International Journal of Computer Trends and Technology
ISSN: 2231 – 2803 / https://doi.org/10.14445/22312803/IJCTT-V70I2P104
Volume 70 Issue 2, 22-28, February 2022

Original Article

K-Medoids Clustering Techniques in Predicting Customers Churn: A Case Study in the E-Commerce Industry

Xiancheng Xiahou[1], Yoshio Harada[2]

[1]*Doctoral candidate, Faculty of Business Management, Osaka Sangyo University, Osaka, Japan*
[2]*Professor, Faculty of Business Management, Osaka Sangyo University, Osaka, Japan*

Received Date: 11 January 2022
Revised Date: 12 February 2022
Accepted Date: 25 February 2022

Abstract — *As the churn of customers can result in the decline of business performance, preventing customer churn becomes an important task for e-commerce enterprises. This paper studies customer churn by using shopping data sets of e-commerce customers and machine learning technology. Firstly, a k-medoids clustering algorithm is adopted to cluster customers and divide them into three types. Then we adopt Logistic Regression, Support Vector Machines and Adaboost models to predict the churn of these three types of customers. The results show that the k-medoids algorithm can accurately identify churn and non-churn customers and that the Adaboost model boasts good prediction performance. The research results can guide enterprises to enhance customer relationship management and formulate appropriate marketing strategies and provide a critical theoretical basis for the research of customer churn prediction.*

Keywords — *Customers churn, E-commerce industry, K-medoids clustering algorithm, Machine learning, Predicting technology.*

I. INTRODUCTION

Along with the development of e-commerce enterprises, the increasing number of e-commerce websites provides customers with more choices. Thus, they can compare products among different e-commerce enterprises[1]. The churn of customers caused by fierce market competition is a practical problem that enterprises must cope with[2]. To cope with this issue, many enterprises' marketing activities shifted from being product-oriented to customer retention and customer churn reduction[3]. Customer relationship management is an important management aspect of enterprises, and advanced statistical methods and machine learning technology must be adopted to support market operation decisions. Customer churn prediction technology is an important tool to help enterprises carry out customer retention measures[4],[5]. The strategic focus of enterprises is to predict customer churn behaviour and analyse the causes, establish a strong customer relationship, and retain churn customers. In the era of big data and artificial intelligence, it is easy for enterprises to store shopping transaction data, establish churn prediction models and identify potential churn customers by using machine learning technology, customer purchase data and shopping behaviour characteristics, which provide a new channel for enterprises to understand customers' consumption behaviours and shopping habits and thus devise countermeasures for customer churn.

Researchers have made great achievements in the research on churn prediction, as reflected in the telecommunications industry[6]-[9], the financial industry[10]-[12] and the e-commerce industry[13]-[16]. Reference[17] used 30,104 customer data of European telecom operators to analyze customers' call behaviour, interaction behaviour between customers and operators, package subscription and demographic characteristics and predicted customer churn through the Logit model. The research results show that the Logit model can effectively predict customer churn. Reference[12] extracted 5456 credit cards and predicted the customer churn problem using Logit and Decision Tree algorithms. The results show that these two algorithms can accurately identify non-churn customers and churn customers and reduce the churn rate of bank customers. Reference[18] used Logit, Decision Tree, Boosting and other models to study the customer churn of a B2B e-commerce platform in Australia. The results can help the enterprise accurately identify churn customers and non-churn customers so as to enhance the market competitiveness of the enterprise, improve the profits of the enterprise and reduce cost. To sum up, in this literature, machine learning algorithms such as Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM) and Random Forest (RF) can effectively predict customer churn and thus play a great role in formulating corresponding marketing strategies for managers.

However, the churn of customers in different industries feature different characteristics, mainly because their consumption habits and methods vary in different industries. The research fields of the above literature are more related to telecommunications and finance but less so to B2C. With the rapid development of Internet technology, shopping websites now provide more commodity types, shopping forms and settlement methods, which makes it easy for consumers to compare products, make shopping choices and complete the shopping[19]. The current e-commerce websites are user friendly with complete functions, including commodity details, commodity price, transportation expenses, delivery and receipt time, payment method, invoice management, communication with merchants, commodity collection, commodity shopping cart, commodity evaluation and other functions. The diversity of these shopping data and the verticality of shopping time increase the difficulty and complexity of churn prediction. Reference[20] adopted the time series method to incorporate the consumption time variables into the prediction modelling, clustered the customers in the time dimension, and at the same time, they incorporated the clustering information into the classification model to explore the consumption behaviour with vertical time dimension variables. Reference[21] used a Machine Learning classifier to predict and model customer churn through time period data and achieved satisfactory prediction results. In addition, e-commerce customers are non-contractual, which means that customers are not bound by consumption agreements and enterprises. And due to the popularity of mobile phones, consumers can change goods, compare and choose among e-commerce enterprises anytime and anywhere[22], which makes it difficult for e-commerce enterprises to accurately identify the reasons and time for the churn of B2C customers and whether it is necessary to continue implementing commodity promotion activities or retention measures. The existing literature provides no unified experience on how to reasonably define churn customers.

It has been rendered quite easy for corporate managers to extract consumers' shopping data through cutting-edge technologies such as big data and machine learning. In order to improve business performance and reduce costs, it has become a new development trend for enterprises to make full use of consumer shopping data to conduct customer behaviour analysis, prediction modelling and marketing strategy improvement. Based on the real data of a B2C e-commerce platform, this paper carries out the research on customer churn prediction through data processing, customer classification, feature analysis and modelling. By using the k-medoids clustering algorithm and LR, SVM and Adaboost algorithms, we conduct an empirical study on customer classification and churn prediction. By analyzing and comparing the experimental results, the most suitable prediction model is determined to provide a theoretical basis for enterprises in formulating marketing strategies so as to promote business performance.

This paper is structured as follows: the next section is a review of the previous work, including the application of clustering algorithm and prediction model. Section 3 introduces the methodology, including the basic principles of LR, SVM and Adaboost prediction algorithms; Section 4 describes the empirical research and prediction evaluation indicators; and section 5 presents the results and discussions; Section 6 draws a conclusion and puts forward some suggestions on customer relationship management.

II. LITERATURE REVIEW

Previous literature on the prediction of e-commerce customer churn has contributed to customer relationship management. Reference[23] studied the data of 80,000 customers of an unidentified e-commerce company. The data variables include customer registration information, transaction information and access information, and adopted the improved SVM for prediction. Reference[24] used the registration information, login, transaction and website log information of 50,000 customers of an e-commerce website to build a model and used the improved SVM for prediction. Reference[25] used Logit, DT, Boosting and other models to study the customer churn of a B2B e-commerce platform and found that the Boosting model has better prediction performance. Reference[26] used the data of 6,432 customers of a software supplier and used the uplift Logit leaf model to predict and achieve good results. Reference[27] used the data of 1,968 customers of a parcel mailing company and used linear SVM for customer prediction, and achieved good results. Reference[28] used an improved SVMacc model for the prediction of 80,000 customers of a fast-moving consumer goods company. The results show that the SVMacc model boasts high prediction accuracy.

The performance of the model can directly affect the formulation of enterprise marketing strategy. When evaluating the performance of the prediction model, the most commonly used evaluation indicators include Accuracy, Precision, Recal1, Sensitivity, Specificity, Receiver Operating Characteristic curve (ROC) and AUC value. Reference[29] used the AUC index to evaluate the effect of the Logit model, and Reference[30] used Accuracy, AUC and Top-decile Lift. Reference[31] and Reference[32] used the Top-decile Lift and Gini coefficients as measurement indicators. Reference[33] used Accuracy, Sensitivity, Specificity and AUC as evaluation indicators. Reference[34] comprehensively evaluated Logit, DT and Boosting models by using ROC and Lift indicators.

A large number of literature research shows that the models used for customer churn prediction are mainly divided into three types: prediction model based on traditional statistics, machine learning-based prediction model and prediction model based on the integrated classifier. Due to different customer attributes and data variables, the prediction results are often inconsistent. In previous literature, customer attributes were mostly contractual customers, that is, customers need to sign contracts with enterprises before consumption, such as customers of the bank, telecom, insurance and so on, while B2C e-commerce customers are non-contractual customers who can consume freely without signing a contract with the company and are thus not constrained by the company, which makes it difficult for companies to predict the loss of e-commerce customers. Therefore, this paper uses B2C e-commerce customer data that adopts k-medoids and three prediction models to conduct empirical research. We compare and analyze the three prediction models and provide an experimental basis for enterprise customer relationship management and marketing strategy formulation.

III. METHODOLOGY

Customer churn is defined as a phenomenon that customers give up using a certain product or service and instead use the product or service of another rival enterprise in the market[35]. Customer churn prediction is by nature an issue of binary classification, namely, churn customers or non-churn customers. In previous literature, researchers used the k-means algorithm to cluster. The centre point of the cluster selected by this algorithm is the centre of gravity of all points in the current cluster. However, since the criterion function of K-means for each cluster is a square error, when extreme outliers appear in the data, it will lead to an error in clustering results. The central point selected by the K-medoids algorithm is a point existing in the current cluster, and the criterion function is the minimum sum of the distances from all other points in the current cluster to the central point. Therefore, K-medoids can weaken the influence of outliers. Although the k-means algorithm consumes more computing time in clustering than K-means does, what enterprises desire is higher clustering accuracy. Therefore, this paper uses a k-medoids algorithm for clustering and completes the prediction task on the basis of clustering.

A. K-medoids

In the R language, the process of K-medoids clustering can be specified as follows:

Step 1: initializing centroids from k randomly selected data points from n objects;

Step 2: identify the remaining object and the centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized;

Step 3: update the centroid of each cluster according to the principle of square difference function value reduction;

Step 4: repeat step 2 and step 3 until each cluster remains constant;

The square difference function is,

$$w(c)=\sum_{j=1}^{k}\sum_{p\in Ci}\left|p-o_j\right|^2 \tag{1}$$

B. Logistic Regression

Logistic Regression is a common method for completing classification tasks in machine learning, and it is a generalized linear model. The function formula is,

$$y=g^{-1}\left(w^{\mathrm{T}}x+b\right) \tag{2}$$

The output label y∈{0，1}, the predicted value $z=wx+b$ generated by the linear regression model is the real value; convert the real value Z to 0/1 value, and the ideal is the unit step function. The function formula is,

$$y=\begin{cases}0, & z<0;\\ 0.5, & z=0;\\ 1, & z>0;\end{cases} \tag{3}$$

That is, if the predicted value z>0, it is judged as a positive example; When z<0, it is judged as a counterexample; When z=0, it can be judged arbitrarily. However, the step function is not a continuous function, so we use the sigmoid function for judgment. The Sigmoid function formula is,

$$y=\frac{1}{1+e^{-z}} \tag{4}$$

C. Support Vector Machines

Support Vector Machines is a fast and reliable linear classifier. After the training samples are given, SVM establishes a hyperplane as the decision surface to maximize the isolation boundary of positive and negative examples. The classifier solves the quadratic optimization problem and applies it to the nonlinear classification through the transformation of the kernel function. SVM transforms the original problem into a dual problem that is easier to solve by the Lagrange multiplier method. The function formula is,

$$L(w,\ b,\ \alpha)=\frac{1}{2}\|w\|^2+\sum_{i=1}^{m}a_i\left(1-y_i\left(W^T x_i+b\right)\right) \tag{5}$$

We convert the original problem into the following objective function for solution:

$$max\sum_{i=1}^{m}a_i-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}a_i a_j y_i y_j x_i x_j \tag{6}$$

$$\sum_{i=1}^{m}\alpha_i y_i=0,\quad \alpha_i \geq 0,\quad i=1,2,\ldots,\ m \tag{7}$$

D. Adaboost

Adaboost is an iterative algorithm whose core idea is to train different weak classifiers for the same training set and then collect the weak classifiers to form a stronger final classifier (strong classifier). Assume the training set sample is,

$$T=(x_1, y_1),(x_2, y_2),\cdots,(x_m, y_m) \tag{8}$$

The input sample weight of the training set in the k-th weak classifier is,

$$D(k)=(w_{k1}, w_{k2},\cdots,w_{km}) \tag{9}$$

$$w_{1i}=\frac{1}{m}; \quad i=1,2,\cdots, m$$

The weighted error rate of the k-th weak classifier Gk(x) on the training set is,

$$e_k=P(G_k(x_i)\neq y_i)=\sum_{i=1}^{m} w_{ki} I(G_k(x_i)\neq y_i) \tag{10}$$

Adaboost is the final strong classifier obtained by weighting several weak classifiers. The strong classifier in the k-th round is,

$$f_k(x)=\sum_{i=1}^{k}\alpha_i G_i(x) \tag{11}$$

IV. EMPIRICAL STUDY

A. Data

The original data of this paper is a data set published by the Alibaba Cloud Tianchi platform[36], with a total of 100,150,807 pieces of data extracted from the purchase behaviour of 987,994 customers. The original data includes five indicators: User ID, Item ID, Category ID, behaviour type and Timestamp. There are four kinds of Behavior types: PV (Page view of an item's detail page, equivalent to an item click), Buy (Purchase an item), Cart (Add an item to the shopping cart) and Fav (Favor an item). Customers' shopping activities took place between November 23, 2017, and December 4, 2017. The data period is divided into two stages: the first 6 days is the observation period, and the rest 6 days are regarded as the verification period. Customers who made purchases more than once during the observation period and purchased more than once again during the verification period are defined as non-churn customers, represented by label=0, while customers who purchased more than once during the observation period but did not make any purchase during the verification period are defined as churn customers, represented by label=1. We use the R language for statistical grouping. First, we group the customers according to user ID, then calculate the purchase times of each customer in the observation period and verification period, and retain the customers who meet the screening conditions. To facilitate calculation, we reserve the data of 30,000 customers for subsequent classification and prediction, including 26,490 churn customers, accounting for 88.3% and 3,510 non-churn customers, accounting for 11.7%.

In order to study customers' shopping behaviours in different periods, the time of shopping behaviour is defined as follows: 00:00-06:00 as daybreak; from 06:00 to 12:00 as AM; 12: 00-18:00 as PM; 18: 00-00:00 as night. There are 23 indicators of data variables, including label, User ID, Category, Buy, PV, Cart, Fav, Daybreak PV, Daybreak Buy, Daybreak Cart and Daybreak Fav; AM PV, AM Buy, AM Cart, AM Fav; PM PV, PM Buy, PM Cart, PM Fav, Night PV, Night Buy, Night Cart, and Night Fav. Among them, the label variable is factor type, which is 0 and 1, respectively. 1 stand for churn customers and 0 non-churn customers. Other variables are numerical.

B. K-medoids Clustering Analysis

We removed two-factor variables and four overlapping numerical variables (Buy, PV, Cart, Fav), and the other 17 variables were used for K-medoids cluster analysis. In this paper, k=3 is taken; that is, the customer types are divided into three types: cluster I, cluster II and cluster III. After clustering, we observed the number of churn and non-churn customers in each type. The distribution of the three types of customers is shown in Fig.1. Cluster I has 17,160 customers, of which the number of non-churn customers is 3157, accounting for 18.39% of the total number of cluster I, the number of churn customers is 14,003, and the churn rate is 81.6%; Cluster II has 10,440 customers, of which the number of non-churn customers is 744, accounting for 7.13% of the total number of cluster II, the number of churn customers is 9,696, and the churn rate is 92.87%; Cluster III has 2,400 customers, of which 157 are non-churn customers, accounting for 6.54% of the total number of cluster III, and 2,243 are churn customers, with a churn rate of 93.45%.

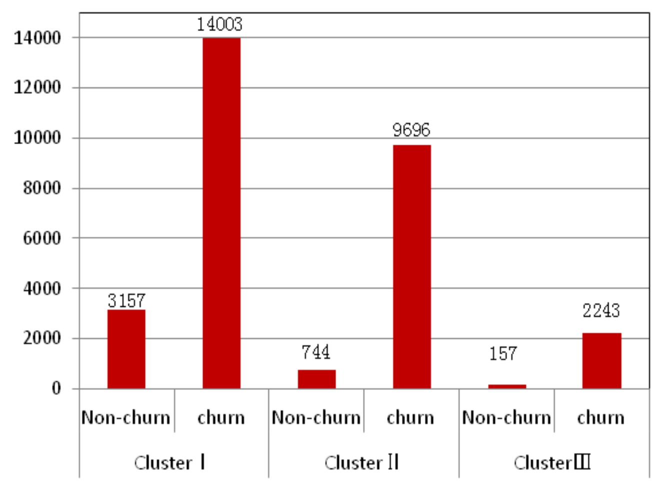

Fig. 1 Number of three types of customers

C. Variable Selection

There are 17 variables in cluster analysis. If there are too many variables, it may lead to miscellaneous information, so we use the RF algorithm to screen the variables. The screened variables will be used as the characteristics of the prediction model in the prediction experiment. When RF filters variables, m features are randomly selected from n variables each time to construct a tree. In this way, we can not only increase the randomness of the tree but also better fit the data and realize optimal classification. We use out-ofbag error (OOB error) to determine the number of features[37]. When constructing each tree, we use different bootstrap samples for the OOB calculation of the training set. To facilitate calculation, the numbers of variables we select are 2, 3, 6, 9, 12, 15 and 17, respectively. The calculated OOB error value is shown in Table 1.

Table 1. OOB error

Number of features	2	3	6	9	12	15	17
OOB error	0.221	0.222	0.222	0.223	0.222	0.223	0.225

Table 1 shows that the difference between OOB errors is only about 0.001. The OOB error is the smallest When the number of variables is 2. Therefore, a Random forest with parameter entries as 2 and three as 500 is established, and then the Gini index of each variable is calculated. Gini index can distinguish the importance of variables. The greater the Gini value, the stronger the importance of variables[38]. The columnar distribution of the importance of each variable is shown in Fig.2, which clearly shows that the Gini values of Night Buy, PM Buy and PM PV are in the top three, 631.287, 594.386 and 565.18, respectively. Therefore, these three variables are selected as the characteristics of the prediction model.

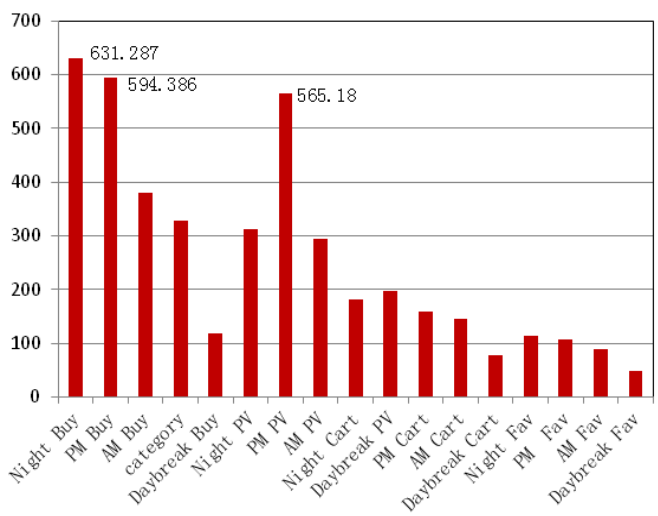

Fig. 2 Importance of random forest variable selection

D. Evaluation Metrics

In this paper, we mainly use Accuracy, Recall and Precision to evaluate the performance of the prediction model and draw the receiver operating characteristic curve (ROC). The model is comprehensively evaluated according to the area under the receiver operating curve (AUC) [39]. The calculation formula of the three indicators is,

$$\text{Accuracy}=\frac{TP+TN}{TP+FN+FP+TN} \tag{12}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{13}$$

$$\text{Precision}=\frac{TP}{TP+FP} \tag{14}$$

Whereby TP stands for True Positive; TN stands for True Negative; FP stands for False Negative, and FN stands for False Negative.

V. RESULTS AND DISCUSSION

We input the data into LR, SVM and Adaboost models respectively for prediction and obtained the Accuracy, Recall, Precision and AUC values. We list TP, TN, FP, FN and various predicted values (Accuracy, Recall, Precision and AUC) in the confusion matrix table (table 2-table 4). Symbols A, R and P in tables 2-4 represent Accuracy, Recall and Precision, respectively. The symbols I, II and III represent cluster I, cluster II and cluster III, respectively. Through empirical research and analysis, we can get the following results:

(1) Cluster I has 17,160 customers, including 3,157 non-churn customers, accounting for 18.39% of the total number of cluster I, and 14,003 churn customers, and the churn rate is 81.6%; Cluster II has 10,440 customers, of which the number of non-churn customers is 744, accounting for 7.13% of the total number of cluster II, and the number of churn customers is 4821, and the churn rate is 92.87%; Cluster III has 2,400 customers, of which the number of non-churn customers is 157, accounting for 6.54% of the total number of cluster III, and the number of churn customers is 2243, and the churn rate is 93.45%.

(2) According to the clustering analysis results in part IV (B), we classify customers into cluster I, cluster II and cluster III. Comparing the churn rates of these three types of customers, we can see that the churn rate of class I customers is 81.6% but compared with cluster II and cluster III customers, the non-churn rate of cluster I customers is 18.39%, while the number of cluster III customers is the least, only 2400, and the churn rate is the highest, reaching 93.45%. This result means that class I customers may be important customers for enterprises and that enterprises should focus on such customers. In addition, k-medoids is very effective in clustering customers and can accurately identify churn customers and non-churn customers. These research results boast good reference value for e-commerce enterprise managers in data analysis and prediction modelling.

(3) Table 2 to table 4 show that the prediction accuracy of the Adaboost model for three types of customers is higher than that of LR and SVM models, and the prediction performance of the Adaboost model is good. By observing the Accuracy, Recall and Precision of the three models, it can be seen that the three indexes of the Adaboost model are 0.9612, 0.9623 and 0.9285, respectively, while those of the LR model are 0.907, 0.908 and 0.904, respectively; The SVM models are 0.9124, 0.9117 and 0.9056, respectively. The results show that the prediction performance of the Adaboost model is better than that of the LR model and SVM model.

(4) We use the AUC index to evaluate the generalization ability of the three models. From table 2 to table 4, it can be seen that the AUCs of the three types of customers in the Adaboost model are 0.9887, 0.9869 and 0.9807, respectively. Compared with the AUCs of the LR model and SVM model, the AUC of the Adaboost model is the largest, indicating that the Adaboost model has good generalization ability; that is, the Adaboost model boasts sound adaptability to other types of data.

Table 2. LR confusion matrix

Actual		Predicted FP	Predicted FN	A (%)	R (%)	P (%)	AUC
Ⅰ	TP	35	27	90.7	85.11	90.88	0.9487
	TN	29	73				
Ⅱ	TP	132	25	90.8	95.31	86.85	0.9225
	TN	58	253				
Ⅲ	TP	660	25	90.4	95.67	77.23	0.8889
	TN	127	368				

Table 3. SVM confusion matrix

Actual		Predicted FP	Predicted FN	A (%)	R (%)	P (%)	AUC
Ⅰ	TP	36	26	91.24	89.35	91.35	0.9793
	TN	29	73				
Ⅱ	TP	132	25	91.17	98.46	87.62	0.9494
	TN	56	256				
Ⅲ	TP	660	25	90.56	98.57	77.44	0.9015
	TN	127	368				

Table 4. AdaBoost confusion matrix

Actual		Predicted FP	Predicted FN	A (%)	R (%)	P (%)	AUC
Ⅰ	TP	36	26	96.12	91.25	95.83	0.9887
	TN	26	75				
Ⅱ	TP	125	32	96.23	92.64	95.14	0.9869
	TN	30	281				
Ⅲ	TP	626	59	92.85	93.75	91.62	0.9807
	TN	56	439				

VI. CONCLUSION

Taking the B2C e-commerce customer data set as the object, this paper studies Customer Churn by using the k-medoids clustering algorithm and three prediction models. The results show that k-medoids can accurately identify churn customers and non-churn customers and determine that cluster I customers are whom enterprises should focus on. Three important features are screened through the RF algorithm, namely, Night Buy, PM Buy and PM PV, indicating that enterprises should formulate targeted marketing strategies according to the shopping time, consumption habits and characteristics of cluster I customers. The empirical results show that the Adaboost model has the highest prediction accuracy and the highest AUC, indicating that the Adaboost model also has good adaptability to other types of data sets. Therefore, we suggest that B2C e-commerce enterprises adopt the Adaboost model to predict customer churn because this model is conducive to enhancing enterprises' customer relationship management and operating performance and reducing cost. However, the results of this paper also have some limitations. For example, the data variables do not include factors such as purchase amount and customer evaluation, which may have an impact on the research of customer churn. Therefore, in the following research, we will choose other data sets to carry out the prediction of multivariable customer data. In addition, the model can also be improved on the basis of the existing model to further improve the accuracy of prediction.

REFERENCES

[1] A.D.Rachid, A.Abdellah, B.Belaid And L.Rachid, Clustering Prediction Techniques In Defining and Predicting Customers Defection: The Case Of E-Commerce Context, International Journal of Electrical and Computer Engineering.8(4) (2018) 2367-2383.

[2] J.Hadden, A.Tiwari, R.Roy and D.Ruta, Computer-Assisted Customer Churn Management: State-of-the-Art and Future Trends, Computers & Operations Research. 34(10) (2007) 2902-2917.

[3] R.C.Blattberg, B.D.Kim and S.A. Neslin, Database Marketing: Analyzing and Managing Customers, New York: Springer. (2008) 495-514.

[4] W.Koen, B.De and A.D.Caigny, Spline-Rule Ensemble Classifiers with Structured Sparsity Regularization for Interpretable Customer Churn Modelling, Decision Support Systems.150 (2021)1-14.

[5] A.D.Caigny, K.Coussement And K.W.D. Bock, A New Hybrid Classification Algorithm for Customer Churn Prediction Based on Logistic Regression and Decision Trees, European Journal of Operational Research. 269 (2018) 760–772.

[6] H.Mohammed, T.Ali, E.Tariq And ATM Saeed, Customer Churn In Mobile Markets: A Comparison Of Techniques, International Business Research. 8(6) (2015) 224-237.

[7] A. Q.Ammar, Ahmed and D. Maheswari Churn Prediction on Huge Telecom Data using Hybrid Firefly Based Classification, Egyptian Informatics Journal. 18 (2017) 215–220.

[8] N.Alboukaey, A.Joukhadar And N.Ghneim, Dynamic Behaviour-Based Churn Prediction in Mobile Telecom, Expert Systems with Applications. 162 (2020)1-17.

[9] J.Kozak, K.Kania, P.Juszczuk And M.Mitręga, Swarm Intelligence Goal-Oriented Approach to Data-Driven Innovation in Customer Churn Management, International Journal of Information Management.60 (2021)1-16.

[10] N.N.Y.Vo, S.Liu, X.Li and G.Xu, Leveraging Unstructured Call Log Data for Customer Churn Prediction, Knowledge-Based Systems. 212 (2021) 1-15.

[11] B.Larivière and D.V.D.Poel, Investigating the Role of Product Features in Preventing Customer Churn, By using Survival Analysis and Choice Modelling: the Case of Financial Services. Expert Systems with Applications, 27(2) (2004)277-285.

[12] G.Nie, W.Rowe and L.L.Zhang, Credit Card Churn Forecasting by Logistic Regression and Decision Tree, Expert Systems with Applications, 38(12) (2011)15273-15285.

[13] Y. L. Deng And Q.Y.Gao, A Study on E-Commerce Customer Segmentation Management Based on Improved K-Means Algorithm, Information Systems and E-Business Management.18 (2020) 497-510.

[14] Y.F.Qiu and C.Li, Research on E-Commerce user Churn Prediction Based on Logistic Regression, In Proc. Of IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC 2017), (2017) 12-15.

[15] M.O.Pratama, R.Meiyanti,H.Morrison, A.Ramadhan, A.N.Hidayanto, Influencing Factors of Consumer Purchase Intention Based on Social Commerce Paradigm, In Proc. International Conference on Advanced Computer Science and Information Systems (ICACSIS),(2017) 73-79.

[16] M. Zhao, Q. J. Zeng, M. Chang, Q. Tong and J.F.Su, A Prediction Model of Customer Churn Considering Customer Value: An Empirical Research of Telecom Industry in China, Discrete Dynamics in Nature and Society.5 (2021) 1-12.
[17] K. Coussement, S. Lessmann and G.Verstraeten, A Comparative Analysis of Data Preparation Algorithms for Customer Churn Prediction: A Case Study in the Telecommunication Industry, Decision Support Systems. 95 (2017) 27-36.
[18] A.T. Jahromi, S.Stakhovych and M.Ewing, Managing B2B Customer Churn, Retention and Profitability, Industrial Marketing Management. 43(7) (2014) 1258-1268.
[19] J.Hadden, A.Tiwari, R.Roy and D.Ruta, Computer-Assisted Customer Churn Management: State-Of-the-Art and Future Trends, Computers & Operations Research.34(10) (2007) 2902-2917.
[20] A.Prinzie And D.V.D.Poel, Incorporating Sequential Information Into Traditional Classification Models By using an Element/Position-Sensitive SAM, Decision Support Systems. 42 (2006) 508-526.
[21] T.G. Itschert And U.W.Thonemann, How Training On Multiple Time Slices Improves Performance in Churn Prediction, European Journal of Operational Research. 295(2021) 664-674.
[22] Y.C.Zhen, P.F.Zhi And M.Sun, A Hierarchical Multiple Kernel Support Vector Machine for Customer Churn Prediction using Longitudinal Behavioural Data, European Journal of Operational Research.223(2012) 461-472.
[23] N.Gordini And V.Veglio, Customers, Churn Prediction And Marketing Retention Strategies, An Application Of Support Vector Machines Based on the AUC Parameter-Selection Technique in B2B E-Commerce Industry. Industrial Marketing Management.62 (2016) 100-107.
[24] X.B.Yu, S.S.Guo, J.Guo and X.R.Huang, An Extended Support Vector Machine Forecasting Framework for Customer Churn in E-Commerce. Expert Systems with Applications.38(3) (2011) 1425-1430.
[25] A.T.Jahromi, S.Stakhovych and M.Ewing, Managing B2B Customer Chum, Retention and Profitability, Industrial Marketing Management. 43(7) (2014) 1258-1268.
[26] A.D.Caigny, K.Coussement, W.Verbeke, K.Idbenjra And M.Phan, Uplift Modelling and its Implications for B2B Customer Churn Prediction: A Segmentation-Based Modelling Approach, Industrial Marketing Management.99 (2021) 28-39.
[27] S. E. Schaeffer and S.V.R.Sanchez, Forecasting Client Retention-A Machine-Learning Approach, Journal of Retailing And Consumer Services.52 (2020) 101918.
[28] N.Gordini And V.Veglio, Customers, Churn Prediction And Marketing Retention Strategies. An Application of Support Vector Machines Based on the AUC Parameter-Selection Technique in B2B E-Commerce Industry, Industrial Marketing Management.62 (2017) 100-107.
[29] C. Kristof, L.Stefan and V.Geert, A Comparative Analysis of Data Preparation Algorithms for Customer Churn Prediction： A Case Study on the Telecommunication Industry, Decision Support Systems. 95 (2017) 27-36.
[30] K.W.D. Bock and D.V.D.Poe, An Empirical Evaluation of Rotation Based Ensemble Classifiers for Customer Churn Prediction, Expert Systems with Applications. 38(10) (2011) 12293-12301.
[31] S.Neslin, S.Gupta, WA.Kamakura, L.Liu and C.Mason, Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Model, Journal Of Marketing Research. 43(2) (2006) 204-211.
[32] N.Holtrop, J.E.Wieringa, M. J. Gijsenberg and P. C.Verhoef, No Future Without the Past? Predicting Churn in the Face if Customer Privacy, International Journal if Research in Marketing. 34(1) (2016) 154-172.
[33] E.Lima, C.Mues And B.Baesen, Monitoring and Backtesting Churn Models, Expert Systems with Applications. 38(1) (2011) 975-982.
[34] A.T.Jahromi, S.Stakhovych and M.Ewing, Managing B2B Customer Chum, Retention and Profitability, Industrial Marketing Management. 43(7) (2014) 1258-1268.
[35] A. Amin, S.Anwar, A. Adnan, M.Nawaz, K. Alawfi, A. Hussain and K.Huang, Customer Churn Prediction in the Telecommunication Sector using a Rough Set Approach. Neurocomputing. 237(10) (2017) 242-254.
[36] (2021) Available Tianchi (Online): Https://Tianchi.Aliyun.Com/Datase (Accessed On 17 December 2021).
[37] L.Breiman, Bagging Predictors: Machine Learning.24 (1996) 123-140.
[38] B.A.Goldstein, E. C.Polley And F.B.S.Briggs, Random Forests for Genetic Association Studies, Statistical Applications in Genetics And Molecular Biology.10 (2011) 32.
[39] X . Fan and T.Ke, Enhanced Maximum AUC Linear Classifier, In Proc. if the International Conference on Fuzzy Systems and Knowledge Discovery(FSKD), (2010) 1540-1544.