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

DOI : 10.36909/jer.16771

Fatma Önay Koçoğlu*[1], Tuncay Özcan[2]

[1]Department of Software Engineering, Faculty of Engineering, Muğla Sıtkı Koçman University, Muğla, 48000, Turkey.

[2]Department of Management Engineering, Faculty of Management, İstanbul Technical University, İstanbul, 34367, Turkey.

* Corresponding Author: fonaykocoglu@mu.edu.tr

ABSTRACT

Customers' behaviors such as tendencies, loyalty status, satisfaction criteria show an alteration day by day due to the changing world. So, these behavior changes should be analyzed very well in every step of the decision-making process. Customer churn analysis is the determination of customers who tend to leave by analyzing the customer data with various methods before this situation occurs. This study aims to develop an Extreme Learning Machine based model for customer churn prediction problem and to determine the model parameters that provide the best performance. Grid search is used for hyperparameter tuning. Also modified accuracy calculation approach has been presented. The churn data set obtained from the UCI Machine Learning Repository has been used. Naive Bayes, k-Nearest Neighbor and Support Vector Machine methods are selected for performance comparison of the model. With a value of 93.1%, the best accuracy measure has been obtained with Extreme Learning Machine. Due to the low number of parameters to be determined and performance evaluation measures that compete with other models' results, it can be said that the Extreme Learning Machine is highly effective and interesting in the solution of the problem.

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

INTRODUCTION

Under changing world conditions, customers' behaviors such as tendencies, loyalty status, satisfaction criteria show an alteration. When the market reaches the saturation point, the number of suitable customers is very limited and the enterprises are focusing on promoting existing customers rather than targeting new customers (Hung et al., 2006). Customer churn is defined as the customer having a break or terminating his/her transactions with the company he/she previously received or performing with another company providing the same service/product (Hung et al., 2006; Kaur et al., 2013). For effective customer churn management, the forecasting model needs to be created to produce more effective and accurate results (Tsai & Lu, 2009).

In the literature, there are various studies using machine learning algorithms for churn prediction. In the study of Adwan et al. (2014) have presented Multilayer Perceptron (MLP) topologies-based Artificial Neural Networks (ANN) models. Al-Shboul et al. (2015) have developed a hybrid method based on the Fast-Fuzzy C-Means and Genetic Programming (GP) methods and aimed to reveal the outliers in the data set. Al-Shboul et al. (2015) have developed a hybrid method based on the Fast-Fuzzy C-Means and Genetic Programming (GP) methods and aimed to reveal the outliers in the data set. Coussement et al. (2017) have aimed to determine the effect of data preprocess on prediction success. Stripling et al. (2018) have introduced a classifier proposal based on optimizing the maximum benefit measure by using the Genetic Algorithm (GA). Jain et al. (2020) have proposed two prediction models based on Logistic Regression (LR) and Logit Boost (LB). In the study of Dias et al. (2020), for the retail banking customers, churn prediction up to 6 months have been taken into account and 6 different machine learning methods have been examined. The features that provide the best prediction performance have been also determined. Jain et al. (2021) have created the models with four algorithms, namely LR, Random Forest (RF), SVM and XGBoost, and datasets from three different fields, namely banking, telecom and IT, have been used. Lalwani et al. (2022)

have obtained various models by using LR, Naïve Bayes (NB), Support Vector Machine (SVM), RF, Decision Trees (DT). Hyperparameter tuning has been also performed. Beeharry & Tsokizep Fokone (2021) have proposed a two-layer flexible voting ensemble model to predict customer churn in telecommunications industries. The effect of whether the data set is balanced or not on the model success has been also examined. Nwaogu & Dimililer (2021) have employed SVM, MLP and Neural Networks (NN) machine learning algorithms. Some of other studies has been summarized 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

In this study, it is aimed to determine the performance of ELM algorithm for the solution of customer churn analysis problem, to perform parameter tuning of ELM method with the grid search optimization, to propose accuracy-based performance evaluation measure with a different perspective. This study differs from other studies in the literature in terms of the hyperparameter tuning performed and the calculation of the new accuracy measure used.

The rest of this paper is as follows in “Methodology” the data, methods, and tools has been detailed, 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 changes under the influence of many factors such as the structure of the data, its size and the number of records related to classes, algorithms, performance validation methods, sampling methods, and feature selection methods. One of the most important factors affecting model performance is the hyperparameters used for algorithms. Hyperparameters are parameters of the algorithm that can be manipulated by the model developer. Hyperparameter tuning to achieve the best model success is an important research topic. In this study, it is aimed both to measure the performance of the ELM algorithm for the determined problem and to obtain the parameters that can maximize the performance. In this direction, grid search optimization has been used. Naive Bayes (NB), k-Nearest Neighbor (KNN), Support Vector Machines (SVM) algorithms have been used for performance comparison of ELM. Due to their high efficiency in detecting the complex structure in the data, methods such as SVM, ANN, GA are widely preferred (Alsumaiei, 2021). Therefore, within the scope of this study, SVM and ANN-based ELM have been preferred. Also, accuracy-based performance evaluation measure has been proposed.

In this direction, firstly, the theoretical explanations of the methods used are given below and then, the steps followed in the study according to CRISP-DM are explained in order.

Extreme Learning Machine

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

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

where Ñ is the number of neurons, Ñ<N, g(x) is activation function, $w_i = \{w_{i1}, w_i, \dots, w_{in}\}^T$ is

weight vector associated with input and hidden neuron i, $\beta_i = \{\beta_{i1}, \beta_i, \dots, \beta_{im}\}^T$ is weight vector associating the hidden neuron *i* with output, b_i is threshold value for hidden neuron i. Under normal conditions, the output of the prediction models and real values indicated by Equation (1) for the class attribute can be different. To maximize network performance, this difference is expected to be 0 (zero) or minimum. In the ELM, it is tried to calculate the output vector that minimizes the error. Moore–Penrose Pseudo inverse method is used for the solution.

Hyperparameter Tuning and Grid Search Optimization

Machine learning algorithms require user-defined parameter values to obtain a balance between accuracy and generalizability (Srivastava & Eachempati, 2021). Determining the optimal hyperparameters for the machine learning model is crucial for the bias-reduced assessment of the predictive power of the model (Schratz et al., 2019). Hyper-parameter optimization is used for the selection of the appropriate parameter values of the algorithms and for better classification performance in this direction (Shankar et al., 2020). There is a need for more automated and guided methods to determine hyper parameter values. Methods used in this direction include Grid Search, Random Search and Bayesian Optimization (Elgeldawi et al., 2021). Grid Search method has been used within the scope of the study. In this method, model training is performed with the combination of all the values in finite intervals determined for the hyperparameters. In ANN models, the number of neurons is performed by trial and error without a specific rule (Munir et al., 2022). All trained models are compared according to the performance evaluation measures.

Application

The study was carried out in accordance with the CRISP-DM steps. Customer churn analysis has been chosen as the research problem of this study.

Data Set

The literature review shows that customer churn analysis applications are performed mostly in telecom and banking sector, respectively. The telecom churn data set from UCI Machine Learning Repository has been preferred for reasons such as the comparability of the results and the ability of the analysis by different researchers. In the data set, there are 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 been thought to not affect the results of the analyzes. 19 attributes have been used in the analyzes, 4 are categorical and 15 are numerical values. The data type and the explanations of the attributes have been given in Table 2. The churn ratio of data set is 14.14% (707 of the records are churn customers and 4293 of them are not churn).

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: Curent status of the customer (0:not churn, 1:churn)	Binary {0, 1}

The descriptive statistics of numerical attributes of the dataset has been given 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 values have been converted to values in the range of [0,1] by using the linear data transformation method as in Equation (5).

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

On the other side, some algorithms work with attributes which have only the numerical value. Usually, categorical attributes have been made ready for analysis with the help of dummy attributes. This process prevents data loss. In the “churn” data set, “area_code” attribute field consists of three categories (only attribute with multiple categories). Therefore, three dummy attribute fields have been added instead of the area attribute field. These are area_code1, area_code2, area_code3.

Performance Validation Methods, Evaluation Measures, and Proposed Evaluation Measure

Hold-out method has been used in order to use the data set for training and testing purposes. 80-20% (training-test) rates was used for the Hold-out method. In order to evaluate the model performances, accuracy, sensitivity, specificity, precision, and F-score values have been taken into consideration and churn customers' class label has been used as the reference value. The confusion matrix and performance evaluation measures have been given in Table 4.

Table 4. Confusion matrix and performance evaluation measures

		Actual Class		
		Positive	**Negative**	**Evaluation Measure**
Predicted Class	**Positive**	True Positive (*TP*)	False Positive (*FP*)	Precision: $\frac{TP}{TP+FP}$
	Negative	False Negative (*FN*)	True Negative (*TN*)	Negative Prediction Value: $\frac{TN}{FN+TN}$
	Evaluation Measure	Sensitivity: $\frac{TP}{TP+FN}$	Specificity: $\frac{TN}{TN++FP}$	Accuracy: $\frac{TP+TN}{TP+FP+FN+TN}$

It is important in the prediction studies, such as churn analysis, that is to identify, in particular, churn customers, rather than identifying the classes of all samples. When calculating the accuracy, the number of correctly predicted classes is checked for all instances, regardless of class. For example, 100 of the records in a data set are churn customers, 900 of them belong to non-churn customers. Consider that the class predictions of 900 records are correct in this data set and only 10 of these accurate predictions are churn. In general, this model has %90 accuracy rate but on the other side, true prediction rate of churn customers is only %10. For this reason, a modified accuracy calculation approach has been presented in this study. The equation of this accuracy measure is given below.

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

According to this equation, customers who have no churn status and are correctly classified by the model are excluded from the accuracy rate formulization.

Parameter Tuning for The Algorithms

The prediction models and the performance values of these models have been obtained for all combinations of parameter values given 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 fonc:** radial based, polynomial and sigmoid	6
ELM	**Noron number:** [1,1000] integers **Activation func:**sigmoid, sin, radial basis, hard-limit, symmetric hard-limit, satlins, tan-sigmoid, triangular basis, positive linear, linear **Treshold values**=0.40, 0.50, 0.60, 0.70	80000

RESULTS

R programming language and RStudio editor have been used in the coding of all models (R Development Core Team, 2008; RStudio Team, 2016). The following R packages have been used for the study: 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 Support Vector Machine - e1071 (Meyer et al., 2015), for Extreme Learning Machine - elmNN (Gosso, 2012).

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

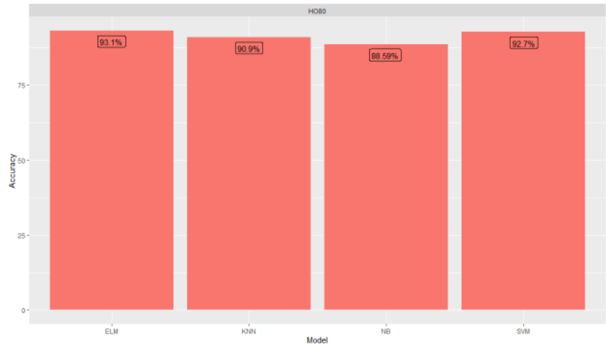
Table 6. Performance evaluation measure values

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**

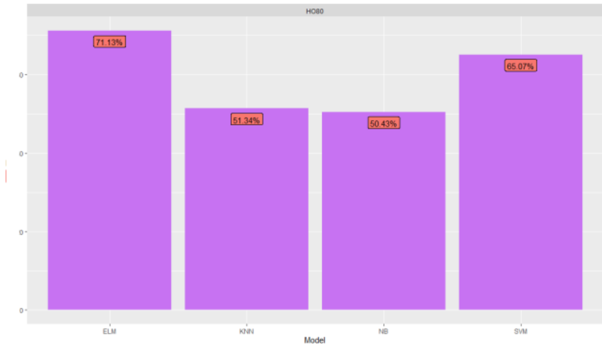
Acc.: Accuracy, AF: Activation function, Alg.: Algorithm, F Sc.: F-score, KF: Kernel function, MA: Modified Accuracy, NN: Number of neuron, Prec.: Precision, P.V.M.: Performans validation method, Sens.: Sensitivity, Spec.:Specificity, TV:Treshold value

According to Table 6, the highest accuracy has been obtained with 93.10%, the highest F-score with 71.13%, sensitivity with 64.89%, and modified accuracy with 55.19% has been obtained by using ELM with 584 neuron number, sinus activation function and 0.5 threshold value. These accuracy, F-score, and modified accuracy values are the best in all model performances regardless of performance validation method.

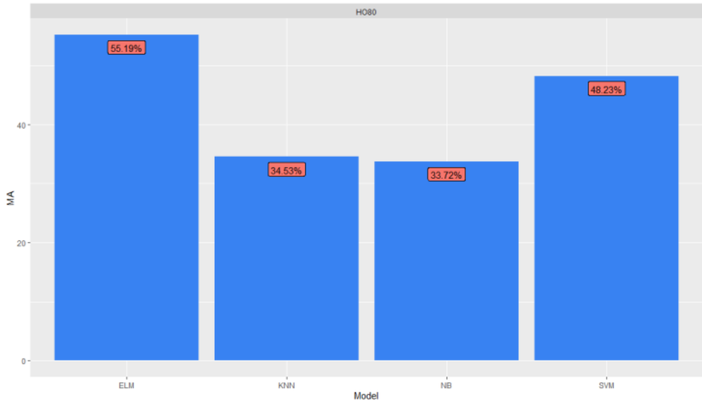
Figure 1 shows the comparative analysis of the prediction models.



a. Accuracy



b. F Score



c. Modified Accuracy

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

DISCUSSION AND CONCLUSION

In this study, it is aimed to develop a hybrid approach based on ELM and Grid Search for customer churn analysis which is considered as a management science problem. The parameters of producing the best solution under the conditions mentioned in the methodology has been determined. ELM's performance has been compared with the other various algorithms had an important place in the literature.

When the results were examined, the best performance according to the accuracy, F-score, and modified accuracy performance evaluation measures according to the hold-out performance validation method has been provided by the ELM among all methods. Considering the issue of unbalanced class distribution and although the churn rate is low, ELM's churn prediction performance is quietly well. According to sensitivity values, churns in all churn customers can be classified with more than 70% prediction ratio. In another respect, when the prediction measure is examined, it is seen that the customers who are classified as a churn are churn with more than 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 the study are at a level that can compete with different models. With the developed models, Accuracy values has been obtained as approximately 85% by Jain et al. (2020), values between 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 can be said that the ELM is quite advantageous considering the results obtained from different models and the less personalized intervention in the parameter tuning compared to other methods.

Also modified accuracy ratio calculation approach has been presented. F score measure is a measure obtained from sensitivity and precision measurements and it provides a wider evaluation according to the accuracy value. When the results are examined, the modified accuracy value produces similar results with the F score measure at the point where the accuracy ratio and the F score measure are different. Therefore, in future studies, modified accuracy calculation can be considered instead of accuracy calculation.

The study was carried out with a data set obtained from a public database. This allows the results obtained to be tested by other researchers and compared with the studies conducted with a similar data set. The analyzes can be repeated with data set balanced with various sampling methods. At this point, the increase in the value of F score is expected.

ACKNOWLEDGEMENT

This research study was supported by Scientific Research Project Coordination Unit of Istanbul University with the project number: 49762 and title: Analytical approaches for solving in churn analysis problem. Also, we would 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

Alsumaiei, A. 2021. Short-term forecasting of monthly water consumption in hyper-arid climate using recurrent neural networks. Journal of Engineering Research 9(3B): 56-69.

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

Munir, A., Khattak, A., Imran, K., Ulasyar, A., Haq, A. U., Khan, A., & Ullah, N. 2022. Artificial neural network based simplified one day ahead forecasting of solar photovoltaic power generation. Journal of Engineering Research 10(1A): 175–189.

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