A data modeling approach for classification problems: application to bank telemarketing prediction

Conference Paper · March 2019

DOI: 10.1145/3320326.3320389

3 authors:

Cédric Stéphane Koumetio Tekouabou
Mohammed VI Polytechnic University
35 PUBLICATIONS **230** CITATIONS
SEE PROFILE

Walid Cherif
School of Information Sciences
29 PUBLICATIONS **325** CITATIONS
SEE PROFILE

Silkan Hassan
39 PUBLICATIONS **207** CITATIONS
SEE PROFILE

A data modeling approach for classification problems: application to bank telemarketing prediction

Stéphane Cédric Koumetio Tekouabou
Laboratory LIMA,
Department of Computer Science,
Faculty of Sciences.
B.P. 20, 24000,
El Jadida, Morocco
ctekouaboukoumetio@gmail.com

Walid Cherif
Laboratory SI2M, Department of Computer Science,
National Institute of Statistics and Applied Economics,
B.P. 6217,
Rabat, Morocco.
chrf.walid@gmail.com

Hassan Silkan
Laboratory LIMA,
Department of Computer Science,
Faculty of Sciences.
B.P. 20, 24000,
El Jadida, Morocco
silkan_h@yahoo.fr

ABSTRACT

In this paper, we present a new data modeling approach for five common classification algorithms to optimize the prediction of telemarketing target calls for selling bank long-term deposits. A Portuguese retail bank addressed, from 2008 until 2013, data on its clients, products and social-economic attributes including the effects of the financial crisis. An original set of 150 features has been explored and 21 features are retained for the proposed approach including label. This paper introduces a new modeling approach that preprocessed separately each type of features and normalize them to optimize prediction performance. To evaluate the proposed approach, this paper compares the results obtained with five most known machine learning techniques: Naïve Bayes (NB), Logistic Regression (LR), Decision Trees (DT), Artificial Neural Network (ANN) and Support Vector Machines (SVM) and it yielded better improved performances for all these algorithms in terms of accuracy and f-measure.

Keywords

Bank telemarketing; Data Mining; Data Modeling, classification algorith; Optimization; ….

NISS19, March 27–29, 2019, Rabat, Morocco

ACM ISBN 978-1-4503-6645-8/19/03…$15.00
https://doi.org/10.1145/3320326.3320389

1. INTRODUCTION

Recently, technological progress deeply affects Marketing domain with now specific demand for specific target. Thus, many companies prefer target marketing campaign instead of mass marketing campaign which become very less effective because of the intensive competition when facing challenges from a rapidly changing market situation [1].

Direct marketing methods such as telemarketing is in the main strategy of many banks and insurance companies for interacting with their customers and make new business [2] [3]. Huge information collected on clients became very important and useful for the target marketing strategy such as in telemarketing [4] [2]. Because of it remoteness characteristic, telemarketing is an operationalized direct marketing through a contact center which consists of an interactive technique to solicit prospective customers via phone, mails, social medias… in order to make direct sales of products or services [5], [6] [4].

The success of telemarketing campaign, is to focus on the quality of prospect data, attempting to understand costumers behavior and predict the expected customers that could have higher probability to be a client by using machine Learning techniques [6] [7]. From one case to another, depending on the selected features, different machine learning techniques can be used.

Machine learning is a subdomain of Data Mining that studies automatic techniques for learning to make accurate predictions based on past observations. Machine Learning uses two types of techniques: supervised learning (classification and regression), which trains a model on known input and output data so that it can predict future outputs, and unsupervised learning (clustering), which finds hidden patterns or intrinsic structures in input data. Classification techniques (Naïve Bayes, k-Nearest Neighbors, Decision Trees, Artificial Neural Network, Support Vector Machines,…) [8]–[10] have been

used in previous works on Portuguese Bank telemarketing campaign prediction.

Based on the results obtained in those previous works on Portuguese dataset, we noticed that the importance of the features may vary and consequently, a consideration of this factor, classifying clients should be adopted. In this sense, this paper introduces a new modeling approach that preprocessed separately each type of features and normalize them to optimize prediction performance. To evaluate the proposed approach, this paper compares the results obtained with five most known machine learning models: Naïve Bayes (NB), Logistic Regression (LR), Decision Trees (DT), Artificial Neural Network (ANN) and Support Vector Machines (SVM).

The rest of this paper is organized as follows: Section II summarizes main works that applied machine learning techniques in bank telemarketing; section III introduces the proposed technique and details the process to adapt for each type of features. Section IV analyzes the obtained results and compares the proposed approach to other machine learning models. Finally, section V concludes this work.

2. BACKGROUND

In previous research of last 10 years, many authors have focused their works on bank telemarketing success prediction with data mining techniques applied mainly on Portuguese bank database which will be used in this paper. The original set of 150 features has been explored and 22 most significant features are retained for the proposed approach.

S. Crone & al. [11] have investigated the influence of different preprocessing techniques of attribute scaling, sampling, coding of categorical as well as coding of continuous attributes on the classifier performance of decision trees, neural networks and support vector machines. Supported by a multifactorial analysis of variance, they have provided empirical evidence that DPP has had significant impact on predictive accuracy. But this work published in 2006 was not on Portuguese dataset as well as those that will be presented in the rest of this section

S. Moro & al. [7], have compared in their first paper four Data mining models (Logistic Regression, k-Nearest Neighbors, Decision Trees, Neural Network, Support Vector Machines) by using two metrics: area of the receiver operating characteristic curve (AUC) and area of the LIFT cumulative curve (ALIFT). Those four models were also tested on an evaluation dataset, using the most recent data (after July 2012) and a rolling window scheme. He has obtained the best results with NN (AUC=0.8 and ALIFT= 0.7), allowing to reach 79% of the subscribers by selecting the half better classified clients. In their next papers [12], he has used Customer Lifetime Value (CLV) to improve predictive performance without even having to ask for more information to the companies they serve. He has used CRISP-DM methodology [13] to show how we can increase campaign efficiency by identifying the main characteristics that affect success and this has been also shown in Z. Chu Thesis [14] .

H. A. Elsalamony [4] has used Multilayer perceptron neural network (MLPNN) and Ross Quinlan has used new decision tree model (C5.0) in his paper on the same bank data base to increase the campaign effectiveness by identifying the main characteristics that affect a success. He has estimated the performances by three statistical measures; classification accuracy, sensitivity, and specificity

However, D. Grzonka & al. [15] have discussed and compared classification methods (decision trees, bagging, boosting, and random forests) and according to their analysis, the effectiveness of previous campaigns has been the most significant parameter to predict consumer decisions to open or not a deposit in the bank. Their best results has been obtained for random forests even though the largest percentage of true positive classifications was obtained for a single decision tree.

While R. Farooqi [16] has attempted to analyze the data mining techniques and its useful application in banking industry like marketing and retail management, CRM, risk management and fraud detection.

In another work, T. Palar & al. [17] have used the same dataset considering two feature selection methods namely information gain and Chi-square methods to select the important features. The methods have been compared using a supervised machine learning algorithm of Naive Bayes and their experimental results have shown that reduced set of features improves the classification performance. B. Famina & E. Sudheep [18] have proposed efficient CRM-data mining framework and have studied two classification models, Naïve Bayes and Neural Networks to show that the accuracy of Neural Network is comparatively better.

Y. Kawasaki & M. Ueki [19] have examined predictive modeling with several sparse regression methods for bank telemarketing success they have concluded that an effective predictive model may help in reducing costs for marketing in companies.

We have noticed that last years, many authors have been interested by this topic of predicting potential future client in telemarketing campaign. Many data mining approaches have been studied and applied at the same dataset giving different performances that remain also optimizable. To improve this performance we will introduce in this paper introduces a new modeling approach that preprocessed separately each type of features and normalize them to optimize prediction performance. To evaluate the proposed approach, this paper compares the results obtained with five most known machine learning models: Naïve Bayes (NB), Logistic Regression (LR), Decision Trees (DT), Artificial Neural Network (ANN) and Support Vector Machines (SVM).

3. THE PROPOSED APPROACH

To evaluate the performance of the proposed approach, the direct marketing campaign of "Portuguese bank dataset" is used and will be describe as follow.

3.1 Dataset

The data is related with direct marketing campaigns (phone calls) of a Portuguese banking institution and provided by The UCI Machine Learning Repository (Lichman, 2013). Data were collected from May 2008 to June 2013 [7]. The total number of training phone contacts database (the sample size) is 41188 and the marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be yes (y = 1) or no(y = 0) for subscribe [6] [19]. The dataset includes 21 variables (including binary label variable y) where numeric and categorical variables are mixed as we can see in Table 1. For more details of the data, see the information available at The UCI Machine Learning Repository.[7]

Table 1. DETAILS OF IMPUT VARIABLES OF PORTUGUESE BANK DATASET

N°	Variable Name	Description	Type
1	Age	Age of the client	Numerical
2	Job	Type of client's job	Categorical
3	Marital	Client's marital status	Categorical
4	Education	Highest education of the client	Categorical
5	Default	Client credit	Categorical
6	Housing	Housing loan	Categorical
7	Loan	Personal loan	Categorical
8	Contact	Type of contact communication with the client	Categorical
9	Month	Last month of the year contracting to the client	Categorical
10	Day of Week	Last day of the week contracting to the client	Categorical
11	Duration	Duration of client contact	Numerical
12	Campaign	Number of contacts performed during this campaign and for this client	Numerical
13	Pdays	Number of days elapsed after the client's last visit	Numerical
14	Previous	Number of contacts performed before this campaign and for this client	Numerical
15	Poutcome	Outcome of the previous marketing campaign	Categorical
16	Emp.var.rate	Employment variation rate	Numeric
17	Cos.price.idx	Consumer price index	Numeric
18	Cons.conf.idx	Consumer confidence index	Numeric
19	Euribor3m	Euribor 3 month rate	Numeric
20	Nr.employed	Number of employees	Numeric
21	Label	Client subscription	Categorical

3.2 The proposed modeling approach

This classification technique consists of 3 big steps which are: preprocessing, normalization and classification. To illustrate these different steps, let's considered some of the attribute figuring in the dataset.

Table 2. INITIAL DATA VALUES

Instance	Age	Marital	Housing	Education	Y
I_1	59	married	no	Professional	**no**
I_2	39	single	yes	Basic.9y	**yes**
I_3	59	married	no	university	**no**
I_4	41	divorced	no	Unknown	**no**
I_5	44	married	no	Basic.4y	**yes**

3.2.1 Preprocessing

The first step of the classification technique consists on defining the type of the feature as the technique differs for either type.

3.2.1.1 Preprocessing 1: Codification

- For numerical features: (Ref. Table 1)

We calculate directly statistic parameters (min, max, mean, variance, standard deviation) of each numerical feature.

- For categorical features:

We distinguish 3 subtypes of features: scaled features, binomial features and nominal features.

-For scaled values (Education, Month, day-of-week):

We substitute items by their ordinal number. After that substitution, we calculate the statistics parameters as for numerical values.

-For Boolean features:

We have only two possibilities yes or no (1/0); success or failure (1/0); telephone/cellular (1/0).

-For Nominal features:

These features are considered as independent features and are directly associated to the decision function for our approach.

The particularity of this approach is how it deal with nominal independent features (job/marital). Each value of these two features V_j for the example i can belong to class C_k if its maximum class frequency is reached on the class in the training set; this frequency is given by the formula:

$$V(C_k)_{ij} \leftarrow \begin{Bmatrix} 1 \; if \; Max \frac{n_k(V_{ij})}{N_k} & k = 0,1 \\ 0 \quad else \end{Bmatrix} (1)$$

-$n_k(V_{ij})$ is the number of V_{ij} variable j in the class C_k

-N_k is the total number of k

So, independent nominal attribute of the testing set takes the value 1 in this class k and 0 in other classes.

For example if we consider the table V:
-The total number of yes $N_{yes} = 2$
-The total number of no $N_{no} = 3$
- The number of "$married$" in the class "no" is: $n_k(married) = 2$
- The number of "$married$" in the class "yes" is: $n_k(married) = 1$

$\frac{n_{no}(married)}{N_{no}} = \frac{2}{3} = 0.66$ *and* $\frac{n_{yes}(married)}{N_{yes}} = \frac{1}{2} = 0.5$

0.66 >0.5; so, the belonging to the class "no" of all marital attributes "married" will be 1 and their belonging to the class "yes" will be 0

3.2.1.2 Preprocessing 2: Missing Attribute Values:

There are several missing values in some categorical attributes, all coded with the "unknown" label. These missing values have been treated by using completion imputation techniques which consist to replace all the missing values v_{ij} of the feature V_j by the average if they are scaled or numeric variable or the mode if they are Boolean or nominal variables inside the class k:

$$(V_{ij})_{k\,miss} \leftarrow \begin{Bmatrix} Mode(V_j)_k \text{ if boolean or nominal} \\ Mean(V_j)_k if \; numeric \; or \; scaled \end{Bmatrix} (2)$$

For example if we consider table II bellow: $V_j = '\; education'$ is scaled variable;

I_4 takes the value Unknown for education and k= 'no', $Mean(V_j)_{no} = \frac{3+2}{2} = 2.5$; The missed value will be replaced by 2.5.

Table 3. TRANSFORMED DATA VALUES

Instance	Age	Marital	Housing	Education	Y
I_1	59	married	0	3	**no**
I_2	39	single	1	1	**yes**
I_3	59	married	0	2	**no**
I_4	41	divorced	0	2.5	**no**
I_5	44	married	0	0	**yes**

3.2.1.3 Normalization

This step which consists to eliminate the impact of the order of magnitude is very important to optimize KNN performance. Considering the two following examples: C_0 which is the center of the first class no, and C_1 which is the center of the second class yes[20], [21]:

Table 4. NON NORMALIZED DISTANCES TO CENTER

	x_1	x_2	x_3	x_4	x_5	y
C_0	39.91	0.163	220.84	984.11	5176.17	no
C_1	40.91	0.152	553.19	792.04	5093.12	yes
E_{21}	27	0.1	698	999	5228.1	

We can easily estimate the similitude of the instance E_{21} to every one of the classes by calculating the Euclidian distances for each center of the two classes which are: $d(E_{21}; C_0) = 480.38$ and $d(E_{21}; C_1) = 286.73$.

E_{21} is then closer to C_1, we can conclude that E_{21} belong to the class yes regardless of the order of magnitude of the features x_j.

In order to overcome such problems, a normalization is adopted for each value v_{ij} of the attribute x_j for the example E_i. It reduces each V_{ij} to the interval [0,1] by the formula:

$$V_{ij} \leftarrow \frac{V_{ij} - min_j(V_{ij})}{max_j(V_{ij}) - min_j(V_{ij})} \quad (3)$$

Table 5. NORMALIZED DISTANCES TO CENTER

	x_1	x_2	x_3	x_4	x_5	y
C_0	0.28	0.61	0.04	0.99	0.80	no
C_1	0.30	0.56	0.11	0.79	0.49	yes
E_{21}	0.12	0.32	0.14	1	1	

With this normalization $d(E_{21}; C_0) = 0.96$ and $d(E_{21}; C_1) = 1.05$. The example E_{21} belongs, after normalization, to the class no as it is closer to C_0 than C_1. Thus, the normalized data values can be presented as follows:

Table 6. NORMALIZED DATA VALUES

Instance	Age	Job	Housing	Education	Y
I_1	0.519	1	0	0.875	**no**
I_2	0.272	0.019	1	0.562	**yes**
I_3	0.519	0.089	0	1	**no**
I_4	0.296	0.215	0	0.688	**no**
I_5	0.333	0.061	0	0.250	**yes**

3.2.2 Classification

The performance of this classification modeling approach depends in particular on the normalization and the choice of machine learning algorithm.

Five most common machine learning models have been compared without and then with normalization, namely: Naïve Bayes (NB), Logistic Regression (LR), Decision Trees (DT), Artificial Neural Network (ANN) and Support Vector Machines (SVM).

4. RESULTS AND ANALYSIS

4.1 Performance Measure

To evaluate the performance of our approach, the first used metric is the accuracy which is common to the majority of authors who previously worked on this database: it expresses the rate of correct predictions; and the second metric is the F1-measure, which is calculated from precision and recall:

$$Accuracy = \frac{a+b}{a+b+c+d} \quad (5)$$

$$precision = \frac{a}{(a+c)} \quad (6)$$

$$recall = \frac{a}{(a+d)} \quad (7)$$

$$FM = \frac{2\,a}{2a+c+d} \quad (8)$$

a : refers to the set of clients that are correctly predicted "yes", b: refers to the set of clients that are correctly predicted "no", c : is the number of false positive and d : is the number of false negatives.

4.2 Results

Five machine learning algorithms have been evaluated in terms of f-measure before and after normalization: DT, LR, NB, ANN and SVM. The following figures summarize the obtained results:

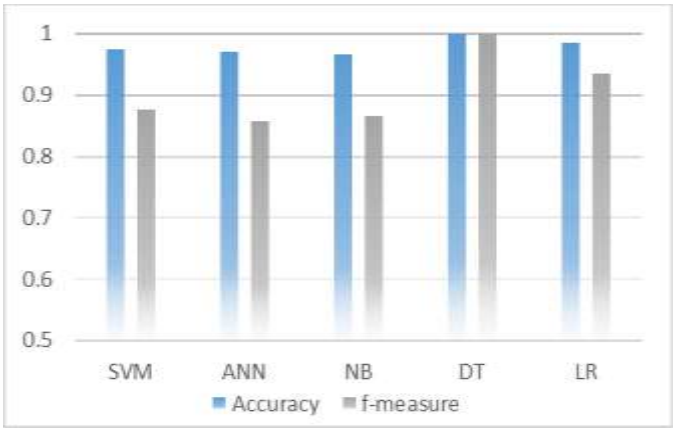

Figure. 1. Results without normalization (accuracy and f1-measure)

Fig.1 shows that without normalization DT [22] and LR performed the highest accuracy with respectively 100% and 98.61% followed by SVM (Linear Kernel, C=1.0, tol=0.001)[23] and ANN with respectively 97.43% [24] and 97.03%; NB with 96.63% gives the lowest score. To optimize this results for some algorithms training and testing datasets have been normalized giving the following results.

After normalization in Fig.2, ANN and DT performed the highest accuracy with respectively 99.07% and 98.98% followed by LR and SVM (Linear Kernel, C=1.0, tol=0.001) with respectively 98.78% and 98.34%; NB with 85.19% gives the lowest accuracy score.

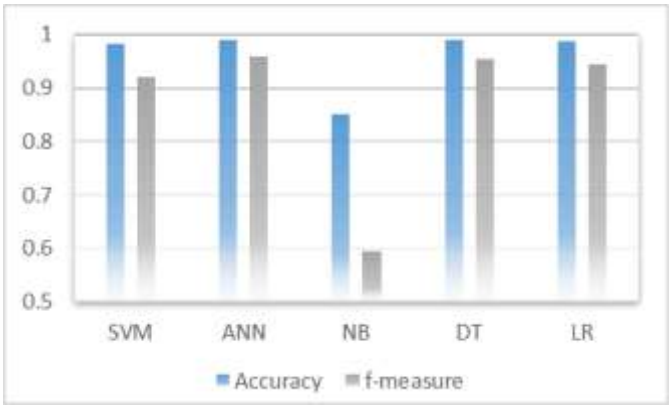

Figure. 2. Results with normalization (accuracy and f1-measure)

In Fig.2 with normalization, ANN and DT performed the highest f-measure with respectively 95.83% and 95.45% followed by LR and SVM (Linear Kernel, C=1.0, tol=0.001) with respectively 94.43% and 91.94%; NB with 59.49% gives the lowest score.

Fig.1 shows that without normalization DT [22] and LR performed the highest f-measure with respectively 100% and 93.53% followed by SVM (Linear Kernel, C=1.0, tol=0.001)[23] and NB with respectively 87.61% [24] and 86.63%; ANN with 85.71% gives the lowest score. To optimize this results for some algoritms training and testing datasets have been normalized giving the following results.

Table 7. COMPARAISON OF BEST PERFORMANCES

	Approach	**Ref. [3]**	**Ref. [2], [4]**	**Ref.[6]**	**Ref. [15]**
Nb features	21	21	17	21	8
Accuracy	100%	93.5	93.23%	92.14%	89%
Best algorithms	DT C5.0	NN	DT C5.0	DT C5.0	Random forests

Table 7 compare the best performance obtained with those five cited authors who worked on the same dataset. With respectively 21, 17, 21, and 8 significant attributes, Ref. [3]; [2], [4]; [6] and [15] achieved their best score of 93.5%; 93.23%; 92.14% and 89% using NN, C5.0, C5.0 and Random Forest respectively. It shows that in many studies, DT C5.0 perform the best performance and this is also verified in this paper where DT C5.0 achieve 100% of correct prediction.

The obtained results show our preprocessing approach has significantly improved the performance of these algorithms compared to the results obtained in previous work on the same database by the previous authors. And even better, the standardization allowed to improve these performances for SVM, LR, ANN even though it was disadvantageous for DT and much more for NB. However, the best performance (100% Accuracy and 100% f-measure) is obtained without normalization for DT.

5. CONCLUSION AND PERSPECTIVES

The applications of data mining techniques are expanding day by day and affect almost every aspect of our daily lives such as industry, medicine, education, finance, etc. This rapid expansion of data mining is a field in with many new research results and new approaches developed.

After overviewing main works that applied machine learning techniques in bank telemarketing, this paper has introduced a modeling approach that preprocessed separately each type of features and normalize them to optimize prediction performance. To evaluate the proposed data modeling approach, this paper shows the results obtained with classification algorithms for bank telemarketing prediction on the Portuguese bank retail dataset by comparing five common machine learning techniques Naïve Bayes (NB), Logistic Regression (LR), Decision Trees (DT), Artificial Neural Network (ANN) and Support Vector Machines (SVM). The results indicate that these performances have been highly improved compared to precedent works (less than 93% best accuracy performances) on this dataset and achieving the better score of 99.07% of accuracy and 95.83% of f-measure by ANN with normalization, and the best score or 100% of f-measure an accuracy with DT without normalization.

Our future works will consist to build own more optimized classification algorithm to more improve these performances in term of f-measure, processing time and resources.

6. REFERENCES

[1] C. T. Su, Y. H. Chen, and D. Y. Sha, "Linking innovative product development with customer knowledge: a data-mining approach," *Technovation*, vol. 26, no. 7, pp. 784–795, 2006.

[2] H. A. Elsalamony and A. M. Elsayad, "Bank Direct Marketing Based on Neural Network," *Int. J. Eng. Adv. Technol.*, vol. 2, no. 6, pp. 392–400, 2013.

[3] V. L. Migue, "Predicting direct marketing response in banking : comparison of class imbalance methods," 2017.

[4] H. A.Elsalamony, "Bank Direct Marketing Analysis of Data Mining Techniques," *Int. J. Comput. Appl.*, vol. 85, no. 7, pp. 12–22, 2014.

[5] R. Vaidehi, "Predictive Modeling to Improve Success Rate of Bank Direct Marketing Campaign," *IJMBS*, vol. 9519, pp. 2230–2232, 2016.

[6] C. Vajiramedhin and A. Suebsing, "Feature selection with data balancing for prediction of bank telemarketing," *Appl. Math. Sci.*, vol. 8, no. 114, pp. 5667–5672, 2014.

[7] S. Moro, P. Cortez, and P. Rita, "A data-driven approach to predict the success of bank telemarketing," *Decis. Support Syst.*, vol. 62, pp. 22–31, 2014.

[8] E. W. T. Ngai, L. Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification," *Expert Syst. Appl.*, vol. 36, no. 2 PART 2, pp. 2592–2602, 2009.

[9] B. B. Bezabeh, "The Application of Data Mining Techniques To Support Customer Relationship Management: the Case of Ethiopian Revenue and Customs Authority."

[10] A. K. Muhammed, "Application of Data Mining In Marketing Application of Data Mining In Marketing," *IJCSN Int. J. Comput. Sci. Netw.*, vol. 2, no. 5, pp. 2277–5420, 2013.

[11] S. F. Crone, S. Lessmann, and R. Stahlbock, "The impact

of preprocessing on data mining: An evaluation of classifier sensitivity in direct marketing," *Eur. J. Oper. Res.*, vol. 173, no. 3, pp. 781–800, 2006.

[12] S. Moro, P. Cortez, and P. Rita, "Using customer lifetime value and neural networks to improve the prediction of bank deposit subscription in telemarketing campaigns," *Neural Comput. Appl.*, vol. 26, no. 1, pp. 131–139, 2014.

[13] S. Moro and R. M. S. Laureano, "Using Data Mining for Bank Direct Marketing: An application of the CRISP-DM methodology," *Eur. Simul. Model. Conf.*, no. Figure 1, pp. 117–121, 2011.

[14] Z. Chu, "Bank Marketing with Machine Learning," 2015.

[15] B. B. D. GRZONKA, G. SUCHACKA, "Application of selected supervised classification methods to bank marketing campaign d," *Inf. Syst. Manag.*, vol. 5, pp. 36–48, 2016.

[16] R. Farooqi, "Effectiveness of Data mining in Banking Industry: An empirical study," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 5, pp. 827–830, 2017.

[17] T. Parlar and S. K. Acaravci, "Using Data Mining Techniques for Detecting the Important Features of the Bank Direct Marketing Data," *Int. J. Econ. Financ. Issues*, vol. 7, no. 2, pp. 692–696, 2017.

[18] B. T. Femina and E. M. Sudheep, "An efficient CRM-data mining framework for the prediction of customer behaviour," *Procedia Comput. Sci.*, vol. 46, no. Icict 2014, pp. 725–731, 2015.

[19] Y. Kawasaki and M. Ueki, "Sparse Predictive Modeling for Bank Telemarketing Success Using Smooth-Threshold Estimating Equations," *J. Jpn. Soc. Comp. Stat.*, vol. 28, pp. 53–66, 2015.

[20] W. Cherif, "Optimization of K-NN algorithm by clustering and reliability coefficients: Application to breast-cancer diagnosis," *Procedia Comput. Sci.*, vol. 127, pp. 293–299, 2018.

[21] W. Cherif, "Hybrid Reliability-Similarity-Based Approach for Supervised Machine Learning," *Procedia Comput. Sci.*, vol. 12, no. 3, pp. 170–175, 2018.

[22] E. Venkatesan and T. Velmurugan, "Performance Analysis of Decision Tree Algorithms for Breast Cancer Classification," vol. 8, no. November, 2015.

[23] X. Zhang, D. Qiu, and F. Chen, "Neurocomputing Support vector machine with parameter optimization by a novel hybrid method and its application to fault diagnosis," *Neurocomputing*, vol. 149, pp. 641–651, 2015.

[24] A. Goyal and R. Mehta, "Performance Comparison of Naïve Bayes and J48 Classification Algorithms," vol. 7, no. 11, 2012.