, O'Connell & Savage 2018.)

Ylipainoisuus tilanteen ennustamisessa voidaan käyttää tehokkaasti veren indeksejä hyväksi koneoppimislähestymistavan avulla. Tämä on osoitettu olevan erittäin tehokas tapa ennakkoida diabetes mellituksen ym. kroonisiin sairaustiloihin johtavan ylipainon kehittymistä. Ylipainon täsmällisen tilan ennustamiseen käytetään uutta koneoppimis tekniikkaa, äärimmäisten oppimiskoneiden (extreme learning machine) soveltamista. Ylipainon täsmällisen tilan tunnistamisen ihmisissä tiedetään olevan erittäin ratkaisevaa terveysriskien ennaltaehkäisyssä ja vähentämisessä johtuen sen aiheuttamista terveysriskeistä, kuten diabetes mellitukseen sairastumisesta. Tämä on tärkeää myös sen takia, että ylipainoisten ihmisten määrä jatkaa kasvuaan maailmassa. (ks. Huiling, Yang, Liu, Wenbin, Liu, Zhang, & Lufeng 2015).

Tutkimuksessa havaittiin, että veri ja biokemiallista mittaamista voidaan käyttää tunnistamaan ylipainotila tehokkaasti ylipainosten ja terveiden koehenkilöiden joukosta. Havainto metodi arvioitiin tarkasti datajoukkoa kohtaan täsmällisyyden, herkkyyden, tarkkuuden ja AUC vertailu ennusteena. Ylipainoon korreloivien tekijöiden tunnistamiseksi tutkittiin myös piirteiden valitsemista. On osoitettu, että on olemassa huomattavia eroja veri- ja biokemiallisissa indekseissä terveiden ja ylipainoisten ihmisten välillä. Äärimmäinen oppimiskone on vaikuttanut lupaavalta, uudelta täsmällisestä metodilta henkilöiden ylipaino tilojen tunnistamisessa. (ks. Huiling ym.2015). Koneoppimisella voidaan antaa tarkasti arvioita eroista ylipainoisen ihmisen ja terveiden ihmisten välillä.

Tutkimuksessa nousi esille myös erinomainen syväoppimiseen perustuva sovellus diabetes potilaiden halvaus riskin ehkäisemiseksi. Tässä sovelluksessa syväoppimiseen pohjautuva kokonaan kierteisten verkko solun keskiontelo luonnehtii diabetes potilaiden halvaus riskiä käyttäen kaulavaltimon ultraääntä. Soluontelon havaitsemiseen onkin tämän takia käyttöön otettu automaattinen syväoppimis perusteinen järjestelmä soluontelon havaitsemiseen. Järjestelmä koostuu kahden syväoppimisjärjestelmän yhteensovittamisesta: koodittamisesta ja koodinavaamisohjelmasta solukalvon lohkoihin jakamiselle. Koodinavaamis ohjelma käyttää täysin kierteisten verkkojen (Fully convolution network) (FCN) kolmen osan näyttöön otto kerroksia solukalvon havaitsemiseen. Kolmelle syväoppimisjärjestelmälle toteutettiin ristiin validointi yhteyskäytäntö, ja olennainen täsmällisyys kolmelle syväoppimisjärjestelmälle koskien 407 skannausta oli 99,61 %, 97,75 % ja 49,89 % kukin erikseen. Vastaava AUC kolmelle syväoppimisjärjestelmälle oli 0,95 %, 0,91 % ja 0,93 %. Tulokset todistavat syväoppimisjärjestelmän loistavasta suorituskyvystä perinteisiin metodeihin verrattuna. (ks. Biswas, Kuppili, Laba, Edla, Suri, Shauma, Chadrado-Godia, Laird, Nicolaides & Suri 2019). Syväoppimisjärjestelmä kykenee tekemään täsmällistä solutason havainnointia, ja arvioimaan sen pohjalta tarkoin diabetekseen liittyvää kaulavaltimon halvausriskiä.

Koneoppimisalgoritmit pystyvät ennustamaan erityisiä valmiuksia ohjaamattomaan oppimiseen sisältyvillä pääkomponenttien analyysillä. Tutkimus osoittaa maailmanlaajuisen sydänsairauksista/oireyhtymistä johtuvien kuolemien syiden olevan suurin syy kuolemiin. Kehityssuunnan jatkuessa lähes 24 miljoonan ihmisen odotetaan kuolevan sydänsairauksiin. Vaikka terveydenhoito kerääkin nykyisin suuren määrän sydänsairaus dataa, jota ei ole kuitenkaan valitettavasti louhittu löytämään piilevää informaatiota tehokkaaseen päätöksentekoon. Tutkimus osoitti, että pääkomponenttien analyysien avulla (Principal Component Analysis, PCA) useiden ohjattujen koneoppimisalgoritmien täsmällisyyttä sydän sairauksien ennustamisessa voidaan edistää käyttäen minimi lukumäärä attribuutteja. (ks. Kanchan ja Mahale 2016.)

Datan louhinnan monien algoritmien lähestymistapaa on hyödynnetty diabetes sairauksien ennustamisessa. Data louhinnalla on useita tärkeitä tekniikoita, kuten luokittelu ja esikäsitteily. Diabeteksen tiedetään olevan hengenvaarallinen sairaus, joka pitäisi torjua ennalta käsin useissa kaupungistuneissa ja kasvavissa kaupungeissa. Data kategorisointi tarkoittaa diabetes potilaiden data joukkoa, joka kerätään data sairaalan säilytyspaikasta ja se koostuu erilaisista ominaisuuksia omaavista esimerkeistä. Data joukosta ovat esimerkkinä kahden kategorian veri ja virtsatestit. Data louhinnasta on tullut tunnettu käytäntö, ja sitä käytetään useissa terveysorganisaatioissa mm. diabeteksen luokitteluun. (ks Kanchan ja Mahale 2016.)

Tutkimuksessa havaittiin, että koneoppimista käytetään ns.diabetes potilaiden ennustavassa analyysissa. Tämä on tapahtunut hyödyntämällä heidän potilas dataa. Datan kerääminen säilyttäminen ja prosessoiminen on välttämätöntä, jotta löydetään hyödynnettävää tietoa merkittävien päätösten tekemiseksi. Tutkimuksessa nousi esille, että diabetes mellitus on ei-tarttuva tauti(Non-Communicable Diseases)(NCD), joka on yksi kriittisimmistä sairauksista, jolla on pitkäaikaisia komplikaatioita ja josta seuraa useita terveysongelmia. Teknologialla on rakennettu järjestelmä, joka säilyttää ja analysoi diabeettista dataa ja ennustaa mahdollisia riskejä oikein. Tutkimuksessa nousi esille ennustava analyysi metodi, joka integroi useita datan louhimis tekniikoita, koneoppimisalgoritmeja ja tilastotietoja, jotka käyttävät nykyistä ja vanhaa dataa saadakseen lisäoivalluksia ja ennustaakseen tulevia riskejä. Koneoppimisalgoritmi Hadoop MapReduce pystytään toteuttamaan Diabetes joukossa virheellisten arvojen ja polkujen löytämiseksi. Tämä mahdollistaa laajalle levinneen diabeteksen riskien ennustamisen ja potilaan riskitasojen mukainen hoitotyyppi voidaan tarjota. (ks. Kalyankar, Shivanande & Dharwakar 2017).

Ilmeni myös, että koneoppimismalleja voidaan soveltaa diabeettisen retinopatian diagnosoimisessa potilaille terveysrekisteri datan kanssa. Nousi esille viisi koneoppimismallin joukkoa diagnosoimaan diabeettista retinopatiaa potilailla terveysrekisteri datan kanssa, ja ne muodostivat hoitometodijoukon. Tutkimus osoitti, että satunnainen metsä (RF) koneoppimismallissa voi saavuttaa 92 % tarkkuuden hyvällä suorituskyvyllä. Syöttö piirteitä analysoimalla ja niiden tärkeyttä luokittelemalla, on havaittu ne altistavat tekijät, jotka panevat alkuun ihmisen diabeettisen retinopatia sairauden(liittyvät munuais ja maksatoimintoihin.) Tutkimus vahvistaa sitä tosiseikkaa, että elektroniseen terveysrekisteri dataan perustuvat sairaus diagnoosi menetelmät tulevat olemaan olennainen osa älykästä terveyden hoitoa ja mobiili terveyden hoitoa. (ks. Sun & Zhang 2019.)

Syväoppimista voidaan käyttää automatisoituun ja nopeaan määrän ikään liittyvään makulaariseen rappeutumisen luokitteluun käyttämällä kierteisiä hermoverkkoja (convolution neural networks). Niiden avulla voidaan tehdä automatisoitu ja nopea kuivan ja määrän ikään liittyvään rappeutumisen luokittelu. Diabeettinen retinopatia ja ikään liittyvä makulaarinen rappeutuminen (AMD) ovat sairauksia, joilla tiedetään olevan vahingollisia vaikutuksia ikääntyneiden silmiin vaikuttamalla verkkokalvon keskiosaan, eli mykiöön. Todistettiin, että kuiva ja märkä tyyppiset rappeutumukset voidaan löytää täsmällisesti oikea-aikaista hoitoa varten hermoverkon avulla. (ks. Serner & Serte 2018.)

Diabeettisen retinopatian seulominen on nykyisin mahdollista toteuttaa käyttämällä reaali-aikaista dataa. Tutkimusten perusteella tiedetään, että diabetes mellitus aiheuttaa komplikaatioita muissa elimissä, kuten verkkokalvoissa. Sairaus, joka hyökkää verkkokalvoon, on diabeettinen retinopatia (DR) ja se saattaa aiheuttaa sosiaaliseen ja henkiseen kuntoon vaikuttavan pysyvän sokeuden. Tätä käyttötarkoitusta varten onkin rakennettu aikaista havainnointia suorittava järjestelmä, jossa alkuperäinen tarkoitus voi olla nopeampaa. Tutkimus osoitti, että kokeessa saavutettiin kuitenkin vain 52 % testaus täsmällisyys ja tutkimuksessa havaittiin järjestelmän antavan erilaisia tuloksia käytettäessä eri dataa. (ks. Fitriati & Muratoko 2016.)

Diabeettisen retinopatian havaitsemiseen voidaan käyttää myös ohjattua oppimista, tutkimus osoitti, että ns. kevyiden laskennallisten tekniikoiden olevan tehokkaita tähän käyttötarkoitukseen. Diabetes mellitus luokitellaan joukoksi aineenvaihdunnallisia sairauksia, jossa henkilöllä on korkea verensokeri. Diabeettinen retinopatia (DR) on yleinen diabeteksen komplikaatio, joka vaikuttaa verkkokalvon verisuoniin ja sen aiheuttaa verkkokalvon hapen puute. Tietokantaan tallennetuissa silmän verkkokalvon kuvissa tapahtuu ulkonäön muutoksia, ja siksi on käytetty erilaisia analyysi metodeja ja louhittu 40 piirrettä. Tästä johtuva epätasapainoinen datajoukko vähentää kuitenkin täsmällisyyttä, joten ylinäytteisyys tehdään ennen luokittelua. Luokittelijat, kuten tukivektorikone (SVM), satunnainen metsä(Random Forest), kaltevuussuhteen/mittausnopeuden vahvistaminen(Gradient boost), Ada Boost, Gaussian Naive Bayesia puolestaan käytetään havainnoimaan diabeettista retinopatiaa. Tutkimus osoittaa, että tukivektorikone antoi hyvän 88,71 % täsmällisyyden, satunnainen metsä tekniikka antoi 83,34 % täsmällisyyden, kaltevuussuhteen vahvistamis algoritmin täsmällisyys puolestaan oli 83,34 % , Ida Boost:in 54,3 % ja Gaussin Naiivi Bayesilaisen metodi puolestaan antoi 37,09 % täsmällisyys arvon. (ks. Labhade, Chouthmol & Oeshmukh 2016).

Tutkimuksessa ilmeni myös, että koneoppimis algoritmia voidaan käyttää jalkamätähaavan piirteiden louhimisessa ja terveen ihon alueiden erojen ymmärtämisessä konenäön näkökulmasta. Tutkimuksessa havaittiin, että ihon kunnan arvioiminen molemmista luokista binäärisessä(kaksiarvoisessa) luokittelussa onnistuu uudella kierteisellä hermoverkko arkkitehtuurilla, DFUNetillä. Sillä havaittiin olevan parempi piirteiden erottelu terveen ihon ja jalkamätähaavan välisten piirteiden erojen tunnistamiseksi. Käyttäen 10 kertaista ristiinvalidointia, DFUNet saavutti korkean AUC pistemäärän 0,961. Hermoverkko pystyikin suoriutumaan jopa paremmin kuin syväoppimis ja koneoppimis luokittelijat. Ilmeni, että uusi hyvä, herkkä ja kehittynyt DFUNet, soveltuu objektiiviseen ja diabeettiseen jalkamätähaavan havaitsemiseen hyvin. Tämä mahdollistaa diabeettisten potilaiden joukossa onnistuneen diabeettisen jalkahoidon, joka puolestaan merkitsee kustannustehokasta, etäistä ja tarkoituksenmukaista terveyden hoidon ratkaisua. (ks.Goyal & Royce 2018.)

Tutkimuksessa nousi esille myös, kuinka kierteistä hermoverkkoja käytetään havaitsemaan ja erottelemaan verkkokalvoon liittyviä retinaalisia verisuonia fundoskooppisissa kuvissa. Diabeettisen retinopatian erityisen progressiivisen vaiheen diagnosoinnin tai luokittelun tiedetään olevan tärkeää diabetes mellituksen ja siihen liittyvän retinopatian arvioimisessa. Useimmat laajentuneet ophthalmoskooppiset, havaittavissa olevat löydökset viittaavat verkkokalvokuvan verisuoniin kansainvälinen kliinisen diabeettisen retinopatian ja diabeettisen makulaarisen turvotus sairauden vakavuuden mittaamisen mukaisesti. Objektivisen ja täsmällisen määritelmän löytämiseksi diabeettisen retinopatian erityisen sairauteen liittyvään luokitteluun, on olennaista automaattisesti havaita ja louhia verkkokalvoon liittyviä, retinaalisia, verisuoni fundoskooppisia kuvia. (ks. Normahani, Agrawal, Bravis, Falinska, Bloomfield, Mehar, Gaulton, Sangster, Ackle, Gamm, Aslam, Stanfield & Usmann. 2019).

Tutkimusryhmän tekemät vertailut useiden kierteisten hermoverkkojen tunnistamien verisuonten fundoskooppisista kuvista osoittivat lähestymistavan tehokkuuden. Ennustavaa analytiikkaa puolestaan käytetään terveydenhoidossa riskien profiloimiseen kroonisissa sairauksissa, ja siinä käytetään Bayesilaista Monitehtävä Oppimis Lähestymistapaa. Potilaan terveystietokannasta saatu kliininen älykkyys voi tukea kliinistä päätöksentekoa henkilöidssä ja ennaltaehkäisevässä hoidossa. Ennustava analyysitutkimus terveydenhoidossa käyttää elektronisia terveystietokantoja ja mahdollistaa riskiprofiloinnin. (ks. Normanhi 2019.)

Tutkimuksessa nousi esille myös se, että kroonista sairauksista kärsivät potilaat kohtaavat monia riskejä suurina epäsuotuisina terveysvaikutus tapahtumina. Voimassa olevat riskimallit tyypillisesti keskittyvät yhden erityisen tapahtumaan eivätkä ne ennusta tuloksia. Bayesilainen Monitehtävä Oppiminen (Bayesian Multitask Learning) (BMTL) on käyttöönotettu suunnittelutiede lähestymistapana parannellun riskiprofiloinnin saavuttamiseksi. Bayesilainen Monitehtävä Oppiminen mahdollistaa terveydenhoidon tarjoajien saavuttaa monipuolinen riski profilointi ja suuri määrä tapahtumia samanaikaisesti. Kokeellisessa arvioissa havaittiin, että BMTL lähestymistapa saavuttaa paremman ennustavan suorituskyvyn verraten vaihtoehtoihin, jossa malli moninkertaistaa tapahtumat erikseen. Tutkimuksessa huomattiin myös, että useissa tapauksissa BMTL lähestymistapa voi suoriutua merkittävästi paremmin kuin olemassa olevat monioppimistekniikat. BMTL lähestymistapa voi luoda merkittävää potentiaalista vaikutusta kliiniseen käytäntöön vähentämällä virheitä ja viiveitä ennaltaehkäisevässä interventioissa. BMTL:ää koskeva tutkimus on antanut myötävaikutusta terveysinformaatio teknologialle, big datalle, ennustavalle analyysille ja suunnittelutiede tutkimukselle. (ks. Normanhi 2019.)

Yhteen liittyneen koneoppimisen (Federated machine learning) (FML) uskotaan tulevaisuudessa pystyvän merkittävästi kehittämään päätöksentekoa diabeteksen seulomisessa, ja tämän arkkitehtuuria on arvioitukin sen toiminallisen suorituskyvyn näkökulmasta perustuen kolmeen teoreemaan ja caseen kliinisestä päätöksenteosta. Uskotaan myös vahvasti, että tietojärjestelmätieteen kehitys itsessään saattaa antaa arvokasta myötävaikutusta etenemistieteeseen. (ks. Thomas, Abraham & Liu 2018).

Tutkimuksessa havaittiin, että koneoppimisen onnistuneessa käyttöönotossa diabeteshoitoon, pitäisi aina määritellä ja analysoida tämän mahdollinen myötävaikutus kliniseen päätöksenteon tukeen terveydenhoito paikassa, eli kuinka se lisää päätöksentekijöiden tietoa ja sen pohjalta tapahtuvaa reagoitua. Ei-tietoon-perustuvien klinisten päätöksenteon tukijärjestelmien suunnittelussa diabetes hoitoon pitäisikin huomioida se, että niiden perimmäinen tarkoitus on nimenomaan lisätä klinikoiden tietoa hoidossa päätöksentekohetkellä, antamalla erilaisia malleja mm. sairaustilan kehityskulusta, diagnosoinnista ja liitännäissairauksien ilmentymisestä eikä antaa välttämättä niinkään automaattisesti suoraan hoitosuosituksia ilman vastuussa olevaa terveydenhoidon ammattihenkilöä. Koneoppiminen voi näin ollen huomattavasti kehittää lääketieteellisen diabetes hoidon laatua hoidon aikana ja sitä ennen antamalla tärkeää klinistä päätöksenteon tukea klinikoille.

7. Diskussio

Tutkimuksen tavoitteena oli löytää toimivia diabeteshoidossa sovellettavia koneoppimis malleja ei-tietämyskantaisen kliinisen päätöksenteon tukijärjestelmän suunnittelumallia varten terveydenhuolto-organisaatioille. Diabeteshoidon lääketieteellisen laadun parantamiseksi kliinisissä tietojärjestelmissä, siinä sovellettavia koneoppimisalgoritmeja etsittiin käyttäen tutkimusmenetelmänä integroivaa kirjallisuuskatsausta. Tutkimuskohteesta voitiin näin ollen muodostaa laaja ja kattava kuvaus. Tutkimuksessa haettiin tarkoin ja täsmällisesti niitä toteutettuja koneoppimismalleja, joita terveydenhuolto-organisaatiot voivat käyttää hyväksi laadukkaassa ja kustannustehokkaassa kliinisen päätöksenteon tukijärjestelmän suunnittelussa. Tutkimuksen tuloksena saatiin integroivan kirjallisuuskatsauksen avulla muodostettua synteesejä koneoppimisen hyödyntämisestä ei-tietopohjaisen kliinisen päätöksenteon tukijärjestelmän suunnittelumallin toteuttamiseksi terveydenhoito paikkoihin. Tämä tapahtui määrittelemällä tärkeimpiä diabeteshoidossa tarvittavan kliinisen päätöksenteon tukijärjestelmän vaatimuksia.

Integroivan kirjallisuuskatsauksen perusteella luodussa optimaalisessa diabeteshoidon kliinisen ei-tietämyskantaisen päätöksenteko tukijärjestelmän suunnittelumallin idean kartoittamisessa havaittiin koneoppimismallien pystyvän huomattavasti edistämään kliinistä päätöksentekoa diabetes hoidossa. Syväoppimiseen pohjautuvat järjestelmät, ohjaamaton oppiminen, ohjattu oppiminen, yhteen liittynyt koneoppiminen, äärimmäiset oppimiskoneet ja korkea-ulottuvuus koneoppiminen ovat tutkimustulosten perusteella huomattavasti pystyneet kehittämään diabetes hoidon laatua.

Syväoppimiseen perustuvat kliiniset päätöksenteko tukijärjestelmät ovat osoittautuneet toimiviksi diabetes hoidossa, ja etenkin syväoppimiseen pohjautuvien keinotekkoisten hermoverkkojen hyödyllisyys diabeettisen retinopatian diagnosoimisessa verkkokalvokuvista. Modifoitu Hopfield neuraalinen hermoverkko (MHNN) on muun muassa antanut abnormalien verkkokalvokuvien luokittelussa tehdyssä kokeessa korkeita herkkyyden ja spesifisyyden arvoja järjestelmässä. Kyseinen neuraali hermoverkko (MHNN) antoi retinopatiaan viittaavien abnormalien verkkokalvokuvien

luokittelussa tehdyssä kokeessa mallille korkeita herkkyyden ja spesifisyyden arvoja kliinisessä kokeessa, ja myös ulkoinen tarkkuusarvo osoittautui mallissa erittäin korkeaksi. Korkeaa herkkyyttä arvoa voidaankin pitää erityisen tärkeänä ja hyödyllisenä onnistuneen sairauden vakavuusasteen diagnosoinnin luokittelussa.

Syväoppimisjärjestelmä on osoittanut erinomaisen hyvää suorituskykyä diabeettisen retinopatian, silmänpainetaudin ja ikään liittyvän makulaarisen rappeutuman havaitsemisessa suurissa, monietnisissä ihmispopulaatioissa. Erityisesti herkkyyttä ja tarkkuus ovat olleet syväoppimisjärjestelmällä korkeita koskien diabeettisen retinopatiaa, ja siihen liittyvien silmäsairauksien tunnistamista. Kliinisestä potilasdatasta voidaan esim. kompleksisella hoidon aiheeseen liittyvällä syväkaareutuvalla neuraali hermoverkolla (Google Net) nykyisin ennakoita hoidon aihetta tarkoin. Tämä hermoverkkosovellus antoi erinomaisia herkkyyttä, tarkkuus ja ROC arvoja. EyePacs-1 ja Messidor-2 syväoppimissovellukset antoivat myös diabeettisen retinopatian ja diabeettisen makulaarisen turvotuksen viittaamisessa korkeita herkkyyttä ja tarkkuus arvoja. Syväoppimiseen pohjautuva älykäs etähoitojärjestelmä Intelligent Diabetic Assistant (IDA) puolestaan kykenee antamaan itsenäisiä päätöksiä koskien diagnooseja ja hoidon priorisointia. Syväoppimiseen pohjautuvista tekniikoista kuvaperusteinen kehysten hyödyntämisen maksimointi ruokavalion virheiden määrittelymiseksi taas on osoittautunut tehokkaaksi tavaksi antaa yksilöllisiä terveellisiä ravintosuosituksia diabeetikoille.

Syväoppimiseen ja kohdealuetietoon perustuvia metodeja on kehitetty olennaisesti diabeettiseen retinopatiaan liittyvien punaisten viottumien havaitsemiseen. Nämä menetelmät kykenevät luokittelemaan verkkokalvokuvien piirteistä diabeettisen retinopatian yhtenäisesti ja täsmällisesti verraten niitä lääkärin manuaalisesti silmänpohjan valokuvista tekemiin oikein ja väärin tehtyihin luokittelu havaintoihin. Tutkimus osoitti, että syväoppimiseen pohjautuvista metodeista, kierteinen hermoverkko oppii piirteitä myös enemmän yhdistämällä käsin sommiteltuja piirteitä. Syväoppimista hyödyntävillä koneoppimissovelluksilla voidaan myös mm. analysoida ruokavalio dataa. Nykyiset automaattiset syväoppimiseen pohjautuvat järjestelmät voivat havainnoida myös jopa solutasolla asti ja voivat näin ollen määrittellä tarkoin diabetespotilaiden halvausriskiä.

Syväoppimiseen perustuvat kliiniset päätöksenteko tukijärjestelmät voivat kaiken kaikkiaan tuottaa klinikoille nopeasti tärkeää tietoa diabeteksen liitännäissairauksien tilasta, tunnistaa niitä täsmällisesti, auttaa toimittamaan liitännäissairauksiin oikeanlaista hoitoa täsmällisesti, luokitella liitännäissairauksien vakavuusasteen, ennustaa diabeteksen olennaisesti liittyvien silmäsairauksien kehityskulkua ja auttaa rakentamaan tehokas seulonta järjestelmä diabeteksen yleisimpien liitännäissairauksien ehkäisyä varten. Syväoppiminen ja tarkka kohdealue tieto ovat tehokkaita diabeteksen diagnosointityökaluja, ja ne mahdollistavat klinikoille diabeettisen retinopatian oireisiin perustuvan vakavuusasteen luokittelun verkkokalvokuvista. Diabeettisen retinopatian vakava-asteinen komplikaatio eli makulaarinen turvotus voidaan luokitella taudin vakavuusasteen eri vaiheisiin suurella koneoppimismallin tarkkuusarvolla perustuen tulehduseritteen sijaintiin ja makulatäplän koordinaatteihin.

Koneoppiminen voi edistää kliinistä päätöksentekoa ennen kaikkea tunnistamalla ja seulomalla diabetekseen liittyviä silmäsairauksia. Diabetekseen liittyvää verkkokalvo-, kardiovaskulaarisia ja neuropaattisten komplikaatioiden riskiin liittyviä komplikaatioita ja diabetekseen liittyvää kuoleman riskiä voidaan Mimic 3-tietokannan, koneoppimisteknologian ja kaksijäsenisen regressiomallinnuksen avulla selvittää luomalla koneoppimismalleja, jotka antavat ennustemalleja luokittelemalla tehokkaasti dataa vain muutaman muuttujan yhdistelmänä. Diabeettisen retinopatian vakava-asteinen komplikaatio eli makulaarinen turvotus voidaan myös luokitella taudin vakavuusasteen eri vaiheisiin suurella koneoppimismallin tarkkuusarvolla perustuen tulehduseritteen sijaintiin ja makulatäplän koordinaatteihin.

Data luokittelutekniikoista Gaussilainen menetelmä on osoittanut olevan paras koneoppimis malli diabetes potilaiden data luokittelussa antaen korkeita tarkkuus, herkkyys ja erityisyysarvoja. Gaussilaiseen dataluokitteluun perustuva koneoppimisjärjestelmä osoittautui parhaaksi järjestelmäksi luokiteltaessa diabetespotilaiden dataa sairauden ennakoimiseksi, tämä voikin antaa paljon kliinistä lisätietoa klinikoiden käyttöön. Ristiinvalidointi koneoppimismetodeilla voidaan puolestaan kehittää kakkostyyppin diabetekseen liittyvien komplikaatioiden riskien

arviointia kehittämällä siihen tarkan kaavan, joka auttaa hoidosta vastuussa olevia päätöksentekijöitä sairauden kokonaistilanteen arvioinnissa. Ristiinvalidointi koneoppimismetodilla voidaan siis arvioida kakkostyyppin diabetekseen liittyviä komplikaatoriskejä tarkoin ja vaikuttaa hoitotulokseen myönteisesti valitsemalla ennustavia muuttujia datajoukosta. Metodilla valitaan riskikaavaan useita potilaiden muuttujia, komplikaatoriskejä ennustaviksi tekijöiksi. Tapahtumapohjainen todistus viitekehys koneoppimismetodi voi tulkita ja ratkaista tehokkaasti uusia kompleksisia lääketieteellisiä diabeteshoidon ongelmia käytännönläheisissä asiantuntijuutta vaativissa ongelmissa. Tämän koneoppimismetodin on kiistatta osoitettu toimivan diabeteshoidon diagnosoinnissa, tehostamalla päätöksentekoa vaikeiden lääketieteellisten käsitteiden sisällön analyysin ymmärtämisessä. Metodi hyödyntää aikaisemmin opittua, sovellusalueen uusien ongelmien ratkaisussa. Tukivektorikone (SVM), satunnainen metsä(Random Forest), kaltevuussuhteen/mittausnopeuden vahvistaminen(Gradient boost), Ada Boost ja Gaussian Naive Bayes ovat puolestaan osoittautuneet parhaiksi dataluokittelijoiksi diabeettisen retinopatian täsmällisessä havaitsemisessa.

Koneoppimisalgoritmeilla ja tutkimuskuvilla voidaan seuloa diabeettista retinopatiaa nykyisin hyvin tehokkaasti. Koneoppimista pystytään soveltamaan aterian jälkeisen veren glykeemisen vastineen täsmällisessä ennustamisessa, joka mahdollistaa diabeetikon yksilöllisen ja suunnitelmallisen dieettihoidon. Koneoppimisen ja datalouhinnan kombinaation avulla voidaan puolestaan edistää diabetes tutkimuksen tekoa, bioteknologiaa ja terveystieteitä. Koneoppimiseen perustuva Mhealth sovellus määrittelee todennäköisyyden joutua diabetes sairauden vaikutuksen alaiseksi. Sovellus antaa potilaille päätöksenteon tukea mahdollisuuksista selvitä ilman lääkäriä tai lääketieteellistä testiä.

Ohjaamattoman koneoppimisen algoritmeista pääkomponenttien analyysi (Principal Component Analysis, PCA) osoitti pystyvän ennustamaan diabetes sairautta. Ohjattu koneoppiminen ennustaa diabeteksen liitännäissairauksista johtuvaa sairaalaan joutumista suurella tarkkuudella. Tilastollista metodia ja hajautettua koneoppimistekniikkaa yhdistävää hybridi metodia puolestaan on käytetty onnistuneesti potilaan hoidollisten tavoitteiden saavuttamisessa koskien diabetesta. Tämä osoittautui parhaaksi koneoppimistekniikaksi ilman kontrollia olleiden diabetes potilaiden kohdalla antamalla korkeita tarkkuusarvoja.

Ohjatuilla koneoppimisalgoritmeilla ennustetaan diabeteksen kehityskulkua, luokitellaan vakavuusaste, kyetään ennustamaan täsmällisesti sairaalaan joutuminen ja voidaan tuottaa tietoa potilasryhmän tilasta. Tutkimuksessa havaittiinkin, että erityisesti ohjattu oppiminen, (jossa neuroverkolle syötetään ennalta luokiteltuja näytteitä) edistää diabeteksen diagnosointiin liittyvää kliinistä päätöksentekoa. Tämä tapahtuu mm. lisäämällä yksityiskohtaista ja asianmukaista tietoa diabeteksen liitännäissairauksista ja antamalla tehokkaita päätöksenteko työkaluja päätöksentekijöiden käyttöön. Ohjatuilla koneoppimisalgoritmeilla ennustetaan siis diabeteksen kehityskulkua, luokitellaan vakavuusaste, kyetään ennustamaan täsmällisesti sairaalaan joutuminen ja voidaan tuottaa tietoa potilasryhmän tilasta. Ohjattu oppiminen mahdollistaa ennustavien mallien luomisen diabeteshoitoon, tällöin on mahdollista esim. identifioida sairaalahoitoa ja lääkitystä lähitulevaisuudessa tarvitsevia potilasryhmiä.

Ohjattu oppiminen mahdollistaa myös diabeettisen retinopatian riskitekijöiden täsmällisen havaitsemisen, ja tämä puolestaan mahdollistaa ennaltaehkäisyohjelmien yhteydessä tehokkaan päätöksenteon. Ohjatulla koneoppimisella voidaan ennustaa diabeteksen täsmällisiä tapahtumia laajasti määritellyistä datajoukoista, joka tekee sairauden reaaliaikaisen seuraamisen päätöksenteon tukijärjestelmissä helpoksi. Ohjattua oppimista käyttämällä voidaan myös toteuttaa toimiva kakkostyyppin diabeteksen seulontajärjestelmä, joka tuottaa tulokset nopeasti klinikoiden käyttöön ilman aikaa vieviä verikokeita. Ohjatun oppimisen avulla voidaan myös hallita lyhytaikaisia ja pitkäaikaisia sairauden komplikaatioita, ja hoitoa voidaan myös kehittää yksilöllisempään

suuntaan ja hoitoon sitoutumattomiin riskipotilasryhmiin voidaan myös kohdistaa interventioita.

Diabeettista munuaissairautta (nefropatiaa) voidaan seuloa tehokkaasti kliinisestä datasta käyttäen ohjatun koneoppimisen algoritmeista satunnaista metsää (Random Forest) ja yksinkertaisia logistista regressiometodeita (Simple Logistic Regression Methods). Muista kroonisista sairauksista, kuten kardiovaskulaarisista sairauksista kärsivien potilaiden joukossa, voidaan koneoppimisen ja elektronisten lääkerekisterien tietokannan avulla nykyisin toteuttaa ennustava malli diabeteksen sairastumista varten, ja malli antoi hyvän tarkkuusarvon. Ohjatun oppimisen algoritmia käyttämällä koneoppimismallilla voidaan myös ennustaa ja ehkäistä hypoglykeemisiä tapahtumia ykköstyypin diabetesta sairastavissa. Ykköstyypin diabeteksen liittyvien biomarkkereiden kehittymisen ennustaminen ja sairauden hoitotoimenpiteiden seuraamiseen kyseisellä tekniikalla kehitetty koneoppimissovellus mallintaa verensokerin dynamiikkaa varsin hyvin.

Ketjujohteiset hermoverkot, myötä kytketty hermoverkko, tukivektorikoneet, gaussilainen prosessi, geneettinen algoritmi ja syvä hermoverkot osoittautuivat tehokkaiksi verensokerin ennustajiksi diabetes potilaiden hoidossa. Diabetes hoidossa käytetään nykyisin paljon moni lineaarisista regressioita ennustemallien kehittämisessä. Havaittiin myös, että diabeteksen liittyvän kaulavaltimon rasvoittumistaudin kehityskulku voidaan ennustaa käyttämällä Fisher Score-perusteista piirre valitsemismetodia ja sen ennustamiseen voidaan käyttää Bayers -perusteista oppimismetodia. Tämä koneoppimismetodi on osoittautunut tehokkaaksi nopeasti edistyvien glukoositoleranssista kärsivien potilaiden tunnistamisessa ja identifioinnissa.

Satunnaisen metsän luokittelijan avulla voidaan empiirisesti tunnistaa vioittumilta vaikuttavia objekteja. Metodi pystyy suorituskkyisemmin arvioimaan viottuman eri lähteitä verrattuna ihmisasiantuntijan arvioon. DeepDR-net syväoppimis algoritmia voidaan käyttää diabeettisen retinopatian olemassaolon ja luokitellun diabeettisen retinopatian vakavuusasteen havaitsemiseen. Biomedikaalisista piirteistä voidaan myös nykyisin ennustaa ja identifioida diabeettista retinopatiaa tehokkaasti käyttäen koneoppimisalgoritmeja. Tukivektorikoneita, päätöksentekopuita, tekoäly hermoverkkoja ja logistisia regressiota käyttäen voidaan rakentaa ennustemalli

koneoppimisalgoritmeista, jotka saavuttivat korkean AUC arvon. Tukivektorikoneella toteutetulla koneoppimismetodilla voidaan identifioida insuliinin käyttö ja diabeteksen kesto uutena tulkinnallisena piirteenä avustamaan kliinisiä päätöksiä riskiryhmien tunnistamisessa. Satunnaisen metsän (Random Forest) suorituskyvyn tutkimus diabeteksen tapahtumien havaitsemiseksi laajan havaintodatan joukosta, on osoittanut myös tarkaksi metodiksi diabetes tapahtumien havaitsemisessa.

Naiivi Bayesilainen antoi korkeita tarkkuusarvoja samoin satunnainen metsä, joka antoi hieman alempia tarkkuusarvoja. Havaittiin myös, että koneoppimismallit voivat olla kehittämässä diabetes hoitoa yksilöllisempään suuntaan. Koneoppimismallilla pystytään nykyisin myös ennustamaan yksilöllisiä eroja diabeetikoiden tulevaisuuden glykeemisessä reagoinnissa, ja sitä on mahdollista soveltaa mainiosti diabeetikoiden hoitoon sitoutumisen täsmällisessä mittaamisessa ja heidän sairaanhoidon tarpeiden kartoittamisessa. Satunnaisen metsän algoritmeja soveltamalla löydetään sairaalaan sitoutumisen ennustajat. Tutkimus osoitti myös, että osallistamalla koneoppimisen täyspainoiseen kehittämiseen, toteuttamiseen ja arvioimiseen hoitohenkilökunta voi maksimoida koneoppimisesta saadut hyödyt.

Ohjaamattoman koneoppimisen algoritmeista Marginaali-ryvästys klusterointi algoritmin toteuttamisella pystyttiin ratkaisemaan diabeettisten ruokatuotteiden kykyä vaikuttaa diabeetikoiden taudinkulkuun. Algoritmi soveltui hyvin myös ennaltaehkäisevien toimenpiteiden tai hyödyllisten diabeteksen hoito- ja parannuskeinoihin analysoimalla niiden sokeria-vapauttavaa-säätö potentiaalia tarkoin. Toisella ohjaamattomalla koneoppimisalgoritmillä, eli Pääkomponenttien Analyysillä(Principal Component Analysis)(PCn) ja BCA-integroidun Dmaic kuuden sigman askeleiden kautta on pystytty myös säätämään diabeettisen potilaan veren sokeripiikkejä.

Koneoppimisella pystytään lisäämään tietoaaineistoa sairaudesta nopeasti ja toimittamaan nopeasti kliinistä dataa sitä tarvitseville klinikoille, ja sairauden vakavuusasteen tasoon perustuvaa luokittelua pystytään myös tekemään erittäin tehokkaasti. Tämä tukee myös reaaliaikaista kliinistä päätöksentekoa liitännäissairauksien ilmentyessä, ja ennustavia biomarkkereita voidaan tunnistaa. Koneoppimisen avulla voidaan ennakoida sairauden puhkeamista ja taudin kehityskulkua tehokkaasti, ja voidaan myös päättää diagnooseista

nopeasti ja priorisoida hoitoa onnistuneesti. Koneoppimisalgoritmeilla ja tutkimuskuvilla voidaan seuloa diabeettista retinopatiaa erittäin tehokkaasti. Koneoppimisella voidaan myös tuottaa sovelluksia diabeteksen itsehoitoon. Nämä edellä mainitut asiat voivat myös luonnollisesti tarjota tehokkaita päätöksenteko työkaluja klinikoille ja vähentää rutiininomaisia tehtäviä hoitoprosessissa. Uusista koneoppimistekniikoista äärimmäisten oppimiskoneiden käyttöönotto voi ennustaa potilaan ylipainotilan kehittymistä ja auttaa myös ennakoimaan terveysriskejä. Korkea-ulottuvuus koneoppiminen puolestaan voi ennustaa täsmällisesti diabetes tapahtumia.

Koneoppiminen on erinomainen työkalu laajasti lisääntyvän diabetes ongelman hallinnassa ja siitä aiheutuvien suurten terveydenhoitomenojen kustannusten vähentämisessä. Tarkasteltaessa sen soveltamista diabetes hoitoon voidaan todeta, että koneoppimismallit mahdollistavat tehokkaan potilas tietoaineiston käsittelyn ja luokittelun terveydenhuolto-organisaatioissa. Vakavan ja lisääntyvän kansanterveydellisen ongelman ratkaisemisessa tarvitaan ennen kaikkea tekoälyä hyödyntäviä ratkaisuja. Koneoppimisteknologian käyttöönotto voi tehostaa huomattavasti tietokoneavusteista päätöksenteon tukea terveydenhuollon ammattihenkilöiden keskuudessa. Diabetes hoidossa tämä voi tarkoittaa mm. klinikoiden kognitiivisen kuormituksen huomattavaa vähenemistä. Tämä johtuu sairauden diagnosoinnin helpottumisesta, taudinkuvan vakavuusasteen havainnoimisen tehostumisesta, sairauden kehityskulun ennustettavuuden täsmenemisestä ja parantuneen koneoppimisella toteutettujen digitaalisen mobiilisovellusten tuoman itsehoito mahdollisuuden myötä myös vähentyneestä työtaakasta.

Ei-tietoon-perustuvan kliinisen päätöksenteon tukijärjestelmän suunnittelussa terveydenhoito paikkaan täytyy myös huomioida rajoitteet, joita koneoppimista hyödyntävillä kliinisillä päätöksenteon tukijärjestelmillä vielä on olemassa. Diabeettisen retinopatian luokittelun tietokoneavusteisessa seulontateknologiassa, pitäisi vielä osoittaa koneoppimisalgoritmien ja tutkimuskuvien kelpoisuus kyseiseen käyttötarkoitukseen suuremmasta tutkittavan populaation tietoaineistosta. Syväoppimista hyödyntävällä älykkäällä etähoitojärjestelmällä on myös havaittu olevan klinikoiden keskuudessa vielä puutteita järjestelmän käytettävyydessä, joita pitäisi huomioida ennen järjestelmän varsinaista suunnittelua. Hermoverkkojen hoidosta antamia arvoja ei saisi koskaan myös

ottaa varsinaisina hoitosuosituksina vaan nämä tulisi tapahtua aina hoitavan lääkärin toimesta. Syväoppimiseen perustuvien sovellusten algoritmien käyttökelpoinen toteutettavuus kliinisiin puitteisiin vaatii vielä myös lisänäyttöä jatkotutkimusten kautta. Pitäisi saada lisänäyttöä siitä, johtaako algoritmin käyttö todellisuudessa parempaan hoitoon ja tuloksiin verrattuna nykyisiin silmätautiopin lääkärin arviointeihin. Ykköstyyppin diabeteksen verensokerin dynamiikan mallinnuksessa on havaittu hybridijärjestelmissä suurimmaksi puutteeksi se, että hiilihydraattien ottoa täytyy arvioida manuaalisesti, jolloin se on altis vakavasti ennustavan suorituskykyyn vaikuttaville virheille. Koneoppimisella toteutetuilla ennustemalleilla täytyisi myös osoittaa olevan kiistaton vaikutus diabeteksen ja siihen liittyvien komplikaatioiden hoidossa, koska koneoppimisella toteutetuilla ennustemalleilla on potilaan hoitotuloksiin vielä varsin niukka tutkimuksellinen näyttö.

Tutkimus rajattiin käsittelemään koneoppimisjärjestelmän soveltuvuutta diabeteshoitoon ei-tietoon-perustuvan kliinisen päätöksenteon tukijärjestelmän vaatimusten määrittämiseksi. Muiden tekoälypohjaisten työkalujen ja ratkaisujen käyttöönottoa tähän käyttötarkoitukseen ei tutkimuksen puitteissa käsitelty, ja aihepiiri pyrittiin rajaamaan täsmälliseksi. Tutkimuksen rajoittavaksi tekijöiksi nousi mielestäni paljolti integroivan kirjallisuuskatsauksen tiedon etsimisvaiheessa oikeanlaisen hakumerkkijonon määrittely relevantin tutkimuskirjallisuuden löytämiseksi tietokannoista. Hakujen mukana saattoi tulla mukaan asiaankuulumatonta, irrelevanttia kirjallisuutta, jota oli karsittava pois katsauksesta. Oikeanlaisen tarkasteluvälin valitseminen tutkimusaineiston hakemisen yhteydessä oli myös vaikeaa, koska oli vaikea tietää, oliko vanhemmissa aikaisemmin julkaistuissa tutkimuksissa jotain tärkeää tutkimuskysymyksiin liittyvää oleellista tietoa. Englanninkielisten tieteellisten, merkityksiltään kompleksisten käsitteiden ymmärtäminen ja oikea tulkitseminen oli toisinaan myös vaikeaa.

Katsauksesta tuli myös varsin laaja, joten tulosten kuvaaminen selvästi ja johdonmukaisesti synteessin avulla osoittautui melko työlääksi. Tutkimusaineistoon valittiin kuitenkin vastikään julkaistuja puolueettoman tieteellisen arvioinnin läpikäyneitä relevantteja katsaus, tutkimus ja konferenssi artikkeleita. Valitut artikkelit olivat julkaistu viimeisen viiden vuoden aikana ja niiden sisältämä tutkimusnäyttö oli saatu aikaan

suurehkolla otannalla, joten tutkimuksen validiteettia ja reliabiliteettia voidaan pitää kuitenkin melko hyvänä.

Mielenkiintoinen jatkotutkimus aihe voisi olla koneoppimisen käyttöönottoon liittyvien suurimpien ongelmakohtien selvittäminen terveydenhuollon ammattihenkilöiden keskuudessa, koska koneoppimisjärjestelmän matalasta hyväksymisasteesta on saatu joissakin terveydenhuolto organisaatioissa tutkimukseen pohjautuvaa näyttöä. Tutkimuksen tulos osoittaa, miten koneoppimismalleja voidaan hyödyntää onnistuneesti diabeteshoidossa toimivan ei-tietämyskantaisen kliinisen päätöksentekojärjestelmän toteuttamisessa. Järjestelmän onnistunut toteuttaminen mahdollistaa tehokkaan kliinisen päätöksenteontuen kompleksisissa ja vaihtuvissa tilanteissa diabetes hoidossa erilaisten siihen liittyvien muuttujien käsittelyssä.

Ei-tietopohjaisten päätöksentekojärjestelmien tarve terveydenhoitopaikoissa lisääntyy tulevaisuudessa johtuen mm. kroonisten sairauksien, kuten kakkostyypin diabeteksen lisääntymisestä populaatiossa. Kliinisen päätöksenteon on kaikilta osin oltava optimaalista koskien sairauden diagnosointia, toimivia interventioita, vuorovaikutusta ja hoidon oikea-aikaista toimittamista potilaille. Tutkimuksessa esiin nousseiden koneoppimismetodien, dataluokittelijoiden ja syväoppimiseen pohjautuvien neuraalisten hermoverkkojen huomioiminen diabeteshoidon toimivan kliinisen päätöksentekojärjestelmän vaatimusten määrittelyssä voi hyödyntää terveydenhuolto organisaatioita toimivan järjestelmän suunnittelussa ja teknisessä toteuttamisessa diabetes hoitoon.

LÄHTEET

Anand, R.S., Stey, P., Jain, S., Biron, D.R., Bhatt, H., Monteiro, K., Feller, E., Ranney, M.L, Sarkar, J.N. & Chen, S. E.(2018).Predicting Mortality in Diabetic ICU patients Using Machine Learning and Se. *Transi Sci Proc.* [online]. [4.12.2019]. Saatavissa: <https://pubmed.ncbi.nlm.nih.gov/29888089-predicting-mortality-in-diabetic-icu-patients-using-machine-learning-and-severity-indices/>

Akagi, D. (2014). A Primer on Deep Learning. [online]. [3.4.2020]. Saatavissa. <https://www.datarobot.com/blog/a-primer-on-deep-learning>

Alavi, M., & D. Leidner. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107-136. [online]. [9.1.2020]. Saatavissa: https://moodle.ufsc.br/pluginfile.php/950622/mod_resource/content/1/MISQ%202001%20Vol%2025%20No.%201%20page%20107%20Alavi%20Leidner.pdf

Athey, S., & G. Imbens (2015). Machine learning methods for estimating heterogeneous causal effects. *stat*, 1050(5). [Verkkodokumentti]. [3.2.2020]. Saatavissa: <https://econpapers.repec.org/paper/eclstabus/3350.htm>

Basu, S. Sussman, J. Berkowiz,S.A.,Hayward,R.& J.S. Yudkin(2015). Development and validation of Risk Equations for Complications Of type 2 Diabetes(RECORDe) using individual participant data from randomized trials. *Lancet Diabetes Endocrinal.* 15(10),788-798.(Oct 2017). [online]. [8.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/28803840>

Baumeister F., & M. R. Leary (1997). Writing narrative literature reviews. *Review of General Psychology*, 1(3), 311-320. [online]. [12.3.2019]. Saatavissa: <https://psycnet.apa.org/doiLanding?doi=10.1037%2F1089-2680.1.3.311>

Bearfield, D. A. & S. E Warren (2008). Writing a Literature Review: The Art of Scientific Literature. Teoksessa: *Handbook of Research Methods in Public Administration*, 61–72. Toim. Kaifeng Yang & Gerald J. Miller. Boca Raton: CRC Press.

Berner, E. S. (2007). Clinical Decision Support Systems. *Berner, Eta S.* [Verkkodokumentti].[12.12.2019].Saatavissa: <https://www.springer.com/gp/book/9781441922236>

Bhattacharjee, J. (2017).Some Key Machine Learning Definitions. *Nineleaps* [Verkkodokumentti].[21.1.2020]. Saatavissa: <https://medium.com/technology-nineleaps/some-key-machine-learning-definitions-b524eb6cb48>

Biswas,M., Kuppili, V., Laba,L., Edla, D.R., Suri, H.S, Shauma,A., Chadrado-Godia, E., Laird, J. R, Nicolaidis, M. & Suri, J.S.(2019). Deep learning fully convolution network for lumen characterization in diabetic patients using carotid ultrasound: a tool for stroke risk. *Medical & Biomedical Engineering & Computing.* [online]. [2.12.2019]. Saatavissa: <https://link.springer.com/article/10.1007/s11517-018-1897-x>

Birmingham, P. (2000). Reviewing the Literature. Teoksessa *Researcher's Toolkit: The Complete Guide to Practitioner Research*, 25–40. Toim. David Wilkinson. [Verkkodokumentti]. [27.1.2020]. Saatavissa: <https://epdf.pub/researchers-toolkit-the-complete-guide-to-practitioner-research.html>

Boell, S. K., & D. Cecez-Kecmanovic. (2014). A hermeneutic approach for conducting literature reviews and literature searches. *Communications of the Association for Information Systems*, 34(1), 257286. [Verkkodokumentti]. [1.2.2020]. Saatavissa: http://skb.unifind.de/publications/2014_CAIS-Boell,Cecez-Kecmanovic-A_Hermeneutic_Approach_to_Literature_Reviews.pdf

Brismi, T.S., Xu, T., Wang, T., Dai, W. & I.C Paschalidis.(2018).Predicting diabetes-related hospitalizations based on electronic health records. *Stat Methods Med Res*, 28(12).3667-3682. [online]. [21.1.2020]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/30474497>

Casanova, R, Saldana, S., Simpson, S.L, Lacy, M. E, Subauste, A,R., Blackshear, C, Wagenknecht, L & A,G, Bertoni.(2016). Prediction of Incident Diabetes in Jackson Heart Study Using High-Dimensional Machine Learning Prediction of Incident Diabetes in the Jackson Heart Study Using High-Dimensional Machine Learning. *PLoS one* 2016,11(10):e0163942. [online]. [12.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5058485/>

Chen, H., Yang, B., Liu, D., Liu, W., Liu, Y., Zhang, Xy. & H. Lu. Using Blood Indexes to predict Overweighth Statuses; An Extreme Learning Machine-Based Approach.(2015). *Plos One*.2015, 10(11):eo143003.[online].[20.1.2020].Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4658146/>

Cheung C.Y., Tang, F., Ting, D.S.W., Tan, G.S. W& T.Y. Wong (2019). Artificial Intelligence in Diabetic Eye Disease Screening. *Asia Pac J ophtalmol.(Philox)*. [online]. [2.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/31016915>

Choi, B, G., RHA.,S,W, Kim,S,W, Kang, J,H, Park,J.Y& Y,K. Noh. (2019). Machine Learning for the Prediction of New-Onset Diabetes Mellitus during 5-year Follow-up in Non-Diabetic Patients withCardiovascularRisks. *Yonsei Med J*. 60 191-199.[online]. [19.1.2020]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/30666841>

Cronin, P., Frances R. & C. Michael (2008). Undertaking a Literature Review: A Step-by-Step Approach. *British Journal of Nursing 17: 1*, 38–43. [online]. [12.3.2019]. Saatavissa: https://www.researchgate.net/publication/5454130_Undertaking_a_literature_review_A_step-by-step_approach

Cui, S., Wang, D., Wang, Y., Yu, P-W & J. Yaochu (2018). An improved support vector machine-based diabetic readmission prediction. *Computer Methods Programs*

Biomed.[online]. [3.12.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/30415712>

Debédát J, Sokolovska N, Coupaye M, Panunzi S, Chakaroun R, Genser L, de Turenne G, Bouillot JL, Poitou C, Oppert JM, Blüher M, Stumvoll M, Mingrone G, Ledoux S, Zucker JD, Clément, K.& J. Aron-Wisnewsy (2018).Long-Term Relapse of Type 2Diabetes After Roux-en-y-Gastric Bypass. [online]. [7.12.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/30082327>

Deo, R.C.(2015). Machine Learning in Medicine. *Circulation*, 132(20), 1920-

30.[online]. [4.9.2018]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/26572668>

Dibbern, J., Goles, T., Hirschheim, R., & B. Jayatilaka (2004). Information systems outsourcing: A survey and analysis of the literature. *The DATABASE for Advances in Information Systems*, 35(4), 6-102. [Verkkodokumentti] [30.12.2019].Saatavissa:

https://www.researchgate.net/publication/220627990_Information_systems_outsourcing_A_survey_and_analysis_of_the_literature

Diabetes Care (2019).[Verkkodokumentti] [30.9.2019].Saatavissa:

https://care.diabetesjournals.org/content/diacare/suppl/2018/12/17/42.Supplement_1.DC_1/DC_42_S1_2019_UPDATED.pdf

Dudovskiy, J. (2018). “Pragmatism Research Philosophy”. *Research methodology*.

[Verkkodokumentti]. [21.4.2019]. Saatavissa: [https://research-](https://research-methodology.net/research-philosophy/pragmatism-research-philosophy/)

[methodology.net/research-philosophy/pragmatism-research-philosophy/](https://research-methodology.net/research-philosophy/pragmatism-research-philosophy/)

Duan T., Rajpurkar, P.,Laird, O.,NG, A,Y.& S. Basu (2019). Circ cardiovascular Qual Outcomes 12 (3), e005010. Clinical Value of Predicting Individual Treatment Effects for Intensive Blood pressure Therapy. *Circ cardiovascular Qual Outcomes* 12 (3),

e005010. [online]. [9.1.2020]. Saatavissa: <https://pubmed.ncbi.nlm.nih.gov/30857410-clinical-value-of-predicting-individual-treatment-effects-for-intensive-blood-pressure-therapy/>

Duodecim. Ebmeds. [Verkkodokumntti]. [13.12.2019]. Saatavissa:

<https://www.ebmeds.org/>

Elements of AI :3 Regressio. [Verkkodokumentti]. [8.1.2020]. Saatavissa:

<https://course.elementsofai.com/fi/4/3>

Elements of AI: 3 Naiivi Bayes luokitin. [Verkkodokumentti]. [8.1.2020]. Saatavissa:

<https://course.elementsofai.com/fi/3/3>

El-Sappagh, S., Elmogy, E & A.M Riad (2015). A Fuzzy-ontology-oriented case - based reasoning framework for semantic diabetes diagnosis. *Artific Intell.Med.*65.179-208.[online]. [15.12.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/26303105>

Emont, S. (2011). *Measuring the impact of Patients portals*: What the literature tells us. Oakland, CA: California Health Care Foundation: [Verkkodokumentti].[6.1.2020].

Saatavissa: <https://www.chcf.org/publication/measuring-the-impact-of-patient-portals-what-the-literature-tells-us/>

Eskola, J.(2001). Laadullisen aineiston analyysi vaihe vaiheelta. Yhteiskuntatieteellinen tietoaarkisto. Jyväskylä. [Verkkodokumentti]. [17.1.2020]. 135-140. Saatavissa: https://www.fsd.uta.fi/menetelmaopetus/kvali/L7_3_2.html

Eskola, J & J, Suoranta.(1998). Johdatus tieteelliseen tutkimukseen. Tampere: Vasta. Paino. [Verkkodokumentti]. [17.1.2020]. Saatavissa: https://www.fsd.uta.fi/menetelmaopetus/kvali/L7_3_2.html

Evans, D. (2008). Overview of Methods. Teoksessa: *Reviewing Research Evidence for Nursing Practice: Systematic Reviews*, 137–148. Toim. Christine Webb & Brenda Ross. Oxford: Blackwell Publishing.[Verkkodokumentti]. [5.3.2020]. Saatavissa: https://tohtori.turkuamk.fi/uploads/2020/04/92b18b03-kirjallisuuskatsaus_20.4.20.pdf

Fagella, D. (2018). Machine Learning in HealthCare.Applications-2018 and beyond. *Techmergence* [Verkkodokumentti]. [5.9.2018]. Saatavissa: <https://www.techmergence.com/machine-learning-healthcare-applications/>

Fink, A. (2005). *Conducting Research Literature Reviews: From the Internet to the Paper*. Thousand Oaks: Sage Publications, Inc.2nd edition. [Verkkodokumentti]. [7.1.2020]. Saatavissa: <https://www.worldcat.org/title/conducting-research-literature-reviews-from-the-internet-to-paper/oclc/55947868>

Fink, A. (2010). *Conducting research literature reviews: From the Internet to paper*. Thousand Oaks, CA: Sage.3rd edition.[Verkkodokumentti]. [8.1.2020]. Saatavissa: https://books.google.fi/books/about/Conducting_Research_Literature_Reviews.html?id=2bKI6405TXwC&redir_esc=y

Fitriati, D. & A. Muratako (2016). Implementation of Diabetic Retinopathy screening using realtime data. International Conference on Informatics and Computing. *IEEE*. [online]. [20.1.2020]. Saatavissa: <https://ieeexplore.ieee.org/document/7905715/authors#authors>

Goyal, M., Reeves, N. D, Davisaon, A.K., Rajbhandari,S., Spragg, J. & M.H. Yap. (2018). DFUNet: Convolutional Neural Networks for Diabetic Foot Ulcer classification.*IEEE Transactions on Emerging Topics in Computational Intelligence*.Electronic ISSN: 2471-2850. [online]. [20.1.2020]. Saatavissa: <https://ieeexplore.ieee.org/document/8464076/authors#authors>

Green, B. N., Claire, D., Johnson & A. Adams (2006). Writing Narrative Literature Reviews for Peer Reviewed Journals: Secrets of the Trade. *Journal of Chiropractic Medicine* 5: 3, 101–117. [Verkkodokumentti]. [2.2.2020]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2647067/>

Grewal, P, S., Oloumi, F., Rubin, U., & Tennant, M.T.S.(2018). Deep learning in Ophthalmology. *Can J Ophthamal*,53(4),309-313 Aug 2018.Review article. [online].[3.12]. Saatavissa: <https://pubmed.ncbi.nlm.nih.gov/30119782-deep-learning-in-ophthalmology-a-review/>

Guishan, V., Peng, L., Coran, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venuagopalan, S., Widner, K., Maradas, T., Cuadros, J., Kim, R., Raman, R., Nelson, P., Mega, L. & Webster, D.R (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photography. *Jama* 316(22) 2402-24010. [online]. [9.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/27898976>

Haaparanta, L & Niiniluoto, I. (1986). *Johdatus tieteelliseen ajatteluun*. Helsinki: Helsingin Yliopisto

Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination*. Thousand Oaks, CA: Sage. [Verkkodokumentti]. [3.2.2020]. Saatavissa: https://www.cuzproduces.com/producinganew/files/resources/HART_Doing%20a%20literature%20review_1988_ch1.pdf

Hastie, T., Tibshirani, R., & J. Friedman (2009). The elements of statistical learning: data mining, inference, and prediction.. *Library of Congress Control Number: 2008941148*. [Verkkodokumentti]. [31.1.2020]. Saatavissa: http://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12.pdf

Hermanth, D. J., Anitha, J., Son, L.H., & J. Mittal (2018). Diabetic retinopathy Diagnosis from Retinal Using Modified Hopfield Neural Network. *J Med Syst*. 42(12),247 [online]. [3.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/30382410>

Herttroijs, D.F., Elisen, A.M., Brouwers, M.C.G.J., Scharper, N.C, Köhler, S., Popa, M.C., Asterodis, S., Heendriks, S.H, Bilo, H.J.& Ruwald.D. (2018). A Risk Score Including Body Mass index, Glycated Haemoglobin and Triglycerides Predicting Future Glycaemic Control in People with Type 2 Diabetes. *Obs Metab*. [online]. [9.12.2019]. Saatavissa: https://pubmed.ncbi.nlm.nih.gov/29095564-a-risk-score-including-body-mass-index-glycated-haemoglobin-and-triglycerides-predicts-future-glycaemic-control-in-people-with-type-2-diabetes/?from_term=Machine+learning+and+predicting+outcome+of+diabetic+care+outcome&from_filter=years.2014-2020&from_page=3&from_pos=8

Higgs, J., Jones, M., Loftus, S & N. Christensen (2008). *Clinical Reasoning in the Health Professions*. Elsevier Limited: All rights reserved. [Verkkodokumentti]. [22.1.2020]. Saatavissa: https://books.google.fi/books?id=yxXXLn1Yco4C&q=clinical+decision&dq=clinical+decision+making+definition&lr=&hl=fi&source=gbs_word_cloud_r&cad=4#v=snippet&q=clinical%20decision&f=false

Hirsjärvi, S., Remes, P. & P., Sajavaara (2004). Tutki ja kirjoita. Kustannusosakeyhtiö Tammi. [Verkkodokumentti]. [17.1.2020]. Saatavissa: https://www.fsd.uta.fi/menetelmaopetus/kvali/L7_3_2.html

Hirsjärvi, S., Remes, P. & P., Sajavaara (2008). Tutki ja kirjoita. Kustannusosakeyhtiö Tammi [Verkkodokumentti]. [17.1.2020]. Saatavissa:

https://www.fsd.uta.fi/menetelmaopetus/kvali/L7_3_2.html

Hsin-Yi. T, Chan. P-Y & E, Chia-Yu Su. (2018). Predicting Diabetic Retinopathy and Identifying Interpretable Biomedical Features Using Machine Learning Algorithms. *BMC Bioinformatics*.2018, 19(suppl 9): 283.[online]. [11.12.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6101083/>

Hu, X., Reaven, A., Saremi, N., Liu, M,A, Abbasi, H, Liu & R,Q. Migrino (2016). Machine learning to predict rapid progression of carotid atherosclerosis in patients with impaired glucose tolerance. *Erasip j Bionform Syst Biol*, 2016 (1), 14.[online].

[19.12.2020]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/27642290>

Huiling G, C, Yang, B., Liu, D., Wenbin, L., Liu, Y., Zhang, X. & H. Lufeng (2015). Using Blood Indexes to Predict Overweight Statuses: An Extreme Learning Machine-Based. *Approach.Plos One*,10(11).[online]. [19.1.2020]. Saatavissa:

https://pubmed.ncbi.nlm.nih.gov/26600199-using-blood-indexes-to-predict-overweight-statuses-an-extreme-learning-machine-based-approach/?from_single_result=Chen+2015+extreme+learning+machine+and+predicting+over+weight

Hwang, K. &M. Chen (2017). Machine Learning for Big Data in Healthcare Applications. Teoksessa: Big Data Analytics for Cloud, IoT and Cognitive Computing learning. UK. *John Wiley & Sons Ltd.* [Verkkodokumentti]. [3.2.2020]. Saatavissa:

https://books.google.fi/books?id=uzdZDgAAQBAJ&pg=PP7&lpg=PP7&dq=Big+Data+Analytics+for+Cloud,+IoT+and+Cognitive+Computing.+UK.+John+Wiley+%26+Sons+Ltd.&source=bl&ots=aA4LNY0dkG&sig=ACfU3U1u0Oi1rUgnqFhGv2wVB_SQC5nyHQ&hl=fi&sa=X&ved=2ahUKEwid457Ru8LnAhXupIsKHaP5ChkQ6AEwAXoEC_AoQAQ#v=onepage&q=Big%20Data%20Analytics%20for%20Cloud%2C%20IoT%20and%20Cognitive%20Computing.%20UK.%20John%20Wiley%20%26%20Sons%20Ltd.&f=false

Häyrynen, K.(2006).Sosiaali ja terveydenhuollon tietotekniikan ja tiedonhallinnan tutkimuspäivät.Päätöksenteontutkimus.*Julkari*. [Verkkodokumentti]. [22.3.2019].

35-40. Saatavissa: <http://www.stakes.fi/verkojulkaisut/tyopaperit/T18-2006-VERKKO.pdf>

Jasperson, J., Carte, T., Saunders, C., Butler, B., Croes, H., & W. Zheng (2002). Review: Power and information technology research: A metatriangulation review. *MIS Quarterly*, 26(4), 397-459. [Verkkodokumentti]. [22.3.2020]. 35-40. Saatavissa:

<https://misq.org/review-power-and-information-technology-research-a-metatriangulation-review.html>

Jauhiainen, J-P.(2019) *Koneoppiminen- Ohjattu Oppiminen taloustieteellisessä tutkimuksessa*. Oulu: Oulun yliopisto. Oulun yliopiston kauppa- ja korkeakoulu. [online].

[31.1.2020]. Pro gradu-tutkielma. Saatavissa: <http://jultika.oulu.fi/files/nbnfioulu-201906052387.pdf>

James, G., Witten, D., Hastie, T., & R. Tibshirani (2013). An introduction to statistical learning. *Springer*. [online]. [2.2.2020]. Saatavissa:

<https://www.bcf.usc.edu/~gareth/ISL/ISLR%20First%20Printing.pdf>

John S., Ram, K., Sivapsakasan, M., & R. Raman (2016). Assessment of Computer-Assisted Screening Technology for Diabetic Retinopathy Screening in India- Preliminary Results and Recommendations from a Pilot Study. *Stud Health Technol Inform* 231, 74-81. [online]. [3.3.2020]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/27782018>

Institute of Medicine Committee on Quality of Health Care in America. (2001).

[Verkkodokumentti]. [30.9.2019]. Saatavissa:

<https://search.ahrq.gov/search?q=Diabetes>

Iiro (2017). Opinnäytetyön rakenne. [Verkkodokumentti]. [15.9.2019]. Saatavissa:

<https://www.scribbr.fi/opinnaytetyon-rakenne/kirjallisuuskatsaus-opinnaytetyo/>

Kalyankar, G.D., Shivanande, R.P., & N.V. Dharwakar (2017). Predictive analysis of diabetic patient data using machine learning and Hadoop. *IEEE*. [online]. [12.1.2020].

Saatavissa: <https://ieeexplore.ieee.org/abstract/document/8058253>

Kanchan, D. B. & M. M. Kishor. (2016). Study of machine learning algorithms for special disease prediction using principal of component analysis. *IEEE*. [online].

[21.1.2020]. Saatavissa:

<https://ieeexplore.ieee.org/abstract/document/7955260?anchor=references>

Kavaktosis, I., Tsave, O., Salifoglor, A., Maglaveras, N., Vlahavas, I. & J. Chowarde (2017). Machine learning and Data Mining Methods in Diabetes Research. *Comput Struct Biotechnical J*. 15:104-116. [online]. [8.12.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/28138367>

Khan, N. S., Muaz, M. H., Kabir, A., & N. M. Islam (2017). Diabetes Predicting mHealth Application Using Machine Learning. *IEEE*. [online]. [22.1.2020]. Saatavissa:

<https://ieeexplore.ieee.org/document/8468885/authors#authors>

Khanji C., Lalonde, L. & M. E. Schnitzer (2019). Lasso Regression for the Prediction of Intermediate Outcomes Related to cardiovascular Disease Prevention Using the Transit Quality Indicators Randomized controlled trial. *Medical Care*. [online]. [2.1.2020].

Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/30439793>

Koivuranta-Vaara, P. (2011). Terveystieteiden laatuopas. *Kuntaliitto*.

[Verkkodokumentti]. [15.3.2019]. Saatavissa:

https://www.google.fi/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=2ahUKEwiIydWkpfrgAhWsw8QBHVLVCEEQFjACegQIChAC&url=http%3A%2F%2Fshop.kuntaliitto.fi%2Fdownload.php%3Ffilename%3Duploads%2Fterveydenhuollon_laatuopas.pdf&usg=AOvVaw0ahOV85jOcsIWtGxEW1Nxl

Kohli, R., & V. Grover (2008). Business value of IT: An essay on expanding research directions to keep up with the times. *Journal of the Association for Information Systems*, (9)1, pp. 23-39. [Verkkodokumentti]. [25.10.2019]. Saatavissa

https://www.researchgate.net/publication/220580471_Business_Value_of_IT_An_Essay_on_Expanding_Research_Directions_to_Keep_up_with_the_Times

Kwon, J. Y., Karim, M.E., Topaz, M. & Currie, L.M. (2019). Nurses “Seeing Forest for the Trees” in the Age of Machine Learning Using Nursing Knowledge to Improve Relevance and Performance.(2019). [online]. [10.12]. Saatavissa.

<https://pubmed.ncbi.nlm.nih.gov/30688670-nurses-seeing-forest-for-the-trees-in-the-age-of-machine-learning-using-nursing-knowledge-to-improve-relevance-and-performance/>

Labhade, J.D, Chouthmol, L,K & S. Oeshmukh (2016).Diabetic retinopathy detection using soft computing techniques. *IEEE*. [online]. [19.1.2020]. Saatavissa:

<https://ieeexplore.ieee.org/document/7877573>

Lebech Cichosz, S., Denker Johansen, M & O. Hejlesen (2015). Towards big data analytics in Management of Diabetes and Its Complications. *Journals of Diabetes Science and Technology*. [online]. [20.1.2020]. Saatavissa:

<https://journals.sagepub.com/doi/10.1177/1932296815611680>

LePine, J. A., & A. Wilcox-King (2010). Editor’s comments: Developing novel theoretical insights from reviews of existing theory and research. *Academy of Management Review*, 35(4), 506-509. [Verkkodokumentti]. [8.2.2020]. Saatavissa:

<https://asu.pure.elsevier.com/en/publications/editors-comments-developing-novel-theoretical-insight-from-review>

LO-Ciganic, W.-H., Donohue, J, M., Thorope, J,M, Perera, S., Thorope, C., Marcum, Z, A. & W.F Gellad (2015). Using Machine Learning to Examine Medication Adherence Theresholds and Risk of Hospitalization.*Med Care* 2015 Aug, 53,(8):720-728.[online]. [3.2.2020]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4503478/>

Magoulas, G.D. & A. Prentza.(2015).Machine Learning in Medical Applications.Department of Informatics, University of Athens, GR-1574 84.Athens.Greece. Department of Electrical and Computer Engineering National Technical University of Athens, GR-15773 Athens, Greece. Department of Informatics. *University of Athens, GR-15784 Athens, Greece*. [online]. [10.3.2019]. Saatavissa: http://www.dcs.bbk.ac.uk/~gmagoulas/ACAI99_workshop.pdf

Maniruzzaman, Md., Kumar, N., Abedin, Md, M., Islam, S, MD., Suri, M,S., El-Baz, A, S., & J.S. Suri (2017). Compative approaches for classification diabetes mellitus data:Machine learning paradigm: *Comput Methods Programs Biomed*, 152, 23-34.2017. [online]. [7.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/29054258>

Martinez-Vernon, A,S., Covington, J,A., Arasarndnam, R,P., Esfahani, S.,O’Conell, N.& R,S. Savage.(2018) An improved machine learning pipeline for urinary volatiles disease detection: Diagnosing diabetes. *Published online 2018.Sep27*. [online]. [19.1.2020]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6160042/>

Moreno, E. M, Lujan, M.J, Fernandez, P. J, Manrique, P.N, Trivino, C.A, Miquel, M.P, Rodriquez, M.A & M.J. Burquiclos (2017). Type 2 Diabetes Screening Test by Means of a Pulse Oximeter. *IEEE.Trans Bio Med*.[online]. [12.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/28113188>

Mulay, P., Joshi, R. & A. Chaudari(2018). Mapping of Six Sigma to Thershold Based Incremental Clustering Algorithm. *IEEE*. [online]. [4.2.2020]. Saatavissa:

<https://ieeexplore.ieee.org/document/8745432>

Normahani, P., Agrawal, R., Bravis, V., Falinska, A., Bloomfield, L., Mehar, Z., Gaulton, D., Sangster, A., Ackle, T., Gamm, C., Aslam, M., Stanfield, N, J & J. Usmann. (2019). Effectiveness of a multicentre training programme to teach point-of-care vascular ultrasound for detection peripheral arterial disease in people with diabetes. *J Foot Ankle Res*(2018) 11.41.[online]. [20.1.2020]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6048877/>

Obermeyr, Z., Phil, M. & J.E. Ezkiel (2016). Big Data, Machine Learning, and Clinical Medicine. *N.Eng. J.Med.*375(13):1216-1219.[online]. [11.3.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5070532/>

Okoli, C., & K. Schabram (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts. Concordia University, Canada*.[online].

[21.1.2020]. Saatavissa: http://aisel.aisnet.org/sprouts_all/348.

Orlando, J.,I, Prokofyeva,E., del Fresno, M.& B. Blaschko (2018). An ensamble deep learning based approach for red lesion detection in fundus images.*Computer Methods and Progress in Biomedicine*. [online]. [11.12.2019]. Saatavissa:

<https://www.sciencedirect.com/science/article/abs/pii/S0169260717307897?via%3Dihub>

PubMed. [online]. [3.3.2019]. Saatavissa:<https://www.ncbi.nlm.nih.gov/pubmed/>

PubMed.Health. [online]. [5.3.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3133896/>

Prahs,.P., Radek, V., Mayer, C., Cvetkova, Y., Cvetkova, N., Helbig, H.& D. Märker (2018). OCT-based Deep Learning Algorithm for the Evaluation of Treatment Indication with Anti-Vascular Endothelial Growth Factor Medications.*Grafes Arch Clin Exophtalmol* 256(1),91-98 jan 2018.[online]. [7.12.2019]. Saatavissa:

<https://pubmed.ncbi.nlm.nih.gov/29127485-oct-based-deep-learning-algorithm-for-the-evaluation-of-treatment-indication-with-anti-vascular-endothelial-growth-factor-medications/>

Pyle, D. & S. J. Cristina (2018). An Excutive Guide to Machine Learning. Cognitive Computing Consortium. [online]. [3.9.2018]. Saatavissa:

<https://cognitivecomputingconsortium.com/an-executives-guide-to-machine-learning/>

Qiu, L., Bhattarya, P.& P. Tuan (2017). Battling Diabetes through Food photography: An Image-based Utility Maximimization Framework for Diet Diagnostics ICIS 2017 Proceedings. *AIS eLibrary*. [online]. [8.2.2020]. Saatavissa:

<https://aisel.aisnet.org/icis2017/IT-and-Healthcare/Presentations/23/>

Ravindran, P. (2016). What is the difference between machine learning and ML algorithm. [online]. [21.2.2020]. Saatavissa: <https://www.quora.com/What-is-the-difference-between-machine-learning-model-and-ML-algorithm/answer/Prashanth-Ravindran?srid=dZyb>

- Ren, F., Cao, P., Zhao., D & C. Wan (2018). Diabetic macular edema grading in retinal images using vector quantification and semisupervised learning. *Technol Health Care* 2018, 26(Supp1); 389-397.[online]. [5.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6004946/>
- Roberts, N., Galluch, P., Dinger, M., & V. Grover (2012). Absorptive capacity and information systems research: Review, synthesis, and directions for future research. *MIS Quarterly*, 36(2), 625-648. [Verkkodokumentti]. [15.1.2020]. Saatavissa: <https://pdfs.semanticscholar.org/274b/fcc1a3de7717423142595ddbe8129f9b91bb.pdf>
- Rodeiquez-Romero.,V, Bergstrom.,R.,F, Decker, B.,S, Lahu, G., Vakilynejad, M. & R.R. Bies. Prediction of Nephropathy in Type 2 Diabetes.(2019). *Clin Transl Sci*, 12 (5), 519-528. Sep 2019.[online]. [13.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/31112000>
- Roos, T. (2016). Johdatus Tekoälyyn. Helsingin Yliopisto.(31-44).[online]. [30.8.2018]. Saatavissa: <https://www.cs.helsinki.fi/u/ttonteri/ai/2017/luento6.pdf>
- Rowe F., (2014). What a literature review is not: Diversity, boundaries and recommendations. *European Journal of Information Systems*, 23(3), 240-250. [Verkkodokumentti]. [13.1.2020]. Saatavissa: <https://www.tandfonline.com/doi/full/10.1057/ejis.2014.7>
- Ruokoniemi, P.& P. Rannanheimo (2018). Data ja tekoäly muuttuvat lääkealaolemmeko valmiita? *Sic! Lääketietoa Fimeasta*.2018:2. [online]. [12.9.2019]. Saatavissa: <http://sic.fimea.fi/verkkolehdet/2018/2018/laakkeet-ja-digitalisaatio-1.0/data-ja-tekoaly-muuttavat-laakealaa-olemmeko-valmiita->
- Rutstein, D.D., Berenberg, W., Chamer, C,Y., Child, C, G., Fishuan, A.,P., Perrin, E, B., Feldman, J, J., Leaverton, P, E., Lane, M. J., Spencer, D.Y.(1976) C.C. Evans (1976). Measuring the Quality of Medical Care -A Clinical Method. *N.Engl.JMed* 1976:294:582-588. [online]. [30.8.2018]. Saatavissa: <https://www.nejm.org/doi/full/10.1056/NEJM197603112941104>
- Russel, C.L.(2005). An overview of the integrative literature review. *Progress in Transplantation*, March. [online]. [15.1.2019]. Saatavissa: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.472.8890&rep=rep1&type=pdf>
- Saba, L., Biswas, M., Harman, S,S., Viskovic, K., Laird, J,R., Guardrado-Godia, E., Nicolalaides, A., Khanna, N, N., Viswanthan, V. & S.S. Sun (2018).Ultrasound-Based Carotid stenosis measurement and risk stratification in diabetic cohort: A deep learning paradigm. *Cardiovasc diagnose therapy* 2019.Oct; 9(5):439-461.[online]. [19.1.2020]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6837906/>
- Salminen, A.(2011) *Mikä kirjallisuuskatsaus*. Vaasan yliopiston julkaisuja. Vaasa. [Verkkodokumentti]. [9.3.2019]. Saatavissa: https://www.univaasa.fi/materiaali/pdf/isbn_978-952-476-349-3.pdf

- Serener, A. & S. Serte (2019) Dry and Wet Age-Related Macular Degeneration Classification Using OCT Images and Deep Learning. *IEEE*. [online]. [22.1.2020]. Saatavissa: <https://ieeexplore.ieee.org/document/8741768>
- Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B, S., Waldstein, S., M., & H. Bogunovic (2018). Artificial intelligence in Retina. Retin Eye Res. 2018. *Prog Retin Eye Res*, 67, 1-29 nov 2019. [online]. [4.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/30076935>
- Schyren, G.(2015). Writing Qualitative IS Literature Reviews- Guidelines Synthesis, Interpretation and Guidance of Research. University of Regensburg. Communications of the Association for Information Systems. [Verkkodokumentti]. [2.3.2020]. 290-320. Saatavissa: <https://aisel.aisnet.org/cais/vol37/iss1/12/>
- Schroderus, R. (2019). Kuinka hyvä on tarpeeksi hyvä luokittelumalli. [Verkkodokumentti]. [1.5.2020]. Saatavissa:<https://lamia.fi/blog/hyva-luokittelumalli1>
- Schwartz, W, B.(1970). Medicine and the computer —The Promise and Problems of Change. Special Article. *NEngl J Med* 1970;283:1257-1264. [online]. [7.12.2020]. Saatavissa: <https://www.nejm.org/doi/full/10.1056/NEJM197012032832305>
- Schwarz, A., Mehta, M., Johnson, N., & W. Chin (2007). Understanding frameworks and reviews: A commentary to assist us in moving our field forward by analyzing our past. *Database*, 38(3), 2950. [Verkkodokumentti]. [12.2.2020]. Saatavissa: <https://dl.acm.org/doi/10.1145/1278253.1278259>
- Starr, D. (2018). Current use cases for Machine Learning in healthcare. *Microsoft Azure*. [online]. [16.8.2018]. Saatavissa: <https://azure.microsoft.com/en-us/blog/current-use-cases-for-machine-learning-in-healthcare/>
- Sun, Y.& D., Zhang (2019). Diagnosis and Analysis of Diabetic Retinopathy Based on Electronic Health Records. *IEEE*. [online]. [18.1.2020]. Saatavissa: <https://ieeexplore.ieee.org/abstract/document/8721113>
- Sutton, R. I., & B.M. Staw (1995). What theory is not. *Administrative Science Quarterly*, 40(3), 371-384. [Verkkodokumentti]. [8.1.2020]. Saatavissa: <https://www.jstor.org/stable/2393788?seq=1>
- Tamhane, S., Rodriquez-Gutierrez, R., Hargraves, I & V.M Montori (2015). Shared decision-making in diabetes care. *Cual Diab Ren*, 15(12), 112. [online]. [12.9.2019]. Saatavissa: <https://pubmed.ncbi.nlm.nih.gov/26458383-shared-decision-making-in-diabetes-care/>
- Thomas, M.A., Abraham, S.D & D. Liu (2018). Federated Machine Learning for Translational Research. Amics. *AIS eLibrary*. [online]. [21.1.2020]. Saatavissa: <https://aisel.aisnet.org/amcis2018/Health/Presentations/34/>
- Ting, D.S.W, Peng, L., Varadrajana, A. V., Keane, P. A., Burlina, P.M., Chiang, M, F., Schemetter, L., Pasquale, L, R., Bressler, N. M., Webster, D.R. , Abramoff, M.& T.Y. Wong (2018). Deep learning in Ophthalmology:The technical and clinical consideration. *Prog Retin Eye Res*. [online]. [2.12.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/31048019>

Ting, D.S.W, Cheung, C.Y.L, Lin, G, Tan, G.S.W., Quang, N.D., Gan, A., Hazmah, H., Franco-Carcia, R., San Yeo, I.Y., Yen Lee, S., Wong, E.Y.I., Sabranayagam, C., Baskaran, M., Ibrahim, F., Tan, N.C., Finkelstein, E.A., Lamoureux, E., Wong, I.Y., Bressler, N.M., Sivosirasad, S., Varma, R., Jonas, J.B., He, M.G., Cheng, C.Y, Cheung, G.R.M., Aung, T., Hsu, W., Li Lee, M. & T.Y. Wong (2017). Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations with Diabetes. *Jama* 318(22), 2211-2223.[online]. [5.12.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/29234807>

Tjoa, S. (2013). Introduction to Deep Learning. Standaford.edu. [online]. [3.4.2020]. Saatavissa:

https://ccrma.stanford.edu/workshops/mir2013/CCRMA_MIR2013_DBN.pdf

Torok, Z., Peto, T., Tukacs, E., Molnac, A.M., Berta, A. Tozeen, J, Haydu, A.,Navy, V., Domokos, B. & A. Csutak.(2016). Combined Methods for Diabetic Retinopathy Screening, Using Retina Photography and Tear Fluid Proteomics Biomarkers. *J Diabetes Res* 2015, 673619.[online]. [13.12.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/26221613>

Torvinen, P. (2016).Diabetes tutkimus: Diabetesepidemia maksaa yli 720 miljardia euroa vuodessa. *Mediuutiset*. [online]. [20.3.2019]. Saatavissa:

<https://www.medi uutiset.fi/uutiset/tutkimus-diabetesepidemia-maksaa-yli-720-miljardia-euroa-vuodessa/0cbe4c6f-d9bb-3890-97c9-8f086f9b226c>

Torraco, Richard J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review* 4: 3, 356–367. [online]. [11.3.2019]. Saatavissa:

<https://journals.sagepub.com/doi/10.1177/1534484305278283>

Tsao, H.Y, Chan P.V &E.C. Su (2018). Predicting Diabetic Retinopathy and Identifying Intrepretable Biomedicer Features Using Machine Learning Algorithms. *BMC Bioinformatics*, 19 (suppl9), 283. [online]. [15.3.2019]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/30367589>

Tuomi, J & A. Sarajärvi.(2002). Laadullinen tutkimus ja sisällöllinen analyysi. Helsinki: Tammi: 95-99. [Verkkodokumentti]. [17.1.2020]. Saatavissa:

https://www.fsd.uta.fi/menetelmaopetus/kvali/L7_3_2.html

Tuomi, J. (2007). *Tutki ja lue: Johdatus tieteilisen tekstin ymmärtämiseen*. Helsinki: Kustannusosakeyhtiö Tammi: ISBN: 978-951-26-5460-4. [Verkkodokumentti].

[18.1.2020]. Saatavissa: <https://jyu.finna.fi/Record/jykdok.1034804>

Tyran, C. K., & M. Shepherd (2001). Collaborative technology in the classroom: A review of the GSS research and a research framework. *Information Technology and Management*, 2(4), 395-418. [Verkkodokumentti]. [13.1.2020]. Saatavissa:

<https://link.springer.com/article/10.1023/A:1011450617798>

Valvira(2015).Tietojärjestelmät. [Verkkodokumentti]. [10.9]. Saatavissa:

https://www.valvira.fi/terveydenhuolto/terveysteknologia/tuotteen_markkinoille_saattaminen/tietojarjestelmat

Vane, L. (2016). *Sairaalan tietojärjestelmät*. Helsinki: Helsingin yliopisto. Tietojen käsittelytieteen laitos. Seminaarityö. [Verkkodokumentti]. [9.1.2020]. Saatavissa:

<https://docplayer.fi/22480909-Sairaalan-tietojarjestelmat.html>

Van Wart, M. (2003). Public-Sector Leadership Theory: An Assessment. *Public Administration Review* 63: 2, 214–228

Vehi J., Contreas, I., Ovido, S., Bialgi, L. & A, Bertachi (2019). Prediction and Prevention of Hydroglycaemic Events in Type-1 Diabetic Patients Using Machine Learning. *Health Informatics*. 2019 jun13. [online]. [19.1.2020]. Saatavissa:

<https://www.ncbi.nlm.nih.gov/pubmed/31195880>

Venetresca M.J. & J.W Mohr (2017). Archival Research Methods. *Wiley Online Library*. [online]. [11.9.2019]. Saatavissa:

<https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781405164061.ch35>

Vähäkainu, P. & P. Neittaanmäki (2017). *Tekoäly terveydenhuollossa*.

[Verkkodokumentti]. [17.9.2018]. Jyväskylä: Jyväskylän yliopisto. Saatavissa:

<https://www.jyu.fi/it/fi/tutkimus/julkaisut/tekes-raportteja/tekoaly-terveydenhuollossa.pdf>

Wager, K.A., Lee, F., W. & J.P. Glaser (2017). Jossey-Bass. *Health Care Information Systems*. [Verkkodokumentti]. [12.1.2020]. Saatavissa: <https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/reader.action?docID=4815068>

Webster, J. & Watson, R.T. Analyzing The Past to Prepare For The Future: Writing a Literature Review. (2002). *MIS Quarterly* Vol. 26 No.2. June 2002.

[Verkkodokumentti]. [3.2.2020]. Saatavissa:

<https://www.jstor.org/stable/4132319?seq=1>

Winter, A., Haux, R., Ammenwerth, E., Brigl, B., Hellrung, N. & Jahn, F. (2011). *Health Information Systems*. Springer. 2nd Edition. [Verkkodokumentti]. [3.1.2020].

Saatavissa: <https://www.springer.com/gp/book/9781849964401>

Wiejssignhe, I., Gamage, L., Perera, J. & Chitaranjan, C. (2019). A Smart Telemedicine System with Deep Learning to Manage Diabetic Retinopathy and Foot Ulcers.

IEEE. [online]. [14.11.2019]. Saatavissa: <https://ieeexplore.ieee.org/document/8818682>

Whittemore, R. & K. Kanlf. (2005). Methodological issues in nursing research: the integrative review: updated methodology. *Journal of Advanced Nursing* 52(2), 546-553. [Verkkodokumentti]. [14.1.2020]. Saatavissa:

<https://pubmed.ncbi.nlm.nih.gov/16268861/>

Whittemore, R. (2008). Rigour in Integrative Reviews. Teoksessa *Reviewing Research Evidence for Nursing Practice: Systematic Reviews*, 149–156. Toim. Christine Webb & Brenda Ross. Oxford: Blackwell Publishing. [Verkkodokumentti]. [5.1.2020].

Saatavissa: <https://www.sciencedirect.com/science/article/abs/pii/S0147956314001897>

- WHO. Cardiovascular diseases (CVDs). [Verkkodokumentti]. [14.5.2019]. Saatavissa: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- WHO. Diabetes mellitus. [Verkkodokumentti]. [Viitattu 17.3.2019]. Saatavissa: <https://translate.google.fi/translate?hl=fi&sl=en&u=https://www.who.int/mediacentre/factsheets/fs138/en/&prev=search>
- WHO. Quality of Care. [Verkkodokumentti]. [Viitattu 2.9.2018]. Saatavissa: http://www.who.int/maternal_child_adolescent/topics/quality-of-care/en/
- Woldargay, A. Z., Årsand, E., Botsis, T., Albery, D., Mamykina, L. & G. Hartgvisen (2019). Data-Driven Blood Glucose Pattern, Classification and Anomalies Detection: Machine Learning Applications in Type 1 Diabetes. [online]. [17.1.2019]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6658321/>
- Wolfswinkel, J. F., Furtmueller, E., & C. P. M. Wilderom (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1), 45-55. [Verkkodokumentti]. [6.2.2020]. Saatavissa: <https://link.springer.com/article/10.1057/ejis.2011.51>
- Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., & Z.H Zhou (2008). Top 10 algorithms in data mining. *Knowledge and information systems*, 14(1), 1-37. [Verkkodokumentti]. [6.2.2020]. Saatavissa: <https://link.springer.com/article/10.1007/s10115-007-0114-2>
- Zeevi, D., Kovem, T., Zmora, N., Israeli, D., Rotschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnil-Sagi, T., Lotan-Pomper, M., Suez, J., Ali Mahdi, J., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Kosower, N., Rein, M., Zillerman-Schapira, G., Dohnalova, L., Pereoner-Fischer, M., Bikovsky, R., Halper, Z., Elinav, E. & E. Segal (2015). Personalized Nutrition by Prediction of Glycemic Responses. *Randomized controlled trial* 163(5), 1079-1094. [online]. [6.2]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pubmed/26590418>
- Zhang, W., Yu, Q., Siddique, B., Divarkaran, A. & N., Sawhney. (2015). "Snap-n-Eat" Food recognition and Nutrition Estimation on a Smartphone". *J. Diabetes Sci Technol* 2015 May, 9(3): 525-533. [online]. [19.1]. Saatavissa: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4604540/>
- Zou, K. H., O'Malley, A. J., & L. Mauri (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115(5), 654-657. [Verkkodokumentti]. [18.2.2020]. Saatavissa: <https://pubmed.ncbi.nlm.nih.gov/17283280/>
- Zorn, T., & N. Campbell (2006). Improving the writing of literature reviews. *Business Communication Quarterly*, 69(2), 172-183. [Verkkodokumentti]. [11.2.2020]. Saatavissa: <https://journals.sagepub.com/doi/10.1177/1080569906287960>
- Zsolt, T., Peto, T., Csoosz, E., Tukacs, E., Molnar, A.M., Berta, A., Tozser, J., Hajdu, A. Nagy, V., Donokos, B. & A. Csutak. (2015). Combined Methods for Diabetic Retinopathy Screening Using Retina Photographs and Tech Fluid Protomics

Biomarkers. *Journal of Diabetes Research*. [online]. [11.12.2019]. Saatavissa: <https://new.hindawi.com/journals/jdr/2015/623619/>