A Time-aware Multi-task Learning Model for Customer Value Prediction in Civil Aviation

Haofei Yang1,3, Youfang Lin1,2,3, Zhihao Wu1,2,3+ and Yiji Zhao1,3

1 Beijing Key Lab of Traffic Data Analysis and Mining
2 Key Lab of Intelligent Passenger Service of Civil Aviation
3 School of Computer and Information Technology, Beijing Jiaotong University

Abstract. The precise prediction of customer value is essential for any successful dynamic customer relationship management (CRM) system. It is also the key for the company to maximizing customer returns. In this research, we concentrate on two main aspects of the work in civil aviation field. Firstly, a reasonable value model is the premise of this prediction issue. Therefore, we propose a parametric customer value model RFUM to estimate customer value in civil aviation. It evaluates customer value from four different attributes and then presents customer value by the weight of the attributes. Secondly, Time-aware Multi-task Value Prediction (TMVP) model is proposed to predict the future value of customer. It employs two supervisory signals of purchase propensity and customer value to better train a specific neural network to automatically learn features. Experiments demonstrate that the RFUM model can more accurately measure the value of customer in civil aviation market and the TMVP model can achieve a more precise regression prediction result. In addition, we also find that increasing the time of a single calculation window can improve the performance markedly.

Keywords: customer value, value prediction, multi-task learning, civil aviation.

1. Introduction

Customer value always plays a vital role in any company's development. It determines whether a company can survive and grow in the fierce market competition. By analysing customer value, enterprises can identify customer consumption power and levels, and distinguish different customer groups. This allows business policy makers to allocate limited resources reasonably, which can maximize business benefits and occupy market share. However, in the face of fierce market competition, it is not enough for companies to adjust market strategies just via analysing customers' historical value. Companies need the ability to predict customers' future value. With this ability, enterprises can predict which user will grow to a high value customer [1]; who may churn in the future [2]; and who can be provided value-added services? All of these can further enhance the market competitiveness of enterprises.

In recent years, China civil aviation market has been increasing at a rapid pace. As is shown in Fig. 1, the Civil Aviation Administration of China said that the transportation of civil aviation passengers has reached 550 million by 2017. Its potential ability for development has attracted 48 airlines to compete for the market share. However, some bad phenomena have caught airlines' attention. For example, the airlines' revenue is not increasing with the improvement of passenger transport volumes, showing a downward trend instead. The number of complaints is rising year by year. In such environments, airlines have to predict the possible change of customers' value to provide appropriate services and adjust the corresponding marketing strategies.

However, how to identify every entities’ value of civil aviation? How to predict the future value of the customer? It is not an easy task to solve these problems. First, the conception of customer value is varies.
with different industries. Some companies pay attention to users' one-time purchasing power, others focus on long-time spending ability. Second, the rationality of the definition of customer value determines whether it can be used to analyse further. Third, customer value is unstable and changing significantly. Customers' purchasing behaviours are not only influenced by users' various prior behaviours, but also affected by users' current work or living environment. Obviously, it would be a more difficult task to predict customers' future value. Although a variety of data mining technologies, including association rule, classification, clustering, regression and visualization, have been used to solve these problems, there still is not a perfect solution [3-8].

Fig. 1: Civil aviation passenger transportation in 2006-2017

In this research, we track and make two things on these problems. The first thing we did is building a parametric value model RFUM (Recency, Frequency, Unit revenue, Miles), where R is the last purchasing time in a specific period; F is the purchasing frequency in a specific period; U is the unit revenue of per kilometre in a specific period; M is the total kilometres. RFUM can be employed in civil aviation to evaluate customer value from the four-dimensional attribute. Then, Parameterization means that we convert the RFUM to a scalar value by using Analytic Hierarchy Process (AHP). The second thing we did is constructing a multi-task value prediction model TMVP, which considers the time dependence of user's behaviour and the relation between customer value and customer's purchasing propensity to monitor customer's future value by automatically learning the hidden feature representation of customer’s behaviours. Experiments demonstrate that both of TMVP are effective in the prediction problem of customer value prediction and purchase propensity.

The rest of this paper is organized as follows. Section 2 reviews the related work of the definition of customer value, value prediction model and a brief description of multi-task learning. Section 3 presents our customer value prediction model in detail. Section 4 gives the detail of experiment data, experiment setting and experiment results. Section 5 briefly concludes our work.

2. Related Work

This work is related to the area of value prediction problems, including basic the measurement model of customer value, customer prediction approaches and the conception of multi-task learning.

2.1. Customer value

The most direct form of customer value is customer lifetime value (CLTV), which is defined as the present value of the future cash flows attributed to the customer during his/her entire relationship with the company [9]. However, the indicator is hard to acquire. For a growing customer, the company can only define his/her historical value, and cannot know the future value of the customer.

Another customer value model is RFM [10], where R (recency) variable is defined as "last purchasing time in a specific period"; F (frequency) variable is defined as "purchasing frequency in a specific period" and M (monetary) variable is defined as "the sum of purchasing rice in a specific period". Although RFM does not explicitly provide a net profit for customer value, Fader et al. [11] indicated it can be used to build customer value. In addition, different industries can measure customer value from different view by using RFM's variants such as such as RFMDR [12], RFTM [13] and FSLC [14] in a more reasonable way. Meanwhile, RFM models show customer value in three-dimensional space or above, which makes it difficult for market decision makers to conduct marketing analysis. Therefore, Analytic Hierarchy Process (AHP) is
employed to convert RFM models into a scalar value [15, 16]. AHP proposed by Satty [17] is used to define the weights of different elements on the process of multi-object.

However, the concept of customer value in civil aviation is to some extent different from that in traditional industries. For example, airlines provide different travel classes: First Class, Business Class and Economy Class, whose prices are generally associated with service quality. Customers A and B are assumed to be frequent travellers of airline recently, where A is the passenger who enjoys the First Class discount fare, and B is the passenger who buys the ticket of Economy Class with original price. RFM will classify users A and B into same group. Obviously, RFM is an inappropriate classification method to correctly judge customer value in civil aviation market.

Therefore, in this research, we propose a revised value model RFUM, which takes the discount information of tickets into account. Experimentally, it can more reasonably measure customer value in civil aviation.

2.2. Value prediction model

There are many models, including probability models [18, 24], econometric models [19, 20], persistence models [21, 22], computer science models [23, 24] and diffusion/growth models [25], have been proposed to predict customer value. Gupta et al. [9] conducted a comprehensive review of related work. Compared with typically favored structured parametric models (logit, probit, or hazard models), machine learning literature have introduced many models with good predictive ability, such as support vector machine [24], multivariate decision tree [26]. However, most of these models focus on a binary classification problem to predict whether a customer would grow into a high-value customer [27, 28], and rarely study the regression prediction of customer value. In fact, the scalar value of customer value is gradually becoming the focus of enterprises’ attention. For example, Vanderveld et al. [7] innovatively proposed a two-stage task model with random forest to predict customers’ revenues, where the first stage is a binary classification to predict customer's purchase propensity, and the second stage predicts the dollar value for users who is predicted to purchase in stage 1. However, the two-stage model splits the correlation between purchasing propensity and payment amount by the means of one-to-one prediction.

For all models mentioned above, they remain little knowledge in the marketing. Obviously, feature extraction will play an important role in the entire predictive project. Therefore, a lot of manual data preprocessing work is required to contain some prior knowledge. Coussement et al. [29] compares the effect of different preprocessing methods on the prediction result, which shows the prediction result of these models depend on the rationality of the feature design. Chamberlain et al. [8] attempted to combine the artificial features with the fractional features automatically learned by using neural embedding to further enhance the prediction accuracy of customer value.

2.3. Multi-task learning

Caruana [30] summarizes the goal of Multi-Task Learning (MTL) succinctly: “MTL improves generalization by leveraging the domain specific information contained in the training signals of related tasks”. MTL has numerous advantage features: Implicit data augmentation, Attention focusing, Eavesdropping, Representation bias, and Regularization. Generally, MTL can effectively improve learning results and prediction accuracy, opposed to traditional single task learning. There are a large amount of literature about the application of MTL. Attenberg et al. [31] applying the principles of MTL, proposed an collaborative spam filtering that facilitates personalization. Chapelle et al. [32] applied MTL in web search ranking. Ruder and Sebastian [33] made a general overview of the current state of multi-task learning and related concept in deep learning. Past empirical work showed MTL has been used successfully across all applications of machine learning, from natural language processing [34] and speech recognition [35] to computer vision [36, 37] and drug discovery [38].

In this research, we propose a sequential multi-task learning model TMVP to predict customer value. Firstly, motivated by the success of deep learning to extract abstract features for customer churn [39], we propose a time-aware network to learn the fractional feature representation of customer value. Secondly,
considering the relationship between the binary classification prediction of customer growth and the regression prediction of customer value, we employ a multi-task mechanism to better learn a hidden features.

3. Methodologies

This section covers the steps undertaken in implementing the previously described objectives. It describes the reasoning behind the RFUM model as well as how it is realized. Furthermore, based on RFUM model, this section covers the idea detail of the proposed TMVP model through taking the time dependency of customer behaviours into account and the potential relation of customer purchase propensity and customer value.

3.1. The RFUM model

For industrial applications, it is challenging to develop a reasonably achievable customer value model. The simplicity, rationality and achievability are essential principles. Therefore, this is why RFM is popular in various industries. Given a customer \( c \), his/her current purchasing \( t \) and paid price \( p_t \), we can define \( C(p_t) = \{ p_t | t - T \leq \tau \} \) as the contextual set of customer's purchasing records, where \( \tau \) is a hyperparameter to control the size of the time window and \( T \) is the timestamp to calculate RFM. Then, RFM can be defined as the set of \( \{ R: \min(T - t), F: \sum_{i=0}^{T} I(p_i > 0), M: \sum_{i=0}^{T} p_i \} \). However, as mentioned in section 2, attributes of RFM model cannot distinguish the potential difference of customer value in civil aviation.

Therefore, a revised parametric value model RFUM is proposed in this work. In detail, we propose a new attribute \( U \) (unit revenue for per kilometre) to reflect the potential difference of customer value. For airlines, they must consider the take-off cost of each flight. Therefore, inspired the role of discount rate in CLTV, the discount rate based on Revenue for Average per Seat Kilometre (RASK, often used to evaluate the economic value for each flight) will be a good index to measure customer value. So, the attribute of \( U \) can calculate as follows:

\[
U = \frac{\sum_{i=1}^{F} RASK_i \times d_i}{\sum_{i=1}^{F} d_i},
\]

where \( F \) is purchase frequency in a specific period, \( r_i \) is the discount rate of \( t_{th} \) purchasing, \( RASK_i \) is the RASK of \( t_{th} \) execution flight, and \( d_i \) is the distance of \( t_{th} \) execution flight.

In addition, flight miles are also an important factor in determining the value of customer. While enjoying the same attribute of \( U \), long-distance passengers are more valuable than short-distance passengers. The product of \( U \) and flight miles can reflect the total amount that a customer contributes. Therefore, the attribute of \( M \) is replaced by a new conception, which can be represented as equation 2.

\[
M = \sum_{i=1}^{F} d_i.
\]

Combining with various attributes mentioned above, the RFUM model is formed naturally, where the concept of \( R \) and \( F \) used to describe the activity of user is as same as that in RFM.

Finally, just as AHP’s success in RFM models, a scalar representation of customer value will be more acceptable by market decision markers. We also utilize AHP procedure to analyze the attributes' weights of RFUM. The customer value can be reached via equation 3.

\[
CV = W_R \times R + W_F \times F + W_U \times U + W_M \times M,
\]

where CV is customer value, \( W_R, W_F, W_U \) and \( W_M \) are weights of RFUM variables.

3.2. The TMVP model

In civil aviation market, the prediction of customer value is also challenging in that the rationality of feature design and the balance of data distribution will affect the prediction result. Customer behaviours are both random and periodic. If time period is short, the customer’s behaviour is like a random behaviour, while the time window is long, his behaviour is like a periodic behaviour. This shows that the hidden information based on time window size and time window length plays an important role in customer value prediction. Naturally, due to the success of recurrent neural network (RNN) in sequence problems, it is possible to extract abstract feature of information to predict customer value.
There are connections between customer churn and customer value. In this research, the definition of customer churn is that customer does not purchase anything in a specific period. Customer churn means customer value will be zero. In other words, if customer value is not zero, it means that customer is not lost in the specific period. That is to say, the two problems are not independent with each other. Compared with the two-stage model \[7\] separating the relation of each other, it is possible to establish a multi-task learning mechanism to learn churn and customer value at the same time. Through training the two task promoting each other, we can get a better hidden feature representation used to predict customer value.

Thus, based on the above discussion, we propose TMVP model to forecast the customer future value. The model structure is shown in Fig. 2, which consists of two parts from bottom to top. In the bottom layer, \(x_t\) is a vector representing the basic attributes extracted from user’s behaviours. Considering that \(x = \{x_1, x_2, \cdots, x_t\}\) can be treated as sequential data, we use Long Short-Term Memory (LSTM) to generate hidden state \(h_t\) as shown in the following,

\[
    h_t = f_\theta(h_{t-1}, x_t),
\]

where \(h_{t-1}\) is the hidden feature vector of the previous moment, \(x_t\) is the input data of the current moment, and \(\theta\) denotes the network to be learned.

![Fig. 2: The architecture of this research proposed to predict customer value, consists of two part. The bottom part receives the sequential data of users’ behaviours and generate hidden feature vectors; The top part is two supervised signals, including the user’s willingness to purchase and the value of the purchase.](image)

Meanwhile, in order to patch the disadvantages of traditional RNN models, a simple version of soft-attention \[40, 41\] is employed in this research. It allows us to consider the dependent characteristic of the model without regard to their distances in the input or output sequences. The weight of each hidden state can be estimated by equation 5.

\[
    \alpha^{<t,i>} = \frac{g(h_i)}{\sum_{i=1}^{T_e} g(h_i)}, \tag{5}
\]

where \(g(\cdot)\) is a logit function to convert the hidden vector \(h_i\) into a numeric value.

Then we can calculate the final hidden features used to predict as follows:

\[
    context^{<t>} = \sum_{t_i=1}^{T_e} \alpha^{<t-i>} h^{<t_i>} \tag{6}
\]

In the top part, there are two supervisory signals used to train the model. The main signal is a customer value expressed by a scalar, which is used to train the top network by minimizing the Mean Square Error (MSE). It is denoted as

\[
    \text{Main}(c, \theta_m) = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{7}
\]

where \(c\) is the final hidden feature vector, \(\theta_m\) is the training parameters for the loss of MSE.

The auxiliary signal is a purchasing propensity, which is a binary classifier to predict yes/no response for whether customer will make a purchase at the next time window. The cross-entropy loss function is employed, written as follows:

\[
    \text{Auxi}(c, \theta_a) = -\frac{1}{N} \sum_{i=1}^{N} p_i \log \hat{p}_i + (1 - p_i) \log (1 - \hat{p}_i) \tag{8}
\]
where $p_i$ is the target probability distribution, $p_i = 0$ for all $i$ except $p_i = 1$ for the target consumer who will trade with the company, $\theta_a$ is the parameter to be learned in auxiliary loss.

Our goal is to learn the parameters $\theta$ in the function $f(\cdot)$ of extracting feature and $\theta_m$ in the function of customer value prediction, while $\theta_a$ is only parameters introduced to propagate the classification signal during training. These parameters are updated by stochastic gradient descent (SGD). The customer value and propensity gradients are weighted by a hyperparameter $\lambda$ in range 0 and 1. The algorithm is described in Table 1.

Table 1: The training detail of TMVP feature learning algorithm.

<table>
<thead>
<tr>
<th>input:</th>
<th>training set $X = {x_1, x_2, \ldots, x_t, y(t)}$, initialized parameters $\theta, \theta_m, \theta_a$, hyperparameter $\lambda$, learning rate $\eta(t)$, $t \leftarrow 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>while not converge do</td>
<td>$t \leftarrow t + 1$, sample training sample $(x_i, y_i)$ from $X$</td>
</tr>
<tr>
<td>calculate $c$ and loss function $\text{Main;} \text{Aux}$</td>
<td>$\nabla \theta_c = \frac{\partial \text{Main}(c, \theta_m)}{\partial \theta_m} + \lambda \frac{\partial \text{Aux}(c, \theta_a)}{\partial \theta_a}$</td>
</tr>
<tr>
<td></td>
<td>$\nabla \theta = \frac{\partial \text{context} \langle c \rangle}{\partial \theta} \cdot \nabla \theta_c$</td>
</tr>
<tr>
<td>update $\theta_c = \theta_c - \eta(t) \cdot \nabla \theta_c$, $\theta = \theta - \eta(t) \cdot \nabla \theta$</td>
<td></td>
</tr>
<tr>
<td>end while</td>
<td></td>
</tr>
<tr>
<td>output $\theta, \theta_m$</td>
<td></td>
</tr>
</tbody>
</table>

4. Experiment

4.1. Data set

We collected a data set\(^2\) from the encrypted civil aviation PNRs provided by the largest civil aviation IT provider in China. After simple data cleaning, the data set contains 2 million passengers’ flight records with 30 basic attributes of two years. As the theory proposed by Yan et al. [42], human activities are related to the region in which they are located and the economic level of the region. Therefore, we also employ an external data of every city’s category, region and Gross Domestic Product (GDP). Table 2 is the detailed statistics of the data.

Table 2: Data set description information

<table>
<thead>
<tr>
<th>Basic data</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>index</td>
<td>Description</td>
</tr>
<tr>
<td>time span</td>
<td>2 years</td>
</tr>
<tr>
<td>users</td>
<td>2 millions</td>
</tr>
<tr>
<td>features</td>
<td>demographics</td>
</tr>
<tr>
<td></td>
<td>purchasing behaviours</td>
</tr>
<tr>
<td></td>
<td>preferences</td>
</tr>
<tr>
<td>External data</td>
<td>City</td>
</tr>
</tbody>
</table>

Data is segmented by a fixed length of time split to generate the input vector and two supervisory signals. Taking a timestamp as the dividing boundary, customer’s historical behaviour data split by the length of time slice is used to construct temporally continuous input vectors, and the data of next time slice is used to generate the predictive goal of customer value and customer churn. Besides, in order to avoid the trained model can only be applied to a fixed time slice range, the training set and test set is built in different time slices. The logic can be demonstrated through Fig. 3.

\(^2\) The data set is available at http://insis.bjtu.edu.cn/file/
4.2. Evaluating customer value

In order to verify that RFUM can correctly distinguish customers with different values, we show the distribution of $U$ on four continuous time slices in Fig. 4. From the visualization results, it can be said that $U$ make the correct distinction among the vast majority of users with a ratio of over 95%. Meanwhile, as shown in Fig. 5, the distribution of the attribute $M$ is consistent with the statistical characteristics of the population activity in the literature [42]. RFUM model is affirmed by experts in civil aviation.

**Fig. 3:** Generating Training and Validation/Test sets using split data.

**Fig. 4:** The distribution of the $U$ of RFUM's attribute on four continuous time slices. The horizontal axis coordinate is the numerical value of $U$, and the vertical axis coordinate is the number of users.

**Fig. 5:** The distribution of the $M$ of RFUM's attribute on four continuous time slices. The horizontal axis coordinate is the number of user miles, and the vertical axis coordinate is the number of users.
Meanwhile, to convert customer value from high dimensional space RFUM to a scalar, AHP procedure is employed in this research. As shown in Table 3, in order to implement the AHP decision-making process, eight industry experts (5 managers and 3 scholars) in the tourism are kindly requested to fill out "1-9" scale pair-wise compare matrix of RFUM model for air travel's industry to collect professional information. According to the AHP's calculation process, the results computed from these questionnaires matches AHP acceptable condition. Therefore, as the RFUM variable's weight, they are \( R = 0.07, F = 0.20, U = 0.51 \) and \( M = 0.21 \). Therefore, the air traveller’s value can be computed by the proposed equation as follows:

\[
CV = 0.07 \times R + 0.20 \times F + 0.51 \times U + 0.21 \times M
\]

(9)

From the equation, we can see that the value of customer is ranging from 0 to 1. The numerical 0 means user does not have any purchase behaviour in a specific period, and the digit 1 represents the consumer is the most important user for the company.

Table 3: The RFUM questionnaire’s core: pair-wise comparison matrix

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>F</th>
<th>U</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>1</td>
<td>1/i</td>
<td>1/j</td>
<td>1/k</td>
</tr>
<tr>
<td>F</td>
<td>1/i</td>
<td>1</td>
<td>1</td>
<td>1/m</td>
</tr>
<tr>
<td>U</td>
<td>1/j</td>
<td>1/l</td>
<td>1</td>
<td>1/n</td>
</tr>
<tr>
<td>M</td>
<td>1/k</td>
<td>1/m</td>
<td>1/n</td>
<td>1</td>
</tr>
</tbody>
</table>

4.3. Evaluating the effectiveness of customer value prediction

In this context, to evaluate the effectiveness of TMVP model, we compared our model with 3 approaches, including 2 methods used in the market and one single-task TMVP model. (1) ARIMA: Autoregressive Integrated Moving Average Model, which is a famous persistence model to predict time series, where we use customer's historical value to predict customer future value. (2) TSRF: Two Stage Random Forest, which is proposed by Ali Vanderveld et al. [7] to predict customer value by using random forest with two stage. (3) TSVP: Time-aware Single-task Value Prediction Model, which has a prediction output of customer value in TMVP. In the process of experiment, the input information of all models comes from four consecutive time slices, each of which is 3 months in length. We find the optimal parameters of TMVP by using 10-fold cross validation. Meanwhile, Two metrics: MSE and Mean Absolute Error (MAE) is used to contrast experimental results. As shown in Table 4, the TMVP model gets a significant improvement of the prediction performance of customer value.

Table 4: The prediction result of customer value with different models: MSE & MAE.

<table>
<thead>
<tr>
<th>MODEL</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>0.052</td>
<td>0.46</td>
</tr>
<tr>
<td>TSRF</td>
<td>0.037</td>
<td>0.31</td>
</tr>
<tr>
<td>TSVP</td>
<td>0.016</td>
<td>0.12</td>
</tr>
<tr>
<td>TMVP</td>
<td>0.014</td>
<td>0.10</td>
</tr>
</tbody>
</table>

In detail, compared with other model, TMVP model reaches a better result with MSE 0.014 and MAE 0.10. Empirically, we think that human behaviour is unstable. However, ARIMA seriously relies on stability and periodicity of the time curve. Therefore, it performs the worst result of 0.052. As an ensemble model, random forest can generally reduce the variance of prediction to get a better predict result. In this experiment, it reach a better result 0.037 better than ARIMA. As a single-task version of TMVP, TSVP attempts to train the unique target of customer value, getting a better result than TSRF, but not as good as TMVP. Therefore, we think the process of feature extraction in TMVP can effectively get a hidden information of users. Furthermore, multi-task learning can capture a better customer's hidden state at a certain point in time.
4.4. Evaluating the effectiveness of classification

For a comprehensive analysis of the validity of the model, we also compare the binary classification result of purchasing propensity between TMVP model and another 3 approaches. (1) MC: the Markov Chain model which compute the next willingness to consume (0 or 1) based on the state transition matrix and current status. (2) TSRF: the two stage task model proposed by Ali Vanderveld et al. [7]. Each stage utilized the machine learning method of Random Forest. The first stage is employed to compare the result in this part. (3) TSWP: Time-aware Single-task Willing Prediction Model, a single output model of TMVP to predict customer’s purchasing propensity. Given the consumption behaviour of the first to fourth time windows, we use the purchase propensity of the fifth time window as the ground truth. Four metrics, including accuracy, precision, recall and F1 are employed in this research.

As shown in table 5, the model of TMVP gets a better result than other methods. As a comparison, TVMP reaches 73% accuracy in one fell swoop, improving 10% compared with random forest model and 2% than TSWP model. In other evaluation metric, TMVP also reaches better results. In other words, the hidden features got by TMVP can be used not only to predict customer value, but also to predict customer purchase intent.

Table 5: Classification performance: accuracy, precision, recall and F1 value.

<table>
<thead>
<tr>
<th>MODEL</th>
<th>Markov</th>
<th>TSRF</th>
<th>TSWP</th>
<th>TMVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>accuracy</td>
<td>58</td>
<td>63</td>
<td>71</td>
<td>73</td>
</tr>
<tr>
<td>precision</td>
<td>56</td>
<td>68</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>recall</td>
<td>62</td>
<td>64</td>
<td>66</td>
<td>67</td>
</tr>
<tr>
<td>F1</td>
<td>59</td>
<td>66</td>
<td>69</td>
<td>71</td>
</tr>
</tbody>
</table>

4.5. Evaluating performance on different size of time window

As shown in table 6, we also train TMVP model on different size of time window. When model trained on another time window of 6 months, in evaluate metrics of accuracy, recall and F1, the results are better than those of 3 months. Especially recall achieves a significant improvement. Actually, it is not hard to understand the phenomenon that personality behaviours would be more time-regulated in a long-term window.

Table 6: Model performs on different size of time window.

<table>
<thead>
<tr>
<th>Time window</th>
<th>3 months</th>
<th>6 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>73</td>
<td>77</td>
</tr>
<tr>
<td>Precision</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Recall</td>
<td>67</td>
<td>83</td>
</tr>
<tr>
<td>F1</td>
<td>71</td>
<td>78</td>
</tr>
<tr>
<td>MSE</td>
<td>0.014</td>
<td>0.014</td>
</tr>
</tbody>
</table>

5. Conclusion and Future Work

Taking everything into account, it can be said that the general aim of this research has been achieved. In this research, we propose a parametric value model RFUM, which can be used in marketing of civil aviation. Furthermore, we investigate a prediction approach TMVP to predict the future value of customer. Given vast historical behaviours of customers, the approach can train purchase willing and customer value at same time, getting a better result.

In the future, we plan to enhance RFUM metric and prediction model in several ways. First, we try to transfer RFUM model to other economic field. There are a lot of flexibility among R, F, U and M. It is not only used to measure customer value in civil aviation market, but also can be utilized in other industries. Second, more reasonable attributes may improve predictive performance. Third, the addition of an unsupervised generative preprocessing architecture, might have improved the prediction result.
6. References


