

## Predicting reaction based on customer's transaction using machine learning approaches

Israa M. Hayder<sup>1</sup>, Ghazwan Abdul Nabi Al Ali<sup>2,3</sup>, Hussain A. Younis<sup>2,4</sup>

<sup>1</sup>Department of Computer Systems Techniques, Qurna Technique Institute, Qurna, Iraq

<sup>2</sup>School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia

<sup>3</sup>Department of Computer Science (Educational Science), University of Basrah, Basrah, Iraq

<sup>4</sup>College of Education for Women, University of Basrah, Basrah, Iraq

### Article Info

#### Article history:

Received Jan 20, 2022

Revised Aug 9, 2022

Accepted Sep 5, 2022

#### Keywords:

Bank marketing

Decision tree

K-nearest neighbors' algorithm

Naive Bayes

Support vector machines

### ABSTRACT

Banking advertisements are important because they help target specific customers on subscribing to their packages or other deals by giving their current customers more fixed-term deposit offers. This is done through promotional advertisements on the Internet or media pages, and this task is the responsibility of the shopping department. In order to build a relationship with them, offer them the best deals, and be appropriate for the client with the company's assurance to recover these deposits, many banks or telecommunications firms store the data of their customers. The Portuguese bank increases its sales by establishing a relationship with its customers. This study proposes creating a prediction model using machine learning algorithms, to see how the customer reacts to subscribe to those fixed-term deposits or offers made with the aid of their past record. This classification is binary, i.e., the prediction of whether or not a customer will embrace these offers. Four classifiers that include k-nearest neighbor (k-NN) algorithm, decision tree, naive Bayes, and support vector machines (SVM) were used, and the best result was obtained from the classifier decision tree with an accuracy of 91% and the other classifier SVM with an accuracy of 89%.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



### Corresponding Author:

Hussain A. Younis

School of Computer Sciences, Universiti Sains Malaysia

11800 USM, Penang, Malaysia

Email: hussain.younis@uobasrah.edu.iq

## 1. INTRODUCTION

Banking advertising comprises advertisements by financial institutions. This category includes, in addition to advertising directed at bank clients, business reports and information pamphlets; statements about the payment of new shares, reports on investment program outcomes, as well as several additional financial announcements may also be included [1]–[3]. Many banks or telecommunication companies store their customers' data to establish a relationship with customers and provide them with the best offers and at the same time be appropriate for the customer with the guarantee that the company will recover their deposits. The Portuguese bank increases its sales by establishing a cordial relationship with its customers. Transaction predictions use k-nearest neighbors' (k-NN) algorithm [4], [5], decision tree [6]–[10], naive Bayes [6], [11]–[13], and support vector machines (SVM) [14] to bank marketing. This study proposes creating a prediction model using machine learning algorithms to see how the customer reacts to subscribing to those fixed-term deposits or offers made through their past data [15]–[19]. This classification is binary, i.e., the prediction of whether or not a customer will participate in these offers. Four classifiers k-NN, decision tree, SVM, and naive Bayes were used.

## 2. LITERATURE REVIEW

As the number of Internet users and businesses grows, many clusters of e-commerce applications appear not to be physically linked to each other in the system but are inter-related in business. Online banking has been in practice since the 1980s, when it was first introduced by four major banks in New York [20], [21]. The study of commercial financial transactions made a short-term expectation using a logistic regressive model and a SVM model. The comparison between them concluded that the SVM model prediction was better than 100% of the logistic regressive model and 97.67%, respectively [22]. The study was based on building an intelligent banking system on Hercules. The results of the study reached that the Hercules architecture's intelligent online banking has greatly improved the intelligence and security of the traditional online banking system. Finally, we summarize and analyze the intelligent online banking system's value and innovation, looking forward to discovering the system's flaws [23]. Research was also conducted based on neural networks for predicting the transactions by automated teller machine (ATM) [24]. Additionally, research was done on the security flaws in online and mobile banking systems [25], banking fraud detection [26], banking apps and online payment systems [27], a strong and secure authentication method [28], and Internet banking user behavior. It is advised that initiatives meant to boost trust in the financial sector receive more attention. Winning customer trust through activities such as the secure processing and transmission of highly secret data could be a helpful step toward retaining electronic customers [29]. The dataset is downloaded from the UCI machine learning repository and is correlated to the Portuguese banking institution's direct marketing campaign. These campaigns were focused on calls over the phone. More than one call to the same customer was also made to reach whether (yes) or not (yes/no) their "term deposit" product was subscribed to. There were 4 datasets with all examples (41,188) and 20 inputs ordered by date from which bank-additional-full.csv is used (from May 2008 to November 2010). There are 20 variables for input and 1 variable for output (desired target). The dataset contained a variety of customer details, including age, job, marital status, education, default, housing, loan, contact, month, day(s)-of-week, duration, campaign, prays, previous, poutcome, em.var.rate, cons.price.idx, cons.conf.idx, euribor3m, nr. employed and one output variable  $y$  denotes whether or not the consumer subscribed to the term deposit. Such datasets were loaded in the Python language for preprocessing to check for any missing values and it was found there were unknown values. However, the dataset is imbalanced [30], [31]. Deep learning, data mining, a robot, a decision tree, and k-NN were among the other studies presented [32]–[38].

## 3. METHOD

### 3.1. Pre-processing

In this step, the data needs to be processed and it is done in four steps: the first step is the drop of unknown data. The second step is to convert features from categorical to numeric; the third step is to balance the data, and the last step is to choose the best feature. In the beginning, we had to clean the data from the null values. We excluded features with 330 unknown attributes, and some categorical data have been converted to numeric.

#### 3.1.1. Dataset unbalanced

When we are using some classification algorithms such as k-NN or naive Bayes, and so on, those algorithms can perform poorly, or they may lead to some overfitting problems. Because we have so many imbalanced data points in our dataset, such that the ratio of one class is 70% and the other class is 30%, the accuracy method will be ineffective. When applying the dataset to the classifiers, we will not get high accuracy. However, before applying classifiers, we must first process the dataset [39]. The difference between the balanced dataset and the unbalanced dataset was significant as it showed the bad performance in the unbalanced dataset and the low accuracy with the delay in finding the best value for  $K$ , which reached 37 with an accuracy of 0.92. Unlike the weighted data, it reached a value of  $K=3$ , with a high accuracy of 94, as shown in Figure 1. To solve this problem, we will use two techniques: random under-sampling and the synthetic minority oversampling technique, as shown in Figure 2.

### 3.2. Classifiers and test evaluation

#### 3.2.1. Random under-sampling method

The method is processing of the minority to make it equal to the majority by sampling fakes from some minority class randomly and then repeating this process until the two classes are equal [40]. However, it may negatively affect the model's performance by copying the noise rows. Moreover, they are known as "naive resampling" approaches because they make no assumptions about the data and utilize no heuristics. This makes them easy to design and quick to execute, which is ideal for highly large and complex datasets.

The optimal number of neighbors is 3  
 The optimal score is 0.94  
 The optimal number of neighbors is 37  
 The optimal score is 0.92

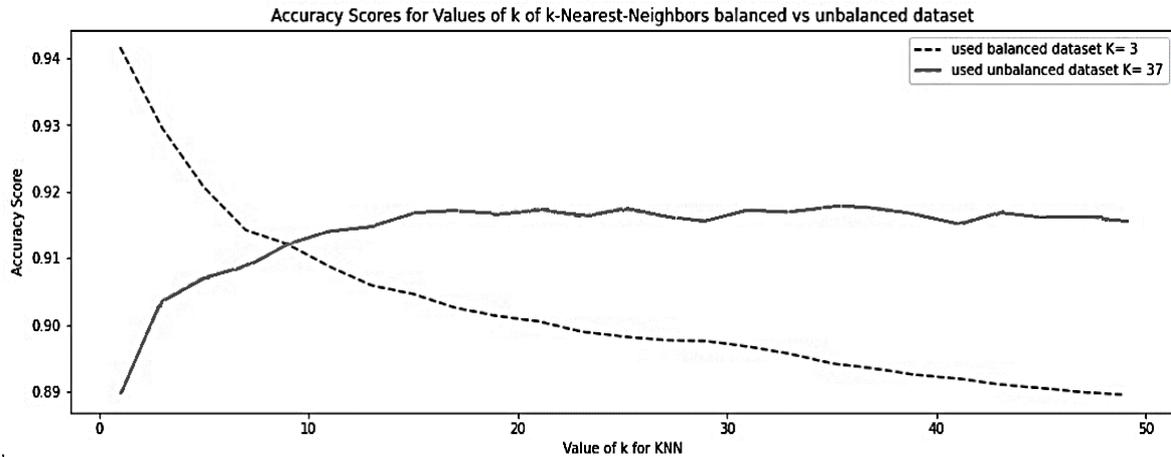


Figure 1. Difference between the balanced and the unbalanced dataset

**3.2.2. Synthetic minority oversampling technique**

Synthetic minority oversampling technique (SMOTE) is attempting to select a point from the minority and then produce a new point using the k-NN method [41]. This is a sound strategy since, in terms of feature space, new samples are quite similar to existing instances from the minority class. But necessitates numerous calculations, as depicted in Figure 3.

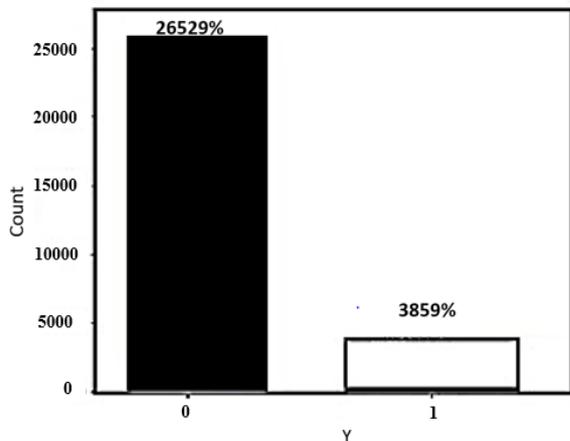


Figure 2. Our dataset unbalance

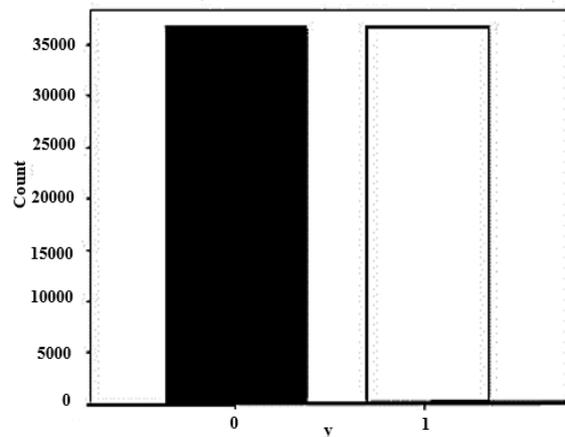


Figure 3. Our dataset balanced

**3.2.3. Experiment to choose of approach dataset unbalanced**

After applying the two techniques, it was found that the best choice was SMOTE because the results were better in generalization and reduced overfitting in the k-NN model. Moreover, the number of samples was sufficient for the purpose of applying this technique, as shown in Figure 4. It has been concluded that the value of  $k=3$  was deduced early in the first cycle, whose high accuracy is 0.93 and that the other method achieves the value of  $k=11$  in the last cycle, whose weak accuracy is 0.87, such that k starts from 3 to 100.

**3.2.4. Features selection from original dataset**

This involves choosing some features to improve performance and reduce the model's prediction time. When employing statistically based feature selection techniques, each input variable's relationship to the target variable is assessed, and the input variables with the most robust relationships are chosen. Although

the choice of statistical measures depends on the data type of both the input and output variables, these techniques can be quick and efficient. Therefore, we will review some techniques that work on choosing the best.

The optimal number of neighbors is 3  
 The optimal score is 0.93  
 The optimal number of neighbors is 11  
 The optimal score is 0.87

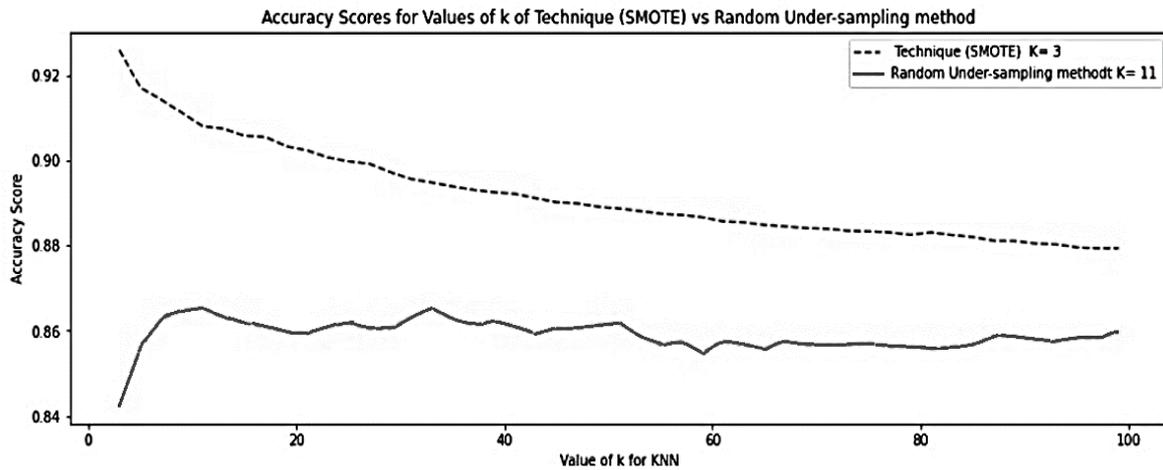


Figure 4. Accuracy values for k-NN with random under-sampling vs SMOTE

Filter methods are based on calculating the score for each feature with a target and then choosing the best score between them. It is not a learning process, but rather the search for related features with a label prior to process learning [42]. Algorithm advantages are not computationally expensive, and they avoid features with least effects on the target [43]–[45].

The wrapper methods create a subset of the dataset by training (machine learning module) and then repeat this training process by adding or removing some features until finding the best combination to achieve the aim (using a greedy algorithm to find the best combination). The common techniques in this approach are forward selection and backward elimination [42]. Quality is an algorithm advantage, but it is computationally costly [46]–[48].

**3.2.5. Experiment features selection**

Two experiments were going to be conducted. The first experiment employed filter selection to obtain the best 14 features, yielding the results shown in Table 1. The second experiment used wrapper methods to get the best 14 features, and we got the results as shown in Table 2. As shown in Figure 5, after applying the k-NN model to both methods, we obtained the highest accuracy in wrapper methods.

Table 1. Filter selection

No	Features
0	age
1	default
2	housing
3	loan
4	month duration
5	duration
6	campaign
7	previous
8	poutcome
9	Emp.vqr.ratec
10	cons.price.idx
11	cons.price.idx
12	Eunbor3m
13	Nr. employed

Table 2. Wrapper methods

No	Features
0	age
1	job
2	education
3	loan
4	contact
5	month
6	Day-of-week
7	duration
8	campaign
9	Emp.vqr.ratecons.conf.idx
10	Emp.var, rate
11	cons.price.idx
12	Cons. price.idx
13	Euribor3m

	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.99	0.86	0.92	7978	0	0.96	0.86	0.91	7978
1	0.88	0.99	0.93	8000	1	0.87	0.96	0.92	8000
accuracy			0.93	15978	accuracy			0.91	15978
macro avg	0.93	0.93	0.93	15978	macro avg	0.92	0.91	0.91	15978
weighted avg	0.93	0.93	0.93	15978	weighted avg	0.92	0.91	0.91	15978

Text(0.5, 51.0, 'Predicted label')

0.9108774565026913

Text(0.5, 51.0, 'Predicted label')

0.9257729377894605

0.9108774565026913

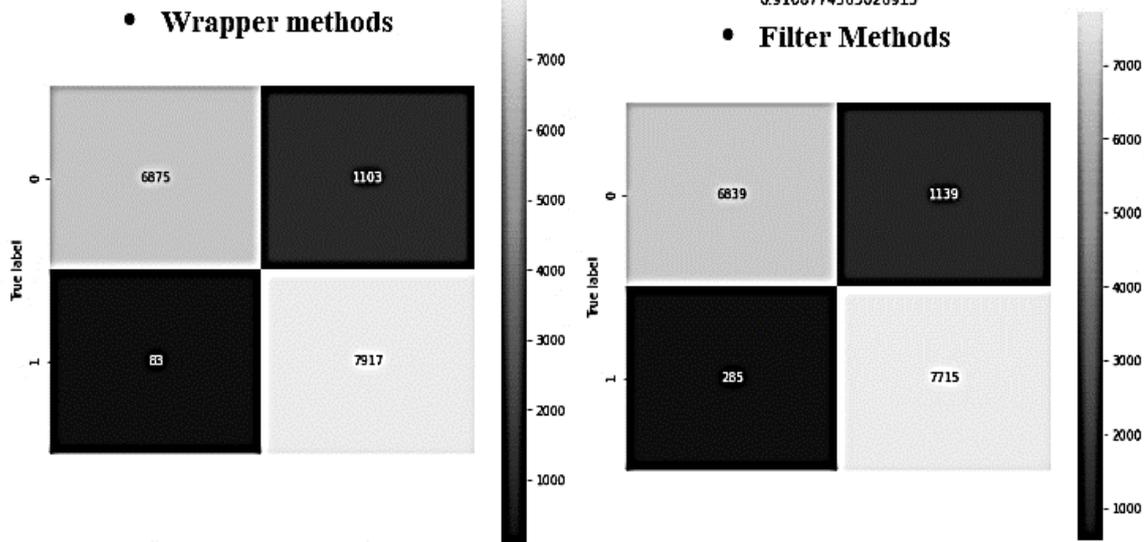


Figure 5. Metrics of accuracy for two methods

### 3.3. Classifiers and test evaluation

#### 3.3.1. Classifiers

After applying data processing in the previous steps, in this step, we implement machine learning algorithms, which are divided into two categories: parametric and non-parametric classifiers. This process aims to find the best models for each of the two sections. Moreover, the best model for each type will be selected.

#### 3.3.2. Non-parametric classifier

It is also called “lazy teaching” as it does not use assumptions in learning. Simply, the samples collected in training data are used [49]. The algorithms that under this technique are k-NN and decision trees. After processing it in the previous steps, we will apply our dataset to the k-NN classifiers and decision trees, compare the results, and choose the best algorithm between them.

### 3.3.3. Parametric classifier

Linear machine learning algorithms are what metric algorithms are called, and linear regression is frequently used in them [50]. This technique employs algorithms such as naive Bayes, SVM, and others. Following the processing in the preceding steps, we will apply our dataset to the naive Bayes classifiers and SVM, compare the results, and select the best algorithm among them.

### 3.3.4. Experiment test evaluation

After balancing the data, we have 53,258 rows, and we will experiment with separating the data 1 to 20% for the testing dataset and 80% for the training dataset. 2 to 30% for the testing dataset and 70% for the training dataset. According to exp experience using the decision tree in Figure 6, it has been found that 20% of testing is better than 30%.

```
{'criterion': 'entropy', 'max_depth': 15, 'min_samples_split': 2}
0.9147263948497854
0.9171360620853674
DecisionTreeClassifier(criterion='entropy', max_depth=15)
{'criterion': 'entropy', 'max_depth': 16, 'min_samples_split': 2}
0.9180867258909273
0.9169170108899737
DecisionTreeClassifier(max_depth=16)
```

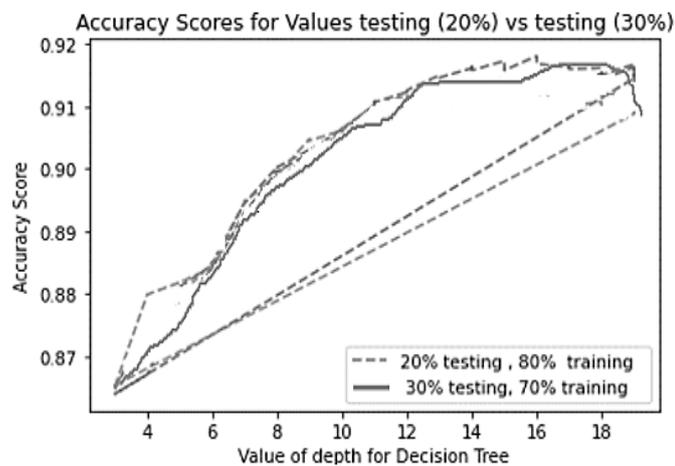


Figure 6. Accuracy scores for values testing (20%) vs testing (30%)

## 4. EXPERIMENTS AND ANALYSIS

The experiment was performed with Python, which contains most of the machine learning algorithms. Before modeling, we need a preliminary exploration of the data set (41,188 rows and 21 features). We have previously explained that the data is not balanced to a large extent, and this problem was addressed using the SMOTE method, which was superior to the other methods. In this step, we will build the models both (k-NN and decision trees) for the non-parametric method, and we will try (Pace, SVM) for the parametric and choose the best of them in the models. The k-fold (5 folds) cross-validation that gave better results was employed in the course of conducting the study.

### 4.1. KMM

We obtained the best accuracy after changing the model settings k-NN, where the number of neighbors was chosen 3, and at the same time, the weight was determined to distance. The k-NN algorithm was chosen is brute with k-fold (5 folds) cross-validation, see Figure 7. Moreover, we got an accuracy of 93, but also the model does not work well in class 0, and values of k equal to 3 are undesirable because it tends to bring about bias.

### 4.2. Decision tree model

The decision tree algorithm was implemented and what was obtained was the highest accuracy after changing the model's settings, where the best criterion was chosen, which was Gini and *min\_samples\_split* was 2. The length of the tree was 14 with k-fold (5 folds) cross-validation. Moreover, we got an accuracy of 92, but also the model does not work well in class 0, as shown in Figure 8.

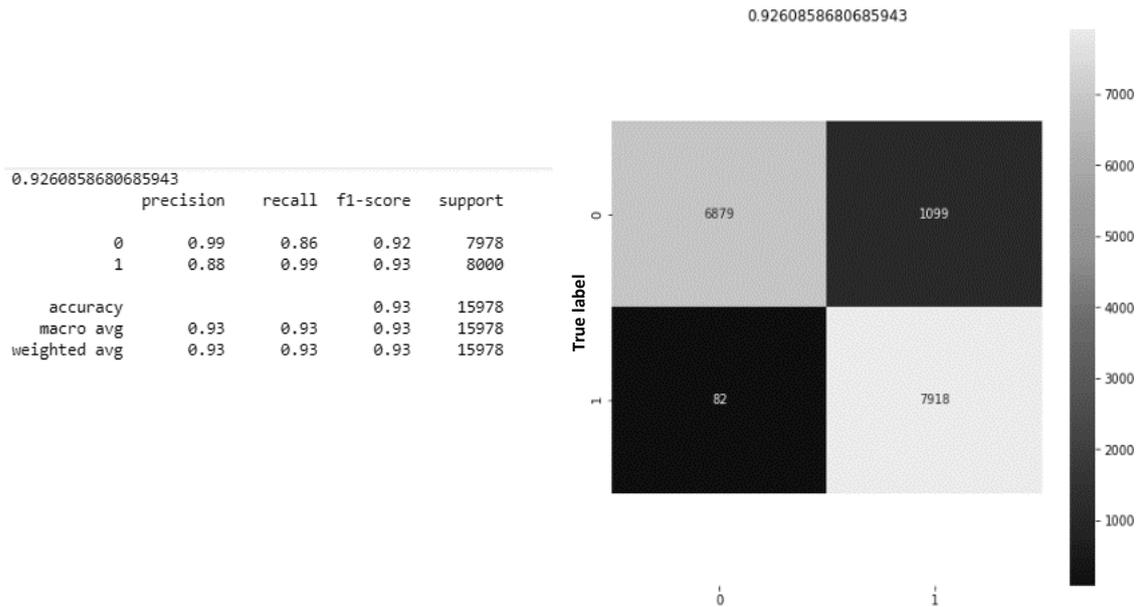


Figure 7. k-NN model and classification report

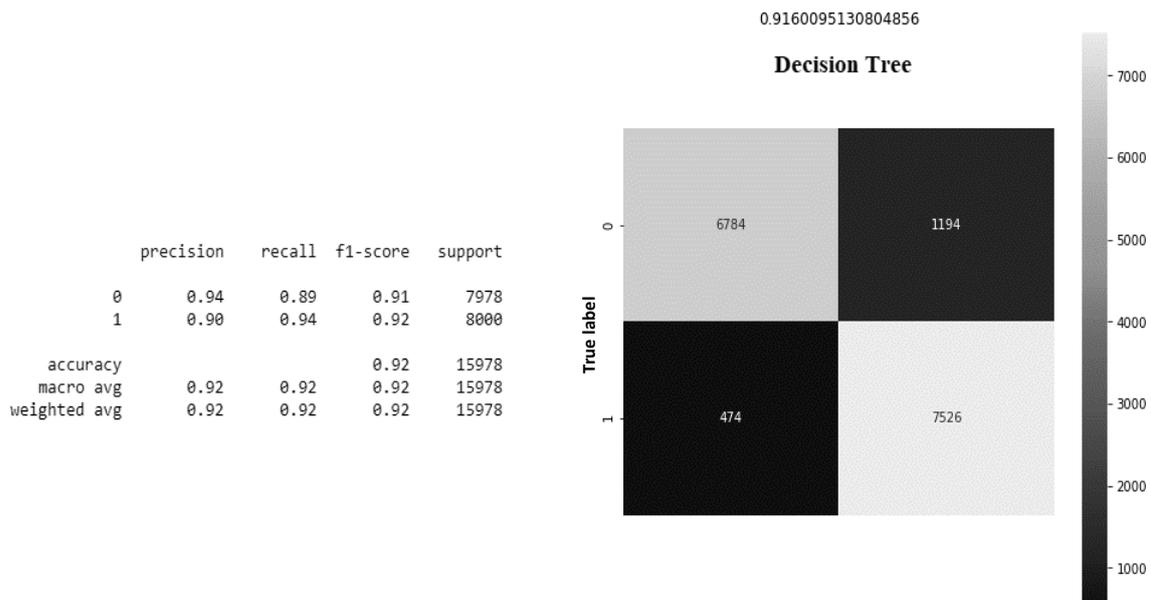


Figure 8. Decision tree model and classification report

### 4.3. Naive Bayes model

Naive Bayes model algorithm is a classification method built on the Bayes theorem and predicated on the idea of predictor independence. It is simply said that the presence of one feature does not depend on the existence of any other feature in the class. The naive Bayes model algorithm was implemented with default values, and we got an accuracy of 81%, as shown in the Figure 9.

### 4.4. SVM model

SVM is a collection of supervised learning techniques used for outliers' identification, regression, and classification. The benefits of SVMs include still useful in situations where the number of dimensions exceeds the number of samples and efficient in high-dimensional environments. SVMs have been implemented by using the class of SVM model from Sklearn, and we got 89% accuracy by using the settings {'C': 1, 'gamma': 0.1, 'kernel': 'rbf'}, as shown in Figure 10.

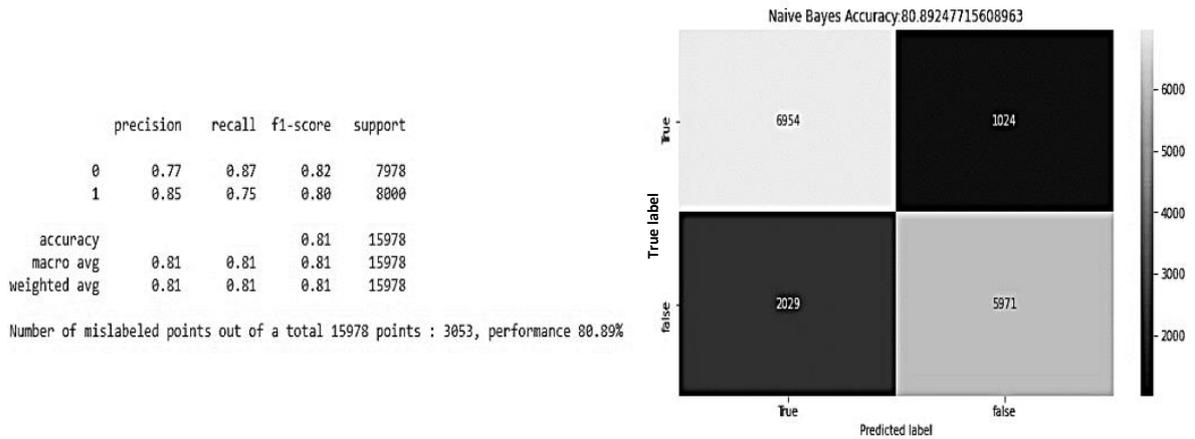


Figure 9. Naive Bayes model and classification report

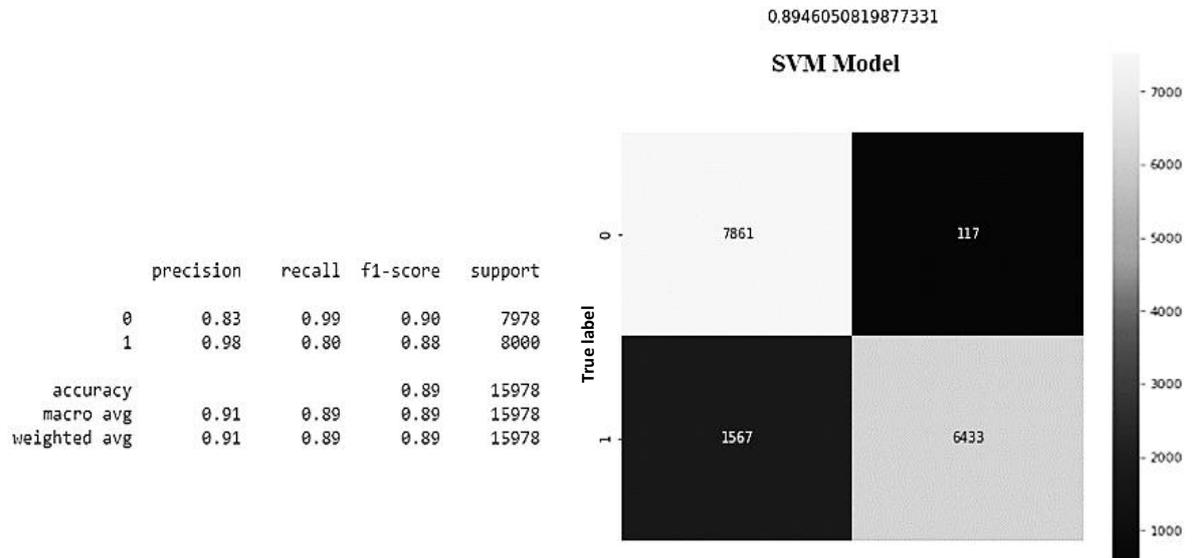


Figure 10. SVM model and classification report

5. RESULTS

That the model is nonparametric, k-NN has 14 features and K=3 value of parameter will be testing accuracy of 92% and score before testing 91%. The model SVM has 14 features and Kernel='rbf' value of parameter will be tested with 89% accuracy before testing 89%. With knowledge, the type of model is nonparametric. The model decision tree that has 14 features and a depth of 14 values of parameters will be tested with an accuracy of 92% and a score of 91% before testing. Finally, naive Bayes with 14 features and defaults of value parameter will have a testing accuracy of 80% and a score before testing of 80%. It is a parametric type of model. Table 3 shows the results of each model and the number of features that were used.

Table 3. Results of experiments models

Variable	Features	Value of parameter	Testing Accuracy	Score before Testing	Type of model
k-NN	14	K=3	92%	91%	nonparametric
SVM	14	Kernel='rbf'	89%	89%	nonparametric
Decision tree	14	Depth=14	92%	91%	nonparametric
Naive Bayes	14	Default	80%	80%	parametric

## 6. CONCLUSION

The algorithm and the model are two terms for processor and the simplification of problems. The model k-NN has 14 features and a K-3 value of parameter will be tested with 92% accuracy. The score before testing is 91%. Knowing that the type of model is nonparametric, the model SVM has 14 features and Kernel="rbf" value of parameter will be tested with 89% accuracy before testing and 89% after testing. The model decision tree that has 14 features and depth=14 value of 14 values of parameters was tested with an accuracy of 92% and a score of 91% before testing. Finally, the naive Bayes with 14 features and value parameter defaults will have a testing accuracy of 80% and a score before testing of 80%. It is a parametric type of model. The results of each model and the number of features that were used. The current assignment uses 14 features rather than the original 20 because they were identified by feature selection techniques that improved the model's performance in terms of time speed. According to the task assigned to us, we had to choose only two algorithms. One is parametric and the other is referred to as non-parametric. The naive Bayes algorithm was selected as a parametric algorithm and got an accuracy of 80%. Moreover, k-NN, SVM's, and decision tree accuracy were (92%, 89%, and 91%) respectively, and all were considered non-parametric algorithms. However, the decision tree was selected as an alternative to the k-NN since the value of  $k=3$  in the k-NN is low and could be sensitive to noise and outliers. For future work, deep learning and clustering can also be used to determine and know the loans. This work can be made in the form of an application that can be used on smart mobile phones in general.

## REFERENCES

- [1] C. I. Mbama and P. O. Ezepue, "Digital banking, customer experience and bank financial performance," *International Journal of Bank Marketing*, vol. 36, no. 2, pp. 230–255, Apr. 2018, doi: 10.1108/IJBM-11-2016-0181.
- [2] J. Mylonakis, "The influence of banking advertising on bank customers: an examination of Greek bank customers' choices," *Banks and Bank Systems*, vol. 3, no. 4, pp. 44–49, 2008.
- [3] H. A. Laskey, B. Seaton, and J. A. F. Nicholls, "Strategy and structure in bank advertising: an empirical test," *International Journal of Bank Marketing*, vol. 10, no. 3, pp. 3–8, Mar. 1992, doi: 10.1108/02652329210015282.
- [4] L. Xiong and Y. Yao, "Study on an adaptive thermal comfort model with K-nearest-neighbors (KNN) algorithm," *Building and Environment*, vol. 202, Sep. 2021, doi: 10.1016/j.buildenv.2021.108026.
- [5] J.-H. Chen, "KNN based knowledge-sharing model for severe change order disputes in construction," *Automation in Construction*, vol. 17, no. 6, pp. 773–779, Aug. 2008, doi: 10.1016/j.autcon.2008.02.005.
- [6] B. Zhang, "Tactical decision system of table tennis match based on C4.5 decision tree," in *2021 13<sup>th</sup> International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Jan. 2021, pp. 632–635, doi: 10.1109/ICMTMA52658.2021.00146.
- [7] N. Rochmawati *et al.*, "Covid symptom severity using decision tree," in *2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE)*, Oct. 2020, pp. 1–5, doi: 10.1109/ICVEE50212.2020.9243246.
- [8] A. A. Supianto, A. Julisar Dwitama, and M. Hafis, "Decision tree usage for student graduation classification: A comparative case study in faculty of computer science Brawijaya University," in *2018 International Conference on Sustainable Information Engineering and Technology (SIET)*, Nov. 2018, pp. 308–311, doi: 10.1109/SIET.2018.8693158.
- [9] E. Papageorgiou, C. Stylios, and P. Groumpos, "A combined fuzzy cognitive map and decision trees model for medical decision making," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2006, pp. 6117–6120, doi: 10.1109/IEMBS.2006.260354.
- [10] Juan Sun and Xi-Zhao Wang, "An initial comparison on noise resisting between crisp and fuzzy decision trees," in *2005 International Conference on Machine Learning and Cybernetics*, 2005, vol. 4, pp. 2545–2550, doi: 10.1109/ICMLC.2005.1527372.
- [11] M. Dahmani and M. Guerti, "Vocal folds pathologies classification using Naïve Bayes networks," in *2017 6<sup>th</sup> International Conference on Systems and Control (ICSC)*, May 2017, pp. 426–432, doi: 10.1109/ICoSC.2017.7958686.
- [12] Y. Benmahamed, Y. Kemari, M. Tegar, and A. Boubakeur, "Diagnosis of power transformer oil using KNN and Naive Bayes classifiers," in *2018 IEEE 2<sup>nd</sup> International Conference on Dielectrics (ICD)*, Jul. 2018, pp. 1–4, doi: 10.1109/ICD.2018.8468532.
- [13] D. Larose, "Naïve Bayes estimation and bayesian networks," in *Data Mining Methods and Models*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006, pp. 204–239.
- [14] K. R. Singh, K. P. Neethu, K. Madhurekaa, A. Harita, and P. Mohan, "Parallel SVM model for forest fire prediction," *Soft Computing Letters*, vol. 3, Dec. 2021, doi: 10.1016/j.socl.2021.100014.
- [15] L.-L. Li, Z.-F. Liu, M.-L. Tseng, K. Jantarakolica, and M. K. Lim, "Using enhanced crow search algorithm optimization-extreme learning machine model to forecast short-term wind power," *Expert Systems with Applications*, vol. 184, Dec. 2021, doi: 10.1016/j.eswa.2021.115579.
- [16] B. Czejdo, S. Bhattacharya, and C. Spooner, "Improvement of protein model scoring using grouping and interpreter for machine learning," in *2019 IEEE 9<sup>th</sup> Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2019, pp. 0349–0353, doi: 10.1109/CCWC.2019.8666524.
- [17] S. Z. Shogrkhodaei, S. V. Razavi-Termeh, and A. Fathnia, "Spatio-temporal modeling of PM2.5 risk mapping using three machine learning algorithms," *Environmental Pollution*, vol. 289, Nov. 2021, doi: 10.1016/j.envpol.2021.117859.
- [18] M. Amini *et al.*, "Overall survival prognostic modelling of non-small cell lung cancer patients Using positron emission tomography/computed tomography Harmonised radiomics features: The quest for the optimal machine learning algorithm," *Clinical Oncology*, vol. 34, no. 2, pp. 114–127, Feb. 2022, doi: 10.1016/j.clon.2021.11.014.
- [19] N. Roy, R. Ahmed, M. R. Huq, and M. M. Shahriar, "User-centric activity recognition and prediction model using machine learning algorithms," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 12, no. 12, pp. 501–510, 2021.
- [20] A. A. B. Ng and N. L. Abdullah, "Security challenges in designing an integrated web application for multiple online banking," in *2010 International Symposium on Information Technology*, Jun. 2010, pp. 1–5, doi: 10.1109/ITSIM.2010.5561291.

- [21] M. J. Cronin, *Banking and finance on the internet*. John Wiley & Sons, 1997.
- [22] L. Sijia, T. Lan, Z. Yu, and Y. Xiuliang, "Comparison of the prediction effect between the logistic regressive model and SVM model," in *2010 2nd IEEE International Conference on Information and Financial Engineering*, Sep. 2010, pp. 316–318, doi: 10.1109/ICIFE.2010.5609308.
- [23] G. Luo, W. Li, and Y. Peng, "Overview of intelligent online banking system based on HERCULES architecture," *IEEE Access*, vol. 8, pp. 107685–107699, 2020, doi: 10.1109/ACCESS.2020.2997079.
- [24] H. Razavi, H. Sarabadani, A. Karimisefat, and J.-F. LEBRATY, "Profitability prediction for ATM transactions using artificial neural networks: A data-driven analysis," in *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)*, Feb. 2019, pp. 661–665, doi: 10.1109/KBEI.2019.8735037.
- [25] N. Yildirim and A. Varol, "A research on security vulnerabilities in online and mobile banking systems," in *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*, Jun. 2019, pp. 1–5, doi: 10.1109/ISDFS.2019.8757495.
- [26] J. Cui, C. Yan, and C. Wang, "ReMEMBeR: Ranking metric embedding-based multicontextual behavior profiling for online banking fraud detection," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 3, pp. 643–654, Jun. 2021, doi: 10.1109/TCSS.2021.3052950.
- [27] S. Duarte, M. Costa, M. Brito, A. Miranda, and M. Au-Yong-Oliveira, "Bank and online payment apps: How beneficial to society in the XXI century?," in *15th Iberian Conference on Information Systems and Technologies (CISTI)*, Jun. 2020, pp. 1–6, doi: 10.23919/CISTI49556.2020.9140938.
- [28] H. Venugopal and N. Viswanath, "A robust and secure authentication mechanism in online banking," in *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, Nov. 2016, pp. 1–3, doi: 10.1109/GET.2016.7916850.
- [29] S. Alimolaei, "An intelligent system for user behavior detection in internet banking," in *2015 4th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, Sep. 2015, pp. 1–5, doi: 10.1109/CFIS.2015.7391642.
- [30] S. Abbas, "Deposit subscribe prediction using data mining techniques based real marketing dataset," *International Journal of Computer Applications*, vol. 110, no. 3, pp. 1–7, Jan. 2015, doi: 10.5120/19293-0725.
- [31] S. Moro, P. Cortez, and P. Rita, "A data-driven approach to predict the success of bank telemarketing," *Decision Support Systems*, vol. 62, pp. 22–31, Jun. 2014, doi: 10.1016/j.dss.2014.03.001.
- [32] H. A. Younis, A. S. A. Mohamed, R. Jamaludin, and M. N. A. Wahab, "Survey of robotics in education, taxonomy, applications, and platforms during COVID-9," *Computers, Materials and Continua*, vol. 67, no. 1, pp. 687–707, 2021, doi: 10.32604/cmc.2021.013746.
- [33] R. M. Duwairi and Z. A. Halloush, "Automatic recognition of Arabic alphabets sign language using deep learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, pp. 2996–3004, Jun. 2022, doi: 10.11591/ijece.v12i3.pp2996-3004.
- [34] E. H. Muktafin, P. Pramono, and K. Kusriani, "Sentiments analysis of customer satisfaction in public services using K-nearest neighbors algorithm and natural language processing approach," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 1, pp. 146–154, Feb. 2021, doi: 10.12928/telkomnika.v19i1.17417.
- [35] H. A. Younis, I. M. Hayder, I. S. Seger, and H. A.-K. Younis, "Design and implementation of a system that preserves the confidentiality of stream cipher in non-linear flow coding," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 23, no. 7, pp. 1409–1419, Oct. 2020, doi: 10.1080/09720529.2020.1714890.
- [36] T. Hlaing, "Feature selection and fuzzy decision tree for network intrusion detection," *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 1, no. 2, pp. 109–118, Sep. 2012, doi: 10.11591/ij-ict.v1i2.591.
- [37] H. Waguih, "A data mining approach for the detection of denial of service attack," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 2, no. 2, Jun. 2013, doi: 10.11591/ij-ai.v2i2.1937.
- [38] B. M. Susanto, "Naive Bayes decision tree hybrid approach for intrusion detection system," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 2, no. 3, Sep. 2013, doi: 10.12928/eei.v2i3.208.
- [39] S. Redkar, S. Mondal, A. Joseph, and K. S. Hareesha, "A machine learning approach for drug-target interaction prediction using wrapper feature Selection and class balancing," *Molecular Informatics*, vol. 39, no. 5, May 2020, doi: 10.1002/minf.201900062.
- [40] M. Mimura, "Using fake text vectors to improve the sensitivity of minority class for macro malware detection," *Journal of Information Security and Applications*, vol. 54, Oct. 2020, doi: 10.1016/j.jisa.2020.102600.
- [41] R. Geetha, S. Sivasubramanian, M. Kaliappan, S. Vimal, and S. Annamalai, "Cervical cancer identification with synthetic minority oversampling technique and PCA analysis using random forest classifier," *Journal of Medical Systems*, vol. 43, no. 9, Art. no. 286, Sep. 2019, doi: 10.1007/s10916-019-1402-6.
- [42] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang, "Benchmark for filter methods for feature selection in high-dimensional classification data," *Computational Statistics and Data Analysis*, vol. 143, Mar. 2020, doi: 10.1016/j.csda.2019.106839.
- [43] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 16–28, Jan. 2014, doi: 10.1016/j.compeleceng.2013.11.024.
- [44] Y. Jiang, X. Liu, G. Yan, and J. Xiao, "Modified binary cuckoo search for feature selection: A hybrid filter-wrapper approach," in *2017 13th International Conference on Computational Intelligence and Security (CIS)*, Dec. 2017, pp. 488–491, doi: 10.1109/CIS.2017.00113.
- [45] A. Suppers, A. van Gool, and H. Wessels, "Integrated chemometrics and statistics to drive successful proteomics biomarker discovery," *Proteomes*, vol. 6, no. 2, Apr. 2018, doi: 10.3390/proteomes6020020.
- [46] V. F. Rodriguez-Galiano, J. A. Luque-Espinar, M. Chica-Olmo, and M. P. Mendes, "Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods," *Science of The Total Environment*, vol. 624, pp. 661–672, May 2018, doi: 10.1016/j.scitotenv.2017.12.152.
- [47] S. M. Kasongo and Y. Sun, "A deep learning method with wrapper based feature extraction for wireless intrusion detection system," *Computers & Security*, vol. 92, May 2020, doi: 10.1016/j.cose.2020.101752.
- [48] H. Das, B. Naik, and H. S. Behera, "A Jaya algorithm based wrapper method for optimal feature selection in supervised classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3851–3863, Jun. 2022, doi: 10.1016/j.jksuci.2020.05.002.
- [49] A. Gupta *et al.*, "A novel approach for classification of mental tasks using multiview ensemble learning (MEL)," *Neurocomputing*, vol. 417, pp. 558–584, Dec. 2020, doi: 10.1016/j.neucom.2020.07.050.
- [50] D. Lestari, R. R. Bintana, and N. Budiman, "Online internship acceptance registration application at bank," *Computer Science and Informatics Journal*, vol. 3, no. 2, pp. 127–138.

**BIOGRAPHIES OF AUTHORS**

**Israa M. Hayder**    is an M.Sc. in the Computer Science Department, College of Computer Science and Information Technology, India. She received the B.Sc. degree from Computer Science Department at Basrah University, Iraq, in 2007. She is currently a lecturer with the Southern Technical University, Iraq. She is recently interested in artificial intelligence data security database systems human computer interaction programming languages. She can be contacted at [israa.mh@stu.edu](mailto:israa.mh@stu.edu).



**Ghazwan Abdul Nabi Al Ali**    received the B.S. degree in Computer Science from Iraq, University of Basra, and the M.S. degree in Computer Science from The University of Science Malaysia. He is currently working as a programmer at the University of Basra. His research interests include software engineering and deep learning. He can be contacted at [ghazwan.alali@uobasrah.edu.iq](mailto:ghazwan.alali@uobasrah.edu.iq).



**Hussain A. Younis**    received the bachelor's degrees from the University of Basrah, Iraq, master's degrees from the Shiats University, India. He is currently pursuing a Ph.D. degree with the School of Computer Sciences, Universiti Sains Malaysia (USM). He is currently a lecturer with the College of Education for Women, University of Basrah. His research interests include artificial intelligence, electronic education, robots, image processing, pattern recognition, QR code, biometrics, and intelligent information systems. He can be contacted at [hussain.younis@uobasrah.edu.iq](mailto:hussain.younis@uobasrah.edu.iq).