



Machine Learning Approach and Model Performance Evaluation for Tele-Marketing Success Classification

Fatma Önay Koçoğlu, Muğla Sıtkı Koçman University, Turkey , İstanbul University-Cerrahpaşa, Institute of Graduate Studies, Turkey*

 <https://orcid.org/0000-0002-1096-9865>

Şakir Esnaf, İstanbul University-Cerrahpaşa, Turkey

 <https://orcid.org/0000-0001-6261-3118>

ABSTRACT

Up to the present, various methods such as data mining, machine learning, and artificial intelligence have been used to get the best assessment from huge and important data resources. Deep learning, one of these methods, is an extended version of artificial neural networks. Within the scope of this study, a model has been developed to classify the success of tele-marketing with different machine learning algorithms, especially with deep learning algorithms. Naïve Bayes, C5.0, Extreme Learning Machine, and Deep Learning algorithms have been used for modelling. To examine the effect of class label distribution on model success, synthetic minority oversampling technique has been used. The results have revealed the success of deep learning and decision trees algorithms. When the data set was not balanced, the deep learning algorithm performed better in terms of sensitivity. Among all models, the best performance in terms of accuracy, precision, and F-score have been achieved with the C5.0 algorithm.

KEYWORDS

Class Label Distribution, Classification, Decision Trees, Deep Learning, Deep Neural Network, Machine Learning, Tele-Marketing Banking

INTRODUCTION

Today's world is going through a period in which the amount of data increases rapidly, and knowledge-based strategies and decisions become more important. This can be seen as a result of the development of information and communication technologies and the new systems where more large-scale data can be stored, and advanced analyses can be performed. The first step in accessing knowledge is storing the data. Today, data is collected from different sources. In order to make this stored data meaningful, it should be analyzed with the methods according to the structure of the data. Although the data analysis started with the use of statistical methods, today, data analysis is also carried out with Machine Learning algorithms as part of Artificial Intelligence (AI).

AI journey, which started with the Alan Turing's Turing test in the 1950s, continues today with driverless vehicles, more effective chat boxes and intelligent robots. AI is the ability to perform mental functions such as thinking, reasoning, and learning by computer or computer-controlled machines

DOI: 10.4018/IJBAN.298014

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

(Nabiyev, 2005). AI is a field that deals with the development of computer systems capable of processing symbols that can be used to solve problems that cannot be solved by algorithmic structures (Becerra-Fernandez et al., 2004). There are two basic steps of AI: Learning and practicing. In the learning step, various experiences are provided in order to determine the response of the computer systems in the face of an event, and in the second step, these systems are provided to produce new human-specific response without any command. Machine Learning methods are used in the learning step. Machine Learning is defined as the branch of computer algorithms developed to be used for transforming the stored data into a smart action by imitating the above-mentioned features of human (El Naqa & Murphy, 2015). Also, Deep Learning has recently become a prominent method of AI and Machine Learning. Deep Learning algorithms has been obtained by the development of Artificial Neural Networks (ANN) in order to make inference at a higher level and increase the predictive performance (Greenspan et al., 2016).

While all these developments are taking place, it is possible to talk about a world where competition conditions change. As a result of the globalizing world, the rivals have been moved from the national to the international level and the competition have become more difficult. At this point, to continue their existence and to determine the effective strategies and decisions, enterprises need to carry out new marketing and R&D activities and to use knowledge as an intellectual capital more effectively. In the banking sector, various campaigns are developed for customers and transactions with customers are wished to be kept active. Accordingly, one of the most important methods used to increase interactions and sales is marketing. With the development of telecommunication technologies, marketing has evolved towards telemarketing (Mustapha & Alsufyani, 2019). Telemarketing is a marketing method in which phone and call centers are used to acquire potential customers, sell products or provide services to existing customers (Kotler & Keller, 2016). However, the expectations, demands and needs of each customer may differ. For this reason, it should be determined which customers will be interested in the developed campaigns. With defining target customers, it is possible for the campaigns to reach higher success with less time and cost. Classification models to be created with machine learning algorithms and recorded data make it easier for tele-marketing managers to make more accurate and faster decisions at this point. Increasing marketing efficiency with data recorded in the database is one of the main issues that are still emphasized and require intensive research (Ghatasheh et al., 2020). Although many researchers carry out various studies with machine learning methods in line with this purpose, changing data set characteristics, new methods developed, different approaches in obtaining models show that this issue is open to research. For this reason, the problem of the success of telemarketing classification has been handled.

With this study, it is aimed to obtain models with high classification performance by using machine learning algorithms to determine whether the telemarketing campaign will be successful and specially to determine the success of the Deep Learning approach for this problem. In this direction, different classification models have been obtained with C5.0 Decision Tree Algorithm, Naive Bayes Algorithm (NB), Extreme Learning Machine (ELM), and Deep Learning based ANNs (DL) and comparative performance evaluation has been presented.

In the next parts of the study, similar studies in the literature, the study process steps, explanations about the used methods, and the model performance values obtained have been included. The results have been evaluated and interpreted in the last section.

BACKGROUND OF THE STUDY

There are many studies conducted with different Machine Learning algorithms in the literature for the problem of tele-marketing success classification. Some of these studies have been summarized below.

Pan and Tang (2014) have carried out a study on the distribution of class labels frequently observed in the data sets. In this respect, they have applied bagging and boosting methods. In the analyses performed with Logistic Regression and Neural Networks, it has been observed that Neural Networks performed better. Shashidhara et al. (2015) have conducted a study on the most suitable

platform for the analysis of marketing data in the banking sector with Machine Learning techniques. In the study, Weka, Scikit Learn, Apache Spark environments have been examined with two different data sets. Logistic Regression and Support Vector Machines algorithms have been used for analysis. Although the results show that the Scikit Learn environment has the fastest results with a certain size data set, the most suitable solution is provided by Apache Spark due to its the parallel programming capability when the large data set is considered. Bahari and Elayidom (2015) have developed models with ANN and Naive Bayes algorithms using banking marketing data. When the results are examined, it is seen that ANNs take a lot of time in terms of processing time even if the evaluation values are close for both algorithms. Ruangthong and Jaiyen (2015) have aimed to investigate the effect of class label distribution on classification success and to improve the results of the analysis with different sampling methods. Synthetic Minority Oversampling Technique (SMOTE) and Rotation Forest Algorithm (Rotation Forest) based model has been proposed. ANNs, Naive Bayes, Bayesian Networks and Decision Tree (C4.5, Random Forest) algorithms have been used for the evaluation of model performance. The results have showed better results with the proposed model. Keles and Keles (2015) have aimed to develop an Intelligent Bank Market Management System (IBMMS) that uses Machine Learning-based methods to manage marketing campaigns efficiently. In order to develop the learning model of the system, decision tree algorithm has been used. The results obtained with the decision tree classification model, an inference engine using the decision structures and the expert opinion have been compared with other methods such as Neural Networks, Logistic Regression, Naive Bayes and Support Vector Machines in terms of accuracy, precision and specificity. Although the results indicate that the neural networks show a slightly better result, when considering the system in general it has been considered appropriate to use decision trees as a predictive model. Serrano-Silva et al. (2016) have applied supervised learning method to predict tele-marketing success. Related analyses have been performed using four different versions of a bank data set. C4.5, k-Nearest Neighbour, Repeated Incremental Pruning to Produce Error Reduction-RIPPER, Neural Networks (Multilayer-Feed Forward), Sequential Minimal Optimization, and Naive Bayes algorithm have been preferred to create models. Although the best performance was achieved with RIPPER, it has been observed that the performance of the algorithm was significantly lower compared to other studies in the literature. This situation has been explained by the fact that the unbalanced class label distribution. Parlar and Acaravcı (2017) have aimed to determine the most important features to increase the efficiency of campaigns. In this respect, information gain and chi-square coefficients have been calculated. Customer data obtained from bank marketing processes has been used. As a result, the most important attribute is the period of being the customer of the bank. Kachwala and Sharma (2017) have stated that customer data should be analyzed by using Machine Learning techniques in order to predict the response of customers to the campaigns so that a correct marketing strategy can be developed. In this direction, the models have been developed with C4.5, Naive Bayes, Stochastic Gradient Descent Algorithm, and ANNs. It has been concluded that the algorithms could predict with the accuracy rate of approximately 80%. Palaniappan et al. (2017), performed the classification performance evaluation of algorithms using bank telemarketing data. Naive Bayes, Random Forest, and Decision Tree algorithms, accuracy, precision and recall rates performance evaluation measures have been used. Koumédio et al. (2018) developed a classification method based on supporting the most important features within the model. The researchers who tested the method based on the telemarketing problem used Naive Bayes, Decision Trees, Artificial Neural Network and Support Vector Machines algorithms for comparison. Hassan et al. (2019), conducted a heuristic and comparative study to evaluate data mining techniques. In the study in which bank telemarketing data was used, attribute extraction has been made, Neural Network, Support Vector Machine, Decision Trees, k-Nearest Neighbour, Naive Bayes and Logistic Regression algorithms have been used. According to the true positive rate, the best result was obtained by Logistic Regression. In the study of Halim et al. (2020), a model has been developed to predict successful calls with the telemarketing data set. In the model, especially the data pre-processing phase has been emphasized, the data set has been passed through various data

cleaning, balancing and normalization processes. For classification ANN has been preferred. The results obtained show that these processes in the data pre-processing step have a positive contribution to the performance of the model. Borugadda et al. (2021) used various machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), Decision Tree (DT) and Logistic Regression (LR) to explore the demand for long-term bank deposits. Desai and Khairnar (2022) propose a hybrid model based on artificial neural network and extreme gradient boosting by examining the comprehensive features associated with a Portuguese bank's customers, products and services. The proposed model was compared with known machine learning methods and successful results were obtained. Feng et al. (2022) proposed a method for predicting the success of bank telemarketing time deposit sales. The proposed dynamic ensemble selection method takes into account average profit maximization as well as forecast accuracy. Compared to known machine learning methods, the proposed method has achieved successful results.

When the literature is evaluated, the studies can be grouped as suggestions of new classification models with different algorithms, balancing class distribution, and determining important features (feature selection).

As can be seen from the literature, classical machine learning algorithms especially Decision Tree, Naïve Bayes and ANN have been used for the models. When using ANNs, a deep neural nets approach has not been used. On the other hand, when Deep Learning studies are examined, it is seen that the studies are carried out with non-structural data such as images and sound, and various deep learning architectures are used. Kim et al. (2015) have adopted the deep learning approach for the telemarketing success classification, but unlike this study, they have benefited from Convolutional Neural Network architecture. In this study, especially ANNs have been deepened without any architecture, and the performance of the model obtained accordingly has been evaluated. Also, Extreme Learning Machine Algorithm, which is a different ANN based algorithm, has been also used for model production.

Ruangthong and Jaiyen (2015), have applied the sampling process over the whole data set. While over sampling methods produce new records for the minority class, sometimes they can repeat the same records. It may occur that these records are in both data sets to be used for model training and testing. In this study, considering this situation, the balancing process has been carried out in two ways, both balancing the data used for model training and balancing the whole data set.

Machine learning model performance varies depending on factors such as algorithms, parameters, pre-processing steps, selected performance validation methods, feature selection methods preferred by the researcher. Model performance, on the other hand, is evaluated according to various measures, process time, and easy applicability of the model on the business side, and the best model is selected as a result of this evaluation. In this study the opportunity to make an evaluation according to all performance evaluation measures that can be obtained with the confusion matrix has been provided.

METHODOLOGY

The study has been carried out in accordance with the Cross Industry Standard Process for Data Mining (CRISP-DM) steps used for both Data Mining and Machine Learning. The CRISP-DM process consists of six steps:

- Step 1: Understanding the problem
- Step 2: Understanding the data
- Step 3: Preparing the data for analysis
- Step 4: Developing the models
- Step 5: Evaluating the model performances
- Step 6: Choosing the best model and application

Accordingly, the problem has been identified first.

Understanding the Problem and the Data

Enterprises need more complex techniques to model customer behavior due to the increasing costs and low response rates of marketing campaigns. Because the ability to predict the reaction of a customer before the campaign will provide a great advantage in the mentioned issues. Within the scope of the study, the problem of telemarketing success classification has been addressed. For this reason, it is aimed to present a model that produces better results than similar studies in the literature according to various performance indicators and to create a different option for managers.

To be more detailed, in order to determine the success of the telemarketing campaign offered by a bank to its customers, it is aimed to

- obtain the success performances of well-known machine learning algorithms,
- use a model to be created with Deep Learning based ANNs,
- perform comparative performance analysis,
- and determine the effect of class label distribution on model performance.

Bank Marketing dataset from UCI Machine Learning Repository has been used in the study (URL1). This data set has been obtained by the Portuguese banking institution as a result of direct marketing campaigns during the campaign period between May 2008 - November 2010. The data set consists of 21 attributes, one of which is class label. The class attribute has two values as “success” and “fail”. Therefore, the problem is handled as a binary classification problem. 41188 records. 36548 records belong to failed and 4640 records (11%) belong to the successful campaign. As can be seen from the rates, class attribute distribution is not balanced. In other words, while the rate of records (or samples) belonging to one class (success) in the whole data set is 11%, the number of records (or samples) belonging to the other class (fail) is 89%.

Data Pre-processing

The most important points for achieving the best performance of the model are that the data must be suitable for the model to be used in the analysis and pre-processed accordingly. The values of the attributes in the data set may have different ranges. Some attributes' values range from 0 to 1, while values for other attributes can take values that are much larger. This may affect the results of the analysis. In this study, the data set has been subjected to linear data transformation. The following equation has been used to convert all values in the data set to values in the range [0-1].

$$x_{normalValue} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

On the other hand, categorical attribute values cannot be used in some algorithms especially based on ANN. In this case, categorical attribute values can be introduced as numerical values to models with the help of dummy attributes. In this study, categorical attribute values in the data set have been transformed into binary values with the help of dummy variables. With the use of dummy attributes, the data set consists of 63 attributes, one of is class label.

The «euribor3m» variable has been removed from the data set because it generated noise. The pdays attribute field shows the time elapsed over the last call with customer for previous campaign and receives numerical values. However, for customers that have not been contacted before, this field is defined as 999. Values have been changed from 999 to 0.

Within the scope of this study, due to the unbalanced distribution of the class attribute included in the data, the necessary configuration has been performed by the Synthetic Minority Oversampling Technique (SMOTE) algorithm. With this method, the number of samples in the minority class, which

has fewer records in the whole data set, is synthetically increased. Thus, the records numbers of the two classes are balanced. Two approaches have been considered for sampling. The first is balancing the class attribute field values only in the training data set during the training phase of the models, and the second is balancing the class attribute field values for the whole data set. This is because when the model is trained with a balanced data set, it is desired to observe how the classification performance will be against an unbalanced test data set.

Modelling

For the solution of the problem, a Deep Learning based classification model has been created according to the different parameter values. For the comparison of the model performance, different classification models have been obtained with C5.0 Decision Tree Algorithm, Naive Bayes Algorithm, and Extreme Learning Machine. Algorithm descriptions have been given below.

Artificial Neural Networks and Deep Learning

Artificial Neural Networks (ANN) is a computer system developed to automatically realize the ability to derive, produce and discover new information through learning, like the functions of the human brain, without any help (Öztemel, 2016). ANN have been designed considering the human biological neural networks. The biological nerve cell consists of synapses, dendrite, cell body and axon structures to collect and reprocess electrical signals and then transmit the new signal to other nerve cells. The operation of ANN is like biological neural networks. Accordingly, the signals from the input layer are collected in the nerve cell (neuron), the nerve cell produces an output of its own with the help of a function, and the output produced is transmitted to the next layer (neuron).

In the process of Machine Learning, each attribute field in the dataset constitutes the input layer. The values from the input layer and the weights determined for these values are converted to the Net Input Information (NII) value by means of the selected addition function. NII is converted to output value by an activation function together with output weights. The resulting output value indicates the class of the sample given as input.

Let (x_p, y_j) be N different and randomly selected examples, $x_j = \{x_{j1}, x_{j2}, \dots, x_{jn}\}^T$ be input, and $y_j = \{y_{j1}, y_{j2}, \dots, y_{jm}\}^T$ be output. The mathematical model of the Single Hidden Layer Feed-Forward Neural Network (SLFN) is as follows:

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i x_j + b_i) = 0 \quad (j=1,2,\dots,N) \quad (2)$$

\tilde{N} : Number of neurons. N is the number of samples in the training data set $\tilde{N} < N$

$g(x)$: Activation function

$w_i = \{w_{i1}, \dots, w_{in}\}^T$ weight vector that associates the hidden neuron i with input

$\beta_i = \{\beta_{i1}, \dots, \beta_{im}\}^T$ the weight vector that correlates hidden neuron i with output

b_i : Threshold value for the process element i

The model's error is calculated by comparing the output values produced by the estimation models obtained with ANNs and the real values that should be. To obtain a correct model, this error is eliminated according to the structure of neural network. The error is distributed backward into the network in feedback neural network which is commonly used. In this direction, the calculations are done backwards, the error rates are calculated, and the weights are re-determined (Öztemel, 2016; Özkan et al., 2018). The formulas for the calculations have been given below. Let \hat{y} be the output value, the total error E is calculated as:

$$E_{total} = \frac{1}{2n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (3)$$

Partial derivatives are obtained according to the following formula in order to measure the effect of each change in weight on the total error.

$$\frac{\partial E_{total}}{\partial w_k} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) x_{ik} \quad (4)$$

w_k represents the new weight and w'_k represents the previous weight value; weights are updated as follows:

$$w_k \leftarrow w'_k - \eta \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) x_{ik} \quad (5)$$

For an ANN, calculations are made based on various variables such as inputs given to the network, architecture of the network, algorithm to be used in the learning step and activation function in order to produce output in the education phase (Özen & Gülseçen, 2016). In ANNs, the representation of the input with different feature vectors, the selection of the features in the network, the fewer neurons and the deepening of the network with more hidden layers provide the transition from the ANN to the Deep Learning structure. Deep Learning has gained importance due to the fact that it does not experience problems which occurs in other ANN algorithms in the discovery of complex structures in high dimensional data and it can be applied in many fields (LeCun et al., 2015). Chen et al. (2015) describe Deep Learning as a type of neural network, typically consisting of more than three layers. The most important difference of Deep Learning from other methods is that feature extraction is done by this method without the need for using a different method. Representative learning is a set of methods that allow a machine to feed raw data and to automatically discover the necessary impressions for classification (LeCun et al., 2015), and Deep Learning is a method of learning from data representation (Şeker et al., 2017). There are various structures used with Deep Learning such as autoencoders, restricted Boltzmann machine, deep belief network, and convolutional networks.

C5.0 Decision Tree Algorithm

One of the methods frequently used in the literature is Decision Trees. A decision tree has a hierarchical structure, and, in this structure, models are created similar to a real tree structure. The model includes a root node, branches, and leaves attached to it. With the decision tree model, a series of rules are formed, and decisions are made according to the branching in the tree structure. Each path from the root node to the leaf node represents a rule (Quinlan, 1987). Unlike other algorithms, pruning is performed in C5.0 algorithm against the excess branching situation that may occur in decision trees algorithms. Pruning is carried out by including the pruned leaves in the upper nodes or by replacing the pruned sub-tree with the tree that uses this pruned sub-tree the most (Silahtaroglu, 2008; Akpınar, 2014). In this regard, Quinlan (1996) proposed the use of the Gain Ratio (GR) value. Accordingly, branching is realized by selecting the attribute value having the highest gain ratio.

Extreme Learning Machine

The method developed by Huang et al. (2004) based on ANNs differs from neural networks in the output layer. The solution of the matrix obtained in this layer is based on Moore-Penrose method. The

solution is obtained by calculating the output weights minimizing the error. Advantageous aspects of the method include the need to define fewer parameters and faster processing. The mathematical model of the method used in this article is same with single hidden layered neural networks.

Naive Bayes Algorithm

Naive Bayes algorithm is a statistical and probability-based algorithm based on Bayesian Theorem. The model of the algorithm is constructed according to the conditional probability equation. Conditional probability is the expression of the probability of occurrence of one event depending on the probability of occurrence of another event.

Performance Validation and Evaluation

Hold-out and cross-validation performance validation methods have been used in the classification models. In the hold-out method, 70%-30% (training-test) and 80%-20% (training-test) ratios have been used to determine the training-test data sets. The results of the cross-validation method have been obtained for 5 and 10 folds.

Five different performance evaluation measures have been used to measure the performance of the models: accuracy, F-score, sensitivity (also known as recall), specificity, and precision.

Tools

R programming language and RStudio editor have been used for coding all models (R Development Core Team, 2008; RStudio Team, 2016). The following packages have been used in this study:

- Read and write the dataset - `xlsx` (Dragulescu & Arendt, 2018)
- Data transformation - `clusterSim` (Walesiak & Dudek, 2017)
- Hold-out - `caret` (Kuhn, 2018)
- Cross validation - `TunePareto` (Müffel et al., 2012)
- Support Vector Machine Algorithm - `e1071` (Meyer et al., 2015)
- Extreme Learning Machine Algorithm - `elmNNRcpp` (Mouselimis & Gosso, 2018)
- Decision Tree Algorithm - `C50` (Kuhn & Quinlan, 2018)
- Naive Bayes Algorithm - `naivebayes` (Majka, 2018)
- SMOTE Algorithm – `DMwR` (Torgo, 2010)
- Deep Learning Algorithm – `h2o` (LeDell et al., 2019)

FINDINGS

In this section, the performance evaluation measure values of the models created within the scope of the study have been given. Some initials have been used in all tables. The explanations for these initials are as follows:

HO70: Algorithm performs with hold-out method with 70% train-30% test data,

HO80: Algorithm performs with hold-out method with 80% train-20% test data,

CV5: Algorithm performs with cross-validation with 5-fold,

CV10: Algorithm performs with cross-validation with 10-fold,

SMOTE_HO70: Algorithm performs with hold-out method with 70% train-30% test data, and the trained data is balanced with SMOTE algorithm

SMOTE_HO80: Algorithm performs with hold-out method with 80% train-20% test data, and the trained data is balanced with SMOTE algorithm

SMOTE_CV5: Algorithm performs with cross-validation with 5-fold, and the trained data is balanced with SMOTE algorithm

SMOTE_CV10: Algorithm performs with cross-validation with 10-fold, and the trained data is balanced with SMOTE algorithm
 SMOTE_V2_HO70: Algorithm performs with hold-out method with 70% train-30% test data, and the whole data is balanced with SMOTE algorithm
 SMOTE_V2_HO80: Algorithm performs with hold-out method with 80% train-20% test data, and the whole data is balanced with SMOTE algorithm
 SMOTE_V2_CV5: Algorithm performs with cross-validation with 5-fold, and the whole data is balanced with SMOTE algorithm
 SMOTE_V2_CV10: Algorithm performs with cross-validation with 10-fold, and the whole data is balanced with SMOTE algorithm

Firstly, all performance evaluation measure values obtained from the models have been classified according to the classification algorithms used for model training.

Findings of C5.0 Decision Tree Algorithm have been given in Table 1.

Table 1. Findings of C5.0

Model	Accuracy	Sensitivity	Specificity	Precision	F-score
HO70	0,9170	0,5309	0,9661	0,6652	0,5905
HO80	0,9144	0,4806	0,9695	0,6667	0,5585
CV5	0,9126	0,5095	0,9638	0,6419	0,5675
CV10	0,9141	0,5119	0,9652	0,6519	0,5726
SMOTE_HO70	0,8779	0,8937	0,8759	0,4775	0,6225
SMOTE_HO80	0,8821	0,8675	0,8840	0,4870	0,6238
SMOTE_CV5	0,8723	0,8696	0,8727	0,4648	0,6056
SMOTE_CV10	0,8756	0,8750	0,8757	0,4725	0,6134
SMOTE_V2_HO70	0,9560	0,9470	0,9636	0,9559	0,9514
SMOTE_V2_HO80	0,9567	0,9491	0,9630	0,9553	0,9522
SMOTE_V2_CV5	0,9565	0,9478	0,9637	0,9561	0,9519
SMOTE_V2_CV10	0,9565	0,9478	0,9637	0,9561	0,9519

According to Table 1, the highest accuracy value is 95.67% and the F-score reached the highest value of 95.22% with the use of SMOTE sampling method (applied to all data set), and hold-out method with 80-20% training-test rates, the highest specificity value is 96.95% obtained by dividing the data set by hold-out method without sampling.

Findings of Naive Bayes Algorithm have been given in Table 2.

According to Table 2, the highest accuracy value with 88.88% and the highest specificity value with 92.87% have been obtained by dividing the data set by hold-out method without sampling. The highest precision rate is 81.49% and the highest F-score value is 79.68%. These values have been obtained by dividing the data set by cross validation method with SMOTE sampling (applied to all data set).

Table 2. Findings of Naive Bayes

Model	Accuracy	Sensitivity	Specificity	Precision	F-score
HO70	0,8874	0,5675	0,9280	0,5003	0,5318
HO80	0,8888	0,5744	0,9287	0,5057	0,5378
CV5	0,8801	0,5612	0,9206	0,4730	0,5133
CV10	0,8801	0,5614	0,9206	0,4732	0,5135
SMOTE_CV5	0,8373	0,6851	0,8566	0,3776	0,4868
SMOTE_CV10	0,8367	0,6856	0,8558	0,3765	0,4860
SMOTE_HO70	0,8466	0,6739	0,8686	0,3943	0,4975
SMOTE_HO80	0,8442	0,6832	0,8647	0,3906	0,4971
SMOTE_V2_HO70	0,8189	0,7825	0,8493	0,8122	0,7971
SMOTE_V2_HO80	0,8168	0,7791	0,8482	0,8105	0,7945
SMOTE_V2_CV5	0,8193	0,7794	0,8524	0,8149	0,7968
SMOTE_V2_CV10	0,8193	0,7794	0,8524	0,8149	0,7968

Findings of Extreme Learning Machine have been given in Table 3.

When the analyses were performed by selecting the number of 50 neurons, sigmoid activation function and separating the training and test data set by 10-fold cross validation, the accuracy (89.65%) has reached the highest value. When the analyses were performed by selecting the number of 150 neurons and sinus activation function, separating the training and test data set by %80-20 ratio, and using SMOTE (applied to all data set), the F-score (71.56%) has reached the highest value. When the SMOTE sampling method was included in the models with hold-out method, an increase in sensitivity and F-score has been observed.

Findings of Deep Learning Algorithm have been given in Table 4.

It can be seen from Table 4, when the training and test data set were separated by 80-30% and balanced with SMOTE method (applied to all data set), the analyses were performed with the RectifierWithDropout activation function, the accuracy, and F-score have reached their highest values with 90.55% and 89.76% respectively. When the SMOTE sampling method was included in the models, a significant increase especially in sensitivity has been observed for all performance validation methods. Generally, the highest accuracy values for all performance evaluation methods have been obtained by using RectifierWithDropout activation function.

In order to make a comparative analysis of the classification models used, Table 1, Table 2, Table 3, and Table 4 have been summarized in Table 5. The methods are ordered according to the best accuracy, the best sensitivity and the best F-score performance evaluation measures of the algorithms.

When the table is examined, it is seen that the highest accuracy value reaches 95.67% with C5.0 decision tree algorithm. The lowest value belongs to the analysis performed with Naive Bayes algorithm and the accuracy is 88.88%. The accuracy value of Deep Learning is 90.55%. When the methods are ordered according to the sensitivity values from the highest value to the lowest value, it is seen that the highest sensitivity value reached to 94.91% with C5.0 decision tree algorithm. The lowest value belongs to the analysis performed by Extreme Learning Machine algorithm and the sensitivity is 57.84%. For the F-score also, the successful ranking of the algorithms has not changed.

Table 3. Findings of Extreme Learning Machine

Model	Act. Func*	NN*	Accuracy	Sensitivity	Specificity	Precision	F Score
HO70	sig	50	0,8960	0,1594	0,9905	0,6829	0,2585
	sin	10	0,7003	0,2641	0,7563	0,1220	0,1669
	radbas	150	0,8890	0,0342	0,9986	0,7619	0,0654
	sig	500	0,8863	0,0014	0,9998	0,5000	0,0028
HO80	sig	50	0,8960	0,1695	0,9913	0,7200	0,2743
	sin	10	0,7004	0,2657	0,7575	0,1257	0,1707
	radbas	150	0,8866	0,0335	0,9986	0,7619	0,0641
	sig	500	0,8838	0,0010	0,9997	0,3333	0,0021
CV5	sig	50	0,8963	0,1524	0,9908	0,6782	0,2488
	sin	10	0,6952	0,2735	0,7487	0,1214	0,1678
	sig	200	0,8960	0,1136	0,9953	0,7558	0,1971
	sig	500	0,8900	0,0336	0,9987	0,6555	0,0625
CV10	sig	50	0,8965	0,1565	0,9905	0,6775	0,2538
	sin	10	0,7007	0,2683	0,7556	0,1223	0,1675
	radbas	150	0,8914	0,0569	0,9973	0,7396	0,1053
	sig	500	0,8909	0,0450	0,9983	0,7075	0,0826
SMOTECV5	sig	10	0,8941	0,1373	0,9901	0,6377	0,2257
	radbas	500	0,7170	0,4662	0,7488	0,3840	0,3111
	sig	500	0,8876	0,0084	0,9993	0,5534	0,0165
SMOTEHO80	sig	10	0,8915	0,1297	0,9915	0,6667	0,2172
	sin	500	0,8725	0,2427	0,9552	0,4158	0,3065
	sin	50	0,7397	0,2793	0,8002	0,1551	0,1994
	sig	150	0,8903	0,0921	0,9951	0,7097	0,1630
	sig	500	0,8837	0	0,9997	0	NaN
SMOTEHO70	sig	10	0,8937	0,1445	0,9899	0,6465	0,2362
	radbas	500	0,8886	0,1808	0,9794	0,5292	0,2695
	sin	150	0,7633	0,3174	0,8205	0,1849	0,2337
	sig	100	0,8882	0,0313	0,9981	0,6769	0,0599
	sig	500	0,8863	0,0014	0,9998	0,5000	0,0028
SMOTE_V2_HO70	sin	200	0,7849	0,5435	0,9861	0,9702	0,6967
	radbas	500	0,7197	0,3874	0,9966	0,9897	0,5568
	sig	500	0,5526	0,0162	0,9995	0,9658	0,0319
SMOTE_V2_HO80	sin	150	0,7910	0,5784	0,9682	0,9381	0,7156
	sin	500	0,7668	0,4912	0,9966	0,9917	0,6570
	sig	500	0,5493	0,0088	0,9996	0,9535	0,0175
SMOTE_V2_CV5	sin	500	0,7810	0,5246	0,9948	0,9886	0,6844
	sin	150	0,7764	0,5355	0,9772	0,9538	0,6839
	sig	500	0,5552	0,0220	0,9995	0,9527	0,0424

*Act Fun: Activation function, NN: Neuron number

Table 4. Findings for Deep Learning

Model	Activation Func.	Acc.*	Sen.*	Spec.*	Pre.*	F-score
HO70	RectifierWithDropout	0,8972	0,6932	0,9234	0,5372	0,6053
	Tanh	0,8885	0,6676	0,9168	0,5073	0,5765
	MaxoutWithDropout	0,8834	0,7680	0,8982	0,4918	0,5996
HO80	RectifierWithDropout	0,9012	0,7312	0,9235	0,5565	0,6320
	Tanh	0,8963	0,6538	0,9282	0,5444	0,5941
	MaxoutWithDropout	0,8842	0,7688	0,8993	0,5007	0,6064
CV5	RectifierWithDropout	0,8982	0,7073	0,9224	0,5364	0,6100
	Tanh	0,8829	0,7644	0,8979	0,4882	0,5954
	MaxoutWithDropout	0,8785	0,7515	0,8946	0,4757	0,5824
CV10	RectifierWithDropout	0,8935	0,7310	0,9141	0,5207	0,6072
	Tanh	0,8901	0,7237	0,9112	0,5089	0,5974
	MaxoutWithDropout	0,8822	0,6864	0,9071	0,4904	0,5454
SMOTE HO70	RectifierWithDropout	0,8525	0,8633	0,8511	0,4265	0,5710
	Tanh	0,7619	0,7964	0,7575	0,2964	0,4320
	MaxoutWithDropout	0,6809	0,7338	0,6741	0,2241	0,3434
SMOTE HO80	RectifierWithDropout	0,8374	0,9100	0,8279	0,4098	0,5651
	Tanh	0,8282	0,8818	0,8212	0,3930	0,5437
	MaxoutWithDropout	0,7764	0,8243	0,7701	0,3201	0,4611
SMOTE CV5	RectifierWithDropout	0,8360	0,8851	0,8297	0,3980	0,5489
	Tanh	0,7975	0,8084	0,7962	0,3366	0,4746
	MaxoutWithDropout	0,7851	0,7938	0,7840	0,3345	0,4626
SMOTE CV10	RectifierWithDropout	0,8355	0,8672	0,8314	0,3952	0,5427
	Tanh	0,7599	0,8392	0,7499	0,3320	0,4692
	MaxoutWithDropout	0,7537	0,8216	0,7451	0,3082	0,4441
SMOTE_V2_HO70	RectifierWithDropout	0,8479	0,8338	0,8597	0,8320	0,8329
	Tanh	0,8745	0,9026	0,8511	0,8347	0,8673
	MaxoutWithDropout	0,7120	0,6134	0,7942	0,7129	0,6594
SMOTE_V2_HO80	RectifierWithDropout	0,9055	0,9116	0,9003	0,8840	0,8976
	Tanh	0,8611	0,8666	0,8565	0,8342	0,8501
	MaxoutWithDropout	0,7960	0,5793	0,9767	0,9539	0,7208
SMOTE_V2_CV5	RectifierWithDropout	0,8843	0,9029	0,8687	0,8512	0,8761
	Tanh	0,8763	0,8965	0,8596	0,8419	0,8682
	MaxoutWithDropout	0,8708	0,8972	0,8489	0,8339	0,8636
SMOTE_V2_CV10	RectifierWithDropout	0,8785	0,8808	0,8766	0,8560	0,8681
	Tanh	0,8589	0,8614	0,8568	0,8347	0,8470
	MaxoutWithDropout	0,7881	0,8056	0,7735	0,7605	0,7779

*Acc.: Accuracy, Sen.: Sensitivity, Spec.: Specificity, Pre.: Precision

Table 5. Summary of performance analysis of classification models.

Performance Measure	Alg.*	Model	Acc.*	Sen.*	Spec.*	Pre.*	F-score
Accuracy	C5.0	SMOTE_V2_HO80	0,9567	0,9491	0,9630	0,9553	0,9522
	DL	SMOTE_V2_HO80	0,9055	0,9116	0,9003	0,8840	0,8976
	ELM	CV10	0,8965	0,1565	0,9905	0,6775	0,2538
	NB	HO80	0,8888	0,5744	0,9287	0,5057	0,5378
Sensitivity	C5.0	SMOTE_V2_HO80	0,9567	0,9491	0,9630	0,9553	0,9522
	DL	SMOTE_V2_HO80	0,9055	0,9116	0,9003	0,8840	0,8976
	NB	SMOTE_V2_HO70	0,8189	0,7825	0,8493	0,8122	0,7971
	ELM	SMOTE_V2_HO80	0,7910	0,5784	0,9682	0,9381	0,7156
F-score	C5.0	SMOTE_V2_HO80	0,9567	0,9491	0,9630	0,9553	0,9522
	DL	SMOTE_V2_HO80	0,9055	0,9116	0,9003	0,8840	0,8976
	NB	SMOTE_V2_CV5	0,8193	0,7794	0,8524	0,8149	0,7968
	ELM	SMOTE_V2_HO80	0,7910	0,5784	0,9682	0,9381	0,7156

*Alg.: Algorithm, Acc.: Accuracy, Sen.: Sensitivity, Spec.: Specificity, Pre.: Precision

DISCUSSION AND CONCLUSION

In addition to technological advances, human expectations, and interest play an important role in increasing the importance of data and knowledge day by day. The first step in accessing knowledge is storing data. In this direction, the data collected from different sources are kept in various data storage environment and various methods are developed for processing this data stack, which is increasing in size and changing in structure. One of these methods is Machine Learning. Machine Learning is a component of AI and is used to simulate human learning ability by computers or computer-aided machines. One of the methods of Machine Learning is ANN which has been developed based on human biological nerve cells. Deep Learning method has been developed to eliminate the problems with ANNs in the use of data which are difficult to process such as image, sound, and text. Accordingly, in Deep Learning method the data is converted to various representations and, feature extraction, estimation, and definition stages are performed.

Machine Learning methods are used to solve problems in different fields such as health, education, finance, and security. One of these problems is classification of tele-marketing success. In this study, the effectiveness of different machine learning algorithms in solving the problem was investigated by using the data of a bank's tele-marketing campaign. In addition, it is aimed to measure the success of Deep Learning algorithms against classical machine learning algorithms and to determine the effect of balancing the data set. Within the scope of the study, a bank's tele-marketing campaign data has been used and the data set has been organized in two different types according to algorithms that can work with both numerical data and mixed data. Since the class attribute field of the data set is unbalanced distributed, the balancing process has been applied using the SMOTE method. The balancing process has been carried out in two ways, both balancing the data used for model training and balancing the whole data set. C5.0, NB, ELM, DL algorithms and as the performance validation method, hold out (with 70-30% and 80-20% training-test rates) and cross validation (5 and 10 folds) have been applied. Accuracy, sensitivity, specificity, precision, and F-score performance evaluation measures have been calculated for all models. All models have been coded with R programming language.

The algorithms have been analyzed both within themselves and comparatively according to accuracy, sensitivity and F-score performance evaluation measures. In the comparative evaluation,

it is seen that the best classification model is obtained with the C5.0 algorithm. C5.0 is followed by DL, NB and ELM algorithms, respectively. In this ranking, it is seen that, in terms of accuracy, by far better results have been obtained with the C5.0 algorithm, especially as a result of balancing the data set. On the other hand, when the distribution of the class attribute field is balanced using the SMOTE algorithm, a decrease in the accuracy values has been observed, while an increase in the sensitivity values has been detected. When examining performance values, it can be said that besides the accuracy value, sensitivity and F-score gain more importance in terms of evaluation. The sensitivity is the success of predicting positive class among all positive classes. Positive classes are successful calls in the data set. That is, positive classes represent customers that a bank will pay attention for its marketing campaign. The accuracy value provides an overall assessment, while the sensitivity provides an assessment for target customers. The findings obtained in this direction show that models using C5.0 and Deep Learning algorithm have been successful by a significant difference (over 90%). Success in sensitivity is interpreted as follows: "Of all successful calls, successful calls can be accurately predicted (above 90%) using C5.0 or DL models." Again, the most unsuccessful algorithm according to the sensitivity measure is ELM. F-score, on the other hand, is a measure calculated according to both precision and sensitivity. Even if the best F-score and sensitivity values among all models belong to the C5.0 algorithm, when the data set is not balanced, the DL algorithm performs better than the C5.0 algorithm in all model validation methods according to sensitivity and F-score. From this point of view, it can be said that DL algorithm can be preferred in unbalanced data sets if resampling method is not used.

For similar problem, Ruangthong and Jaiyen (2015) reached accuracy value of 91.24% in their study, the highest accuracy value reached by Shashidhara et al. (2015) was 90.81% and the highest sensitivity value was 94.48%. In the study of Kim et al. (2015), who performed Deep Learning application with CNN, the highest accuracy value was reached with 76.70%. Keles and Keles (2015) achieved 82.96% accuracy and 87% sensitivity with the decision support system they developed. In the models that Kachwala and Sharma (2017) obtained using different data sets and different classification algorithms, accuracy values in the range of 90-95% were obtained. When the results obtained with this study are evaluated, it is seen that better results are obtained in terms of performance indicators compared to similar studies in the literature.

Another visible result of the study is the effect of balancing data set on model performances. As mentioned before, balancing has been examined from two aspects. It is thought that the better the model is trained, the better the model classification performance will be. For this reason, it has been tested whether it would be sufficient to perform the class attribute field balancing process only for the training data set. The results obtained show that if the test data set as well as the training data set is balanced similarly, improvement can be achieved in all performance indicators. Otherwise, there may be improvements especially in the sensitivity measure, but it has been observed that there is not enough improvement in values such as accuracy and F-score. In this direction, it is thought that the efficiency of the data balancing process will be at a certain level for real life problems. At this point, the success of the Deep Learning algorithm, which can produce better results than other algorithms with an unbalanced data set, is also important.

According to the indicators, it can be concluded that C5.0 and DL algorithms perform better for solving the problem. On the other hand, especially for DL and ELM algorithms, it is necessary to specify more parameters than other algorithms. Therefore, it will be possible to improve the performance of these algorithms with more sensitive parameter tuning. Considering the structure of the data set, it can be said that the data set is more suitable for C5.0 or NB. Therefore, achieving higher success for these algorithms can be considered normal. If the models obtained are evaluated in terms of an enterprise, the models to be selected according to the determined performance indicators are at a level to help the enterprise to make successful calls.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

- Akpinar, H. (2014). *Data: Veri madenciliği veri analizi*. Papatya Bilim Üniversitesi Yayıncılık.
- Bahari, T. F., & Elayidom, M. S. (2015). An efficient crm-data mining framework for the prediction of customer behaviour. *Procedia Computer Science*, 46, 725–731. doi:10.1016/j.procs.2015.02.136
- Becerra-Fernandez, I., Gonzalez, A., & Sabherwal, R. (2004). *Knowledge management: Challenges, solutions and technologies*. Prentice Hall.
- Borugadda, P., Nandru, P., & Madhavaiah, C. (2021). Predicting the success of bank telemarketing for selling long-term deposits: An application of machine learning algorithms. *St. Theresa. Journal of the Humanities and Social Sciences*, 7(1), 91–108.
- Chen, Y., Zhao, X., & Jia, X. (2015). Spectral–spatial classification of hyperspectral data based on deep belief network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6), 2381–2392. doi:10.1109/JSTARS.2015.2388577
- Desai, R., & Khairnar, V. (2022). Hybrid prediction model for the success of bank telemarketing. In *Intelligent Sustainable Systems* (pp. 693–710). Springer. doi:10.1007/978-981-16-2422-3_54
- Dragulescu, A. A., & Arendt, C. (2018). *xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files*. R package version 0.6.1. Retrieved from <https://CRAN.R-project.org/package=xlsx>
- El Naqa, I., & Murphy, M. J. (2015). What is machine learning? In *Machine Learning in Radiation Oncology* (pp. 3–12). Springer International Publishing. doi:10.1007/978-3-319-18305-3_1
- Feng, Y., Yin, Y., Wang, D., & Dhamotharan, L. (2022). A dynamic ensemble selection method for bank telemarketing sales prediction. *Journal of Business Research*, 139, 368–382. doi:10.1016/j.jbusres.2021.09.067
- Ghatasheh, N., Faris, H., AlTaharwa, I., Harb, Y., & Harb, A. (2020). Business analytics in telemarketing: Cost-sensitive analysis of bank campaigns using Artificial Neural Networks. *Applied Sciences (Basel, Switzerland)*, 10(7), 1–15. doi:10.3390/app10072581
- Greenspan, H., Van Ginneken, B., & Summers, R. M. (2016). Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5), 1153–1159. doi:10.1109/TMI.2016.2553401
- Halim, K. N. A., Jaya, A. S. M., & Fadzil, A. F. A. (2020). Data pre-processing algorithm for neural network binary classification model in bank tele-marketing. *International Journal of Innovative Technology and Exploring Engineering*, 9(3), 272–277. doi:10.35940/ijitee.C8472.019320
- Hassan, D., Rodan, A., Salem, M., & Mohammad, M. (2019, November). *Comparative study of using data mining techniques for bank telemarketing data*. Paper presented at the Sixth HCT Information Technology Trends (ITT), Ras Al Khaimah, UAE. doi:10.1109/ITT48889.2019.9075123
- Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2004). Extreme learning machine: A new learning scheme of feedforward neural networks. In *Proceedings IEEE International Joint Conference on Neural Networks* (pp. 985-990). IEEE.
- Kachwala, T., & Sharma, L. K. (2017). Comparative study of supervised learning in customer relationship management. *International Journal of Computer Engineering & Technology*, 8(6), 77–82.
- Keles, A., & Keles, A. (2015). IBMMS decision support tool for management of bank telemarketing campaigns. *International Journal of Database Management Systems*, 17(5), 1–15. doi:10.5121/ijdms.2015.7501
- Kim, K. H., Lee, C. S., Jo, S. M., & Cho, S. B. (2015). Predicting the success of bank telemarketing using deep convolutional neural network. In *SoCPaR 2015: Proceedings of Seventh International Conference of Soft Computing and Pattern Recognition* (pp.314-317). IEEE. doi:10.1109/SOCPAR.2015.7492828
- Kotler, P., & Keller, K. L. (2016). *A framework for marketing management*. Pearson Education Limited.
- Koumédio, C. S. T., Cherif, W., & Hassan, S. (2018). Optimizing the prediction of telemarketing target calls by a classification technique. In *Proceedings of 6th International Conference on Wireless Networks and Mobile Communications* (pp. 1-6). IEEE. doi:10.1109/WINCOM.2018.8629675

- Kuhn, M. (2018). *Caret: Classification and regression training*. R package version 6.0-80. Retrieved from <https://CRAN.R-project.org/package=caret>
- Kuhn, M., & Quinlan, R. (2018). *C50: C5.0 Decision trees and rule-based models*. R package version 0.1.2. Retrieved from <https://CRAN.R-project.org/package=C50>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. doi:10.1038/nature14539 PMID:26017442
- LeDell, E., Gill, N., Aiello, S., Fu, A., Candel, A., Click, C., Kraljevic, T., Nykodym, T., Aboyoun, P., Michal Kurka, M., & Malohlava, M. (2019). *h2o: R Interface for 'H2O'*. R package version 3.22.1.1. Retrieved from <https://CRAN.R-project.org/package=h2o>
- Leung, K. M. (2007). *Naive bayesian classifier*. Retrieved from <http://cis.poly.edu/~mleung/FRE7851/f07/naiveBayesianClassifier.pdf>
- Majka, M. (2020). *High performance implementation of the naive bayes algorithm*. Retrieved from <https://github.com/majkamichal/naivebayes>
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2015). *e1071: Misc functions of the department of statistics, probability theory group* (Formerly: E1071). TU Wien, R package version 1.6-7. Retrieved from <https://CRAN.R-project.org/package=e1071>
- Mouselimis, L., & Gosso, A. (2018). *elmNNRcpp: The extreme learning machine algorithm*. R package version 1.0.1. Retrieved from <https://CRAN.R-project.org/package=elmNNRcpp>
- Murphy, K. P. (2006). *Naive bayes classifiers*. Retrieved from <https://www.cs.ubc.ca/~murphyk/Teaching/CS340-Fall06/lectures/naiveBayes.pdf>
- Müssel, C., Lausser, L., Maucher, M., & Kestler, H. A. (2012). Multi-objective parameter selection for classifiers. *Journal of Statistical Software*, 46(5), 1–27. doi:10.18637/jss.v046.i05
- Mustapha, S. M. F. D. S., & Alsufyani, A. (2019). Application of Artificial Neural Network and information gain in building case-based reasoning for telemarketing prediction. *International Journal of Advanced Computer Science and Applications*, 10(3), 300–306. doi:10.14569/IJACSA.2019.0100339
- Nabiyev, V. (2005). *Yapay zekâ*. Seçkin Yayıncılık.
- Özen, Z., & Gülseçen, S. (2016). R ile geri yayılım yapay sinir ağı algoritması kullanarak bir sınıflandırma uygulaması. In M. E. Balaban & E. Kartal (Eds.), *R ile Veri Madenciliği Uygulamaları* (pp. 73–100). Çağlayan Kitabevi.
- Özkan, Y., Koçoğlu, F. Ö., & Erol, Ç. (2018). Prediction of student performans by deep learning algorithm. In *FL 2018: Proceedings of 7th International Conference On Innovations In Learning For The Future: Digital Transformation in Education* (pp.25-25). İstanbul, Turkey: İstanbul University Press.
- Öztemel, E. (2016). *Yapay sinir ağları*. Papatya Bilim Üniversitesi Yayıncılığı.
- Palaniappan, S., Mustapha, A., Foozy, C. F. M., & Atan, R. (2017). Customer profiling using classification approach for bank telemarketing. *International Journal on Informatics Visualization*, 1(4-2), 214-217.
- Pan, Y., & Tang, Z. (2014). Ensemble methods in bank direct marketing. In *ICSSSM 2014: 11th International Conference on Service Systems and Service Management* (pp.1-5). IEEE.
- Parlar, T., & Acaravcı, S. K. (2017). Using data mining techniques for detecting the important features of the bank direct marketing data. *International Journal of Economics and Financial Issues*, 7(2), 692–696.
- Quinlan, J. R. (1987). Generating production rules from decision trees. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence* (pp.304-307). Morgan Kaufmann.
- Quinlan, J. R. (1996). Improved use of continuous attributes in C4.5. *Journal of Artificial Intelligence Research*, 4, 77–90. doi:10.1613/jair.279
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. Retrieved from <https://www.R-project.org>

RStudio Team. (2016). *RStudio: Integrated development for R*. RStudio, Inc. Retrieved from <https://www.rstudio.com/>

Ruangthong, P., & Jaiyen, S. (2015). Bank direct marketing analysis of asymmetric information based on machine learning. In *JCSSE 2015: Proceedings of 12th International Joint Conference on Computer Science and Software Engineering* (pp. 93-96). IEEE. doi:10.1109/JCSSE.2015.7219777

Şeker, A., Diri, B., & Balık, H. H. (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Journal of Engineering Sciences*, 3(3), 47–64.

Serrano-Silva, Y. O., Villuendas-Rey, Y., & Yáñez-Márquez, C. (2016). Telemarketing success: Evaluation of supervised classifiers. *Research in Computing Science*, 118(1), 157–166. doi:10.13053/rcs-118-1-15

Shashidhara, B. M., Jain, S., Rao, V. D., Patil, N., & Raghavendra, G. S. (2015). Evaluation of machine learning frameworks on bank marketing and Higgs datasets. In *ICACCE 2015: Proceedings of Second International Conference on Advances in Computing and Communication Engineering* (pp.551-555). IEEE. doi:10.1109/ICACCE.2015.31

Silahtaroglu, G. (2008). *Kavram ve algoritmalarıyla temel veri madenciliği*. Papatya Yayıncılık Eğitim.

Torgo, L. (2010). *Data Mining with R, Learning with Case Studies*. Chapman and Hall/CRC. doi:10.1201/b10328

Walesiak, M., & Dudek, A. (2017). *clusterSim: Searching for Optimal Clustering Procedure for a Data Set*. R package version 0.45-2. Retrieved from <https://CRAN.R-project.org/package=clusterSim>

Fatma Önay Koçoğlu is an Assistant Professor of Software Engineering Department at Muğla Sıtkı Koçman University. She graduated from Istanbul University (IU) Mathematics Department and Sakarya University Industrial Engineering Department. With her dissertations on Data Mining, she received her MSc degree in 2012 and PhD in 2017 in Informatics from IU. She worked as research fellow at the Laboratoire de recherche en informatique de Sorbonne Université, for the academic year of 2021-2022. She worked as a Research Assistant at IU Informatics Department for 11 years and has several national and international papers published. Her research areas are Artificial Intelligence, Data Mining, Machine Learning, Mathematical Modelling, and Optimization.

Sakir Esnaf is a Professor of Department of Industrial Engineering at Istanbul University-Cerrahpasa. He graduated from Istanbul Technical University, Department of Industrial Engineering in 1986. He received his M.Sc and Ph.D degrees in Production Management from Istanbul University. He worked as a postdoctoral research fellow at the Intelligent Fuzzy Systems Laboratory, Department of Mechanical and Industrial Engineering at University of Toronto, Canada. His current research topics are facility location, fuzzy clustering, metaheuristics, big data analytics, theory of constraints, and revenue management.