An Extreme Learning Machine Model for Venue Presence Detection

Author(s): Khan, Wiqar; Raza, Asif; Kuusniemi, Heidi; Elmusrati, Mohammed; Espinosa-Leal, Leonardo

Title: An Extreme Learning Machine Model for Venue Presence Detection

Year: 2023

Version: Accepted manuscript

Please cite the original version:

Khan, W., Raza, A., Kuusniemi, H., Elmusrati, M. & Espinosa-Leal, L. (2023). An Extreme Learning Machine Model for Venue Presence Detection. In: Björk, K.-M. (ed.) *Proceedings of ELM 2021: Theory, Algorithms and Applications*, 144–151. Proceedings in adaptation, learning and optimization, vol. 16. Cham: Springer. https://doi.org/10.1007/978-3-031-21678-7_14

A Extreme Learning Machine Modeling for Venue Presence Detection

Wiqar Khan[1[0000−0002−5350−8663]], Asif Raza[2], Heidi Kuusniemi[3[0000−0002−7551−9531]], Mohammed Elmusrati[3], and Leonardo Espinosa-Leal[2[0000−0001−6861−8024]]

[1] Cloud and Network Services Nokia, Espoo, 02610, Finland,
`Wiqar.Khan@Nokia.com`

[2] Department of Business Management and Analytics, Arcada University of Applied Sciences, Jan-Magnus Janssons plats 1, 00560 Helsinki, Finland
{`leonardo.espinosaleal`,`asif.raza`}`@arcada.fi`

[3] University of Vaasa, Vaasa, Finland
{`Heidi.Kuusniemi`,`Mohammed.Elmusrati`}`@uwasa`

Abstract. Value-added services allocation or denial in a particular venue for a given user is of high significance. It will get more prominent as we move to 5G and 6G networks' roll out as we will get other means to have better aids. In this paper, Extreme Learning Machines (ELM) model performance is compared with Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Logistic Regression (LR), and Random Forest (RF) models for venue presence detection. The input data is collected from the number of UEs (User Equipment) simultaneously placed inside and outside a venue and kept longer. UEs logs essential data such as received reference signal received power for serving cells and neighbor candidate cells, as others. Our findings show that all models perform above 95% for a count of zero, one, and two neighbors.

Keywords: Venue Presence Detection · Equipment logs · Extreme Learning Machines · LTE

1 Introduction

User location in a mobile network has always been of high interest for different use cases. A critical aspect of it has been the value-added services being enabled for the users by the venue owner [10]. Indoor positioning has been more demanding because of the multi-path propagation of communication signals. Most of the venues are considered indoor; therefore, indoor positioning is the subject of interest. Different techniques have been deployed for positioning in indoor facilities. Such techniques range from indoor satellite techniques [12] to mobile and WIFI networks-based techniques. A good survey about selected indoor positioning is presented in [7].

As we move towards 5G and 6G networks, we will have more use cases for the in-door positioning, and the interest level in increasing the accuracy will get

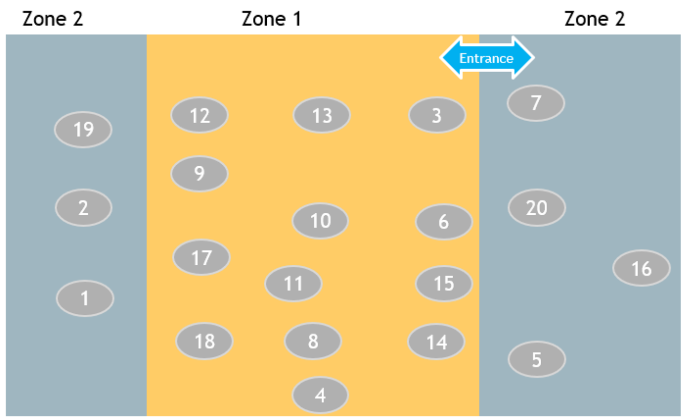

Fig. 1. Scheme of the Venue and the User Equipment distribution.

higher. In 5G and 6G Networks, we will further mean to have better aids. After such advancements, the use cases governed by approaches mentioned in [10] will be more critical than before. The deployment of machine learning algorithms with different conditions and deliberation is the way forward with these in mind.

The data is collected from 20 User Equipments (UEs) for a prolonged duration. The details are mentioned in [11]. On the collected data, we have deployed Extreme Learning Machine (ELM) [9, 6] and compared the obtained results with other machine learning methods.

2 Analysis

This section will cover the core concepts of RSRP (Reference Signal Received Power), PCI (Physical Cell Identity) [4], serving cell and neighboring cells in LTE (Long Term Evolution) network [14]. It also defines the data set used and the cleaning process applied in this paper.

2.1 LTE Network received data set

In an LTE network, there is a serving cell where the UE is connected and served by a 'serving cell' while there are candidate cells called 'neighboring cells' that the UE would connect based on the needs defined per network configurations. In case of need fulfilling the criteria, there could be a switch over from the current serving cell to a neighboring cell hence a more suitable neighboring cell will turn into the serving cell while the serving cell will be considered one of the neighboring cells unless there is again a possible switch over.

Any cell is identified by its Physical Cell Identify called PCI. When a UE scans in an LTE network, it records different parameters' values for available PCIs. The main parameter used for cell selection is RSRP. The RSRP [1] values are obtained for PCIs of the mobile network using the LTE network. There are comprehensive studies available regarding Machine Learning for WIFI indoor positioning, e.g., the authors in [5]. However, we are using RSRPs of the LTE network, and our studied case consists of indoor and outdoor UEs in the same case study as one side of the venue walls to outdoor where UEs were placed for data collection (see Fig.1).

In this work, we have considered neighboring cells as well for our ML algorithms as they provide valuable info about the location where we could get given neighboring cells RSRPs in specific ranges. We use the term neighboring cell depth as the indicator of how many neighboring cells are taken into use and its RSRPs logged by each UE. For our comparison work, we have taken iterations for:

- No neighboring cell (*zero*), i.e., only serving cell and its RSRP data,
- one neighboring cell (*one*) in addition to the serving cell along with respective RSRPs, and
- two neighboring cells (*two*) in addition to the serving cell along with respective RSRPs.

Each UE has its own data logging, and the number of samples collected varies for each UE. It is important to mention that each UE is static at its position during the whole duration of the data collection. The variation of signal levels varies over time based on the mobility around the UEs. The data collection is done in a real-life environment, i.e., when people are moving around, UEs are placed there for a sustained longer duration of around 8-10 hours.

3 Material and Methods

3.1 ELM

Extreme Learning Machines (ELM) [2] is an emergent machine learning technique, which is used for both classification and regression problems. In this paper, ELM will be compared against the traditional machine learning approaches (KNN, LR, SVM, RF) to compare the model test time and accuracy. ELM is well suited for a more extensive data set. In the study [2] author mentioned that "Unlike traditional learning theories and learning algorithms, ELM theories show that hidden neurons need not be tuned in learning and their parameters can be independent of the training data." The ELM architecture benefits from model structure selection and regularization, overfitting, and handing imbalance data.

This study aims to observe the impact of the number of neighbors inclusion in the performance of ELM and other machine learning techniques. Calculations were performed using the state-of-the-art python library for machine learning *scikit-learn* [13], and a new and python compatible extreme learning machine library named *scikit-ELM* [3]. Moreover, with scikit-ELM, we extensively searched the optimal hyperparameters using a randomized gridsearch.

3.2 Parameters used for ELM

The range of parameters used for the ELM to identify the best parameters on which the model produces the best result are consigned in Table 1. The respective ranges/possible values are given for each parameter.

Table 1. ELM parameters and its ranges

ELM parameters	Ranges
alpha	0.00001- 1000
ufunc	'tanh', 'sigm', 'relu', 'lin'
density	0.01, 0.99
No. of neurons	5-400

The optimized parameter *alpha* is the coefficient for the regularization term L^2, *ufunc* the transformation function applied to hidden layer output for adding non-linearity to the algorithm, *density* is the term that controls the density of the random connection weights between input and hidden layers.

4 Results and Discussion

4.1 Comparison of ELM against ML algorithms

The detailed comparison result of ELM against conventional ML algorithms is compiled in Table 2. ELM performance is relative. Model test time mentioned in the Table 2 shows that ELM performed better than SVM and KNN. Whereas results indicate that RF performed better in both time spent on model testing and accuracy where neighbor's count is considered zero.

Since the input features increase in the dataset, by adding more neighbors, the time taken by the SVM and LR is significantly increased for model testing where the neighbor count is *one* and *two*, but ELM time spent on model testing remain consistent, which shows that ELM architecture is performing well by spending less time in model testing and produce better results then the fore mentioned algorithms, this also proves that ELM is a better approach where data in imbalance. However, the RF performed better in all the approaches in the data set by both times spent on model testing and accuracy.

4.2 Impact of neighbors' inclusions and ELM parameters

ELM output values are obtained as we take into account further neighbors. The iterations were repeated with the inclusion of neighbors' cells and respective RSRP values. The summary of the results is shown in Table 3. It shows the values of the computed parameters for different neighbors' counts for ELM. As the neighbors' cells depth increases, i.e., as we consider neighbors, the ELM respective accuracy for test results gets better.

Table 2. ELM and reference ML algorithms results.

Count of Neighbors	ML technique	testing_score	testing_time(s)
zero	KNN	96.00	34.733
	Logistic Regression	94.50	0.525
	Random Forest	96.00	2.076
	Linear SVM	94.50	30.128
	ELM	95.04	10.9
one	KNN	97.60	22.142
	Logistic Regression	94.10	0.719
	Random Forest	96.90	2.458
	Linear SVM	94.50	30.902
	ELM	95.65	11.5
two	KNN	99.20	44.898
	Logistic Regression	94.00	0.73
	Random Forest	98.70	2.705
	Linear SVM	94.40	67.46
	ELM	96.65	10.4

Table 3. Neighbors' cells depth and ELM parameters set results. The score is obtained on the test set.

neighbors	Score (%)	Alpha	Density	Original features	No of neurons	ufunc
zero	95.0420	0.0031	0.0280	TRUE	338	relu
one	95.6505	0.0062	0.6060	TRUE	397	relu
two	96.6519	0.0015	0.5165	FALSE	380	relu

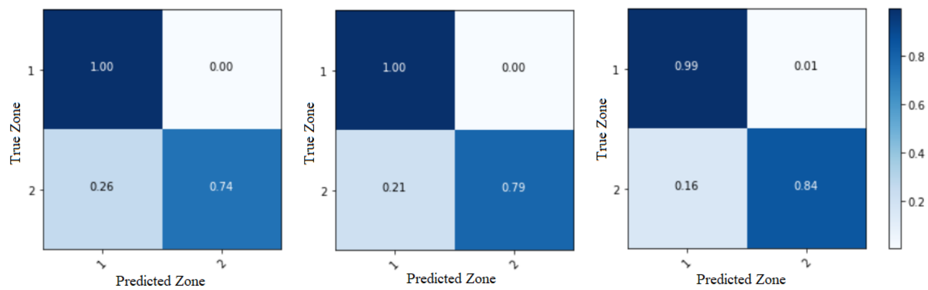

Fig. 2. Confusion Matrices for ELM calculations at different count of neighbors: *left: zero*, *middle: one* and, *right: two* cells respectively.

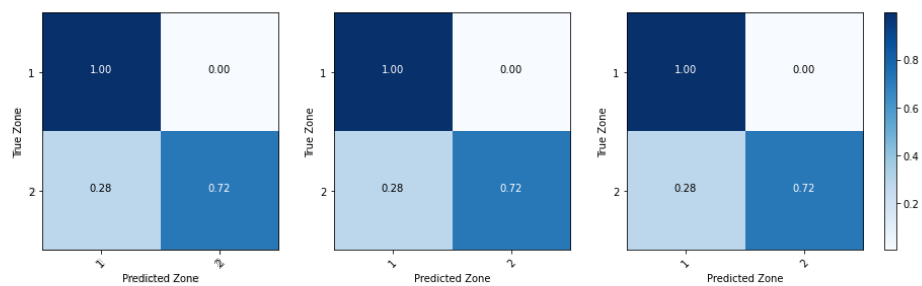

Fig. 3. Confusion Matrices for SVM calculations at different count of neighbors: *left: zero*, *middle: one* and, *right: two* cells respectively.

4.3 Confusion Matrices

The obtained confusion matrices for the different machine learning models are depicted in Fig. 2 for ELM, Fig. 3 for SVM, Fig. 4 for RF and, Fig. 5 for LR. In all confusion matrices. The obtained results show that even though the accuracy is almost similar for all trained models, RF and ELM have fewer false negatives than LR and SVM. The best results are obtained for a count of two neighboring cells.

5 Conclusion and further work

To conclude, ELM is much faster than other techniques used in this study. ELM can handle big data without the need for GPU or larger RAM size. The training and testing time both were fast for this method in comparison to the other tested techniques. In general, ELM is the algorithm that presents the highest ratio among accuracy and prediction time, being also effective in predicting fewer false negatives. Moreover, the advantage of using ELM is that the final models perform excellent with non-linear data and that models can be trained with new data without the need of retraining the whole model each time [8, 3].

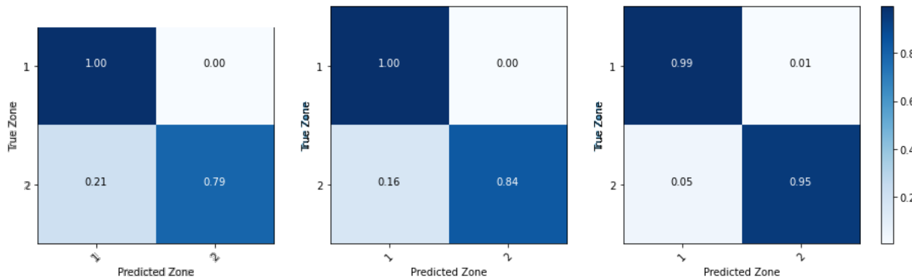

Fig. 4. Confusion Matrices for RF calculations at different count of neighbors: *left: zero*, *middle: one* and, *right: two* cells respectively.

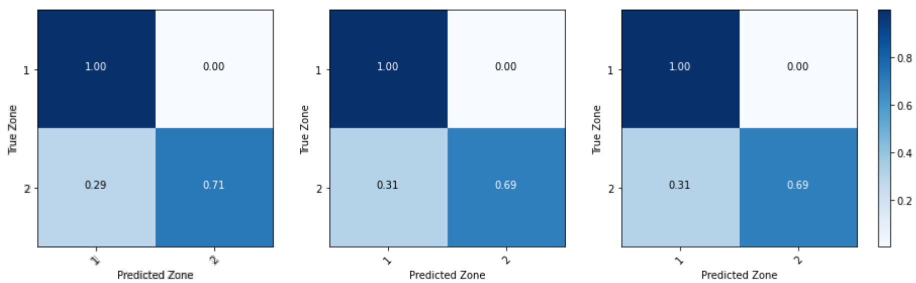

Fig. 5. Confusion Matrices for LR calculations at different count of neighbors: *left: zero*, *middle: one* and, *right: two* cells respectively.

In this paper, we considered data sets for up to two neighbors. Further research aims to extend it for a higher number of neighbors when the data is available. Moreover, it will be worth deploying the same techniques for 5G networks with beamforming at hand.

6 Acknowledgments

The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources.

References

1. 3GPP: Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; Measurements. Technical Specification (TS) 36.214, 3rd Generation Partnership Project (3GPP) (01 2018), `https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2440`, version 14.4.0 Release 14
2. Akusok, A., Björk, K.M., Miche, Y., Lendasse, A.: High-performance extreme learning machines: a complete toolbox for big data applications. IEEE Access **3**, 1011–1025 (2015)

3. Akusok, A., Leal, L.E., Björk, K.M., Lendasse, A.: Scikit-elm: an extreme learning machine toolbox for dynamic and scalable learning. In: International Conference on Extreme Learning Machine. pp. 69–78. Springer (2019)
4. Amirijoo, M., Frenger, P., Gunnarsson, F., Kallin, H., Moe, J., Zetterberg, K.: Neighbor cell relation list and physical cell identity self-organization in lte. In: ICC Workshops-2008 IEEE International Conference on Communications Workshops. pp. 37–41. IEEE (2008)
5. Bozkurt, S., Elibol, G., Gunal, S., Yayan, U.: A comparative study on machine learning algorithms for indoor positioning. In: 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA). pp. 1–8. IEEE (2015)
6. Cambria, E., Huang, G.B., Kasun, L.L.C., Zhou, H., Vong, C.M., Lin, J., Yin, J., Cai, Z., Liu, Q., Li, K., et al.: Extreme learning machines [trends & controversies]. IEEE intelligent systems **28**(6), 30–59 (2013)
7. Davidson, P., Piché, R.: A survey of selected indoor positioning methods for smartphones. IEEE Communications Surveys & Tutorials **19**(2), 1347–1370 (2016)
8. Espinosa-Leal, L., Akusok, A., Lendasse, A., Björk, K.M.: Extreme learning machines for signature verification. In: International Conference on Extreme Learning Machine. pp. 31–40. Springer (2019)
9. Huang, G.B., Zhu, Q.Y., Siew, C.K.: Extreme learning machine: a new learning scheme of feedforward neural networks. In: 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541). vol. 2, pp. 985–990. Ieee (2004)
10. Khan, G.W., Nyman, K.E.O.: Venue owner-controllable per-venue service configuration (jun 2019), uS Patent 10 320 973 B2
11. Khan, W., Keskinen, M., Raza, A., Kuusniemi, H., Elmusrati, M.: Using machine learning for in-out decision accuracy for venue owner definable services. In: 2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA). pp. 1–6. IEEE (2021)
12. Kuusniemi, H., Bhuiyan, M.Z.H., Ström, M., Söderholm, S., Jokitalo, T., Chen, L., Chen, R.: Utilizing pulsed pseudolites and high-sensitivity gnss for ubiquitous outdoor/indoor satellite navigation. In: 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN). pp. 1–7. IEEE (2012)
13. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: Machine learning in python. the Journal of machine Learning research **12**, 2825–2830 (2011)
14. Zulhasnine, M., Huang, C., Srinivasan, A.: Efficient resource allocation for device-to-device communication underlaying lte network. In: 2010 IEEE 6th International conference on wireless and mobile computing, networking and communications. pp. 368–375. IEEE (2010)