Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, No. 2, September : 2023

ANALYZING ONLINE SHOPPERS' PURCHASE INTENTION USING ENSEMBLE LEARNING

Dr. Jhimli Adhikari, Professor, Dept. Of Computer Science, Narayan Zantye College, Bicholim, Goa-403 529, India

Abstract
The customer's purchase intention can be predicted by analyzing the history of the customers. In this study, we have analyzed the data of online shoppers for building a better prediction model to predict their purchase intention. Initially exploratory data analysis is performed on the dataset. We have used different ensemble algorithms such as Random Forest, Gradient Boosting, XGBoost and LightGBM to predict whether a customer, visiting the website of an online shop, will end up with a purchase or not. Later we have performed ensemble methods to boost up the performance of the algorithms using SMOTE to overcome class imbalance. Lastly model performance evaluation is done using parameter tuning. Study has shown that XGBoost model predicts with 89.97% accuracy on imbalanced data, whereas Random Forest displays 93% accuracy after using SMOTE to predict the customer's purchase intention. Moreover, XGBoost shows the highest accuracy, which is 93.54% after parameter tuning. This predictive model could be used in the product recommendation system of shopping websites and guide e-commerce companies to customize various preferential policies and services.

Keywords: Ensemble learning, Machine learning, Online shopper, Prediction.

I. Introduction

In last two years online shopping has gained increasing popularity with the effect of pandemic [1]. Online customers often browse pages of e-commerce sites to select items / products before they actually place orders. However, most of the time customers visiting these e-commerce sites may not make any purchase at all. This could be for various reasons e.g. price of product or window shopping. This activity of users not only leaves considerable data on e-commerce platforms, but also only a small quantity of data is converted into purchasing behaviours. This massive information left by users on the e-commerce platform determines the willingness of future users on the e-commerce platform and predict consumer purchasing behaviours. It is important because the retention measures like recommending suitable products can be taken to convert potential customers into purchasers. This information can help businesses to better cater to customer preferences and help both the business and customers by recommending products, specific to each customer and therefore increasing sales for the businesses. An in depth analysis of online purchasing behaviors of consumers enables e-commerce to have a better understanding of customers' psychology, and then work out better business strategies to increase sales [2]. For example, if a company is able to identify potential customers in advance, it can take profit-making actions, e.g. membership, coupons and bundle sales to target customers with repeat purchases intention.

In order to achieve this goal, an ensemble learning method is applied to predict the purchase intention of consumers. In this regard this paper introduces several ensemble learning algorithms such as Random Forest, Gradient Boosting, XGBoost Model and LightGBM to predict the consumers' purchasing behaviours on e-commerce sites. The objective of this research study is to build a model that can predict customer purchase intention as accurately as possible. With the help of Python tools, this article studies and analyses the characteristics of the purchase intention data of consumers, which can provide enterprises with a new method for analysing and predicting the purchase behaviours on e-commerce platform. The predictive model used in this study could be used in the recommendation system of shopping websites and can also be used to guide e-commerce companies to customize various preferential policies and services, so as to quickly and accurately stimulate the purchase intention of more potential consumers.

After presenting related work in section 2, research methodology is described in section 3. Several machine learning algorithms, criteria for evaluating the performance of the proposed methods, and competing classification methods for comparisons are described in section 4. Section 5 present data characteristics, data analysis and result evaluation about the model used in this study. Conclusion and future work is presented in section 6.

II. Related work

Purchase prediction has been studied widely in the literature. Consumer purchase prediction is a binary classification problem. Recent studies have proposed various frameworks to analyse purchase prediction using users' previous session features in real-time environment [3]. Here embedding session features are incorporated into machine learning models to improve the performance of models in predicting users' purchase intention. Zeng et al. [4], proposed a purchase prediction model to analyse user behaviour during a festival in China. The model shows that if a product is interesting to a user, the user is more likely to spend more time on it.

In a virtual environment, Wu et al. [5] proposed a purchase behaviour prediction model which identifies click patterns rather than session features. Experimental results shows that learned features from click patterns can improve purchase prediction as good as a classification model trained using session features.

Mokryn et al. [6] applied logistic regression, Bagging, Decision Tree algorithms to investigate the effect of temporal features (time, product trendiness, etc.) on purchase prediction performance. Authors have shown that Bagging performed best when temporal features are applied. Mootha et al. [7] proposed a stacking ensemble technique, which makes use of Multi-Layer Perceptron's to detect the intention of a user on whether a product would be purchased or not.

A real-time online shoppers' behavior prediction system is proposed by Batti et al. [8] which predicts the visitor's shopping intention as soon as the website is visited. In this article, authors worked on session and visitor information and investigated naïve Bayes classifier, C4.5 decision tree and random forest algorithms. Further, oversampling technique was used to improve the performance and the scalability of each classifier.

Sakar et al. [9] predicted the purchasing intention of the visitor using aggregated pageview data along with user information. The extracted features are used to random forest (RF), support vector machines (SVMs), and multilayer perceptron (MLP) classifiers as input. Noviantoro et al. [10] applied feature selection technique to filter unrepresentative features by analysing the visitor's clickstream data. Authors examined several supervised machine learning algorithms to identify more accurate prediction performance. In this article we used several ensemble algorithms to boost up the performance of the algorithms using SMOTE (Synthetic Minority Over sampling Technique) and parameter tuning.

III. Research Methodology

Ensemble learning methods in machine learning [11] combine the decisions from multiple learning algorithms to form strong learners to improve the overall performance. Hence, the performance of ensemble learning methods is always better than performance obtained from any of the constituent learning algorithm alone. In this paper four different models (Random Forest, Gradient Boosting, XGBoost model and LightGBM) are used for predicting online shoppers purchase intention. Each model is examined in terms of their accuracy, precision, recall rates, F1 and AUC score. AUC value ranges from 0 to 1 and the model whose prediction is better is close to 1. Python programming language is used to implement all models and ROC curves are plotted for each model. In the following section we introduce the basic ideas of ensemble learning.

3.1 Ensemble learning

The idea behind Ensemble learning is to build a predictive model that seeks better performance by combining the predictions from multiple models [12]. Thus, it helps to decrease the variance, bias and improve predictions [13]. An ensemble method is a technique which uses multiple independent similar or different models / weak learners to derive an output or make some predictions [14]. Our approach exploits ensemble learning classifiers to predict the purchase of shoppers' intention. We evaluated the performance of four ensemble learning classifiers such as Random Forest, Gradient Boosting, XGBoost and LightGBM models discussed in section 4.1. Models were implemented using Scikit-learn in Python [15].

3.2 Ensemble Methods

1) Stacking: The idea of stacking is to train a number of weak learners and then combine them to find a meta model to predict the result [16, 17]. Here the dataset is divided into two groups. The first groups are used to train all the learners. Then the second group is used to find the prediction from those learners. Then the meta model is developed comparing the prediction and the actual result of the second group.
2) Bagging: Bagging is a short form of bootstrap aggregation. In bagging several bootstrap replicas are obtained by choosing randomly with replacement from the training data set [16, 17]. Each of the training subsets of data set is used to train different learners of the same type. The learners are then combined together by using the strategy of maximum vote. This method reduces the variance when the base learners are combined together.
3) Boosting: Boosting is a method of converting a weak learner to a strong learner. In boosting first a tree is generated where each of the data points is given an equal weight [16, 17]. Then after evaluating the first tree the weight of those data points which are difficult to classify are increased and the weight of those who are easy to classify is decreased. The second tree is generated using the weight of the first tree. The third tree is generated by calculating the errors from the first and second tree and predicts the result. This process is repeated for a certain number of iterations. In each iteration, the subsequent errors of the previous trees are decreased.

3.3 Ensemble learning algorithms

The purchasability problem is formulated as a binary classification problem. The current study applies four different classification and boosting algorithms as follows:
Random Forest: An ensemble learning method that uses a combination of tree predictors where each tree depends on the value of a random vector. Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, it predicts the final output. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

- Gradient Boosting: It is a stage-wise additive model that generates learners during the learning process (i.e., trees are added one at a time, and existing trees in the model are not changed). The contribution of the weak learner to the ensemble is based on the gradient descent optimisation process. The calculated contribution of each tree is based on minimizing the overall error of the strong learner.
- XGBoost: In this algorithm, decision trees are created in sequential form. Weights play an important role in XGBoost. Weights are assigned to all the independent variables which are then fed into the decision tree which predicts results. The weight of variables predicted wrong by the tree is increased and these variables are then fed to the second decision tree. These individual classifiers / predictors then ensemble to give a strong and more precise model. Major benefits of

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, No. 2, September : 2023

XGBoost are that its highly scalable / parallelizable, quick to execute, and typically outperforms other algorithms [18].

- LightGBM: LightGBM extends the gradient boosting algorithm by adding a type of automatic feature selection as well as focusing on boosting examples with larger gradients. This can result in the efficiency of the model and reduces memory usage.

IV. Experimental results and discussions

The proposed system is implemented on Online Shoppers Purchasing Intention Dataset [19], provided by UCI's Machine Learning Repository using Python 3.8.2 programming language with a processor of Intel Core i5-8300H CPU @ 2:30 GHz and RAM of 8 GB running on Windows 10. This dataset is licensed and available for general usage. Table 1 shows a brief description of the dataset. However, this dataset has been pre-processed further to meet the needs of this study. Here, the data was consolidated in a way that each row corresponds to a session from a different user, with 12,330 instances. The attribute that indicates whether the purchase has been made or not is used as the class label. 'Purchase' (True or False) is used as target column. For the model architecture, 70% of total records (12,330) are used for training (8,631) and the remaining 3,699 (30%) for testing.

4.1 Data characteristics

The dataset was formed in such way that each session would belong to a different user in one year period to avoid any tendency to a specific campaign, special day, user profile, or period. The primary purpose of the dataset is to predict the purchasing intentions of a visitor to this particular store's website. Data was collected in Excel file uploaded in Jupyter notebook and analysed with Python 3.8.2 software. The dataset has 10 numerical attributes, 7 categorical attributes and one label attribute shown in Table 2 and 3 respectively. 84.5% (10,422) data belong to negative class samples (did not end with shopping) and the rest (1,908) are positive class samples (ending with purchases) of 12,330 sessions.

Table 1: Data file description

Data source	Number of instances	Number of Features	Number of Target Classes	Size of data	File type
www.UCI.com	12,330	11	2	1882 KB	Excel spreadsheet

Dataset consists of 3 different page categories such as "Administrative", "Informational" and "Product Related". Additionally, duration information of the user for these pages is also available. The dataset has "Bounce Rate", "Exit Rate" and "Page Value" features which are measured by Google Analytics. The "Special Day" feature represents the closeness of the site visiting time to a specific special day. For instance, let's take a look at Christmas day. This value takes a nonzero value between December 20 and December 29. Zero value is considered before and after this date unless it is close to another special day, and its maximum value is 1 on December 25.

Table 2: Descriptive statistics of 10 numerical attributes of this study

Attribute name	Attribute description	Min	Max	SD
Administrative	pages viewed by visitor regarding account administration	0	27	3.32
Administrative_Duration	the total amount of time spent by a visitor (in seconds) on the pages	0	3398	176.78
Informational	pages viewed by visitor regarding information of e-commerce, corresponding and purchase procedure	0	24	1.26
Informational_Duration	the total amount of time spent by a visitor (in seconds) on pages regarding information e-	0	2549	140.75

	commerce, corresponding and purchase procedure			
ProductRelated	pages viewed by visitor regarding product related pages	0	705	44.47
ProductRelated_Duration	the total amount of time spent by a visitor (in seconds) on the pages regarding product-related pages	0	63973	1913.67
BounceRates	the average value of bounce rate of pages viewed by the visitor	0	0.2	0.04
ExitRates	the average value of exit rate of pages viewed by the visitor	0	0.2	0.05
PageValues	the average value of contribution unique pages viewed by the visitor	0	1.0	0.19
special_day	proximity of website visits with special day	0	361	18.55

As dataset has categorical features we need to encode before processing. This is done using the LabelEncoder from sklearn.preprocessing. When we look at the categorical values we could see in Table 3 there are 8 different types of operating Systems, 13 browser types, 9 different regions, 20 traffic types and 3 visitor types.

Table 3: Descriptive statistics of 7 categorical attributes and label attribute

Attribute name	Attribute description	Number of categorical values
operating-systems	operating system used by the visitor	8
browser	browser used by the visitor	13
region	geographic region data based on IP-based session visited by the visitor	9
Traffic type	traffic source type which brings the visitor to the Web site (e.g., direct, search engine, display ads, SMS, etc.)	20
Visitor type	visitor category of session visited by the visitor as ‘‘new visitor,’’ ‘‘return visitor,’’ or ‘‘other’’	3
weekend	time of session visited by visitor wheatear “weekday” or “weekend”	2
month	the month when the session visited by the visitor	12
purchase-decision	the class label indicating whether the visit was completed with a purchase or not	2

Class imbalance [20] is a major problem in purchase prediction from session logs since the majority of the sessions end without purchase. When we count these two classes we can see that there are 10422 incomplete transactions and 1908 completed transactions in Figure 1.

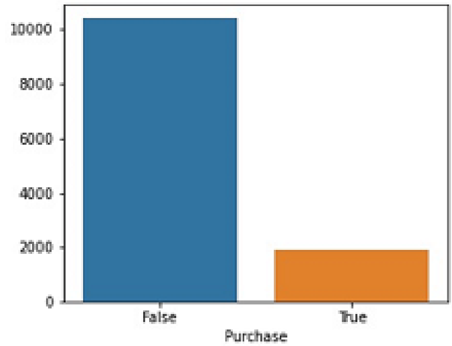

Figure 1. Distribution of visitors purchase

4.2 Exploratory data analysis

A bar chart is shown in Figure 2 to provide a visualization of the purchase among the months of the dataset. The majority of purchasing happened in March, May, November and December. In other words maximum purchase happened at beginning and end of a year. Generally, users tend to research before they purchase a product. Thus, we could see that returning visitors have more intention to purchase a product. On weekdays the volume of transactions is higher than that of the weekend. There it shows no significant difference between the trends of two customer categories, since the reasons like seasons apply to everyone.

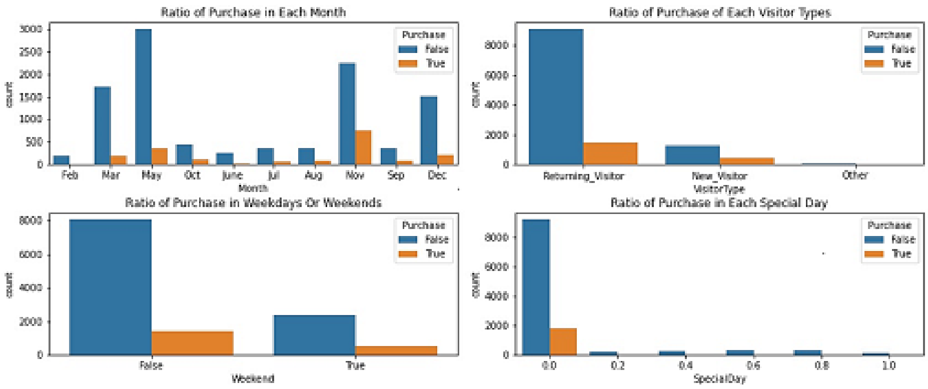

Fig. 2. Visualizations of data characteristics

Figure 3 shows which feature is most likely to be correlated with the 'Purchase' attribute. Resultantly, Page Values has the highest correlation (approx 0.5) and BounceRates, ExitRates and SpecialDay have negative correlation with Purchase.

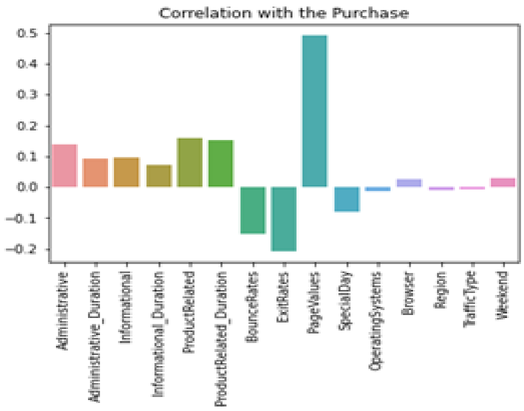

Fig. 3. Correlated features with purchase

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, No. 2, September : 2023

Figure 4 depicts the frequency of operating systems, browser, region and Traffic type categorical data.

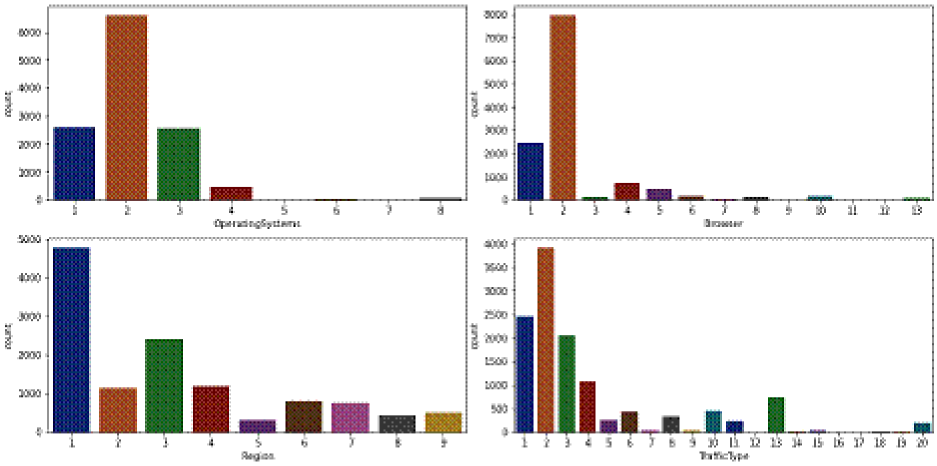

Fig. 4. Frequency of categorical attributes

Generally all the features (independent variables) are not correlated to each other in many of the Machine Learning algorithms. Let's check this by Pearson Correlation and is shown in Figure 5.

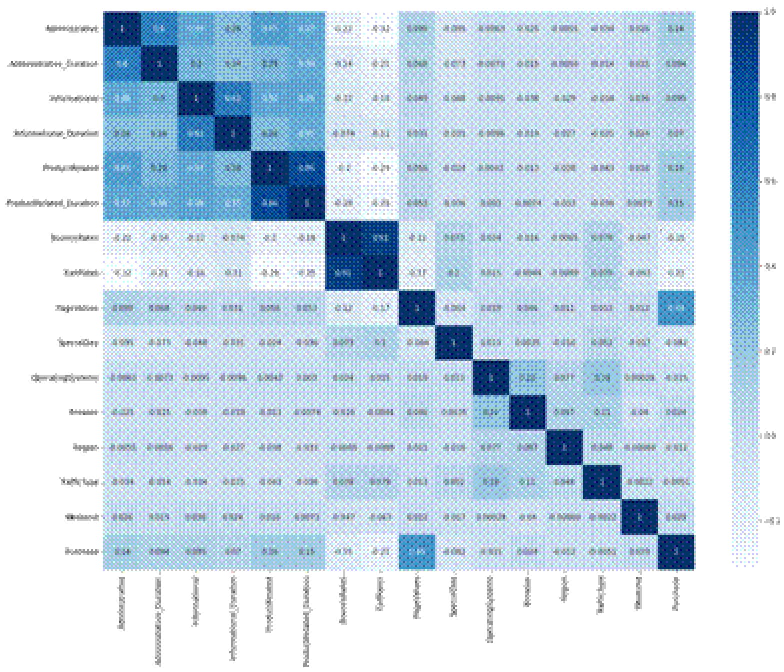

Fig. 5. Feature Correlation of the dataset

From the above figure, it is clear that administrative data are correlated. Information, Product Related, and Rates (Exit and Bounce) have similar Characteristics. Page Value seems to have a stronger correlation with the Purchase.

4.3 Ensemble Model Performance Analysis

There are 4 different types of models tested on the preprocessed dataset. Each model is examined in order of their precision, recall, accuracy rates, AUC and F1 score. Their feature importance is mentioned and ROC curves are also plotted. We split data into train and test datasets. Train dataset size is 70% and test dataset size is 30%. Table 4 reports the results of the machine learning algorithms obtained with the default settings in Scikit-Learn.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, No. 2, September : 2023

Table 4: Performance of ensemble algorithms on Class Imbalanced Dataset

Model	Feature importance	Precision	Recall	F1-score	Accuracy rate	AUC
Random Forest	Page values	0.72	0.56	0.63	89.86	0.91
Gradient Boosting Model	Page values	0.71	0.58	0.64	89.84	0.93
XGBoost Model	Page values	0.72	0.58	0.64	89.97	0.93
LightGBM Model	Exit rate	0.70	0.59	0.64	89.86	0.93

Random Forest model is successful for predicting the non-transactions as it shows accuracy rate 89.86 but the recall rate is only 56%. This shows that even if the model's accuracy rate is high, the model can make wrong predictions. Page value is the most important feature for this model with 0.39. Gradient Boosting Model is also successful as the accuracy rate is 89.84. This model is also successful for predicting the non-transactions and the recall rate is 0.58 which is slightly better than the Random Forest model but not enough for the prediction. Here AUC is 0.93 (Figure 6) which is again slightly better than the Random Forest model (0.91) and the most important feature is Page value like the Random Forest model. The accuracy rate of XGBoost Model is 89.97. Again from table we could see this model is not good at predicting users with transactions. Although its' prediction rate is higher than the previous two models but the recall rate is again low i.e. 0.58. If we check the ROC curve the AUC is 0.93 (Figure 6) which is the same with Gradient Boosting and the most important feature is Page Value like the other two models. LightGBM model has the accuracy rate 89.86. As can be seen in the Table 4 the model is not good at predicting users with transactions, but the true positive rate is higher than the XGBoost model. In Table 4 it can be seen that the recall rate is again 0.59 which is the highest rate among all four models. If we check the AUC score it is 0.93 which is the same with Gradient Boosting and XGBoost models (Figure 6). Here the most important feature is Exit Rate which is different from the other three models.

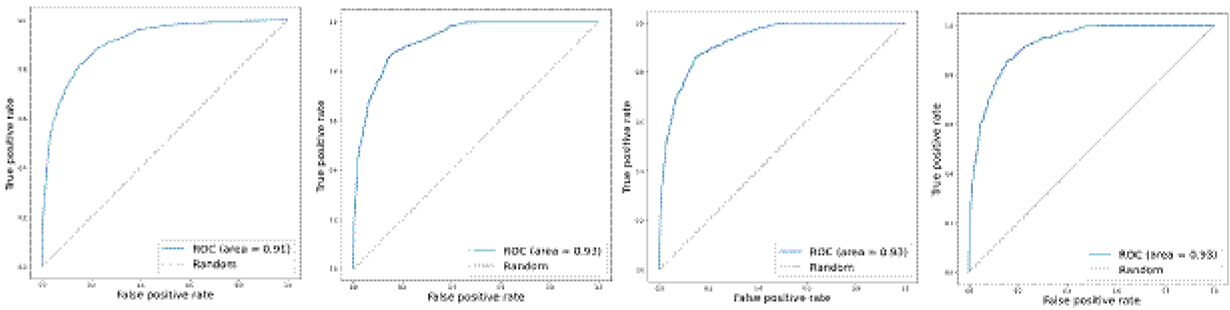

Fig 6. ROC curve of Random Forest, Gradient Boosting, XGBoost and LightGBM

4.4 Class imbalance

Since the machine learning classification algorithm assumes that data are evenly distributed among different classes, predictive performance decreases in the case of unbalanced data [21]. The over-sampling method is a method of replicating samples of a minority class or creating new samples to balance them. SMOTE (Synthetic Minority Over sampling Technique) introduced by Chawla et al. [22] is considered a relatively effective over-sampling technique as a method of balancing by randomly inserting the nearest neighbor pair in a minority class.

4.5 Overcoming Class Imbalance using SMOTE Techniques

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, No. 2, September : 2023

In this study, the target classes are heavily imbalanced (10,422 sessions with Purchase = False and 1,908 sessions with Purchase = True). To prevent this situation and to increase the model's performance the data is transformed into balanced data. To do so an oversampling method SMOTE is applied on the dataset. This method generates random data and increases the transaction so that the data becomes balanced. In Table 5 we evaluate whether oversampling data leads to a better model or not.

Table 5: Comparison between ensemble algorithms after SMOTE

Model	Feature importance	Precision	Recall	F-score	Accuracy	AUC
Random Forest	Page values	0.92	0.94	0.93	93	0.98
Gradient Boosting	Page values	0.91	0.92	0.92	87.13	0.97
XGBoost Model	Page values	0.87	0.87	0.87	86.79	0.97
LightGBM Model	product-related duration	0.92	0.83	0.93	91.11	0.97

Figure 7 shows ROC curve of the classifiers used in prediction for oversampled data

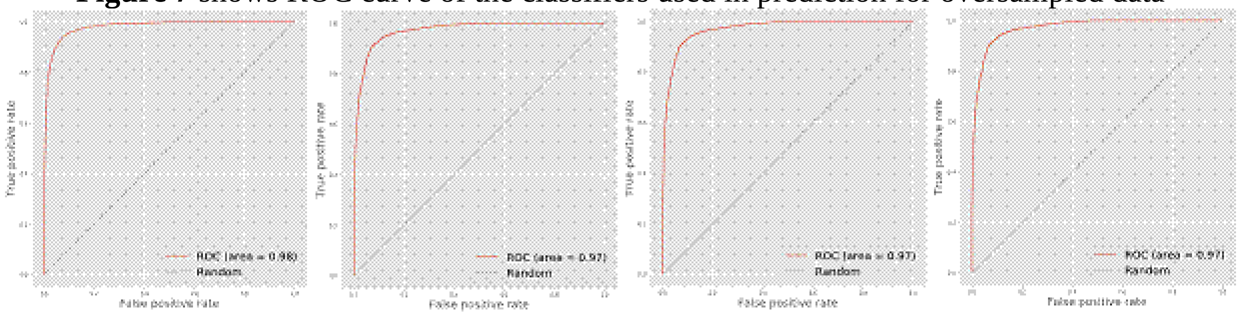

Fig. 7. ROC curve for oversampled Random Forest, Gradient Boosting, XGBoost and LightGBM model

These results prove that the second experiment outperformed the first experiment. The second experiment yielded better results in all evaluations than the first experiment. As can be seen, from Table 5 the AUC is 0.98 for Random Forest model which is close to 1. This shows the oversampling method is increasing the accuracy of the classification. Unlike Random Forest there is slight decrease of the accuracy rate for the Gradient Boosting Model from 89.84 to 87.13 after applying SMOTE method. Although the accuracy rate decreased, the recall rate was normalized and became 0.92 and the AUC metric is increased to 0.97 from 0.93. After the SMOTE method is applied, the accuracy rate for the XGBoost Model is 86.79 which was 89.97 earlier. Although the accuracy rate decreased the recall rate was normalized and became 0.87 and the AUC metric is increased to 0.97 from 0.93. After the SMOTE method is applied, the accuracy rate for LightGBM Model is 91.11 which was 89.86 before. The recall rate is normalized and becomes 0.83. The AUC metric is increased to 0.97 from 0.93. In this case before applying SMOTE exit rate was the most important feature but after oversampling product-related duration becomes the most important feature. We see that Random Forest scored the highest accuracy among the four models.

4.6 Model Evaluation after parameter tuning

Hyperparameter tuning is searching the hyperparameter space for a set of values that optimizes model architecture. For all four models the hyper parameter fine tuning is applied. In Table 6 the best parameters and best accuracies are listed. After the model tuning process we can see that the best model for oversampled dataset is XGBoost with accuracy rate of 93.54. Accuracy rate of Gradient Boosting is also very close to XGBoost i.e. (93.53). Thus, one could perform statistical analysis on these two models, to further explore the best Machine Learning model in future.

Table 6. Performance comparison after tuning the models

Model	Parameters	Best Parameters	Best Accuracy
Random Forest	params = {'max_depth': [3, 6, 10, 20, None], 'max_features': ['auto', 'sqrt'], 'n_estimators': [100, 500, 1000]}	{'max_depth': None, 'max_features': 'auto', 'n_estimators': 1000}	93.03
Gradient Boosting	params = {'max_depth': [3, 6, 10], 'learning_rate': [0.01, 0.1, 0.2], 'n_estimators': [10, 100, 500, 1000]}	{'max_depth': 10, 'learning_rate': 0.2, 'n_estimators': 1000}	93.53
XGBoost	params = {'max_depth': [3, 6, 10, 20, None], 'learning_rate': [0.01, 0.1, 0.2], 'n_estimators': [100, 500, 1000]}	{'max_depth': 10, 'learning_rate': 0.1, 'n_estimators': 500})	93.54
LightGBM	params = {'max_depth': [-1, 1, 5, 10], 'num_leaves': [20, 30, 40]}	{'max_depth': -1, 'num_leaves': 40}	92.98

V. Conclusion and future work

In this article ensemble learning algorithms Random Forest, Gradient Boosting, XGBoost Model and LightGBM are used to predict the consumers' purchasing behaviours. This study reveals the performance of model on imbalanced data. Performance is evaluated on F1 score, accuracy and AUC. Second, the oversampling method is most suitable for the data imbalance problem that occurs in the context of online consumer behavior. Thus, SMOTE, which is the over-sampling method is adopted in this article. Data analysis shows that the predictive power of the ensemble models after applying SMOTE is superior to the algorithms. The predictive model used in this study can be used in the recommendation system of shopping websites and can also be used to guide e-commerce companies to customize various preferential policies and services, so as to quickly and accurately stimulate the purchase intention of more potential consumers. In future work the customer purchase intention prediction can be combined to the e-commerce product recommendation system. Further work could be done to see if the recommending product based on customer intention could have an impact to increase sales or not.

References

[1] Rao, Y., Saleem, A., Saeed, W., & Ul Haq, J. (2021). Online consumer satisfaction during COVID-19: perspective of a developing country. Frontiers in Psychology, 12, 751854.

[2] García-Salirrosas, E.E.; Acevedo-Duque, Á.; Marin Chaves, V.; Mejía Henao, P.A.; Olaya Molano, J.C. Purchase Intention and Satisfaction of Online Shop Users in Developing Countries during the COVID-19 Pandemic. Sustainability 2022, 14, 6302. https://doi.org/ 10.3390/su14106302

[3] Martínez, A., Schmuck, C., Pereverzyev, S., Pirker, C., & Haltmeier, M. (2020). A machine learning framework for customer purchase prediction in the non-contractual setting. European Journal of Operational Research, 281(3), 588–596. https://doi.org/10.1016/j.ejor.2018.04.034.

[4] Zeng, M., Cao, H., Chen, M., & Li, Y. (2019). User behaviour modeling, recommendations, and purchase prediction during shopping festivals. Electronic Markets, 29(2), 263–274. https://doi.org/10.1007/s12525-018-0311-8

[5] Wu, Z., Tan, B.H., Duan, R., Liu, Y., Mong Goh, R.S. (2015). Neural modeling of buying behaviour for E-commerce from click ing patterns. In Proceedings of the international ACM recom mender systems challenge 2015 (p. 12). https://doi.org/10.1145/2813448.2813521.

[6] Mokryn, O., Bogina, V., Kuflik, T. (2019). Will this session end with a purchase? Inferring current purchase intent of anonymous visitors. Electronic Commerce Research and Applications, 34, 100 836. https://doi.org/10.1016/j.elerap.2019.100836.

[7] S. Mootha, S. Sridhar and M. S. K. Devi, "A Stacking Ensemble of Multi Layer Perceptrons to Predict Online Shoppers' Purchasing Intention," 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2020, pp. 721-726, doi: 10.1109/ISRITI51436.2020.9315447.

[8] Baati, K., Mohsil, M. (2020). Real-Time Prediction of Online Shoppers' Purchasing Intention Using Random Forest. In: Maglogiannis, I., Iliadis, L., Pimenidis, E. (eds) Artificial Intelligence Applications and Innovations. AIAI 2020. IFIP Advances in Information and Communication Technology, vol 583. Springer, Cham. https://doi.org/10.1007/978-3-030-49161-1_4

[9] Sakar, C.O., Polat, S.O., Katircioglu, M. et al. Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. Neural Comput & Applic 31, 6893–6908 (2019). https://doi.org/10.1007/s00521-018-3523-0

[10] Noviantoro, T., and Huang, J. P. (2021). Applying Data Mining Techniques to Investigate Online Shopper Purchase Intention Based on Clickstream Data. Review of Business, Accounting, & Finance, 1(2), 130-159.

[11] Zhou, Z.-H. (2012). Ensemble Methods: Foundations and Algorithms (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b12207

[12] C. Zhang and Y. Ma (eds.), Ensemble Machine Learning: Methods and Applications, DOI 10.1007/978-1-4419-9326-7 1, © Springer Science+Business Media, LLC 2012

[13] Pretorius, A., Bierman, S. and Steel, S.J., 2016, November. A bias-variance analysis of ensemble learning for classification. In Annual Proceedings of the South African Statistical Association Conference (Vol. 2016, No. con-1, pp. 57-64). South African Statistical Association (SASA).

[14] Panagiotis Pintelas Ioannis E. Livieris (eds.), Ensemble Algorithms and Their Applications, Mdpi, 2020, ISBN 978-3-03936-958-4 (Hbk); ISBN 978-3-03936-959-1, https://doi.org/10.3390/books978-3-03936-959-1

[15] Ensemble Learning Algorithms With Python: Make Better Predictions with Bagging, Boosting, and Stacking, Brownlee, J., https://books.google.co.in/books?id=IUkrEAAAQBAJ, 2021, Machine Learning Mastery

[16] Bonaccorso, G. (2017). Machine Learning Algorithms. United Kingdom: Packt Publishing, 9781785889622.

[17] Faul AC. 2020. A concise introduction to machine learning. Boca Raton: CRC Press, ISBN 9780815384106, by Chapman & Hall, 334 Pages.

[18] Tianqi Chen and Carlos Guestrin, KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2016) Pages 785–794, https://doi.org/10.1145/2939672.2939785

[19] UCI Machine Learning Repository. Available :https://archive.ics.uci.edu/.

[20] Krawczyk B. Learning from imbalanced data: open challenges and future directions. Prog Artif Intell. 2016; 5(4): 221–32. https://doi.org/10.1007/s13748-016-0094-0

[21] Amalia Luque, Alejandro Carrasco, Alejandro Martín, Ana de las Heras, The impact of class imbalance in classification performance metrics based on the binary confusion matrix, Pattern Recognition, Volume 91, 2019, Pages 216-231, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2019.02.023

[22] Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer,W.P.: Smote: synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357 (2002)