Journal of Engg. Research Vol. 11 No(3A) pp.100-111 **DOI:10.36909/jer.16771**

A grid search optimized extreme learning machine approach for customer churn prediction

Fatma Önay Koçoğlu* and Tuncay Özcan**

**Department of Software Engineering, Faculty of Engineering, Muğla Sıtkı Koçman University, Muğla, 48000, Turkey.*

***Department of Management Engineering, Faculty of Management, İstanbul Technical University, İstanbul, 34367, Turkey.*

** Corresponding Author: fonaykocoglu@mu.edu.tr*

Submitted **:** 15-12-2021

Revised **:** 22-02-2022

Accepted **:** 22-03-2022

ABSTRACT

Behavioral traits of customers, such as loyalty status and satisfaction criteria are subject to alterations due to the rapidly changing world. Therefore, these behavioral changes should be analyzed efficiently at every step of the decision-making process. Customer churn analysis involves the determination of customers who tend to leave a situation before it occurs, by analyzing customer data using various methods. The aim of this study is to develop an extreme learning machine-based model to analyze the customer churn prediction problem and determine the parameters that improve the performance of the model. Grid search is used for hyperparameter tuning. In addition, a modified accuracy calculation approach is presented. In this study, we have developed various models based on Naïve Bayes, k-nearest neighbor, and support vector machine methods and provided a comparison of the performance of each model. According to the results obtained, an accuracy of 93.1% was achieved using the proposed extreme learning machine model.". In addition, the proposed model is highly effective in solving the research problem because, the number of parameters to be determined is less, thus reducing its competition with other models.

Keywords: Churn Analysis; Extreme Learning Machine; Grid Search; Parameter Tuning.

INTRODUCTION

Due to the rapidly changing world, the behavioral traits of customers, such as loyalty status and satisfaction criteria also change. When the market reaches the saturation point, the number of suitable customers is limited. Therefore, enterprises focus on promoting existing customers rather than targeting new customers (Hung et al., 2006). A customer churn is defined as a customer having a break or terminating transactions with the company he/she previously received or performed with another company providing the same service/product (Hung et al., 2006; Kaur et al., 2013). For effective customer churn management, a forecasting model must be created to generate effective and accurate results (Tsai & Lu, 2009).

Various studies using machine learning algorithms for churn prediction have been described in the literature. Adwan et al. (2014) have presented multilayer perceptron (MLP) topologies-based artificial neural networks (ANN) models. Al-Shboul et al. (2015) developed a hybrid method based on fast fuzzy C-means and genetic programming (GP) methods to reveal the outliers in the dataset. Coussement et al. (2017) determined the effects of data preprocessing based on the prediction success. Stripling et al. (2018) introduced a classifier that optimizes the maximum benefit measure using genetic algorithm (GA). Jain et al. (2020) proposed two prediction models based on logistic regression (LR) and LogitBoost. In the study by Dias et al. (2020), churn prediction up to six months was considered for retail banking customers, and six different machine learning methods were examined and features that provided the best predictive performance were determined. Jain et al. (2021) developed models using four types of algorithms, namely LR, random forest (RF), Support Vector Machine (SVM), and

XGBoost. In addition, they used datasets from three different fields, namely banking, telecommunications, and IT. Lalwani et al. (2022) developed various models using LR, Naïve Bayes (NB), support vector machines (SVM), RF, and decision trees (DT) in addition to hyperparameter tuning is also performed. Beeharry and Fokone (2021) proposed a two-layer flexible voting ensemble model to predict customer churn in the telecommunications industry. The effect of balancing the dataset on model success was also examined. Nwaogu and Dimililer (2021) employed SVM, MLP, and neural network (NN) machine learning algorithms. A summary of literature studies in provided in Table 1.

Table 1. Summary of the literature studies on churn prediction

Study	Sector	Classification Methods	Performance Measures
Coussement & De Bock (2013)	Online Betting Site	DT, RF, Generalized Additive Models	Lift
Sharma and Panigrahi (2013)	Telecom	NN	Accuracy
Kim et al. (2014)	Telecom	Spreading Activation Model (SPA), Louvain Algorithm, LR, NN	AUC, Lift, Success Rate
Hudaib et al. (2015)	Telecom	K-Means, Hierarchical Clustering, ANN, DT	Accuracy, Churn Rate
Gajowniczek et al. (2016)	Telecom	LR, NN, DT, SVM	AUC, Lift
Dolatabadi & Keynia (2017)	Unknown	DT, NB, SVM, NN	Accuracy
De Caigny et al. (2018)	Multiple	DT, RF, LR, Logit Leaf Model (LLM)	AUC, Lift, Processing Time
Ahmad et al. (2019)	Telecom	RF, DT, XGBoost, Gradient Boosted Machine	AUC
Srivastava & Eachempati (2021)	Employee Churn	Deep NN, RF, Gradient Boosting	Accuracy
Asghar et al. (2021)	Telecom	DT, LR, NN, NB, PSO	Accuracy
Vo et al. (2021)	Financial	Gaussian Naïve Bayes (NB), LR, RF, Extreme Gradient Boosting, Bidirectional Long Short-Term Memory	AUC, Accuracy, Recall, F-Measure, Precision

The aim of this study is to solve the customer churn analysis problem using ELM algorithm and simultaneously evaluate its performance using parameter tuning with grid search optimization. In addition, an accuracy-based performance evaluation measurement is proposed from a different perspective. This study differs from other studies in terms of the hyperparameter tuning performed and the calculation of the new accuracy measure used.

The rest of this paper is organized as follows: The section on “Methodology” describes the data, methods, and tools used in detail. In “Results” comparative results has been presented, and “Discussion and Conclusion” include assessment of results and recommendations for future studies.

METHODOLOGY

The performance of a machine learning model depends on various factors, such as such as the structure of the data, its size, and the number of records related to the classes, algorithms, performance validation methods, sampling methods, and feature selection methods. One of the most important factors affecting the model performance is the hyperparameters used in the algorithms. Hyperparameters are the parameters of an algorithm that can be manipulated by a model developer. Hyperparameter tuning to achieve an optimal model successfully is an important research topic. The aim of this study is to solve the defined problem using ELM algorithm and simultaneously evaluate its performance. In addition, grid search optimization is used to obtain the parameters that can improve the performance of ELM algorithm. Algorithms, such as NB, KNN, and SVM are implemented to compare their performance with that of ELM algorithm. Owing to their high efficiency in detecting complex structures in data, methods, such as SVM, ANN, and GA are widely preferred (Alsumaiei, 2021). Therefore, we have also implemented SVM and ANN-based ELM algorithm within the scope of this study. In addition, we have proposed an accuracy-based performance evaluation measure.

First, the theoretical explanations of the methods used are given below followed by a detailed explanation of the steps followed in the study according to CRISP-DM.

Extreme Learning Machine

ELM, developed by Huang et al. (2004), is a method based on ANN; however, it differs from NN in some characteristics. To accelerate the learning process, the input weights and threshold values were generated randomly and analytical methods were used to calculate the output weights. Let (x_i, y_j) be different and arbitrarily selected examples where $x_i = \{x_{i1}, x_i, \dots, x_{in}\}^T$ is the input and $y_i = \{y_{i1}, y_{i2}, \dots, y_{im}\}^T$ is the output, thus indicating a mathematical model for the ELM.

$$\sum_{i=1}^{\tilde{N}} \beta_i g\left(w_i x_j + b_i\right) = y_j \ (j=1,2,\dots,N) \tag{1}$$

where Ñ is the number of neurons, Ñ<N, g(x) is the activation function, $w_i = \{w_{i1}, w_i, \dots, w_{in}\}^T$ is the weight vector associated with input and hidden neuron *i*, $\beta_i = \{\beta_{i1}, \beta_i, \dots, \beta_{im}\}^T$ is the weight vector associating the hidden neuron *i* with output, b_i is threshold value for hidden neuron *i*.

The outputs of the prediction models and real values calculated from Equation (1) for the class attributes are different. To maximize the network performance, this difference is expected to be zero or minimum. The objective of the ELM is to calculate the output vector to minimize errors Furthermore, Moore–Penrose pseudoinverse method is used to obtain the solution.

Hyperparameter Tuning and Grid Search Optimization

Machine learning algorithms require user-defined parameter values to achieve a balance between accuracy and generalizability (Srivastava & Eachempati, 2021). Determining the optimal hyperparameters for a machine learning model is crucial for a bias-reduced assessment of the predictive power of the model (Schratz et al., 2019). Hyperparameter optimization is used for the selection of appropriate parameter values for the algorithms, and hence better classification performance (Shankar et al., 2020). However, the number of automated and guided methods to determine hyperparameter values is limited. For this purpose, methods, such as grid search, random search, and Bayesian optimization, are used (Elgeldawi et al., 2021). In the grid search method, the model is trained by combining all the values at finite intervals determined for the hyperparameters. In ANN models, the number of neurons is determined by random trial and error (Munir et al., 2022). All trained models were compared according to the performance evaluation measures.

Application

The methods proposed in this study are performed according to the steps outlined by CRISP-DM. In addition, customer churn analysis is chosen as the research problem to be addressed.

Data Set

The literature review indicates that applications based on customer churn analysis are mostly implemented in the telecom and banking sectors. The telecom churn dataset is obtained from the UCI machine learning repository because the results can be analyzed and compared effectively by different researchers. The dataset consists of 21 attribute fields, one of which is a class attribute, and 5000 records. “State” and “phone” attribute fields have been excluded from the data set because they have negligible effect on the results of the analysis. Nineteen attributes were used in the analyses of which 4 were categorical and 15 were numerical values. The data types and explanations of the attributes are listed in Table 2. The churn ratio of the dataset is 14.14% (707 records are of churn customers and 4293 are not churn customers).

Table 2. Overview of the attributes in the data set used in this study

Attribute	Data Type	Attribute	Data Type
1-account_length: Active duration of an account	Numerical	10-total_eve_minutes: Total usage in minutes during evening time	Numerical
2-area_code: Code of the customer's area	Categorical {0, 1, 2}	11-total_eve_calls: Total calls during evening time	Numerical
3-international_plan: International Plan activated or not	Binary {0, 1}	12-total_eve_charges: Total charge for evening time	Numerical
4-voice_mail_plan: Voice Mail Plan activated or not	Binary {0, 1}	13-total_night_minutes: Total usage in minutes during night time	Numerical
5-number_vmail_message: Number of Voice Mail Messages	Numerical	14-total_night_calls: Total calls during night time	Numerical
6-number_csc: Number of Customer Service Calls made	Numerical	15-total_night_charge: Total charge for night time	Numerical
7-total_day_minutes: Total usage in minutes during day time	Numerical	16-total_intl_minutes: Total usage in minutes for international calls	Numerical
8-total_day_calls: Total calls during day time	Numerical	17-total_intl_calls: Total international calls	Numerical
9-total_day_charge: Total charge for day time	Numerical	18-total_intl_charge: Total charge for international calls	Numerical
		19-Churn: Current status of the customer (0:not churn, 1:churn)	Binary {0, 1}

Descriptive statistics for the numerical attributes of the dataset are presented in Table 3.

Table 3. Descriptive statistics of numerical attributes

	Att.1	**Att.2**	**Att.3**	**Att.4**	**Att.5**
Min	1.0	0.0	0.0	0.0	0.0
1st Qu.	73.0	0.0	143.7	87.0	24.43
Median	100.0	0.0	180.1	100.0	30.62
Mean	100.3	7.755	180.3	100.0	30.65
3rd Qu.	127.0	17.0	216.2	113.0	36.75
Max	243.0	52.0	351.5	165.0	59.76
	Att.6	**Att.7**	**Att.8**	**Att.9**	**Att.10**
Min	0.0	0.0	0.0	0.0	0.0
1st Qu.	166.4	87.0	14.14	166.9	87.0
Median	201.0	100.0	17.09	200.4	100.0
Mean	200.6	100.2	17.05	200.4	99.92
3rd Qu.	234.1	114.0	19.90	234.7	113.0
Max	363.7	170.0	30.91	395	175.0
	Att.11	**Att.12**	**Att.13**	**Att.14**	**Att.15**
Min	0.0	0.0	0.0	0.0	0.0
1st Qu.	7.51	8.5	3.0	2.3	1.0
Median	9.02	10.30	4.0	2.78	1.0
Mean	9.018	10.26	4.435	2.771	1.57
3rd Qu.	10.56	12.0	6.0	3.24	2.0
Max	17.77	20.0	20.0	5.4	9.0

Att.1: account_length Att.2: number_vmail_messages Att.3: total_day_minutes Att.4: total_day_calls Att.5: total_day_charge	Att.6: total_eve_minutes Att.7: total_eve_calls Att.8: total_eve_charge Att.9: total_night_minutes Att.10: total_night_calls	Att.11: total_night_charge Att.12: total_intl_minutes Att.13: total_intl_calls Att.14: total_intl_charges Att.15: number_csc

Data Preprocessing

In this study, all numeric values in the data set were converted to values in the range of [0,1] using the linear data transformation method as shown in Equation (5).

$$x_{normal\ value} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (5)$$

However, some algorithms work with attributes having only numerical values. Therefore, for every categorical attribute, a corresponding dummy variable is generated for the purpose of analysis. In the “churn” data set, “area_code” attribute field consists of three categories (only attribute with multiple categories). Therefore, three dummy attribute fields, namely area_code1, area_code2, area_code3 were added instead of an area attribute field.

Performance Validation Methods, Evaluation Measures, and Proposed Evaluation Measure

For training and testing of data set, holdout method is used with 80 – 20% (training – test) rates. To evaluate model performance, accuracy, sensitivity, specificity, precision, and F-score values are considered, and churn customers’ class labels are used as reference values. The confusion matrix and performance evaluation measures are presented in Table 4.

Table 4. Confusion matrix and performance evaluation measures

<table>
<tr><td></td><td></td><td colspan="2">Actual Class</td><td></td></tr>
<tr><td></td><td></td><td>Positive</td><td>Negative</td><td>Evaluation Measure</td></tr>
<tr><td rowspan="2">Predicted Class</td><td>Positive</td><td>True Positive (TP)</td><td>False Positive (FP)</td><td>Precision: $\frac{TP}{TP+FP}$</td></tr>
<tr><td>Negative</td><td>False Negative (FN)</td><td>True Negative (TN)</td><td>Negative Prediction Value: $\frac{TN}{FN+TN}$</td></tr>
<tr><td></td><td>Evaluation Measure</td><td>Sensitivity: $\frac{TP}{TP+FN}$</td><td>Specificity: $\frac{TN}{TN++FP}$</td><td>Accuracy: $\frac{TP+TN}{TP+FP+FN+TN}$</td></tr>
</table>

In prediction studies such as churn analysis, it is important to identify churn customers rather than identifying the classes of all samples. When calculating the accuracy, the number of classes predicted accurately is checked for all instances, regardless of the class. For example, 100 of the records in a dataset are churn customers, of which 900 of them belong to non-churn customers. Consider that the class predictions of 900 records in the dataset are correct and only 10 of these accurate predictions are churned. In general, this model has an accuracy rate of 90 %; however, the true prediction rate for churn customers is only 10%. Therefore, a modified accuracy calculation approach is presented in this study. The equation for measuring accuracy is given below:

$$Modified\_Accuracy = MA = \frac{TP}{TP+FN+FP} \qquad (6)$$

According to this equation, customers with no churn status and those correctly classified by the model are excluded from the accuracy rate formulation.

Parameter Tuning for The Algorithms

The prediction models and performance values of these models are obtained for all combinations of the parameter values listed in Table 5.

Table 5. Summary of the prediction models created in the study

Algorithm	Parameters	Total Number of Models
KNN	***k***=3,5,7	6
NB	-	2
SVM	**Kernel function:** radial based, polynomial and sigmoid	6
ELM	**Neuron number:** [1,1000] integers **Activation function:** sigmoid, sin, radial basis, hard-limit, symmetric hard-limit, satlins, tan-sigmoid, triangular basis, positive linear, linear **Threshold values**=0.40, 0.50, 0.60, 0.70	80000

RESULTS

In this study, we have used R programming and RStudio editor to code all the models (R Development Core Team, 2008; RStudio Team, 2016). The following R packages are used for coding: For reading and printing data set xlsx (Dragulescu, 2014), for data transformation - clusterSim (Walesiak & Dudek, 2017), for hold-out performance validation method- caret (Kuhn, 2016), for k-nearest neighbor algorithm -class (Venables & Ripley, 2002), for SVM- e1071 (Meyer et al., 2015), and for ELM- elmNN (Gosso, 2012).

The values of the performance evaluation measurements are given in Table 6.

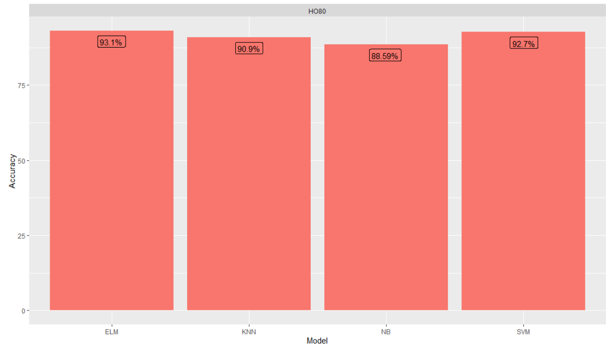
Table 6. Performance evaluation measure values

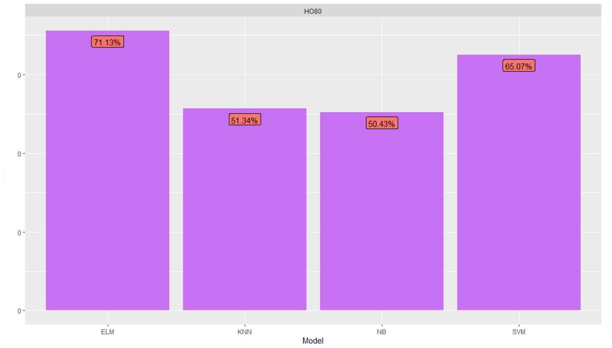
Alg.	Parameter	Acc.	Sens.	Spec.	Prec.	F Sc.	MA
KNN	*k*=5	0.9090	0.3664	**0.9908**	0.8571	0.5134	0.3453
NB	-	0.8859	0.4113	0.9639	0.6517	0.5043	0.3372
SVM	KF: polynomial	0.9270	0.5191	0.9885	**0.8718**	0.6507	0.4823
ELM	NN:584, AF: sin, TV:0.5	**0.9310**	**0.6489**	0.9735	0.7870	**0.7113**	**0.5519**

Acc.: Accuracy, AF: Activation function, Alg.: Algorithm, F Sc.: F-score, KF: Kernel function, MA: Modified Accuracy, NN: Number of neurons, Prec.: Precision, P.V.M.: Performance validation method, Sens.: Sensitivity, Spec.: Specificity, TV: Threshold value

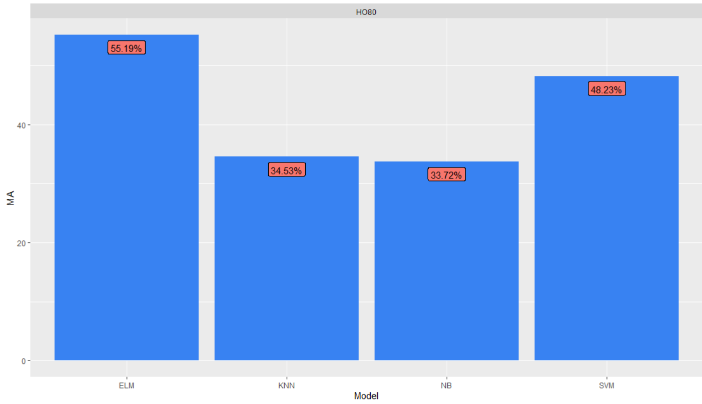
As shown in Table 6, highest accuracy of 93.10%, highest F-score of 71.13 %, sensitivity of 64.89 %, and modified accuracy of 55.19 % were obtained using ELM with 584 neurons, sinus activation function, and 0.5 threshold value. These values of parameters (accuracy, F-score, and modified accuracy) ensure improved model performance regardless of the performance validation method.

Figure 1 shows the comparative analysis of the prediction models.





a. Accuracy b. F Score



c. Modified Accuracy

Figure 1. Performance of the models for different measurements (Models are ELM, KNN, NB, and SVM respectively from left to the right for each figure)

DISCUSSION AND CONCLUSION

Customer churn analysis is management science problem. In this study, we addressed this problem using a hybrid approach based on ELM and grid search. We evaluated certain parameters to find an optimal solution that can improve the performance of the proposed model. The performance of ELM was then compared with that of various other algorithms, which plays an important role in the literature.

Based on the evaluation of parameters, namely accuracy, F-score, and modified accuracy using holdout performance validation method, the performance of ELM was better compared to that of other models. Considering the issue of unbalanced class distribution, although the churn rate is low, the churn prediction performance of ELM is better. According to the sensitivity values, churn for all churn customers can be classified with a prediction ratio of more than 70 %. Alternatively, when the prediction measure is examined, it is seen that customers who are classified as churn are exhibit a churn greater than an 80% ratio in reality. Therefore, if the class distribution is balanced, it can be observed that these rates will improve further. When compared with the performance indicators obtained in the studies examined in the literature, the results obtained within the scope of this study are at a level that can compete with different models. Using the developed models, accuracy values of approximately 85% were obtained by Jain et al. (2020),–75–85% by Lalwani et al. (2022), and 88–94% values by Asghar et al. (2021). In this context, when the method is examined, it was observed that ELM is quite advantageous considering the results obtained from different models and less personalized intervention in parameter tuning compared to other methods.

In addition, a modified accuracy ratio calculation approach was presented. The F-score is obtained from sensitivity and precision measurements provided a wider evaluation according to the accuracy value. When the results were examined, the modified accuracy value generated results similar to the F-score measure when ratios of accuracy and F-score measure were different. Therefore, in future studies, modified accuracy calculations should be considered instead of accuracy calculation.

The study was conducted using a dataset obtained from a public database. Therefore, the results obtained in this study can be validated by other researchers and compared with studies conducted with a similar dataset. The analyses conducted in this study can be repeated with a balanced dataset using various sampling methods, resulting in an increase in F-score

ACKNOWLEDGEMENT

This study was supported by the Scientific Research Project Coordination Unit of Istanbul University with project number49762 and titled Analytical approaches for solving churn analysis problems. We would also like to thank Prof. Ş. Alp Baray (PhD) for his contribution.

REFERENCES

Adwan, O., Faris, H., Jaradat, K., Harfoushi, O., & Ghatasheh, N. 2014. Predicting customer churn in telecom industry using multilayer preceptron neural networks: Modeling and analysis. Life Science Journal 11(3), 75–81.

Ahmad, A. K., Jafar, A., & Aljoumaa, K. 2019. Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data 6(28): 1-24.

Al-Shboul, B., Faris, H., & Ghatasheh, N. 2015. Initializing genetic programming using fuzzy clustering and its application in churn prediction in the telecom industry. Malaysian Journal of Computer Science 28(3): 213–20.

Asghar, M. N., Chaudhry, M. U., Rehman, A. U., & Yaseen, A. 2021. Next-wave of e-commerce: Mobile customers churn prediction using machine learning. Lahore Garrison University Research Journal of Computer Science and Information Technology 5(2): 62–72.

Beeharry, Y., & Tsokizep Fokone, R. 2021. Hybrid approach using machine learning algorithms for customers' churn prediction in the telecommunications industry. Concurrency and Computation: Practice and Experience e6627.

Chu, B.-H., Tsai, M.-S., & Ho, C.-S. 2007. Toward a hybrid data mining model for customer retention. Knowledge-Based Systems 20(8): 703–18.

Coussement, K., & De Bock, K. W. 2013. Customer churn prediction in the online gambling industry: The beneficial effect of ensemble learning. Journal of Business Research 66(9): 1629–36.

Coussement, K., Lessmann, S., & Verstraeten, G. 2017. A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. Decision Support Systems 95: 27–36.

De Caigny, A., Coussement, K., & De Bock, K. W. 2018. A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. European Journal of Operational Research 269(2): 760–72.

Dias, J., Godinho, P., & Torres, P. 2020. Machine learning for customer churn prediction in retail banking. In O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. A. C. Rocha, E. Tarantino, C. M. Torre, & Y. Karaca (Eds.), Computational Science and Its Applications. 576–589. Springer International Publishing.

Dolatabadi, S. H., & Keynia, F. 2017. Designing of customer and employee churn prediction model based on data mining method and neural predictor. 2nd International Conference on Computer and Communication Systems (ICCCS), Poland.

Dragulescu, A. A. 2014. xlsx: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files (R package version 0.5.7) [Computer software]. https://CRAN.R-project.org/package=xlsx

Elgeldawi, E., Sayed, A., Galal, A. R., & Zaki, A. M. 2021. Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis. Informatics 8(4): 79.

Gajowniczek, K., Orłowski, A., & Ząbkowski, T. 2016. Entropy based trees to support decision making for customer churn management. Acta Physica Polonica A 129: 971–79.

Gosso, A. 2012. ElmNN: Implementation of ELM (Extreme Learning Machine) algorithm for SLFN (Single Hidden Layer Feedforward Neural Networks) (R package version 1.0) [Computer software]. https://CRAN.R-project.org/package=elmNN

Huang, G. B., Zhu, Q.-Y., & Siew, C.-K. 2004. Extreme learning machine: A new learning scheme of feedforward neural networks. IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), Hungary.

Hudaib, A., Dannoun, R., Harfoushi, O., Obiedat, R., & Faris, H. 2015. Hybrid data mining models for predicting customer churn. International Journal of Communications, Network and System Sciences 08(05): 91-96

Hung, S.-Y., Yen, D. C., & Wang, H.-Y. 2006. Applying data mining to telecom churn management. Expert Systems with Applications 31(3): 515–524.

Jain, H., Khunteta, A., & Srivastava, S. 2020. Churn prediction in telecommunication using Logistic Regression and Logit Boost. Procedia Computer Science 167: 101–112.

Jain, H., Yadav, G., & Manoov, R. 2021. Churn prediction and retention in banking, telecom and IT sectors using machine learning techniques. In S. Patnaik, X.-S. Yang, & I. K. Sethi (Eds.), Advances in Machine Learning and Computational Intelligence (pp. 137–156). Springer.

Kaur, M., Singh, D. K., & Sharma, D. N. 2013. Data mining as a tool to predict the churn behaviour among indian bank customers. International Journal on Recent and Innovation Trends in Computing and Communication 1(9): 720-725.

Kim, K., Jun, C.-H., & Lee, J. 2014. Improved churn prediction in telecommunication industry by analyzing a large network. Expert Systems with Applications 41(15): 6575–84.

Kuhn, M. 2016. caret: Classification and regression training (R package version 6.0-73) [Computer software]. https://CRAN.R-project.org/package=caret

Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. 2022. Customer churn prediction system: A machine learning approach. Computing 104: 271-94.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. 2015. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (R package version 1.6-7) [Computer software]. https://CRAN.R-project.org/package=e1071

Nwaogu, V. C., & Dimililer, K. 2021. Customer churn prediction for business intelligence using machine learning. 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Turkey.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org

RStudio Team. (2016). RStudio: Integrated Development for R, RStudio, Inc. http://www.rstudio.com

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., & Brenning, A. 2019. Performance evaluation and hyperparameter tuning of statistical and machine-learning models using spatial data. Ecological Modelling 406: 109–120.

Shankar, K., Zhang, Y., Liu, Y., Wu, L., & Chen, C.H. 2020. Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. IEEE Access 8: 118164–73.

Sharma, A., Panigrahi, D., & Kumar, P. 2013. A neural network based approach for predicting customer churn in cellular network services. International Journal of Computer Applications 27(11): 26-31.

Srivastava, P. R., & Eachempati, P. 2021. Intelligent employee retention system for attrition rate analysis and churn prediction: An ensemble machine learning and multi-criteria decision-making approach. Journal of Global Information Management 29 (6): 1–29.

Stripling, E., vanden Broucke, S., Antonio, K., Baesens, B., & Snoeck, M. 2018. Profit maximizing logistic model for customer churn prediction using genetic algorithms. Swarm and Evolutionary Computation 40: 116–130.

Tsai, C.-F., & Lu, Y.-H. 2009. Customer churn prediction by hybrid neural networks. Expert Systems with Applications 36(10): 12547–12553.

Venables, W. N., & Ripley, B. D. 2002. Modern applied statistics with S (Fourth Edition). Springer**.**

Vo, N. N. Y., Liu, S., Li, X., & Xu, G. 2021. Leveraging unstructured call log data for customer churn prediction. Knowledge-Based Systems 212: 106586.

Walesiak, M., & Dudek, A. 2017. clusterSim: Searching for Optimal Clustering Procedure for a Data Set (R package version 0.45-2) [Computer software]. https://CRAN.R-project.org/package=clusterSim