

Will they take this offer? A machine learning price elasticity model for predicting upselling acceptance of premium airline seating

Saravanan Thirumuruganathan^a, Noora Al Emadi^a, Soon-gyo Jung^a, Joni Salminen^{a,b},
Dianne Ramirez Robillos^c, Bernard J. Jansen^{a,d,*}

^a Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar

^b School of Marketing and Communication, University of Vaasa, Vasa, Finland

^c School of Statistics, University of the Philippines, Diliman, Philippines

^d College of Information Sciences and Technology, The Pennsylvania State University, United States

ARTICLE INFO

Keywords:

Upselling
Price elasticity
Recommender systems
Knowledge engineering
Intelligent systems
Machine learning

ABSTRACT

Employing customer information from one of the world's largest airline companies, we develop a *price elasticity model* (PREM) using machine learning to identify customers likely to purchase an upgrade offer from economy to premium class and predict a customer's acceptable price range. A simulation of 64.3 million flight bookings and 14.1 million email offers over three years mirroring actual data indicates that PREM implementation results in approximately 1.12 million (7.94%) fewer non-relevant customer email messages, a predicted increase of 72,200 (37.2%) offers accepted, and an estimated \$72.2 million (37.2%) of increased revenue. Our results illustrate the potential of automated pricing information and targeting marketing messages for upselling acceptance. We also identified three customer segments: (1) *Never Upgrades* are those who never take the upgrade offer, (2) *Upgrade Lovers* are those who generally upgrade, and (3) *Upgrade Lover Lookalikes* have no historical record but fit the profile of those that tend to upgrade. We discuss the implications for airline companies and related travel and tourism industries.

1. Introduction

A crucial business technique for increasing revenue is upselling [1], which in the airline industry means enticing customers who already have tickets to purchase upgraded, premium tickets (e.g., economy to business). Upselling from a basic to a premium category is difficult because many customers tend to be price-sensitive [2,3]. Therefore, inherent in upselling is the price elasticity of demand (a.k.a. price sensitivity), involving the decision of what price points to recommend for upgraded premium tickets to maximize both the customers' acceptance rate and the revenue from dynamic marketing offers. Price elasticity of demand is a measure of the sensitivity of demand to a change in price [4–6]. Practically, this must often be achieved with sparse customer information and with datasets of high cardinality [7], which is true for the research reported here.

In this research, in collaboration with one of the world's largest airline companies, we developed the *PRice Elasticity Model (PREM)*, a machine learning (ML) implementation for this challenging upselling context. PREM is aimed at (a) identifying customers who are most likely

to accept offers for seat upgrading and (b) determining the optimal pricing for these offers. In a competitive, tight-margin vertical like the travel industry, rival companies are perpetually evaluating various strategies to alter customer behavior for long-term engagement [8], forecast demand, and revenue generation, which are common goals for businesses [9] across many industries. The company's current upselling approaches are rule-based (i.e., number of prior upgrade offers, the time between offers sent, status level, etc.) and not data-driven, so the heuristics involved lack the needed customer context [10]. With PREM, the current research shows empirical evidence of more effective customer targeting while demonstrating sophisticated algorithmic techniques to enhance upselling performance by employing complex customer demographics and behavioral attributes [11].

Specifically, we use a rich dataset consisting of more than 64.3 M flight records, more than 14.1 M sent email offers, and about 194,000 purchased upgrades during three years from 2017 to 2019. Using the methodology discussed below, PREM (i) identifies customers who will most likely decline an upgrade offer, (ii) identifies customers who will most likely accept an upgrade offer, and (iii) determines the optimal

* Corresponding author.

E-mail address: jjansen@acm.org (B.J. Jansen).

<https://doi.org/10.1016/j.im.2023.103759>

Received 7 September 2021; Received in revised form 20 January 2023; Accepted 24 January 2023

Available online 1 February 2023

0378-7206/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

price range for those customers who are most likely to accept an offer.

This task is challenging as the acceptance or rejection of upgrade offers relies on complex interactions among the offer, the customer, the booking, and the destination (as shown in Fig. 1).

Our premise was based on the constructs as shown in Fig. 1, taken from discussions with experts on revenue, customer experience, and e-commerce within the company, using semi-structured interviews [12]. These dimensions provide a theoretical and practical paradigm for PREM [13,14] concerning the customer and the product. Prior work has also shown the influence of these three constructs in the airline industry [15], with each construct having subattributes such as departure date and price of the booking. Prior work has also shown that whether the customer is, for example, a leisure or business traveler affects the offer acceptance [16] and also the destination [16,17]. Customer perceptions of the airline service have also been noted as important for continued bookings [17]. While booking is often the only primary preference known about the customer [18], it can also involve subattributes such as the quality of the airline website [19]. As discussed later, our premise of the constructs of customer, booking, and destination can be seen to affect the decision to accept or reject the upgrade offer.

Consequently, implementing PREM requires overcoming several technical and practical challenges, making the research novel relative to ML work in the airline industry and the price elasticity of demand domains. First, as typical for large-scale customer datasets [20,21], airline booking data is sparse, meaning that most customers are not exposed to most of the available products [22]. Second, the data is imbalanced, which is typical for online marketing datasets in which a conversion is a rare event as the vast majority of customers exposed to an offer do not take the offer [23]. For example, in the current dataset, the airline sent about 14.1 M upgrade offers, of which less than 194 K (1.43%) were accepted, which is about 0.3% of all flight bookings. Third, as customary in large-scale customer datasets in various fields [24], the dataset required is highly noisy, with countries and airport codes specified in different notations and customer demographics often unreliable or missing. These factors make building and implementing a robust ML model a challenging task. However, our model directly addresses these challenges. Additionally, human factors and corporate realities affect revenue generation; as a consequence, the PREM model is designed with these multiple variate factors in mind.

Several motivational factors for this research in the upselling domain are common to large-scale customer datasets in various industries. Thus, addressing these issues has wide-reaching implications beyond the context chosen for this study. The company data used for this study is standard for most major airline companies with customers, flights, dates, destinations, prices, and associated information. As such, it may be seen as representative of many major airline companies, of which

there are hundreds, that offer international bookings. These factors include improving the currently employed heuristics approach to upselling, working with sparse and noisy data, and deriving business value from ML modeling. According to our understanding, these predicaments are commonly encountered in many companies provisioning upselling offers to their existing customers, so our research broadly impacts multiple industry domains. The current research provides theoretical and practical implications for understanding upselling in the travel domain [25].

2. Literature review

Pricing is an essential decision task of many industries [26], including the travel and tourism sector. Prior work on ticket purchasing [27,28] and related work [29] has identified appropriate timing as a key factor for purchasing airline tickets [30,31]. Typical approaches apply a regression model and a carefully chosen threshold to determine whether to buy based on estimated price increases. Prior work has used various features such as price information, related itineraries, summary statistics of price, and trip-related features such as holiday travel [32], and incorporating these aspects for customer engagement and trust is essential for information systems in the e-commerce sector [33]. While pricing is a complex issue, upselling [34] to premium products has further nuanced customer aspects involving service, quality, and price among product levels.

Companies in the airline industry are motivated to identify optimal pricing, as doing so can increase profit, assist in acquiring new customers, and support retaining customers [35]. Airline companies are also interested in developing long-term engagement by providing offers that are cognitively, emotionally, behaviorally, and socially relevant [36] to their customers. Prior research has shown that companies focus on long-term value and not just the revenue generated by customers for each transaction [37]. Therefore, it is practical for businesses to target their upselling efforts [38] only to those who are most likely to accept the offer without diluting the premium service [39] with too many discounts. As such, dynamic pricing strategies are increasing in popularity [40]. Dynamic pricing aims to adjust ticket prices based on demand, customer price sensitivity, seasonality, and destination [41]. However, we are unaware of any existing work determining pricing for the personalized upgrade offers to a company's targeted customers. Thus, upselling research is novel and impactful for the airline travel industry and related verticals such as hotels and cruises.

An airline company is typically incentivized to categorize customers [42] based on price sensitivity [43]. In practice, airlines charge different prices for tickets, including efforts to upsell in economy class [44]. This is often done based on static criteria such as trip purpose (business vs.

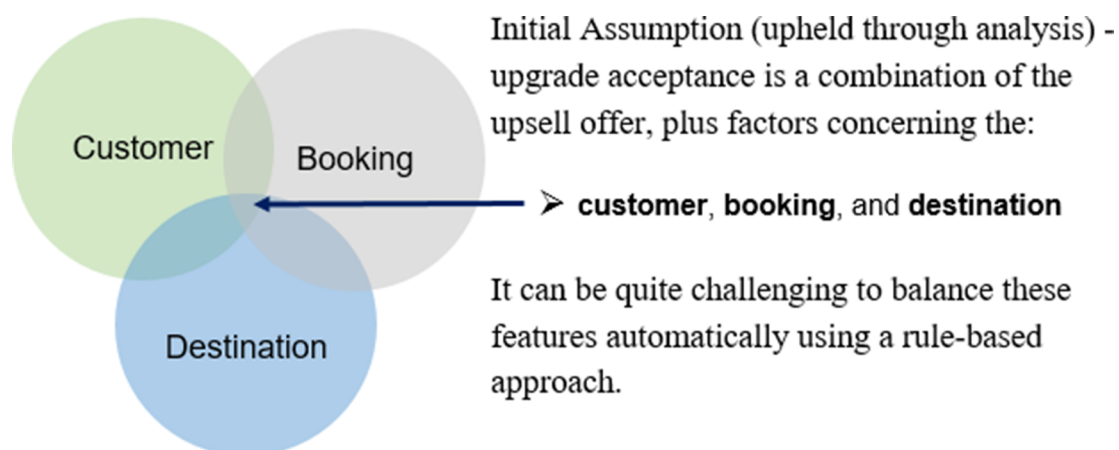


Fig. 1. Constructs of PREM. Acceptance may be seen as the interaction among the customer, booking, and destination when given an upgrade offer. These dimensions were proffered by domain experts and provided a theoretical and practical paradigm for PREM.

leisure), passenger income estimate, type of traveler (tourist vs. regular), and so on. This price discrimination shows that the price varies after controlling for ticket price factors [45]. Furthermore, other work shows that customers making purchases on the weekend are likely to be more price sensitive [46]. Although there is prior work on customer segments [47], little research focuses on premium upselling efforts or repeating customer loyalty.

Big data is increasingly being leveraged in travel research [48]. However, a particular issue with the upselling of bookings is that customer engagement data is sparse and often focuses on minute behaviors [49,50]. Data sparsity is a common aspect of online customer data, and while data sparsity is challenging in itself, in an operational environment in the travel domain, there are additional challenges to finding solutions that are workable with customer-facing and real-time business systems.

While there has been previous work on related topics [51,34,52], most of these studies utilize either a small amount of data or customer samples that are not representative of many companies in the travel industry. For example, the U.S. Department of Transportation provides ticket price information for a small subset (10%) of itineraries for U.S. domestic flights. Other works have obtained data through scraping websites [53]. However, these datasets tend to cover short periods, resulting in findings that may either not represent particular segments [54] or posit models unsuitable for larger datasets. To our knowledge, there has been limited analysis of upselling (including the pricing of offers) in the travel industry, and the lack of benchmark data is identified as a critical issue in developing ML models for pricing [15].

In contrast to the limited datasets utilized in prior work in the upselling field [55], the research presented here was conducted in collaboration with one of the largest airlines in the world, with substantial real-world relevance. We leverage a substantial amount of booking data covering millions of bookings over three years, so the dataset is extensive. However, this real-world data raised several practical challenges prohibiting the use of prior work, which necessitated PREM's novel architecture design. Consequently, this research offers an innovative contribution to the information and management domain.

3. Research goals

In this research, we pursue two research objectives (ROs): **RO1:** (a) *Identify customers who are likely to accept a premium upgrade offer, and (b) identify customers who are likely to decline an upgrade offer*; **RO2:** *Identify the price elasticity of those customers likely to accept an upgrade offer.*

These ROs are critical to the airline business regarding revenue, customer retention, and segmentation. Because most customers book economy class and few book premium class, upselling can benefit airline companies and passengers if some of their customers are exposed to upgrade offers. Intuitively, there are three customer categories: (a) the *frugal travelers* who are likely to decline even the cheapest reasonable upgrade offer; (b) the *bargain hunters* who will almost always accept the upgrade offer as they do not have to pay the full price; (c) and the *messy middle* who would generally not book business class, but could be persuaded with an appealing offer. Based on our analysis of historical data, most travelers (as much as 82%) routinely decline the upgrade offers, and only a very tiny sliver (less than 1%) routinely accepts them.

We model RO1 as a binary classification (i.e., yes/no) with a probability for each outcome (e.g., *Customer X will accept/decline an upgrade offer with a probability of 0.9*). Despite their seeming similarity, RO1a and RO1b are not mirror images of each other. If we consider three hypothetical upgrade offers for a given booking, which features Low Offer for \$500, Middle Offer for \$1000, and a High Offer for \$1500: RO1b estimates the probability of a customer declining the Low Offer, the cheapest upgrade offer. On the other hand, RO1a seeks to estimate the probability that a customer will accept the High Offer, the most expensive offer. These customers and customer decisions are of interest to the airline for different reasons. However, airlines often have business

rules within their organization to avoid cannibalization, such as if a customer who has recently accepted an upgrade offer, then do not send another offer to the customer even if they are likely to accept it, with the rationale that this incentivizes the customer to pay full price in the future.

By clearly segregating customers who routinely decline or accept an upgrade offer, we derive subsets of the population of interest. In order to then customize upgrade offers for these customers, it is essential to understand their price sensitivity, which is the goal of RO2. The outcome of RO2 will be a set of probabilities for different upselling revenue ranges (e.g., *Customer X will accept an upgrade offer of \$500 with a probability of 0.6, \$1000 with a probability of 0.3, and \$2000 with a probability of 0.2*). The output of RO2 can be used to map an optimal strategy for targeted upgrade offers that maximize revenue for the customer.

4. Methodology for model development

4.1. Heuristic upselling process by the company

Our collaboration airline company has more than 100 flights to more than 100 destinations on all major continents. The company serves more than 25 million customers annually, and our data covers a significant proportion of these customers over multiple years. The company uses a hub-and-spoke model (where flight routes are organized as a series of routes connected by a single hub airport), has international flights, competes on service and revenue, faces governance issues common in the industry, has a frequent flyer program, and is a member of a major airline alliance. As such, it represents a typical major airline company that offers bookings to multiple destinations to an international customer base. Fig. 2 illustrates the company's rule-based process for upselling offers via email marketing channels.

In Fig. 2, the customer books a ticket in economy class. The company determines if the customer is eligible to receive an upgrade offer, which is sent via an email message containing the booking details and the upgrade offer price. Upon receiving the offer, the customer either accepts or ignores the proposed offer. Fig. 3 offers an example of an upgrade offer email message (branding removed).

4.2. Airline booking and upgrade data

The dataset used for PREM development consists of more than 64 million trip booking records of customers traveling between 2017 and 2019. The dataset contains information concerning the price of the

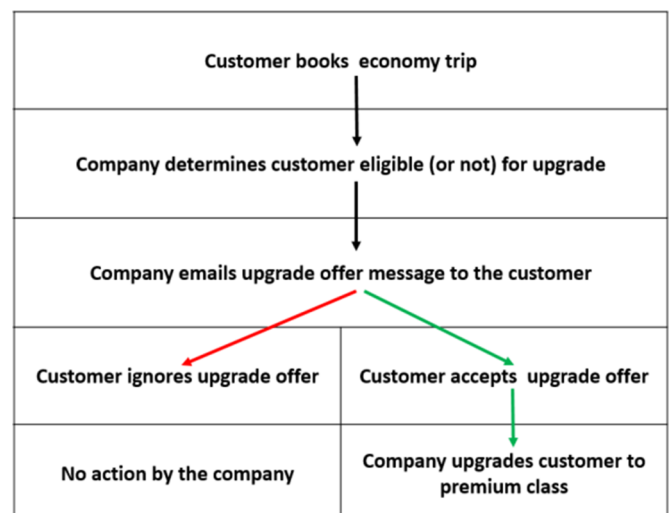


Fig. 2. Overview of upselling offer ticket purchasing flow for the airline.

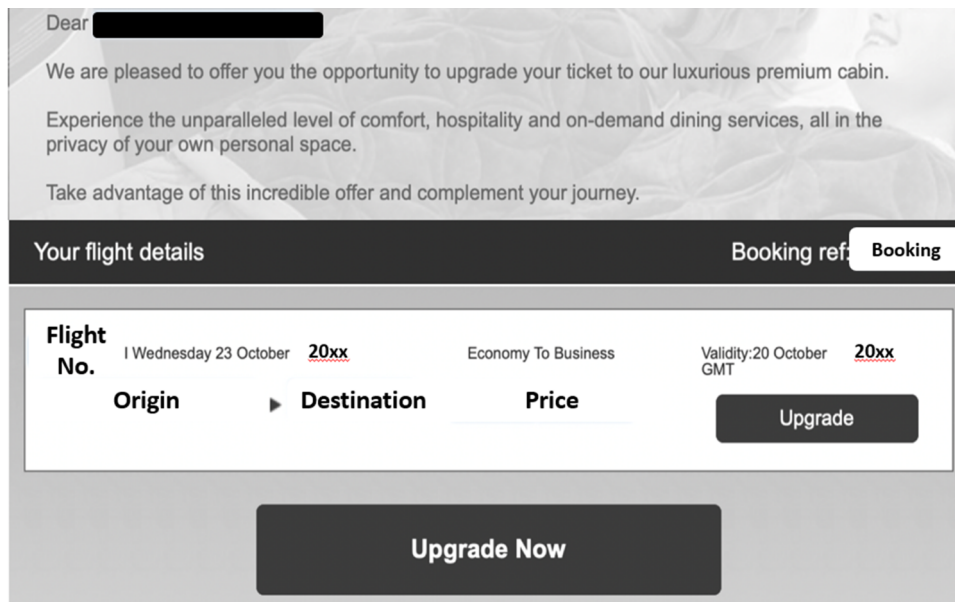


Fig. 3. Example of an upgrade offer email sent to a passenger, with branding removed.

upgrade offer, which customers were sent upgrade offers, and which customers did or did not accept the update offers. This data is valuable for analyzing customer upselling and price elasticity dynamics in a major and competitive industry, and findings have implications for other travel-related domains. More than 14 million customers received upgrade offers (with a reach rate of 22%), and over 194,000 customers accepted the upgrade (with a conversion rate of 1.43%), as shown in Table 1.

The PREM model relies on variables available to most airline companies, allowing the model to be reused. The original dataset consists of 23 variables, including customer demographics and booking information detail. The key variable categories are as follows:

- *Booking information:* cabin class, booking class (granular categories within cabin class), family fare, flight date, and booked date
- *Customer demographics:* age, gender, and nationality
- *Trip information:* airport code, city, and country of source and destination
- *Upgrade details:* original and upgraded cabin, offer acceptance, and the offer price

When working with this data, three significant challenges were imbalance, sparsity, and noise. Regarding data imbalance, as shown in Table 1, fewer than 194,000 (1.43%) of more than 14 M contacted customers accepted the upgrade offers. Concerning sparsity, there is a severe lack of travel information, even in this large dataset. For example, in the data, more than 90% of customers take fewer than three trips per year, which reduces the customer information needed

Table 1
Upselling dataset with bookings, upgrade offers sent, and accepted.

Year	Bookings	Offers sent	Offers accepted	Reach rate	Conversion rate
2017	15.2 M (23.6%)	2.5 M (17.7%)	42 K (21.1%)	16%	1.67%
2018	30.9 M (48.1%)	7.5 M (53.2%)	99 K (510%)	24%	1.32%
2019	18.2 M (28.3%)	4.1 M (29.1%)	53 K (27.3%)	23%	1.29%
All	64.3 M (100%)	14.1 M (100%)	194 K (100%)	22%	1.43%

for sophisticated targeting. Also, the same customer could behave differently to the same upgrade offer based on the destination and/or the booking (e.g., the travel month of the year, whether they are traveling with a companion or family, etc.). Finally, the data was noisy, with inconsistent codes and missing values for nearly all attributes. We conducted a limited data preprocessing involving a value standardization of codes representing the airports and countries. Addressing these data noise issues is a key component of our work.

4.3. PREM model development

A common approach to addressing this topic of upselling would be to tackle this as a straightforward ML classification problem where, given a particular customer, the model predicts the likelihood that they will accept an upgrade offer. Then, upgrade offers could be sent to customers with the highest likelihood of accepting the offer. Our preliminary investigations found that directly applying state-of-the-art ML techniques, such as supervised or semi-supervised learning, gives poor results. A key challenge is that a single-stage model must tackle various challenges, such as data sparsity, fragmentation, noise, and a severe imbalance between accepted and declined upgrade offers. As we show later in our experimental results section, this produces a lower degree of accuracy and would result in a significant revenue loss.

A further consideration is that the airline company is more focused on impactful metrics such as revenue and customer retention [56,57] than on the traditional objectives of ML models, such as accuracy [58]. ML models often do not account for these business objectives [55]. For example, a model could obtain a very high conversion rate through very cheap upgrade offers, which is an excellent algorithmic performance but not a good revenue performance for the company. Relatedly, an ML model could be built whose output could be modified by airline employees, but as addressed in previous work, algorithmic selection for various business goals is seen as a challenging undertaking [59].

These observations led us to design a staged pipeline architecture for PREM. Our proposed approach consists of five components. The first component consists of feature engineering, where we construct a wide variety of features based on three dimensions – the customer, the trip, and the upgrade offer. Our second component constructs feature embedding using denoising autoencoders [60]. We found that production data obtained by federating multiple sources is often quite noisy and

requires extensive reconciliation to obtain clean data. Data cleaning would be a recurring expense, as the airline data would not be standardized. So, instead, our novel approach is resilient to various data errors, which we believe would be of interest to practitioners. Our third component is a cost-sensitive classifier based on our observation that misclassification cost is asymmetric. Missing a customer who would have upgraded is much more expensive than sending an upgrade offer to a customer who ignores it. Our proposed approach is based on Elkan’s cost-sensitive thresholding approach [61]. Once potential customers are obtained, the fourth component of a personalized upgrade offer model is used to understand their price sensitivity. The final component is the revenue maximizer, which selects appropriate customers and offers prices that are likely to appeal to them while maximizing the revenue for the airline.

A key appeal of our proposed approach is that it is extensible and also permits extensive interventions by human decision makers [62] in the airline company, which in this context is the addition or removal of customers from the eligible list. Although our method does not require human intervention, our partners require multiple avenues for customizing PREM outputs. For example, the airline might want to add/remove some customers from the output of the third stage of PREM (the cost-sensitive classifier) in order to incentivize new customers or penalize repeated users of upgrade offers. Also, the airline might want to use an alternate segmentation instead of the one provided by the fourth component of PREM or add some business constraints in the integer program-based optimization formulation. This customization and the integration of business requirements are not possible in a single-stage model. We also note that the revenue maximizer combines ideas derived from decision theory (i.e., expected revenue) and optimization theory (i.e., linear assignment problem). As we empirically show in the experiments, the multistage approach provides better results than a single-stage end-to-end approach using multiobjective optimization with a constraints-based approach. The overall PREM process and components are presented in Fig. 4.

As an overview of Fig. 4, the three most critical problems are noise, sparsity, and imbalance. We tackle each of these problems through two components – feature embeddings and cost-sensitive classification.

First, we use denoising autoencoders for learning the embeddings. Specifically, we train the encoder by artificially injecting noise and penalizing the model if it does not produce an accurate embedding. This ensures that, during testing time, even if the data is noisy (such as with missing or incorrect values), the embeddings will be more robust.

Second, we again use embeddings for handling sparsity. Recall from Section 5.1 that we used 100 features. However, when we represent them in a format that can be fed to an ML model, each trip becomes (approximately) 2000 dimensional vector. To see why, consider two essential features – source and destination. If the airline flies to 250 destinations, we need a 250+250 one-hot encoded vector to represent these two factors. Similarly, we bucketed age into eight categories. So, age is represented as an eight-dimensional vector and so on. By performing an embedding, we reduced the dimensionality from ~2000 to 256 and ensured that similar trips were closer to each other in the embedding space.

We did evaluate other encoding approaches, including dimensionality reduction techniques such as PCA. As discussed below in Table 2 (Results of feature embedding analysis), our embedding algorithm outperforms other approaches. While we did not invent this embedding algorithm, we are the first, to our knowledge, to apply it to airline data to tackle noise and sparsity issues. This applied study draws from computer science to solve a problem with real-world impact.

Finally, we used the cost-sensitive classification approach for handling class imbalance. Given that similar trips are closer to each other in the embedding space, we observed that using traditional cost-sensitive classification is sufficient for handling class imbalance. In

Table 2
Results of feature embedding analysis.

Feature embedding	F1 score	Revenue capture
Denoising AutoEncoder	83.9	100%
Traditional AutoEncoder	77.2	93%
One-Hot Encoding	44.3	62%
Label Encoding	47.8	56%
PCA	53.8	66%

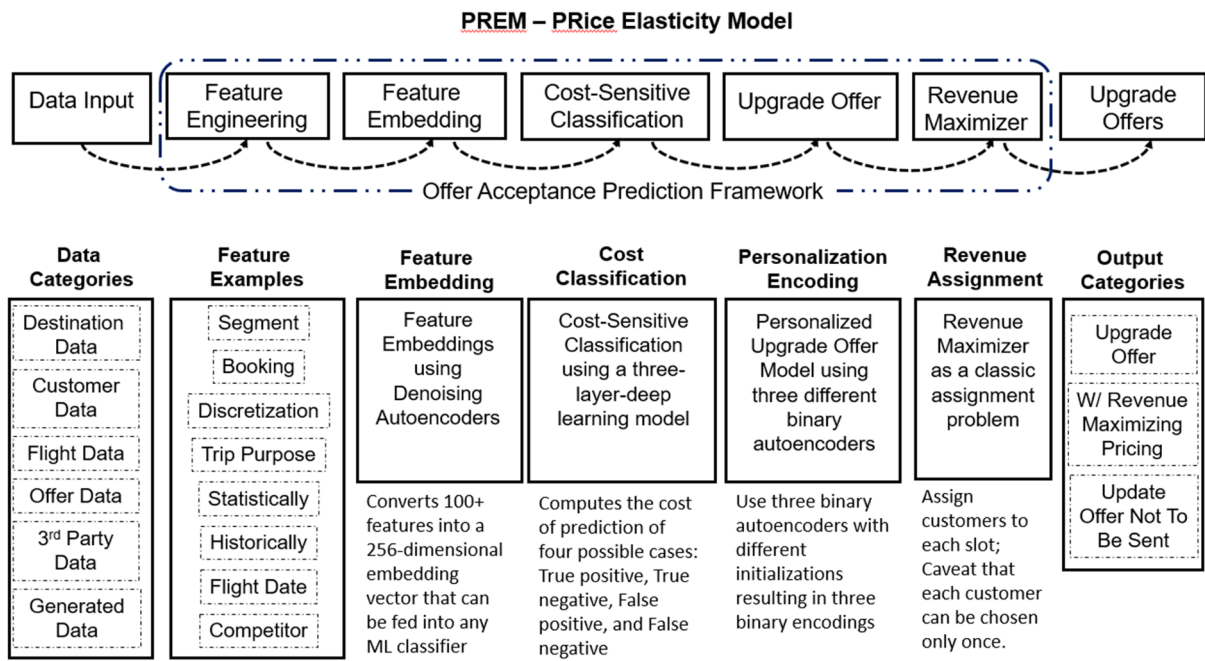


Fig. 4. Price Elasticity Model (PREM) visualization, from data input, to offer acceptance prediction, to output, showing the major steps and processes performed at each stage.

Table 4 (Results of classifier/feature analysis), we also see that this approach outperforms other major approaches for handling class imbalance. Specifically, two approaches (SMOTE and GAN) work by generating synthetic data.

5. PREM components

5.1. First component – feature engineering

We performed feature engineering (see Fig. 4) to identify more than 100 derived features. These features are common to most major airline companies or available via open application programming interfaces (APIs). Here, we highlight some of the more useful derived features.

- **Booking-based features:** Number of stops, the distance between the source-destination, total time taken, the number of passengers in the booking, whether traveling with a family, the number of children traveling in the booking, the number of days between booking and flight dates, and whether the flight is to/from the home country of the passenger.
- **Competitor-based features:** Number of flights between source and destination on the same day from the same airline company and its competitors.
- **Discretization-based features:** We partitioned customer ages into U.S. census groupings (i.e., 13 and younger, 13–17, 18–24, 25–34, 35–44, 45–54, 55–64 years, and 65 years and older). We also unified the cabin class values from a single letter code to the equivalent cabin class.
- **Flight date-based features:** weekday or weekend flight, holiday flight based on whether the date is adjacent to a holiday according to the nationality of the customer, or an overnight flight or not.
- **History-based features:** number of trips and upgrades from prior trips.
- **Segment-based features:** proportion of the segment that upgraded. For example, if the customer is a male from the USA aged 40–50, there is an average acceptance rate for passengers who are (a) male, (b) aged 40–50, and (c) from the U.S., overall and for each attribute.
- **Statistical price metrics-based features:** average price of the same itinerary during different periods (one month, one week, and one day before the flight date), the average price of competitors, and maximum and minimum price. These features were computed using the Rakuten SkyScanner API.¹
- **Trip Purpose-based features:** inferred personal or business trip: For determining the purpose of the trip, we used a heuristic provided to us by the airline. One of the features of the trip is the POS (point of sale). Often, leisure-based trips are booked by individuals or through travel agencies. On the other hand, business trips are booked by the travel departments within the organization. Often, these departments use special tokens to obtain favorable enterprise rates. Hence, by using this heuristic, we could plausibly infer the purpose of the trip. Of course, this heuristic is not foolproof but offers a reasonable assumption.

These features cover different data types (i.e., categorical, ordinal, numerical), which poses an additional challenge when passing the data as input to an ML classifier.

5.2. Second component – feature embedding using denoising autoencoders

Popular machine classifiers accept the trip input as a real-valued vector. Common approaches for converting a vector of mixed data types to a real-valued vector, such as one-hot encoding and label encoding, were not effective with our dataset. One-hot encoding dramatically increases the dimensionality of the data [63], thus

exacerbating the data sparsity and scarcity problems. Specifically, when representing the trip information using one-hot encoding, each trip becomes (approximately) a 2000-dimensional vector. To see why it is important to consider the two key features of source and destination. If the airline flies to 250 destinations, then we need a $250+250=500$ one-hot encoded vector to represent just these two factors. Additionally, we bucketed age into eight categories. So, age is represented as an eight-dimensional vector and so on. Label encoding induces the model to artificially learn an order among the values of a categorical attribute, which is inappropriate for our scenario. We evaluated other methods, such as a dimensionality reduction of the one-hot encoded data using PCA. The Results section of this paper shows that such approaches are sensitive to categorical attributes with large domain cardinality.

PREM uses an alternative approach based on feature embedding (see Fig. 4). Given a mixed data type vector, we output a corresponding embedding as compact, low-dimensional, dense, and real-valued representations of the original tabular data. PREM converts over 100 features (which required ~ 2000 dimensions in one-hot encoding) into a 256-dimensional embedding vector that can be fed into any machine classifier. Embeddings also have an additional desirable property, where if two bookings are similar in the raw tabular format, the distance between corresponding real-valued embeddings will be small. The Results section shows that this property improves the classifier's generalization capability.

PREM uses denoising autoencoders to obtain the embedding-based representations. Autoencoders [64] are used for learning effective encoding for data in a purely unsupervised manner. One can think of autoencoders as a form of dimensionality reduction. Given an input vector x , the autoencoder learns an intermediate representation that could be used to reconstruct x . Autoencoders consists of two components—an encoder and a decoder. The encoder converts the input vector x into a latent representation $h(x)$. The decoder takes the latent representation and reconstructs the original vector. The dimensionality of the latent representation $h(x)$ is smaller than that of x , so any essential information needed to reconstruct x from $h(x)$ is learned. One could train an autoencoder that takes raw tabular input (that could require ~ 2000 dimensions), converts it to an embedding of 256 dimensions, and reconstructs it (approximately) to the original vector. Accordingly, we see that one could use the intermediate representation $h(x)$ as input to ML classifiers instead of x .

We use a variant of autoencoders called denoising autoencoders [65, 66]. Denoising autoencoders take as input a corrupted version of the input vector x denoted as x' . It encodes it to an intermediate representation in order to decode it (approximately) as the uncorrupted version of x . The key insight is that if the decoder can reconstruct an approximately correct version of x from the latent representation $h(x')$ of the corrupted input, then it has learned the “essential” information about x and is not fooled by the corruption. The rationale for using a denoising autoencoder is that it is more robust to the errors inherent in real-world data, such as data that is often obtained by integrating multiple systems of customer analytics, flights, revenue management, and so on.

Hence, it is helpful to learn a robust embedding method for dealing with such noise. Accordingly, we use *swap noise* for categorical data. For example, suppose that Attribute A has three possible values: X, Y, and Z. Consider a tuple t that has Y as the value for A. We will replace its value under the swap noise model with another value (such as X or Z). A noise swap parameter controls the number of categorical cells that are corrupted. Empirically, we found that 15% seems to perform well in practice. Hence, we randomly pick 15% of the training dataset's categorical attributes and replace them with some other attribute domain values for implementation. Naturally, this process was not conducted on the test set to avoid information leakage. Furthermore, embedding learning is an unsupervised method that does not require any labeled data. Hence, we use all the bookings in our dataset to learn an effective embedding method. Once the denoising autoencoder has been trained, it can be used to obtain the embeddings for arbitrary bookings.

¹ <https://english.api.rakuten.net/skyscanner/api/skyscanner-flight-search>

Our approach might initially seem to be counter intuitive as we add more noise during training. As we show later in the experiments, this approach results in a robust classifier for various types of errors and recovers from them. This approach aligns with the emerging paradigm of robust learning that seeks to build models that do not assume that the input is clean. Furthermore, the denoising autoencoder has been widely used in tasks such as inpainting that can repair damaged photographs. PREM follows in those footsteps and develops similar techniques for tabular datasets. We strongly believe that approaches similar to ours and robust to errors will become increasingly widely used. To the best of our knowledge, in the context of the presented paper, we are the first to apply it to airline data to tackle noise and sparsity issues.

5.3. Third component – cost-sensitive classification

We then trained an ML model to predict whether the customer would accept an offer or not. A challenge that needs to be tackled is the asymmetry in cost, wherein this model should conservatively perform its classification. In other words, predicting a customer who could have upgraded as one who would *not upgrade* must be penalized more than sending an offer to a customer with a low probability of upgrading. We achieved this penalty through cost-sensitive classification [61] (see Fig. 4 above).

A key notion for the effective implementation of PREM is identifying and filtering out customers who are unlikely to upgrade. By eliminating these customers upfront, the later PREM stages can focus on more promising customer candidates using the model components described in the following sections. This approach might again seem counterintuitive (i.e., identifying customers to eliminate rather than to keep). However, this approach is an integral part of the success of PREM. In practice, customers who rarely upgraded vastly outnumber those who nearly always upgraded, which is only about $\sim 0.3\%$ of customers in the dataset. Identifying these customers can be done more effectively using ML than identifying customers with a high probability of accepting an offer.

We computed the cost of prediction for each booking in the training data using a cost function heuristic. Note that there are four possible cases for upselling prediction accuracy: True positive (T.P.), True negative (T.N.), False positive (F.P.), and False negative (F.N.). Consider a customer who had an upgrade with an offer of 1000 USD. If the model correctly predicted this, the action is provided with a 1000 (T.P.) reward, but if it does not, it is penalized by 1000 (F.N.). Alternatively, suppose that the customer did not upgrade. Then, correctly predicting the actions gets a reward of 1 (T.N.), whereas incorrectly predicting it gets a 50 (F.P.) penalty. This penalty is needed to prevent the model from predicting everyone as possibly upgrading, and the penalty value trades off the cost of sending a spurious email and can be varied based on an organization's preferences. Assigning different penalties for misclassification costs could produce different classifiers. By default, our misclassification cost for a false negative is \$1000. We computed the upgrade premium as the difference between the upgrade offer and the economy price for all successful upgrades. For example, if the economy price was \$500 while the upgrade offer was \$1500, the premium is \$1000. We analyzed the historical data, finding that the median of a successful upgrade was \$1089, which we rounded to \$1000 for ease of communication with stakeholders.

A three-layer-deep learning (DL) model provided the highest accuracy. The classifier outputs a value between 0 and 1, with higher values indicating a higher likelihood of accepting the offer. By default, we set the cutoff to 0.5, so that all customers above this threshold are passed on to the next stage. We also provide this list to the airline company's human decision makers so that they can apply other thresholds. For example, the airline could have a policy that a customer could not be offered multiple upgrades within a given period.

Notice that we use a threshold of 0.5 to filter out customers who are unlikely to upgrade. This threshold is appropriate as we use a cost-

sensitive variant calibrated based on different costs for false positives and negatives. This recalibration ensures that 0.5 of a cost-sensitive classifier is equivalent to some arbitrary threshold domain experts must manually identify for the traditional variant. Such a threshold also ensures that we avoid accidentally eliminating customers who could have potentially upgraded. Additionally, the output of this stage could be postprocessed by humans to achieve higher reliability.

5.4. Fourth component – personalized upgrade offer

From the classifier, we are given a list of candidate customers who could accept an upgrade offer, and we know how to personalize the upgrade offer within the price buckets that the customers will accept. The personalized upgrade component predicts the customer's likelihood of upgrading for each of these buckets (i.e., convert). We achieve this by hashing the booking through binary autoencoders (see Fig. 4).

There are two challenges in estimating the conversion likelihood: (a) multiple possible labels for customers and (b) a lack of information about individual customers. For the first issue, superficially, estimating the likelihood might seem to be a multiclass classification problem where the goal is to predict a single label for the input tuple among K classes. However, it is possible that the same customer would have a nonzero likelihood for different buckets in our case. Consider a customer who is potentially offered three upgrade offers, which are a Low Offer at \$251–\$500, a Middle Offer at \$501–\$1000, and a High Offer at \$1001–\$1500. If the customer is willing to take any offer of less than \$1000, this individual will accept Low Offer and Middle Offer. Thus, our personalized model gives a score of Low Offer: 1.0, Middle Offer: 1.0, and High Offer: 0.0. We can see that the scores need not add up to 1.0. The revenue maximizer takes this information and sends the Middle Offer to Customer A.

The second issue is a lack of information about individual customers means that most customers (more than 90%) took less than three trips per year and less than ten times over the three years. The fraction of times they accepted an upgrade offer was even smaller than the number of trips. Therefore, one cannot predict that any given customer would accept the offer as the data is insufficient for this granular analysis.

A way out of this conundrum is to split the customers into a set of segments in order to measure the likelihood that a customer would accept an offer. For example, a simplistic segmentation could be based on gender. We could group all the bookings based on gender and compute the proportion of times that customers from each bucket upgraded when offered one of the five upgrade offers. Our evaluation found that demographic-based segmentation is promising but not optimal, as these approaches do not leverage many factors drawn from the booking-related features. Of course, one could perform a more granular analysis using additional demographic attributes such as age group, nationality, source, destination, etc. However, in our initial trials, naively including such attributes sparsified the segmentation and reduced the accuracy. This suggests a trade-off between the granularity of the segment and the number of customers that fall into each of these buckets. Notably, increasing the granularity decreases the number of customers (especially for rare segments) and vice versa. Our experiments evaluate different segmentation techniques and show their efficacy, which we discuss in the Results section.

Intuitively, a binary encoder takes a tabular data input x and converts it into a latent representation $h(x)$. At this stage, we ensure that the vector $h(x)$ is binary. For example, if we set the size of the latent dimension to two, then each tuple is represented as a two-dimensional binary vector with a value of 00, 01, 10, or 11. We can see that the binary encoding of the tabular data plays a role in customer segmentation. We found that this approach is superior to naïve customer demographics-based segmentation as it uses more than 100 features for identifying the binary encoding. Hence, if two customers are very similar, they would likely have the same binary encoding. Our experiments set the latent dimension to eight, splitting the customers into 256

data-driven segments. Given a set of customers, we can partition each into one of 256 distinct buckets using the binary autoencoder. Once the segmentation is done, we can compute the likelihood of arbitrary upgrade offer buckets for customers using binary encoding. Then, we can estimate the fraction of customers who will upgrade, given an upgrade price offer.

We use three different binary autoencoders with different initializations that result in three different binary encodings in practice. Therefore, we compute the likelihood for each of the three encodings and compute the average. For example, the first autoencoder assigns the same binary encoding for Customer A and Customers B, C, and D. For the second autoencoder, Customer A has the same encoding as, say, Customers C, X, and Y. For the third autoencoder, let the other Customers be B, X, and Z. Let the likelihood of customers accepting the offer be 0.50, 0.25, and 0.25. Then, the average likelihood for Customer A to accept the offer is 0.33.

To do this, we must define a set of mutually exclusive upgrade offer buckets. Suppose the economy class ticket is \$500, and the potential upgrade offer buckets are \$501–\$1000, \$1001–\$1500, and so on. A key preprocessing step for PREM is to normalize the buckets based on the economy class ticket price. Specifically, we compute the ratio of a bucket's lower and upper bounds to the economy class ticket. Hence, the normalized upgrade offer bucket could be $(\$501/\$500) - (\$1000 - \$500)$, $(\$1001/\$500) - (\$1500 - \$500)$, and so on. The normalized prices allow us to use the same binary autoencoders for different flights, where the price of an economy class ticket and upgrade offers could be very different. In practice, we computed the ratio using an equi-depth histogram [67] of overall accepted-offer prices. In other words, for each of the 194 K offers that were upgraded, we computed the ratio of the upgraded price to the economy class price. We then partitioned these values into five distinct bucket ranges so that the number of accepted offers in each bucket is approximately equal. Increasing the bucket sizes results in a more granular analysis and is orthogonal to the algorithm design. While the airline could set the number of buckets to any number, we found that the value of *five* provided the highest accuracy for the model.

5.5. Fifth component – revenue-maximizer

This component of PREM sends upgrade offers to customers so that the expected revenue is maximized (see Fig. 4). Suppose that there are N candidate customers. The personalized upgrade-offer model produces the likelihood of each of these N customers accepting a set of K upgrade offers (such as \$501–\$1000, \$1001–\$1500, etc.). Suppose there are M available business class seats; hence, the airline wants to send out the email to $C \times M$ for some constant C . If there are five business class tickets, the airline may want to send out an email to 50 customers at most. These 50 customers need to be chosen in a way that maximizes the expected revenue.

To do this, PREM computes the expected revenue for each customer as the product of the likelihood and the median value of the range. For example, if the customer accepts [\$501–\$1000], with a likelihood of 0.9, then the expected revenue of $\$750 \times 0.9 = \675 . Of course, the airline company can optionally select a scalar upgrade offer within each of these buckets. Regardless, if there are N customers and K upgrade offer buckets, this results in $N \times K$ different assignments. We must choose a subset of $C \times M$ customers to send emails to. Another constraint is that each customer can be sent only one upgrade email. Otherwise, the customer will always use the cheapest offer to upgrade. We can see that this corresponds to a classic assignment problem [68] in which we need to assign customers to each of the $C \times M$ slots, with the additional caveat that each customer can be chosen only once. We formulate this as an integer linear programming (ILP) problem in which the objective

function is to maximize the cumulative expected revenue for each of the selected $C \times M$ customers. We solve this formulation using the GLPK library.² The output is a set of $C \times M$ customers and their corresponding upgrade offers.

6. Results of PREM evaluation

We use two metrics to evaluate the PREM experimental results, the F1 score and Revenue Capture. We use these measures for individual component ablation analysis and then for overall model performance. The *F1 score* is the harmonic mean of precision and recall. *Precision* refers to the proportion of customers who upgraded compared to those whom the model identified as potential upgraders. We use a strict definition of a match where the customer and upgrade offer range must match. *Recall* measures the proportion of customers who upgraded and were identified by our model. The F1 score provides holistic information about both of these measures.

As the second parameter used for evaluation, *Revenue Capture* measures the ratio of the revenue obtained by the airline's current heuristic model to PREM. This is an appropriate metric for our evaluation that was constrained to be conducted on a historical dataset. Hence, our primary goal was to show that PREM would not result in the loss of any income. A ratio of 100% means that PREM would have obtained *at least* the same revenue as the current approach and potentially much more due to better targeting. We evaluate PREM's architectural components through an ablation analysis, showing that these components are critical for enhanced performance results.

6.1. Ablation analysis

6.1.1. Quality of feature embeddings

We evaluate the quality of the embeddings produced by denoising autoencoders. We benchmark with alternative approaches for obtaining the embeddings of traditional autoencoders that do not apply the denoising of swap noise and one-hot and label encoding, as shown in Table 2.

The results presented in Table 2 indicate that denoising autoencoders provide the best results. This is because this approach increases the robustness of the learned embeddings when compared to traditional embeddings. The swap noise increases the accuracy of the model by as much as 7%. The compactness of the embeddings (of both denoising and traditional autoencoders) is a key driver of performance. This can be seen by comparing the autoencoder output to traditional feature encodings, such as one-hot and label encoding, that exhibit a steep drop. The one-hot encoding suffers from an underfitting problem due to the large feature size. On the other hand, label encoding gives poor results because the values from the domain of categorical attributes do not have any inherent order. We also evaluated the use of Principal Component Analysis (PCA) for representing the mixed data in a compact vector to PREM's denoising autoencoder, as PCA represents a standard method for converting high-dimensional data to a more compact, low-dimensional representation. From Table 2, we can observe that the quality of the denoising autoencoder embeddings is superior to this approach, and the most likely reason is that we use a nonlinear model for learning the embeddings, whereas PCA is a linear model.

6.1.2. Impact of embedding size

A key hyperparameter is embedding size, which determines the vector outputted size by the autoencoder, representing the trip information as a fixed-length vector. We evaluate its impact in the next experiment. When using the autoencoder, one could set the size of the latent representation to an arbitrary number. Traditionally, this size is often set as a power of two. Table 3 shows the results of PREM for

² <https://www.gnu.org/software/glpk/>

Table 3

Results of embedding size analysis. As shown, 256 embeddings perform the best.

Embedding size	F1 score	Revenue capture
256	83.9	100%
512	81.8	100%
128	79.1	92%
64	70.2	77%
32	68.2	73%

different embedding sizes. We can see that an embedding size of 256 represents a sweet spot. Reducing the size results in lesser accuracy and revenue capture, while increasing it could potentially result in over-fitting the model.

6.1.3. Impact of the classifiers

In the next experiment, we vary the classifier used for learning the offer acceptance model. The goal of this experiment is twofold. First, we wish to show that our learned embeddings are of high quality and give good results, regardless of the classifier. Second, we wish to show that the improved performance is (mostly) due to the embeddings.

From [Table 4](#), we can see that a DL model using embeddings provides the best results. However, one could also use other traditional classifiers, such as logistic regression, support vector machine, or random forest, and pay only a minor penalty in the F1 score and almost none in revenue capture. Using the original data without the embeddings shows a steep drop in the F1 score for both DL-based and non-DL-based methods.

As shown in [Table 4](#), we evaluated alternate approaches for handling imbalanced data [69], such as the Synthetic Minority Oversampling Technique (SMOTE), oversampling, and the use of generative models such as Generative Adversarial Networks (GANs). We used random forests as the downstream classifier. From [Table 4](#), we can see that traditional approaches such as oversampling are outperformed by the GAN, where we generate synthetic data for the rare classes so that the training data is balanced. Our embedding-based approach outperforms each of these approaches.

6.1.4. Personalized upgrade model

The personalized upgrade model uses a binary autoencoder to segment customer bookings into K segments so that similar customers have similar embeddings. Once the segmentation is obtained, PREM estimates the response rate for each upgrade offer bucket. We consider two other segmentation approaches that also result in K segments. The first is based on K-Means that cluster all bookings into K distinct clusters, with $K = 7$. Our other baseline is a decision tree that tries to partition bookings using the Gini criterion. As shown in [Table 5](#), our approach gives the best results.

We ran K-Means for various K and chose the optimal one using the Elbow method. We observed that $K = 7$ provided the best results. Among the baselines, K-Means outperforms the decision tree-based approach. This is not surprising as K-Means computes the distance between two tuples based on all features, while its depth limits the decision tree. For

Table 4

Results of classifier/feature analysis. Deep learning/embeddings perform the best.

Classifier/Feature	F1 score	Revenue capture
DL / Embeddings	83.9	100%
Logistic Regression / Embeddings	74.7	97%
Support Vector Machines / Embeddings	78.3	98%
Random Forest / Embeddings	79.4	98%
Random Forest / Original data	56.3	73%
DL / One-hot encoding of original data	44.3	62%
Random Forest / SMOTE	53.4	46%
Random Forest / Oversample 1:10	30.2	48%
Random Forest / Oversample 1:100	42.3	59%
Random Forest / GANs	56.4	62%

Table 5

Results of customer segmentation analysis.

Customer segmentation	F1 score	Revenue capture
Binary AutoEncoder	83.9	100%
K-Means	64.9	79%
Decision tree	53.6	73%

example, if the decision tree has a depth of two, it uses, at most, two attributes (e.g., gender = M and nationality = USA).

6.1.5. Revenue maximizer

Next, we evaluate the revenue maximizer component. Recalling that we use an integer programming approach to select the best customers and the corresponding offers, a natural alternative is to use a greedy baseline that works as follows. First, the greedy algorithm computes each customer's expected revenue for each offer. Then, the Revenue Maximizer picks the best among them. For example, if Customer A accepts an offer of \$500 with a 0.5 probability and \$1000 with a 0.2 probability, then the expected revenues are \$250 and \$200, respectively. So, the Revenue Maximizer identifies that the best offer for Customer A is \$500, with expected revenue of \$250. Suppose there is another customer, Customer B, whose expected revenue is \$300, and the model needs to select one customer. Here, the greedy algorithm will pick Customer B. As shown in [Table 6](#), the revenue maximization of PREM (ILP) outperforms the greedy algorithm baseline.

6.1.6. PREM overall

We propose a novel multistage approach based on feature development, the offer acceptance model, a personalized upgrade model, and a revenue maximizer. It is worth investigating if one needs such a multistage model in the first place. Specifically, two questions are of interest: (a) *What would have happened if PREM used a single-stage classifier?* and (b) *What would have happened if PREM skipped the offer cost classification component or the revenue maximizer?* The results presented in [Table 7](#) address these questions.

As shown in [Table 7](#), the proposed PREM approach containing separate and sequential stages provides the best results. If one squeezes all of these stages into a single stage (i.e., if one trains a single classifier that takes all the bookings of a flight as input and returns the list of users and upgrade offers as output), then that gives the worst result. Splitting tasks into the different stages of an ML system clearly improves the performance in this context. Instead of the classifier trying to handle multiple objectives, each classifier in the PREM approach is targeted and focused on a single task, resulting in superior overall performance.

[Table 7](#) also indicates that all three components contribute positively. If one skips the revenue maximizer, then the accuracy drops marginally, but there is a steeper drop in revenue capture. Similarly, if one skips the offer classification and runs the upgrade offer determination for every user on the flight, there is a steep drop in accuracy but a smaller drop in revenue capture.

7. Discussion and implications

7.1. Discussion

For this research, we developed PREM as an ML information system for modeling upselling in an international airline company. The PREM model provides several advantages. It uses a multistage approach,

Table 6

Results of revenue maximization analysis using ILP.

Revenue maximization	F1 score	Revenue capture
ILP	83.9	100%
Greedy	49.2	56%

Table 7
Results of ablation analysis.

PREM variant	F1 score	Revenue capture
PREM (All components)	83.9	100%
PREM (1 component)	42.1	65%
Only Classification and Upgrade	80.3	74%
Only Upgrade and Revenue Maximizer	62.3	82%

employing a novel unsupervised embedding based on denoising autoencoders that are robust to various data errors. PREM also estimates the probability of a customer accepting many mutually exclusive upgrade offers. Estimating this probability is nontrivial and advantageous as it relies on customers' demographics and booking behavior, substantiating our initial premises. Furthermore, PREM uses an innovative binary autoencoder for reliably segmenting the customers in the presence of errors. Finally, PREM uses an ILP formulation to determine the subset of customers to send an offer and maximize the revenue.

A key design principle of PREM was to avoid relying on company-specific details of the airline's behavior. Each PREM component is designed to be generic and applicable to any airline company. The constructed features are based on data present with every airline company has or that can be obtained via publicly available APIs. PREM is also designed to be modular and extensible. Each component can be transparently upgraded with minimal impact on other components. Furthermore, the output of each component can be postprocessed (if necessary) by the airline employees to encode additional dynamic business requirements.

Returning to our research objectives, for RO1 (Identify customers who are likely to accept/decline a premium upgrade offer) and RO2 (Identify the price elasticity of those customers likely to accept an upgrade offer), the personalized selection of upgrade prices and marketing messages allows a much more granular way to target customers than simple rule-based approaches that a company often takes. Our experimental results show that PREM can successfully identify customers who are unlikely to accept upgrade offers, which is in itself a customer satisfaction enhancement [70], by not sending them unwanted emails. An analysis of historical data shows that PREM would have sent approximately a million fewer nonrelevant customer email messages, increased accepted offers, and possibly have significantly improved revenue. The results from PREM show promising outcomes for using ML to effectively leverage customer information [71] to increase company revenue by improving the offer-targeting process. We believe that this research stimulates a more rigorous understanding of customer upselling behavior in the airline industry and other fields.

Our analysis shows several nonobvious relationships among PREM factors that impact a customer's upgrade behavior [72], which has implications in a wide range of airline industry-related domains and situations, such as hotels or cruise bookings. The approach is suitable for upselling scenarios in any business domain that deals with multifaceted data such as customer attributes (i.e., different gender, age, and income levels) and destination attributes (i.e., different cities or countries). The complexity of this data can be handled with PREM while maintaining the business objective of maximizing revenue. Similar application domains can include, for example, hotels (e.g., businesses such as Hotels.com or Airbnb), experienced service providers, event ticket sellers (e.g., concerts or sports), and other service industries that provide tiered or dynamic pricing. The basic requirements would be (a) multifaceted data about customers and goods/services and (b) the need for upselling.

The results showcased so far are based on historical data. However, we also conducted a simulation exercise to understand the potential of PREM. Our partner airline currently contacts around 22% of its customers, with a conversation rate of 1.4%. So, we are interested in two key metrics: (a) the approximate increase in accepted upgrade offers and (b) the revenue improvement in conversion rate by reducing the number of contacts to customers who would decline and identifying customers

who would have upgraded but who had been missed by the heuristic currently used by the partner airline. One critical insight is that the denoising autoencoder used by PREM is also a generative model [73]. As an autoencoder learns to capture the structure of the data-generating density, one can generate samples from a trained model in an unbiased manner. The generated synthetic booking information from the autoencoder reflects real data. We generated 64.3 M synthetic trips commensurate with the statistics from Table 1. We used PREM to estimate the upgrade offers. The contacted customers probabilistic accepted or declined upgrade offer was examined in line with the behavior of their corresponding segmentation. As stated in Section 5.4, we used binary autoencoders to segment the customers and obtain the probability of acceptance within each segment. Then, we assumed that our synthetic customer would have a similar acceptance probability. PREM implementation results in 1.12 M (7.94%) fewer nonrelevant email messages being sent to customers who most likely would not upgrade. PREM also produces a 72.2 K (37.2%) increase in the number of upgrade offers accepted by sending them to customers who would have most likely accepted but did not receive the offer. Finally, given the median upgrade offer of \$1000, we can estimate an increase in revenue of \$72.2 M (37.2%) for the company.

7.2. Takeaways

7.2.1. Identification of upselling customer groupings

In this research, we identified three generic behavioral segments concerning customers' upgrade acceptance. These segments are as follows:

- Never Upgrades:** Customers usually never take an upgrade offer, and therefore the company should never send them an upgrade offer as it is nonrelevant for the customer in this segment. Approximately 84% of customers receiving offers are in this segment. A naïve approach to upgrade offers could send indiscriminate emails being sent to this group. PREM avoids such emails, resulting in a reduction of 1.12 million emails (38%) annually that can be used for other offers. The impact on revenue is unknown as it depends on other offers sent to these customers and the return on those offers. One would expect an increase in customer experience with fewer irrelevant emails.
- Upgrade Lovers:** Customers who usually never book premium unless offered an upgrade and almost always take an upgrade. Thus, the company should always target this segment as it is revenue enhancing and the upgrade offers are relevant to this segment. Approximately 1.3 percent of customers (receiving offers or not) are in this segment.
- Upgrade Lovers Lookalikes:** New customers without an upgrade offer history but who fit the profile of those customers who usually accept upgrade offers. These customers should most likely be identified upon booking, and a decision made whether or not to send them an upgrade offer. The impact on revenue is unknown as it depends on the number of customers in these segments and the number of offers sent. However, the majority of new customers will be *Never Upgrades Lookalikes*.

PREM can use these customer segments in the following ways. Using our feature vector, our classification stage component identifies and filters the *Never Upgrades* customers. Since they greatly outnumber the *Upgrade Lovers* customers, identifying them is often a comparatively easier task. However, we designed a conservative cost-sensitive approach that seeks not to miss any potential *Upgrade Lovers* customers. Missing out on a potential customer is economically much worse than sending an upgrade offer to a customer who is unlikely to accept it. The output of the classification stage is a superset of primarily *Upgrade Lovers* customers. PREM uses the upgrade offer component to estimate the likelihood for the *Upgrade Lover* customers to respond to various

upgrade offers and uses the revenue maximizer to choose the best upgrade offer for revenue-boosting. A key insight is that we use an ILP-based approach for identifying the best *Upgrade Lovers* customers. Finally, the *Lookalikes* can be considered in the same way that the cold-start problem is recognized in recommender systems. By creating combined demographic-behavioral segmenting through denoising and binary autoencoders, PREM can match these customers to one of the existing segments in order to make meaningful upgrade offers.

7.2.2. Developing robust ml models

A key challenge with real-world data is that it is often noisy and replete with errors. A common approach would be to expend a considerable amount of time on data cleaning. However, we found that the sources of data error are very complex, resulting in a combating rear-guard action. In many cases, many data errors were not fixable, such as errors in customer demographics. However, rather than expend futile effort on data cleaning, we sought to build a robust ML model for these errors. Specifically, we relied on three insights. First, most airline companies have a large amount of historical data. Instead of using them directly, we learn a more generic embedding model. Second, we intentionally built the model to be robust to various noises. Third, we used limited supervised data, such as those who accepted offers, to improve the embedding of important customers preferentially.

7.2.3. Building business-aware ml models

As typical for companies deploying customer information systems [7], the airline company is more interested in optimizing revenue that cannot be meaningfully mapped to traditional ML metrics, such as accuracy. Our final PREM component is not an ML model. Instead, it is an ILP-based optimization formulation that uses the output of ML models to make decisions according to business constraints. Therefore, the PREM architecture demonstrates how hybrid approaches that combine ML with optimization models allow us to get the best of both worlds by modeling customer behavior through ML models and then making appropriate decisions subject to business constraints using traditional optimization techniques.

Also, ML models such as PREM do not operate in a vacuum. Most companies have an existing workflow, and any upselling ML model must integrate well with that workflow. Hence, PREM cannot operate as a monolithic system that just outputs a set of customers to contact. It is much preferable to output intermediate results that the end-user could modify. We found that building a single model to predict the best customers to send upgrade offers to is suboptimal as this approach has lower accuracy and is not modular enough to accommodate the workflow of airline operations. The stage-by-stage approach used by PREM often results in simpler models that are easier to maintain, understand, and, if necessary, replace. The staged approach is also beneficial for iterative improvement during development and faster debugging during the deployment stage.

8. Limitations, strengths, and future research

As with all research, some limitations must be acknowledged. First, we use data from one company for one specific upselling product, so the findings may not transfer directly to other domains and products. However, the airline company is international, and the dataset is quite large and robust, spanning multiple years, which is a strength of the current study. Obtaining several datasets (especially from rival firms) would be nearly impossible for one research team. Nevertheless, given the system's nondependency on company-specific variables, we suspect the PREM model is transferable to other companies (i.e., other airline companies), domains (i.e., hotels, cruise, tourism), and products (i.e., rooms, experiences). However, future research would need to be carried out to investigate this premise.

Second, there might be cannibalization aspects inherent in the implementation, which is a potential downside of discounted upselling

offers [74]. Cannibalization occurs when customers start to expect a discount offer and therefore do not directly purchase the premium product. In our exploratory analysis, we investigated the possibility of this effect. Based on the low acceptance rate of these upselling offers, perhaps because economy travelers consider the offers risky [75], cannibalization is not an issue. However, this aspect needs further investigation, as prior research has reported that companies can over-promote offers [76]. Naïvely sending more upgrade offers to certain customer segments could result in potential cannibalization, where the customers learn to game the system. If these customers figure out that they will get an upgrade offer, they might wait rather than directly purchase the business class seat. These results in an intriguing consumer dynamic that we plan to investigate in future research.

Third, we examined the acceptance or nonacceptance of the upgrade offer as a binary classification problem, which poses a limitation as we know that customer behavior may be more complex. Although binary classification is a good starting point, future research could explore multiple classes for this problem and other techniques [77], including causal models to estimate the individual treatment effect (ITE) of price on acceptance. For example, future research could explore the effects of only offering upgrades to specific customers at specific times, leveraging the limited-quantity scarcity (LQS) and limited-time scarcity (LTS) aspects of customer behavior [78]. Finally, we did not consider the impact of email marketing messages' attributes in great detail, which would be a good element to explore in conjunction with a complete field study of PREM. Although we did consider the offer price that is contained in the marketing messages, future research could investigate the impact of the email marketing message on upselling offers.

9. Conclusion

Upselling acceptance consists of the interplay among booking, customer, and destination. To investigate upselling acceptance at scale, we tackled the problem of predicting upgrade willingness in the airline industry in relation to revenue. We developed a modular machine learning architecture that achieved substantial performance gains over the currently used rule-based approach using real customer data representing millions of bookings over three years. The empirical findings on our dataset indicate that customer populations can be divided into three segments based on their likelihood to accept the upgrade offer or not. Our work has implications for enhancing conversion rates and revenue via automated pricing and targeting.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Guillet Basak Denizci, Online upselling: moving beyond offline upselling in the hotel industry, *Int. J. Hosp. Manag.* 84 (2020), 102322 (2020).
- [2] Daifeng Li, Kaixin Lin, Xuting Li, Jianbin Liao, Ruo Du, Dingquan Chen, Andrew Madden, Improved sales time series predictions using deep neural networks with spatiotemporal dynamic pattern acquisition mechanism, *Inf. Process. Manag.* 59 (4) (2022), 102987, <https://doi.org/10.1016/j.ipm.2022.102987> (July 2022).
- [3] Kevin Williams, Dynamic Airline Pricing and Seat Availability, Social Science Research Network, Rochester, NY, 2020, <https://doi.org/10.2139/ssrn.3611696>.
- [4] Jehangir Bharucha, Effect of price elasticity of demand on airline tickets, *Asian J. Res. Bus. Econ. Manag.* 6 (2016) 1, <https://doi.org/10.5958/2249-7307.2016.00052.9> (January 2016).
- [5] Amy Gallo, A refresher on price elasticity, *Harv. Bus. Rev.* (2015). Retrieved March 24, 2022 from, <https://hbr.org/2015/08/a-refresher-on-price-elasticity>.
- [6] Roy D. Voorhees, John Coppett, New competition for the airlines, *Trans. J.* 20 (4) (1981) 78–85 (1981).
- [7] Dmitri Goldenberg, Pavel Levin, Booking.com Multi-destination trips dataset, in: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New

- York, NY, USA, 2021, pp. 2457–2462, <https://doi.org/10.1145/3404835.3463240>. Retrieved March 26, 2022 from.
- [8] Farnoosh Khodakarami, Yolande E. Chan, Exploring the role of customer relationship management (CRM) systems in customer knowledge creation, *Inf. Manag.* 51 (1) (2014) 27–42, <https://doi.org/10.1016/j.im.2013.09.001> (January 2014).
- [9] Xiaochun Lei, Ummul Hanan Mohamad, Aliza Sarlan, Mishal Shutaywi, Yousef Ibrahim Daradkeh, Hazhar Omer Mohammed, Development of an intelligent information system for financial analysis depend on supervised machine learning algorithms, *Inf. Process. Manag.* 59 (5) (2022), 103036, <https://doi.org/10.1016/j.ipm.2022.103036> (September 2022).
- [10] Verónica Feder Mayer, Glauber Eduardo de Oliveira Santos, Osiris Ricardo Bezerra Marques, Option framing for upselling tourism services: does cognitive availability prevent irrational choices? *Tour. Econ.* (2020), 1354816620949722 (2020).
- [11] Foster Provost, Tom Fawcett, *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking* (1st Edition ed.), O'Reilly Media, Beijing, 2013.
- [12] Shazia Jamsheed, Qualitative research method-interviewing and observation, *J. Basic Clin. Pharm.* 5 (4) (2014) 87–88, <https://doi.org/10.4103/0976-0105.141942> (September 2014).
- [13] Steven N. Durlauf, A framework for the study of individual behavior and social interactions, *Sociol. Methodol.* 31 (1) (2001) 47–87 (2001).
- [14] Lifang Peng, Weiguo Zhang, Xiaorong Wang, Shuyi Liang, Moderating effects of time pressure on the relationship between perceived value and purchase intention in social E-commerce sales promotion: considering the impact of product involvement, *Inf. Manag.* 56 (2) (2019) 317–328, <https://doi.org/10.1016/j.im.2018.11.007> (March 2019).
- [15] Juhar Ahmed Abdella, Nazar Zaki, Khaled Shuaib, Fahad Khan, Airline ticket price and demand prediction: a survey, *J. King Saud Univ. - Comput. Inf. Sci.* (2019), <https://doi.org/10.1016/j.jksuci.2019.02.001> (February 2019).
- [16] Rodrigo Acuna-Agost, Eoin Thomas, Alix Lhéritier, Price elasticity estimation for deep learning-based choice models: an application to air itinerary choices, *J. Revenue Pricing Manag.* 20 (3) (2021) 213–226, <https://doi.org/10.1057/s41272-021-00308-z> (June 2021).
- [17] Letizia Lo Presti, Giulio Maggiore, and Mara Mattana. 2019. Tourist approaches to sustainable destination: evidence from Eco-Kibbutz customer experiences. 125–128.
- [18] Alejandro Mottini, Alix Lhéritier, Rodrigo Acuna-Agost, and Maria A. Zuluaga. 2018. Understanding customer choices to improve recommendations in the air travel industry. Vancouver, Canada., 28–32.
- [19] Mary Magdaline EnowMbi Tarkang, Uju Violet Alola, Ruth Yunji Nange, Ali Ozturen, Investigating the factors that trigger airline industry purchase intention, *Curr. Psychol.* (2020), <https://doi.org/10.1007/s12144-020-00815-z> (May 2020).
- [20] Yang Wang, Lixin Han, Adaptive time series prediction and recommendation, *Inf. Process. Manag.* 58 (3) (2021), 102494 (May 2021).
- [21] Yanwu Yang, Panyu Zhai, Click-through rate prediction in online advertising: a literature review, *Inf. Process. Manag.* 59 (2) (2022), 102853, <https://doi.org/10.1016/j.ipm.2021.102853> (March 2022).
- [22] Hee-Woong Kim, Atreyi Kankanhalli, Hyun-Lyung Lee, Investigating decision factors in mobile application purchase: a mixed-methods approach, *Inf. Manag.* 53 (6) (2016) 727–739, <https://doi.org/10.1016/j.im.2016.02.011> (September 2016).
- [23] Bernard J. Jansen, Theresa B. Clarke, Conversion potential: a metric for evaluating search engine advertising performance, *J. Res. Interact. Mark.* 11 (2) (2017) 142–159 (January 2017).
- [24] Paul Covington, Jay Adams, Emre Sargin, Deep neural networks for YouTube recommendations, in: *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*, Association for Computing Machinery, New York, NY, USA, 2016, pp. 191–198, <https://doi.org/10.1145/2959100.2959190>. Retrieved May 30, 2021 from.
- [25] Sami Khenissi, Boujelbene Mariem, Olfa Nasraoui, Theoretical modeling of the iterative properties of user discovery in a collaborative filtering recommender system. Fourteenth ACM Conference on Recommender Systems, Association for Computing Machinery, New York, NY, USA, 2020, pp. 348–357.
- [26] Hoon S Choi and Charlie Chen. 2019. The effects of discount pricing and bundling on the sales of game as a service: an empirical investigation. 20, 1 (2019), 21–34.
- [27] Eyal Bialogorsky, Eitan Gerstner, Dan Weiss, Jinhong Xie, The economics of service upgrades, *J. Service Res.* 7 (3) (2005) 234–244, <https://doi.org/10.1177/1094670504271148> (2005).
- [28] Saravanan Thirumuruganathan, Soon-gyo Jung, Dianne Ramirez Robillos, Joni Salminen, Bernard J. Jansen, Forecasting the nearly unforecastable: why aren't airline bookings adhering to the prediction algorithm? *Electron. Commerce Res.* 21 (1) (2021) 73–100 (2021).
- [29] William Groves and Maria Gini. 2013. Optimal airline ticket purchasing using automated user-guided feature selection. In *Twenty-Third International Joint Conference on Artificial Intelligence*.
- [30] Chih Chien Chen, Zvi Schwartz, Timing matters: travelers' advanced-booking expectations and decisions, *J. Travel Res.* 47 (1) (2008) 35–42 (2008).
- [31] Chih Chien Chen, Zvi Schwartz, Room rate patterns and customers' propensity to book a hotel room, *J. Hosp. Tour. Res.* 32 (2008) 287–306 (April 2008).
- [32] Konstantinos Tziridis, Th Kalampokas, George A. Papakostas, Kostas I. Diamantaras, Airfare prices prediction using machine learning techniques, in: *2017 25th European Signal Processing Conference (EUSIPCO)*, IEEE, 2017, pp. 1036–1039.
- [33] Shu-Hsien Liao, Yu-Chun Chung, Wen-Jung Chang, Interactivity, engagement, trust, purchase intention and word-of-mouth: a moderated mediation study, *Int. J. Services Technol. Manag.* 25 (2) (2019) 116–137 (January 2019).
- [34] Liyin Jin, Yanqun He, Haiyan Song, Service customization: to upgrade or to downgrade? An investigation of how option framing affects tourists' choice of package-tour services, *Tour. Manag.* 33 (2) (2012) 266–275, <https://doi.org/10.1016/j.tourman.2011.03.005> (April 2012).
- [35] Raditha Hapsari, Michael D. Clemes, David Dean, The impact of service quality, customer engagement and selected marketing constructs on airline passenger loyalty, *Int. J. Qual. Service Sci.* 9 (1) (2017) 21–40 (January 2017).
- [36] Shiri D. Vivek, Sharon E. Beatty, Robert M. Morgan, Customer engagement: exploring customer relationships beyond purchase, *J. Mark. Theory Practice* 20 (2) (2012) 122–146, <https://doi.org/10.2753/MTP1069-6679200201> (April 2012).
- [37] V. Kumar, Lerzan Aksoy, Bas Donkers, Rajkumar Venkatesan, Thorsten Wiesel, Sebastian Tillmanns, Undervalued or overvalued customers: capturing total customer engagement value, *J. Service Res.* 13 (3) (2010) 297–310 (August 2010).
- [38] Steven M. Shugan. 2018. Strategic use of product enhancements: upgrades, add-ons, extras, and accessories. In *Handbook of Research on New Product Development*. 207–226.
- [39] Devon Delvecchio, Sanjay Puligadda, The effects of lower prices on perceptions of brand quality: a choice task perspective, *J. Prod. Brand Manag.* 21 (6) (2012) 465–474, <https://doi.org/10.1108/10610421211264946> (2012).
- [40] Ahmet Melih Selçuk, Zeynep Müge Avsar, Dynamic pricing in airline revenue management, *J. Math. Anal. Appl.* 478 (2) (2019) 1191–1217 (October 2019).
- [41] Xiaofeng Pan, Soora Rasouli, Harry Timmermans, Investigating tourist destination choice: effect of destination image from social network members, *Tour. Manag.* 83 (2021), 104217, <https://doi.org/10.1016/j.tourman.2020.104217> (April 2021).
- [42] Dario Bonaretti, Marcin Bartosiak, Tsz-Wai Lui, Gabriele Piccoli, Daniele Marchesani, "What can I(S) do for you?": how technology enables service providers to elicit customers' preferences and deliver personalized service, *Inf. Manag.* 57 (6) (2020), 103346, <https://doi.org/10.1016/j.im.2020.103346> (September 2020).
- [43] Rasha HJ Dierckx, *Airline Price Discrimination: A Practice of Yield Management or Customer Profiling?* University of Twente, 2013. B.S. thesis.
- [44] Yao Cui, A. Yeşim Orhun, Izak Duenyas, How price dispersion changes when upgrades are introduced: theory and empirical evidence from the airline industry, *Manage. Sci.* 65 (8) (2018) 3835–3852, <https://doi.org/10.1287/mnsc.2018.3117> (November 2018).
- [45] Steven L. Puller, Lisa M. Taylor, Price discrimination by day-of-week of purchase: evidence from the U.S. airline industry, *J. Econ. Behav. Organ.* 84 (3) (2012) 801–812 (December 2012).
- [46] Chien-Huang Lin, Jyh-Wen Wang, Distortion of price discount perceptions through the left-digit effect, *Mark. Lett.* 28 (1) (2017) 99–112 (March 2017).
- [47] Bernard J. Jansen, Soon-gyo Jung, Dianne Ramirez Robillos, Joni Salminen, Too few, too many, just right: creating the necessary number of segments for large online customer populations, *Electron. Commer. Res. Appl.* 49 (2021), 101083 (2021).
- [48] Bing Pan, Yang Yang, Forecasting destination weekly hotel occupancy with big data, *J. Travel Res.* 56 (7) (2017) 957–970 (September 2017).
- [49] Yen-Liang Chen, Yi-Hsin Yeh, Man-Rong Ma, A movie recommendation method based on users' positive and negative profiles, *Inf. Process. Manag.* 58 (3) (2021), 102531, <https://doi.org/10.1016/j.ipm.2021.102531> (May 2021).
- [50] Michal Kompan, Ondrej Kassak, Maria Bielikova, The short-term user modeling for predictive applications, *J. Data Semant.* 8 (1) (2019) 21–37 (March 2019).
- [51] Noora Al Emadi, Saravanan Thirumuruganathan, Dianne Ramirez Robillos, Bernard Jim Jansen, Will you buy it now?: Predicting passengers that purchase premium promotions using the PAX model, *J. Smart Tour.* 1 (1) (2021) 53–64 (2021).
- [52] Valarie A. Zeithaml, Katrien Verleye, Isabella Hatak, Monika Koller, Alexander Zauner, Three decades of customer value research: paradigmatic roots and future research avenues, *J. Service Res.* 23 (4) (2020) 409–432 (November 2020).
- [53] Stacey Mumbower, Laurie A. Garrow, Matthew J. Higgins, Estimating flight-level price elasticities using online airline data: a first step toward integrating pricing, demand, and revenue optimization, *Trans. Res. Part A: Policy Pract.* 66 (2014) 196–212, <https://doi.org/10.1016/j.tra.2014.05.003> (August 2014).
- [54] Andreas Knorr, Big data, customer relationship and revenue management in the airline industry: what future role for frequent flyer programs? *Rev. Integr. Bus. Econ. Res.* 8 (2) (2019) 14 (2019).
- [55] Floris Devriendt, Darie Moldovan, Wouter Verbeke, A literature survey and experimental evaluation of the state-of-the-art in uplift modeling: a stepping stone toward the development of prescriptive analytics, *Big Data* 6 (1) (2018) 13–41 (2018).
- [56] Seth Siyuan Li, Elena Karahanna, Online recommendation systems in a B2C E-commerce context: a review and future directions, *J. Assoc. Inf. Syst.* 16 (2) (2015), <https://doi.org/10.17705/1jais.00389> (2015) Article 2.
- [57] Ji Wu, Liqiang Huang, Jianliang Leon Zhao, Zhongsheng Hua, The deeper, the better? Effect of online brand community activity on customer purchase frequency, *Inf. Manag.* 52 (7) (2015) 813–823, <https://doi.org/10.1016/j.im.2015.06.001> (November 2015).
- [58] Pablo Sánchez, Alejandro Bellogín, On the effects of aggregation strategies for different groups of users in venue recommendation, *Inf. Process. Manag.* 58 (5) (2021), 102609, <https://doi.org/10.1016/j.ipm.2021.102609> (September 2021).
- [59] Duncan Simester, Artem Timoshenko, Spyros I. Zoumpoulis, Targeting prospective customers: robustness of machine-learning methods to typical data challenges, *Manage Sci* 66 (6) (2019) 2495–2522 (2019).

- [60] Pinar Ezgi Çöl and Şeyda Ertekin. 2021. Feature dimensionality reduction with variational autoencoders in deep Bayesian active learning. In *2021 29th Signal Processing and Communications Applications Conference (SIU)*, 1–4. DOI:<https://doi.org/10.1109/SIU53274.2021.9477979>.
- [61] Charles Elkan, The foundations of cost-sensitive learning, in: *Proceedings of the 17th International Joint Conference On Artificial intelligence - Volume 2 (IJCAI'01)*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2001, pp. 973–978.
- [62] John L.(John Luis) Wilson. 1995. The value of revenue management innovation in a competitive airline industry. Thesis. Massachusetts Institute of Technology. Retrieved March 29, 2022 from <https://dspace.mit.edu/handle/1721.1/11558>.
- [63] Lan Huong Nguyen, Susan Holmes, Ten quick tips for effective dimensionality reduction, *PLoS Comput. Biol.* 15 (6) (2019), e1006907 (June 2019).
- [64] Pierre Baldi. 2012. Autoencoders, unsupervised learning, and deep architectures. In *Proceedings of ICML Workshop on Unsupervised and Transfer Learning, JMLR Workshop and Conference Proceedings*, 37–49.
- [65] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, Pierre-Antoine Manzagol, Extracting and composing robust features with denoising autoencoders, in: *Proceedings of the 25th International Conference On Machine Learning (ICML '08)*, Association for Computing Machinery, New York, NY, USA, 2008, pp. 1096–1103.
- [66] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, Pierre-Antoine Manzagol, Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion, *J. Mach. Learn. Res.* 11 (2010) 3371–3408 (December 2010).
- [67] M. Muralikrishna, David J. DeWitt, Equi-depth multidimensional histograms, *SIGMOD Rec.* 17 (3) (1988) 28–36, <https://doi.org/10.1145/971701.50205> (June 1988).
- [68] Christos H. Papadimitriou, Kenneth Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Courier Corporation, 1998.
- [69] Guo Haixiang, Li Yijing, Jennifer Shang, Gu Mingyun, Huang Yuanyue, Gong Bing, Learning from class-imbalanced data: review of methods and applications, *Expert Syst. Appl.* 73 (2017) 220–239 (May 2017).
- [70] Badrul Sarwar, George Karypis, Joseph Konstan, John Riedl, Analysis of recommendation algorithms for e-commerce, in: *Proceedings of the 2nd ACM Conference on Electronic Commerce (EC '00)*, Association for Computing Machinery, New York, NY, USA, 2000, pp. 158–167, <https://doi.org/10.1145/352871.352887>.
- [71] Silvia Basile, Cristian Consonni, Matteo Manca, Ludovico Boratto, Matching user preferences and behavior for mobility, in: *Proceedings of the 31st ACM Conference on Hypertext and Social Media (HT '20)*, Association for Computing Machinery, New York, NY, USA, 2020, pp. 141–150, <https://doi.org/10.1145/3372923.3404839>.
- [72] Yao Cui, Izak Duenyas, Ozge Sahin, Pricing of conditional upgrades in the presence of strategic consumers, *Manage. Sci.* 64 (7) (2018) 3208–3226, <https://doi.org/10.1287/mnsc.2017.2783> (July 2018).
- [73] Yoshua Bengio, Li Yao, Guillaume Alain, Pascal Vincent, Generalized denoising auto-encoders as generative models, in: *Proceedings of the 26th International Conference On Neural Information Processing Systems - Volume 1 (NIPS'13)*, Curran Associates Inc., Red Hook, NY, USA, 2013, pp. 899–907.
- [74] Bruce L Alford, Abhijit Biswas, The effects of discount level, price consciousness and sale proneness on consumers' price perception and behavioral intention, *J. Bus. Res.* 55 (9) (2002) 775–783 (September 2002).
- [75] Sandra M. Forsythe, Bo Shi, Consumer patronage and risk perceptions in Internet shopping, *J. Bus. Res.* 56 (11) (2003) 867–875 (2003).
- [76] Praveen K. Kopalle, Carl F. Mela, Lawrence Marsh, The dynamic effect of discounting on sales: empirical analysis and normative pricing implications, *Mark. Sci.* 18 (3) (1999) 317–332, <https://doi.org/10.1287/mksc.18.3.317> (1999).
- [77] Robin M. Gubela, Stefan Lessmann, Szymon Jaroszewicz, Response transformation and profit decomposition for revenue uplift modeling, *Eur. J. Oper. Res.* 283 (2) (2020) 647–661 (2020).
- [78] Yi Wu, Liwei Xin, Dahui Li, Jie Yu, Junpeng Guo, How does scarcity promotion lead to impulse purchase in the online market? A field experiment, *Inf. Manag.* 58 (1) (2021), 103283, <https://doi.org/10.1016/j.im.2020.103283> (January 2021).

Saravanan Thirumuruganatha is a Scientist at the Qatar Computing Research Institute, Hamad bin Khalifa University. His-current research focuses on various aspects of machine learning.

Noora Al Emadi is a Research Associate at the Qatar Computing Research Institute and is pursuing a Ph.D. degree at the College of Science and Engineering, Hamad bin Khalifa University. Her current research focuses on computational challenges in information design.

Soon-gyo Jung received a B.E. degree in computer software from Kwangwoon University and an M.S. degree in electrical and computer engineering from Sungkyunkwan University, Suwon, Korea. He works at the Qatar Computing Research Institute as an engineer, developing a system that generates personas automatically from online analytics and social media data.

Joni Salminen is a Scientist at the Qatar Computing Research Institute, Hamad bin Khalifa University, a Faculty Member at the Turku School of Economics, University of Turku, and a Faculty Member at the School of Marketing and Communication, University of Vaasa, Vasa, Finland. His-current research focuses on data-driven personas, including computational challenges, information design, and value creation with humanized data representations.

Dianne Ramirez Robillos is a graduate of the School of Statistics, University of the Philippines. She has extensive practical experience in the travel and tourism industries.

Bernard J. Jansen is a Principal Scientist at the Social Computing Group, Qatar Computing Research Institute, and a Professor at the College of Science and Engineering, Hamad bin Khalifa University.