IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 23, 2020, accepted January 31, 2020, date of publication February 3, 2020, date of current version February 11, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.2971354*

# An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine

**ALTYEB ALTAHER TAHA AND SHARAF JAMEEL MALEBARY**

Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh 21911, Saudi Arabia

Corresponding author: Altyeb Altaher Taha (aaataha@kau.edu.sa)

This work was supported by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, Saudi Arabia, under Grant DF-680-611-1441.

**ABSTRACT** New advances in electronic commerce systems and communication technologies have made the credit card the potentially most popular method of payment for both regular and online purchases; thus, there is significantly increased fraud associated with such transactions. Fraudulent credit card transactions cost firms and consumers large financial losses every year, and fraudsters continuously attempt to find new technologies and methods for committing fraudulent transactions. The detection of fraudulent transactions has become a significant factor affecting the greater utilization of electronic payment. Thus, there is a need for efficient and effective approaches for detecting fraud in credit card transactions. This paper proposes an intelligent approach for detecting fraud in credit card transactions using an optimized light gradient boosting machine (OLightGBM). In the proposed approach, a Bayesian-based hyperparameter optimization algorithm is intelligently integrated to tune the parameters of a light gradient boosting machine (LightGBM). To demonstrate the effectiveness of our proposed OLightGBM for detecting fraud in credit card transactions, experiments were performed using two real-world public credit card transaction data sets consisting of fraudulent transactions and legitimate ones. Based on a comparison with other approaches using the two data sets, the proposed approach outperformed the other approaches and achieved the highest performance in terms of accuracy (98.40%), Area under receiver operating characteristic curve (AUC) (92.88%), Precision (97.34%) and F1-score (56.95%).

**INDEX TERMS** Credit card fraud, electronic commerce, machine learning, optimization methods.

## I. INTRODUCTION

THE MIGRATION The migration of business to the Internet and the electronic monetary transactions that occur in the continuously growing cash-less economy have made the accurate detection of fraud a significant factor in securing such transactions. Credit card fraud occurs when a thief uses credit card information to complete purchase processes without permission from the credit card owner.

The large-scale use of credit cards and the lack of effective security systems result in billion-dollar losses to credit card fraud [1]. Because credit card firms are typically unwilling to announce such facts, it is difficult to obtain a precise approximation of the losses. However, certain data regarding the financial losses caused by credit card fraud are publicly accessible. The use of credit cards without strong security causes billion-dollar financial losses [2]. Global financial losses due to credit card fraud amounted to 22.8 billion US dollars in 2017 and are expected to continuously increase [3]; by 2020, the number is expected to reach 31 billion US dollars [4].

There are two categories of credit card fraud: application fraud [5] and behavior fraud [6]. Application fraud refers to fraudulent credit card applications. Such fraud occurs when a fraudster initiates a new credit card process using false identity details and the issuer accepts the request. Behavior fraud occurs after a credit card is correctly issued and denotes credit card transactions that involve fraudulent behavior. Credit card fraud detection has been significant issue for credit card users and financial organizations. Because detecting even a

The associate editor coordinating the review of this manuscript and approving it for publication was Feng Xia.

small number of fraudulent transactions would protect large amounts of money, credit card fraud has also become a significant problem for researchers.

For various reasons, fraud detection is considered a challenge for machine learning [7] because, for example, the distribution of data continually evolves over time due to new attack approaches and seasonality and because a very small percentage of all credit card transactions are fraudulent.

This paper proposes an intelligent approach for detecting fraudulent credit card transactions that uses an optimized light gradient boosting machine. In the proposed approach, a Bayesian-based hyperparameter optimization algorithm is intelligently integrated to tune the parameters of the light gradient boosting machine algorithm. The proposed approach is primarily concerned with discriminating between legitimate and fraudulent credit card transactions.

The main contribution of our research is an intelligent approach for detecting fraud in credit card transactions using an optimized light gradient boosting machine in which a Bayesian-based hyperparameter optimization algorithm is utilized to optimize the parameters of the light gradient boosting machine. The performance of the proposed intelligent approach is evaluated based on two real-world data sets and compared with other machine learning techniques using performance evaluation metrics.

The remainder of the paper is structured as follows. Related research is reviewed in the second section. Section three describes our proposed intelligent approach for credit card fraud detection, and in section four, the results of experiments are discussed. Finally, the study' s conclusions are summarized in section five.

## II. RELATED WORK

The potential social and economic importance of detecting fraudulent credit card transactions has increased the number of relevant research efforts in the literature. This section reviews several significant studies. More comprehensive reviews can be found in [8]–[11].

There are two main approaches for detecting fraudulent credit card transactions using machine learning algorithms: supervised learning algorithms and unsupervised learning algorithms. In supervised learning algorithms, historical credit card transactions are labeled as legitimate or fraudulent. Then, supervised learning algorithms start learning using these data to create a model that can be used to categorize new data samples. In contrast, unsupervised learning algorithms are based on the direct classification of credit card transactions using patterns that are considered normal. Then, the algorithm classifies transactions that do not conform to such patterns as fraudulent credit card transactions.

Both supervised learning [8], and unsupervised learning [12] algorithms have been utilized for credit card fraud detection. The most popular algorithms for the detection of credit card fraud use supervised learning and employ labeled transactions for classifier training. Fraudulent credit card transactions are detected by classifying features extracted from credit card transactions [6].

A number of classification algorithms have been utilized to detect fraudulent credit card transactions. A probabilistic neural network (PNN), logistic regression (LOR) and genetic programming (GP) have been employed for classifying fraud in credit card transactions [13]. A data set of 202 Chinese firms was used, and t-statistics were applied to select the important features. The results revealed that PNN outperformed the other approaches [13]. Bayesian belief networks (BNNs) and decision trees (DTs) were used in [14] to detect fraud in financial transactions. Here, a data set of financial transactions collected from 76 Greek industrial companies was used. The data set included 38 financial transactions confirmed to be fraudulent by assessors. The BNNs obtained the highest accuracy (90.3%), whereas the DTs achieved an accuracy of 73.6% [14].

A self-organizing map (SOM) was used to generate a model for unsupervised credit card fraud detection. The advantages of this method are that because the SOM model does not require prior information, the model is updated continuously by adding new credit card transactions; the disadvantage may be the difficulty of detecting fraudulent credit card transactions with high accuracy [15]–[17].

Recently, deep learning has become a powerful component of machine learning and achieved promising results in several fields, such as image processing [18]. Jurgovsky *et al.* [19] utilized a long short-term memory (LSTM) frame to detect credit card fraud as a sequence classification issue in the supervised learning category. Kraus and Feuerriegel [20] utilized deep-learning approaches to support financial decisions. Fiore *et al.* [21] proposed a scheme to make synthetic instances based on generative adversarial networks to enhance credit card fraud detection performance by solving the issue of the imbalanced data set.

Carcillo *et al.* [33] implemented a hybrid approach that utilizes unsupervised outlier scores to expand the set of features of the fraud detection classifier. Their main contribution was to implement and assess various levels of granularity for outlier score definition. Their experimental results indicate that their proposed approach is efficient and enhances detection accuracy.

Carcillo *et al.* [34] also introduced the SCAlable Real-time Fraud Finder (SCARFF), which incorporates big-data techniques (Cassandra, Kafka and Spark) in a machine learning method to address nonstationarity, imbalance, and feedback latency. The results of experiments based on a large data set of real credit card transactions demonstrated that the framework is efficient, accurate and scalable.

Saia and Carta [35] proposed a new approach to credit card fraud detection based on a model defined using a discrete Fourier transform converted to utilize frequency patterns. The approach has the advantages of treating imbalanced class distribution and cold-start issues by considering only past legitimate transactions, thus decreasing the data heterogeneity problem.

Yuan *et al.* [36] introduced a novel framework that combines deep neural networks and spectral graph analysis for fraud detection. They developed and assessed two neural networks for fraud detection: a deep auto encoder and a convolutional neural network. Experimental results indicated that their proposed approach is effective for fraud detection.

In [37], Saia presented a novel credit card fraud detection method based on the discrete wavelet transform, which was utilized to construct an evaluation model capable of overcoming problems related to the imbalanced distribution of credit card fraud data sets. The experimental results indicated that the performance of the proposed approach was comparable to that of state-of-the-art approaches, such as random forests. West and Bhattacharya [38] presented a comprehensive review of financial fraud detection approaches using computational intelligence techniques. In addition, they identified research gaps that were not addressed by other review articles.

Ensemble classifiers associate what is currently learned from new samples from previously attained knowledge. Dhankhad *et al.* [42] applied many supervised machine learning algorithms to identify fraudulent credit card transactions using a real-world data set. Then, they used these algorithms to implement a super classifier based on ensemble learning approaches. Their results indicated that the ensemble approach achieved the best performance. Dal Pozzolo *et al.* [43] designed two fraud detection systems based on an ensemble method and a sliding-window method, respectively. The study revealed that the winning strategy involved training two separate classifiers and then aggregating the outcomes. Based on experiments on a large data set, the results indicated that the proposed approach improved fraud alert precision.

Bio-inspired algorithms offer global solutions to the optimization problems. The combining bio-inspired optimization algorithms with machine learning models may enhance the performances of the machine learning models because it has the ability to deduct the best solutions for the optimization problem. Therefore, machine learning models have been coupled with bio-inspired optimization techniques, Kamaruddin and Ravi [47] developed a hybrid approach of Particle Swarm Optimization and Auto-Associative Neural Network for credit card fraud detection.

## III. PROPOSED INTELLIGENT APPROACH FOR CREDIT CARD FRAUD DETECTION

The overall framework of the proposed intelligent approach for credit card fraud detection is illustrated in figure 1.

The proposed intelligent approach for credit card fraud detection consists of four major steps, which are explained in the following subsections. The experiment was performed using an Intel Core i7 processor with 8GB RAM. The proposed approach and other machine learning techniques were implemented and tested using Python.

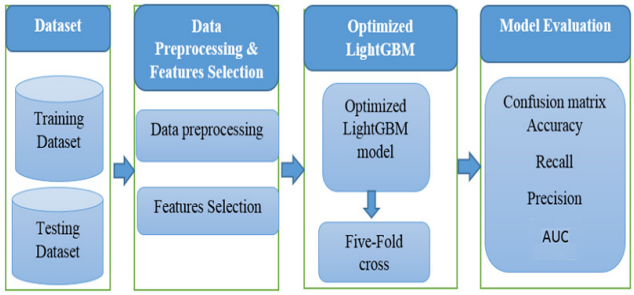


**FIGURE 1.** Overall framework of the proposed intelligent approach for credit card fraud detection.

**TABLE 1.** Summary of the analyzed data sets.

| Dataset | Total No. of transactions | No. of legitimate transactions | No. of fraudulent transactions | No. of Features | Ref. |
|---|---|---|---|---|---|
| Data set 1 | 284,807 | 284,315 | 492 | 31 | [22] |
| Data set 2 | 94,683 | 92,589 | 2,094 | 20 | [39] |

### A. DATA SET AND DATA PREPROCESSING

To develop different experiments for evaluating the proposed approach and demonstrating its generality, we consider two different real-world data sets.

The first data set consists of 284,807 credit card transactions made by the credit card owners in September 2013 in Europe. Of the 284,807 transactions in the data set, 492 were fraudulent; the positive class (i.e., the fraudulent transactions) represents 0.172% of all transactions [22]. The data set includes 31 features. The first 28 features (i.e., V1 to V28) are the principal components obtained using principal components analysis (PCA). The basic reason is to maintain data privacy. ''Time'' and ''Amount'' are the only two features that are not transformed using PCA.

The second data set is the UCSD-FICO Data Mining Contest 2009 Dataset [39], which is a real data set of e-commerce transactions. The objective was to detect anomalous e-commerce transactions. The data set consists of 94,683 transactions, 2,094 of which are fraudulent. The data set was collected from 73,729 credit cards during a period of 98 days. It contains 20 fields, including class, and the fields labels are as follows: amount, hour1, state1, zip1, custAttr1, field1, custAttr2, field2, hour2, flag1, total, field3, field4, indicator1, indicator2, flag2, flag3, flag4, flag5, and Class. The Class feature in the two data sets is the classification variable, which is 1 in the case of credit card fraud and 0 otherwise. The data sets are summarized in table 1, where we present the total number of transactions, the number of legitimate transactions, the number of fraudulent transactions, the number of features in the data set and the references for downloading each data set.

Because the total number of fraudulent transactions is much less than the number of legitimate transactions, the data

distribution is unbalanced, i.e., skewed towards legitimate observations. It is well known that the performance of various machine learning methods decreases when the analyzed data set is unbalanced [32]. To obtain more accurate results, a cross validation procedure is employed in this paper to train and test the model in each subset of the two data sets; then, the average of all the noted metrics is calculated over the data set [40]. Other techniques, such as resampling, have also been suggested to address imbalanced data sets [23].

### B. FEATURE SELECTION

Selecting significant and important features is critical for the effective detection of credit card fraud when the number of features is large [24]. LightGBM [41] utilizes the information gain (IG) method to select the most important features and thus decrease the dimensionality of the training data. Information gain functions by extracting similarities between credit card transactions and then awarding the greatest weight to the most significant features based on the class of legitimate and fraudulent credit card transactions. Because of its computational efficiency and leading performance in terms of precision [44], information gain is employed as a feature selection method in the proposed approach.

### C. THE OPTIMIZED LIGHT GRADIENT BOOST CLASSIFIER

This section explains the proposed intelligent approach for detecting fraudulent credit card transactions using an optimized light gradient boosting framework based on tree learning algorithms. In the proposed approach, a Bayesian-based hyperparameter optimization algorithm is intelligently integrated to tune the parameters of the LightGBM algorithm. The high-performance LightGBM algorithm can quickly handle large amounts of data and the distributed processing of data. It was developed as an open source project by Microsoft. The Light Gradient Boosting algorithm is explained in figure 2.

The LightGBM algorithm includes several parameters, termed hyper parameters. The hyper parameters have a significant impact on the performance of LightGBM algorithm. They are typically set manually and then tuned in a continuous trial and error process. In this paper, a Bayesian based hyperparameter optimization algorithm is intelligently integrated in the proposed approach to tune the hyper parameters of the LightGBM algorithm. The tuned hyper parameters include the following: ''num_leaves'', which is the number of leaves per tree, ''max_depth'', which denotes the maximum depth of the tree, and ''learning_rate''.

The proposed approach is based on the LightGBM algorithm, which can bundle unique features into a single bundle; then, the feature-scanning algorithm can be considered for creating same-feature histograms based on the feature bundles [41]. The computation complexity of the proposed approach was calculated based on theoretical time complexity as follows: computation complexity = (n∗m), where m denotes the number of data set samples, and m denotes the number of bundles that is less than the number of features in the data set. LightGBM uses the gradient-based one side sampling (GOSS) method to preserve the accuracy of the information gain estimation; GOSS retains instances with large gradients (e.g., larger than a predefined threshold or those among the top percentiles) and randomly samples those instances with small gradients [41].

The LightGBM algorithm
Input:
Training data: D = {(χ1, y1), (χ2, y2), ..., (χN, yN)}, χi ∈χ, χ ⊆ R, yi∈{−1,+1}; loss function: L(y, θ (χ));
Iterations:
M; Big gradient data sampling ratio: a; slight gradient data sampling ratio: b;
1: Combine features that are mutually exclusive (i.e., features never simultaneously accept nonzero values) of χi, i = {1, ...,N} by the exclusive feature bundling (EFB) technique;
2: Set $\theta_0(\chi) = arg\ min_c \sum_i^N L(y_i, c)$;
3: For m = 1 to M do
4: Calculate gradient absolute values:
$$r_i = \left|\frac{\partial L(y_i,\theta(x_i))}{\partial \theta(x_i)}\right|_{\theta(x)=\theta_{m-1(x)}}, i = \{1, \ldots, N\}$$
5: Resample data set using gradient-based one-side sampling (GOSS) process:
$topN = a \times len(D)$; $randN = b \times len(D)$;
$sorted = GetSortedIndices(abs(r))$;
$A = sorted[\ 1 : topN]$; $B = RandomPick(sorted[\ topN : len(D)]\ , randN)$;
$\acute{D} = A + B$;
6: Calculate information gains:
$$V_j(d) = \frac{1}{n}\left(\frac{\left(\sum_{x_i\in A_l} r_i + \frac{1-a}{b}\sum_{x_i\in B_l} r_i\right)^2}{n_l^j(d)} + \frac{\left(\sum_{x_i\in A_r} r_i + \frac{1-a}{b}\sum_{x_i\in B_r} r_i\right)^2}{n_r^j(d)}\right)$$
7: Develop a new decision tree $\theta_m(x)'$ on set $D'$
8: Update $\theta_m(\chi) = \theta_{m-1}(\chi) + \theta_m(\chi)$
9: End for
10: Return $\tilde{\theta}(x) = \theta_M(x)$

**FIGURE 2.** The LightGBM algorithm.

### D. MODEL EVALUATION USING PERFORMANCE METRICS

To evaluate the performance of the proposed approach for credit card fraud detection, a cross validation test is applied. The k-fold cross-validation (CV) method is utilized to systematically and carefully assess the performance of the proposed approach for credit card fraud detection. K-Fold CV is a statistical analysis approach that has been widely employed by researchers to assess the performance of the machine learning classifier [25]. In this research, we conduct a 5-fold CV test to assess the performance of the proposed approach. The two analyzed data sets have imbalance in classes: there are more normal than fraudulent transactions. In this case, to achieve more accurate estimates, cross validation is used to train and test the model in each subset of the two data sets; then, the average of all the noted metrics is calculated over the data set [40].

Each data set is divided randomly into five separate subsets of equal size. At each step of validation, a single subset (20% of the data set) is reserved as the validation data set for testing the performance of the proposed approach, while the remaining four subsets (80% of the data set) are employed as the training data set. This process is then repeated five

times until each subset has been used. The average of the performances of the five test subsets is calculated, and the final result is the total performance of the proposed approach on a 5-fold CV test.

To assess the performance of the proposed approach, several measures are considered, including the Confusion Matrix, Precision, Recall, Accuracy (ACC), AUC and F1-score. The metrics are defined based on the confusion matrix.

The Confusion Matrix for measuring credit card fraud detection performance uses the following terms [26]:

TP (i.e., true positive) refers to the number of fraudulent credit card transactions properly classified.

FP (i.e., false positive) denotes the number of legitimate fraud credit card transactions classified as fraud. FN (i.e., false negative) denotes the number of fraudulent credit card transactions classified as normal. TN (i.e., true negative) refers to the number of normal credit card transactions correctly classified. The measures that were used are defined as follows.

$$\begin{aligned} \text{Accuracy} &= (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \\ \text{Recall} &= \text{TP}/(\text{TP} + \text{FN}) \\ \text{Precision} &= \text{TP}/(\text{TP} + \text{FP}) \\ \text{F1}-\text{measure} &= 2 \times (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) \end{aligned}$$

Precision and Recall are important metrics used with unbalanced data when combined together (i.e.,F-score). Precision indicates the correctness of the suitability of the result scale and proximity to the expected solution, while Recall is a measure of the number of relevant results. A high Recall score reflects a low false negative (FN) rate, while high Precision reflects a low false positive (FP) rate. High scores for Precision and Recall indicate that the classifier restores results with high accuracy and recovers most of the positive results [27]. Therefore, the Precision-Recall curve reveals a complete picture of the accuracy of the classifier and is robust even in an imbalanced data set [28]. We consider the AUC value as a general performance measure in addition to the above measures. AUC is a graphical plot of the false positive rate (FPR) and the true positive rate (TPR) at different possible levels. Because it is independent of a cutoff value, AUC is considered a better overall performance indicator than accuracy. A model with better overall performance has an AUC value close to one.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

To investigate the effectiveness of the proposed approach for credit card fraud detection, the 5-fold CV procedure is conducted using two real-world data sets. The proposed approach is trained with parameters optimized using the Bayesian-based hyperparameter optimization algorithm.

Table 2 shows a performance evaluation of the proposed approach based on the 5-fold CV procedure using the two real-world data sets. To obtain a reliable performance comparison, the cross validation procedure is employed because it uses each data set part (i.e., fold) for testing and training processes. The proposed approach achieved average AUC values of 90.94% and 92.90% for data set 1 and data set 2, respectively, which indicates the ability of the proposed approach in distinguishing between legitimate and fraudulent credit card transactions. In addition, the proposed approach achieved an average accuracy of 0.98% for the two data sets, which is the ratio of correctly predicted credit card transactions to the total number of transactions. Attaining the same accuracy score using two different data sets indicates the considerable consistency of the proposed approach. The proposed approach also obtained average Recall scores of 40.59% and 28.33% for data set 1 and data set 2, respectively, which indicates the ability of the proposed approach to correctly detect more than 40% of the suspicious credit card transactions with a low false negative percentage in data set 1. In addition, the proposed approach achieved average Precision scores of 97.34% and 91.72% for data set 1 and data set 2, respectively, which is the ratio of correctly classified fraudulent transactions to classified fraudulent transactions. In addition, the proposed approach achieved average F1-scores of 56.95% and 43.27% for data set 1 and data set 2, respectively. The F1-score indicates the balance between Precision and Recall. Therefore, this score takes both false negatives and false positives into account. It is an important measure, particularly if the number of legitimate and fraudulent credit card transactions is not balanced. The overall performance of the proposed approach is highly consistent based on the 5-fold CV procedure using the two data sets. Using cross validation, the data set 1 results based on the various performance metrics are highly consistent with the results for data set 2.

**TABLE 2.** Performance evaluation of the proposed approach based on 5-fold cross validation using two real-world data sets.

| Data set | Fold No | AUC | Accuracy | Recall | Precision | F1 score |
|---|---|---|---|---|---|---|
| Data set 1 | 1 | 0.9116 | 0.9832 | 0.3788 | 0.9758 | 0.5458 |
| | 2 | 0.9090 | 0.9833 | 0.3824 | 0.9760 | 0.5494 |
| | 3 | 0.9066 | 0.9829 | 0.3647 | 0.9770 | 0.5306 |
| | 4 | 0.9098 | 0.9829 | 0.3665 | 0.9707 | 0.5316 |
| | 5 | 0.9100 | 0.9877 | 0.5373 | 0.9675 | 0.6903 |
| | AVG | **0.9094** | **0.9840** | **0.4059** | **0.9734** | **0.5695** |
| Data set 2 | 1 | 0.9428 | 0.9840 | 0.2912 | 0.9531 | 0.4461 |
| | 2 | 0.9326 | 0.9838 | 0.2936 | 0.9226 | 0.4453 |
| | 3 | 0.9275 | 0.9834 | 0.2792 | 0.9092 | 0.4269 |
| | 4 | 0.9141 | 0.9834 | 0.2768 | 0.9061 | 0.4239 |
| | 5 | 0.9281 | 0.9833 | 0.2760 | 0.8950 | 0.4217 |
| | AVG | **0.9290** | **0.9835** | **0.2833** | **0.9172** | **0.4327** |

For better visualization of the performance evaluation results, the AUC curves based on the 5-fold cross validation procedure using the two data set metrics were plotted in line charts, as shown in Fig. 3 and Fig. 4. The Recall-Precision curves are presented in Fig. 5 and Fig.6.

To further evaluate its performance and robustness, the proposed approach is compared with state-of-the-art machine learning algorithms, including random forest, logistic

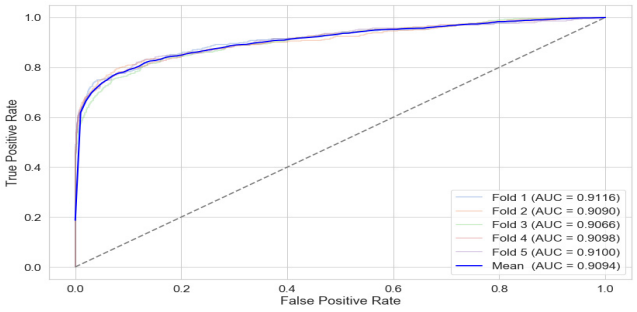


**FIGURE 3.** AUC curve of the proposed approach based on the 5-fold CV procedure for data set 1.

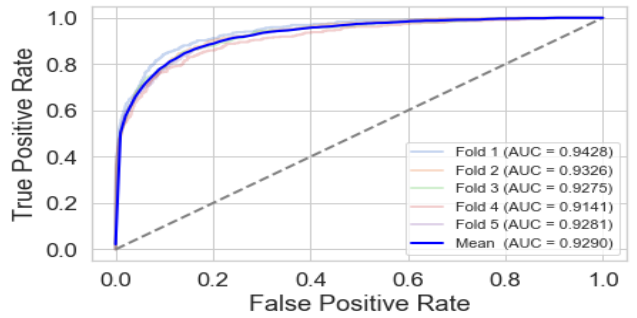


**FIGURE 4.** AUC curve of the proposed approach based on the 5-fold CV procedure for data set 2.

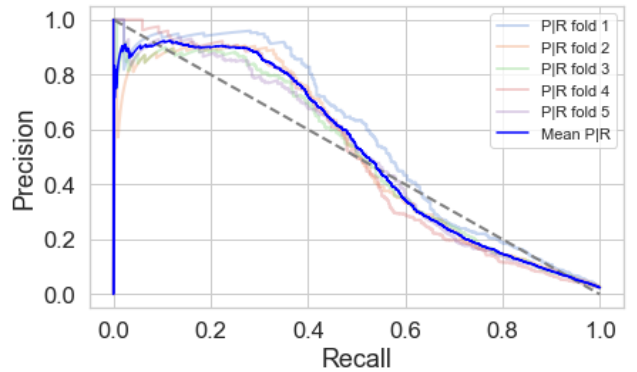


**FIGURE 5.** Precision-Recall curve of the proposed approach based on the 5-fold CV procedure for data set 1.

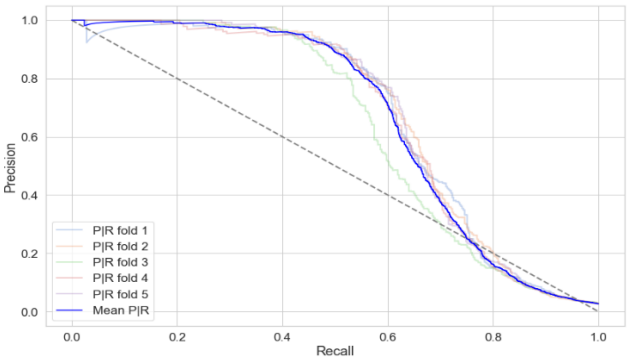


**FIGURE 6.** Precision-Recall curve of the proposed approach based on the 5-fold CV procedure for data set 2.

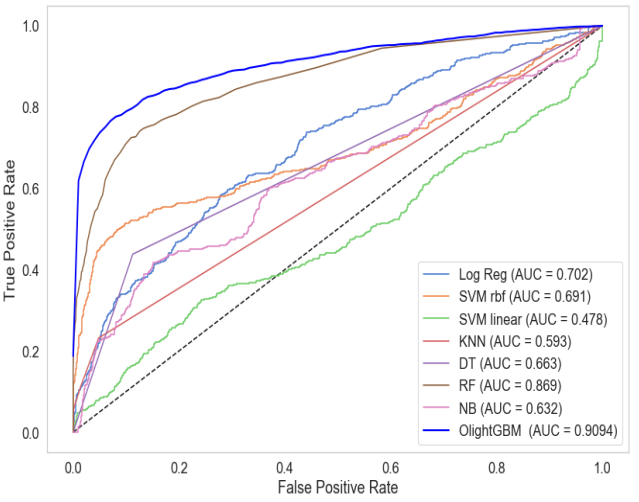


**FIGURE 7.** AUC curves of the proposed approach and other machine learning algorithms for data set 1.

regression, the radial support vector machine, the linear support vector machine, k-nearest neighbors, decision tree, and naïve bayes. Performance is evaluated in terms of AUC, Accuracy, Precision, Recall and F1-score using two real-world data sets (table 3).

As shown in table 3, the proposed approach achieved the highest AUC scores: 90.94% and 92.88% for data set 1 and data set 2, respectively. The RF algorithm achieved the second-highest AUC scores: 86.90% and 90.70% for data set 1 and data set 2, respectively. The SVM linear algorithm achieved the lowest AUC scores: 47.80% and 70.90% for data set 1 and data set 2, respectively. Table 3 shows that the proposed approach attained the highest Accuracy (98.40% and 98.35% for data set 1 and data set 2, respectively), while the NB algorithm attained the lowest Accuracy (85% and 95.92% for data set 1 and data set 2, respectively). The SVM Rbf algorithm achieved the highest Recall values: 86.95% and 82.14% for data set 1 and data set 2, respectively.

The proposed approach achieved the highest Precision scores: 97.34% and 91.72% for data set 1 and data set 2, respectively. In addition, the proposed approach attained the highest F1-score (56.95% for data set 1), while the RF algorithm achieved an F1-score of 45.45% for data set 2, slightly outperforming the proposed approach, which achieved an F1-score of 43.27% for the same data set.

The AUC curve is an important and useful estimation of overall performance and a general measure of the accuracy of fraud investigation [29]. A higher AUC curve indicates better prediction performance.

Fig. 7 and Fig. 8 present the AUC curves of the proposed approach and other machine learning algorithms for the two real-world data sets. Fig. 7 and Fig.8 confirm the results shown in table 3. The AUC curve of the proposed approach is located nearest to the figures' top-left corner, suggesting that the proposed approach for credit card fraud detection achieved better performance for the two data sets. Overall,

**TABLE 3.** Performance evaluation comparing the proposed approach with various machine learning algorithms for two real-world data sets.

| Data set | Approach | AUC | Accuracy | Recall | Precision | F1-score |
|---|---|---|---|---|---|---|
| Data set 1 | Log Reg | 0.7020 | 0.9685 | 0.0101 | 0.0015 | 0.0030 |
| | SVM Rbf | 0.6910 | 0.9733 | 0.8695 | 0.0458 | 0.0871 |
| | SVM Linear | 0.4780 | 0.9709 | 0.0106 | 0.0021 | 0.0014 |
| | KNN | 0.5930 | 0.9690 | 0.1498 | 0.1123 | 0.1284 |
| | DT | 0.6630 | 0.9560 | 0.0799 | 0.4587 | 0.1375 |
| | RF | 0.8690 | 0.9779 | 0.6547 | 0.2522 | 0.3642 |
| | NB | 0.6320 | 0.8500 | 0.0111 | 0.0045 | 0.0064 |
| | **Our Approach** | **0.9094** | **0.9840** | **0.4059** | **0.9734** | **0.5695** |
| Data set 2 | Log Reg | 0.8570 | 0.9781 | 0.7457 | 0.0839 | 0.1509 |
| | SVM Rbf | 0.7410 | 0.9781 | 0.8214 | 0.0438 | 0.0833 |
| | SVM Linear | 0.7090 | 0.9778 | 0.1075 | 0.0231 | 0.0833 |
| | KNN | 0.7150 | 0.9756 | 0.5507 | 0.2175 | 0.3119 |
| | DT | 0.7130 | 0.9667 | 0.3725 | 0.4312 | 0.4151 |
| | RF | 0.9070 | 0.9789 | 0.8168 | 0.3148 | 0.4545 |
| | NB | 0.8480 | 0.9592 | 0.1475 | 0.1812 | 0.1626 |
| | **Our Approach** | **0.9288** | **0.9835** | **0.2833** | **0.9172** | **0.4327** |

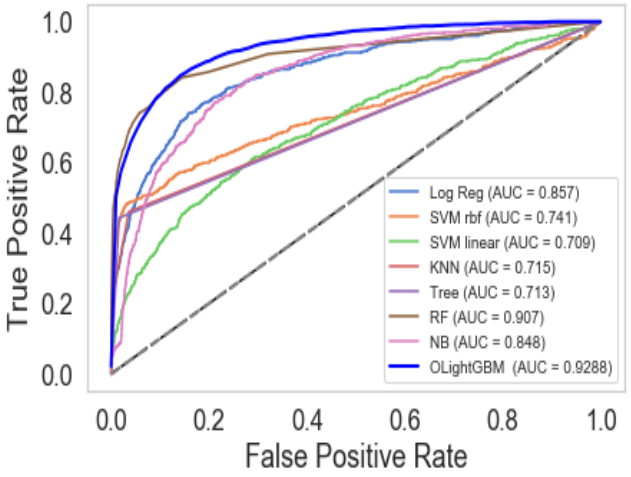


**FIGURE 8.** AUC curves of the proposed approach and other machine learning algorithms for data set 2.

these results demonstrate the effectiveness of the proposed approach.

The precision and recall curve is commonly used to compare classifiers in terms of precision and recall. It is a two-dimensional graph in which the precision rate is plotted on the y-axis and recall is plotted on the y-axis.

The precision-recall curve provides a full picture of the performance of the classification and is stable even in imbalanced data sets [28].

Fig. 9 and Fig.10 present clear visualizations of the precision-recall curves of the proposed approach and the other machine learning algorithms. The precision-recall

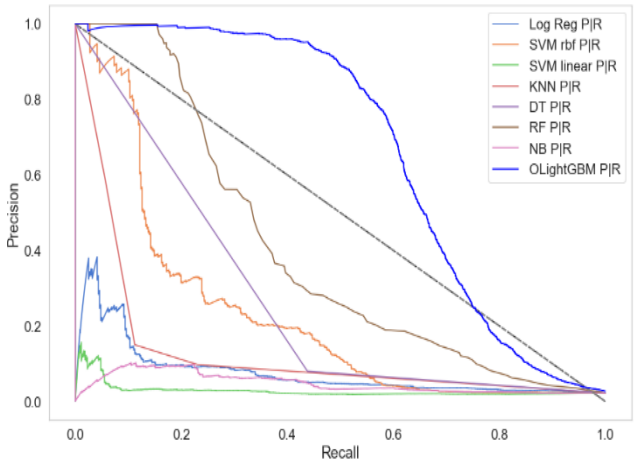


**FIGURE 9.** Precision-recall curves of the proposed approach and other machine learning algorithms for data set 1.

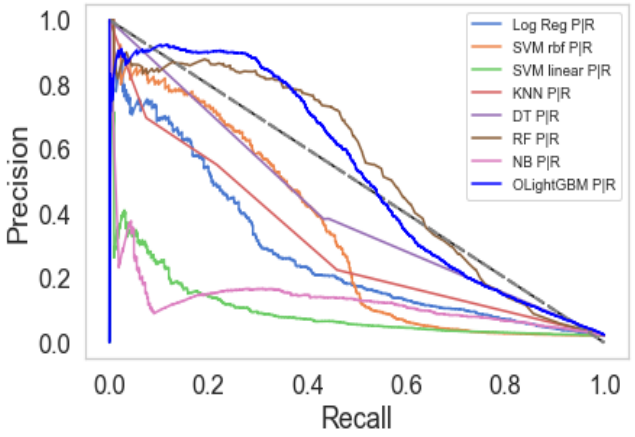


**FIGURE 10.** Precision-recall curves of the proposed approach and other machine learning algorithms for data set 2.

**TABLE 4.** Performance comparison between the proposed OLightGBM and other methods based on the AUC metric.

| Approach | AUC |
|---|---|
| LightGBM | 90.62% |
| Catboost | 87.86% |
| Proposed approach | 92.88% |

curve of the proposed approach is located nearest to the figures' top-right corners, suggesting that the proposed approach for credit card fraud detection achieved better performance for the two data sets.

An additional performance comparison is conducted by comparing the proposed approach with current machine learning techniques and previous studies (tables 4 and 5).

As shown in table 4, the proposed approach outperforms all the other approaches. The proposed approach achieved the highest AUC (92.88%), while Catboost achieved the lowest Accuracy (87.86%). The results reveal that the proposed algorithm is superior to other classifiers. The results also highlight the importance and value of adopting an efficient parameter optimization strategy for enhancing the predictive

**TABLE 5. Performance comparison of the proposed approach with other methods based on the accuracy metric.**

| Approach | Accuracy |
|---|---|
| Concept Drifts Adaptation [31] | 80% |
| Local Outlier Factor [4] | 97% |
| Isolation Forest [29] | 95% |
| Random Forest [45] | 95.5% |
| ANN [46] | 92.86% |
| Proposed approach | 98.40% |

performance of LightGBM models because it increases the AUC of the proposed approach by 2.26%.

Table 5 shows a performance comparison between the proposed approach and other research outcomes based on achieved accuracy for the same data set. The proposed approach obtained the highest Accuracy (98.40%), while the Concept Drifts Adaptation [31] achieved the lowest Accuracy (80%).

## V. CONCLUSION

The detection of credit card fraud is significant to the improved utilization of credits cards. With large and continuing financial losses being experienced by financial firms and given the increasing difficulty of detecting credit card fraud, it is important to develop more effective approaches for detecting fraudulent credit card transactions.

This paper proposes an intelligent approach for detecting fraud in credit card transactions using an optimized light gradient boosting machine (OLightGBM). We conducted several experiments using two real-world data sets. The performance of the proposed approach was evaluated through comparison with other research outcomes and state-of-the-art machine learning algorithms, including random forest, logistic regression, the radial support vector machine, the linear support vector machine, k-nearest neighbors, decision tree, and naive bayes. The experimental results indicate that the proposed approach outperformed the other machine learning algorithms and achieved the highest performance in terms of Accuracy, AUC, Precision and F1-score. The results reveal that the proposed algorithm is superior to other classifiers. The results also highlight the importance and value of adopting an efficient parameter optimization strategy for enhancing the predictive performance of the proposed approach.

## ACKNOWLEDGMENT

This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant no. (DF-680-611-1441). The authors, therefore, gratefully acknowledge DSR technical and financial support.

## REFERENCES

[1] X. Zhang, Y. Han, W. Xu, and Q. Wang, "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture," *Inf. Sci.*, May 2019. Accessed: Jan. 8, 2019.

[2] N. Carneiro, G. Figueira, and M. Costa, "A data mining based system for credit-card fraud detection in e-tail," *Decis. Support Syst.*, vol. 95, pp. 91–101, Mar. 2017.

[3] B. Lebichot, Y.-A. Le Borgne, L. He-Guelton, F. Oblé, and G. Bontempi, "Deep-learning domain adaptation techniques for credit cards fraud detection," in *Proc. INNS Big Data Deep Learn. Conference*, Genoa, Italy, 2019, pp. 78–88.

[4] H. John and S. Naaz, "Credit card fraud detection using local outlier factor and isolation forest," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 4, pp. 1060–1064, Sep. 2019.

[5] C. Phua, R. Gayler, V. Lee, and K. Smith-Miles, "On the communal analysis suspicion scoring for identity crime in streaming credit applications," *Eur. J. Oper. Res.*, vol. 195, no. 2, pp. 595–612, Jun. 2009.

[6] R. Bolton and D. Hand, "Statistical fraud detection: A review," *Stat. Sci.*, vol. 17, no. 3, pp. 235–249, Aug. 2002.

[7] P. A. Dal, G. Boracchi, O. Caelen, C. Alippi, and G. Bontempi, "Credit card fraud detection: A realistic modeling and a novel learning strategy," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3784–3797, Sep. 2017.

[8] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," *Decis. Support Syst.*, vol. 50, no. 3, pp. 602–613, Feb. 2011.

[9] N. Sethi and A. Gera, "A revived survey of various credit card fraud detection techniques," *Int. J. Comput. Sci. Mobile Comput.*, vol. 3, no. 4, pp. 780–791, Apr. 2014.

[10] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques," *Int. J. Syst. Assurance Eng. Manage.*, vol. 8, no. S2, pp. 937–953, Nov. 2017.

[11] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in *Proc. ICCNI*, Lagos, Nigeria, Oct. 2017, pp. 1–9.

[12] M. Carminati, R. Caron, F. Maggi, I. Epifani, and S. Zanero, "BankSealer: A decision support system for online banking fraud analysis and investigation," *Comput. Secur.*, vol. 53, no. 1, pp. 175–86, Sep. 2015.

[13] P. Ravisankar, V. Ravi, G. R. Rao, and I. Bose, "Detection of financial statement fraud and feature selection using data mining techniques," *Decis. Support Syst.*, vol. 50, no. 2, pp. 491–500, Jan. 2011.

[14] E. Kirkos, C. Spathis, and Y. Manolopoulos, "Data mining techniques for the detection of fraudulent financial statements," *Expert Syst. Appl.*, vol. 32, no. 4, pp. 995–1003, May 2007.

[15] D. Olszewski, "Fraud detection using self-organizing map visualizing the user profiles," *Knowl.-Based Syst.*, vol. 70, pp. 324–334, Nov. 2014.

[16] J. T. Quah and M. Sriganesh, "Real-time credit card fraud detection using computational intelligence," *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1721–1732, Nov. 2008.

[17] V. Zaslavsky and A. Strizhak, "Credit card fraud detection using self organizing maps," *Inf. Secur.*, vol. 18, p. 48, Jan. 2006.

[18] L. Wang, T. Liu, G. Wang, K. L. Chan, and Q. Yang, "Video tracking using learned hierarchical features," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1424–1435, Apr. 2015.

[19] J. Jurgovsky, M. Granitzer, K. Ziegler, S. Calabretto, P.-E. Portier, L. He-Guelton, and O. Caelen, "Sequence classification for credit-card fraud detection," *Expert Syst. Appl.*, vol. 100, pp. 234–245, Jun. 2018.

[20] M. Kraus and S. Feuerriegel, "Decision support from financial disclosures with deep neural networks and transfer learning," *Decis. Support Syst.*, vol. 104, pp. 38–48, Dec. 2017.

[21] U. Fiore, A. D. Santis, F. Perla, P. Zanetti, and F. Palmieri, "Using generative adversarial networks for improving classification effectiveness in credit card fraud detection," *Inf. Sci.*, vol. 479, pp. 448–455, Apr. 2019.

[22] *Credit Card Fraud Dataset*. Accessed: Sep. 4, 2019. [Online]. Available: https://www.kaggle.com/mlg-ulb/creditcardfraud/data

[23] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study," *Intell. Data Anal.*, vol. 6, no. 5, pp. 429–449, 2002.

[24] A. C. Bahnsen, D. Aouada, A. Stojanovic, and B. Ottersten, "Feature engineering strategies for credit card fraud detection," *Expert Syst. Appl.*, vol. 51, pp. 134–142, Jun. 2016.

[25] R. D. Kumar, "Statistically identifying tumor suppressors and oncogenes from pan-cancer genome-sequencing data," *Bioinformatics*, vol. 31, no. 22, pp. 3561–3568, 2015.

[26] I. Mekterović, L. Brkić, and M. Baranovi, "A systematic review of data mining approaches to credit card fraud detection," *WSEAS Trans. Bus. Econ.*, vol. 15, p. 437, Jan. 2018.

[27] M. A. Al-Shabi, "Credit card fraud detection using auto encoder model in unbalanced datasets," *J. Adv. Math. Comput. Sci.*, vol. 33, no. 5, pp. 1–16, 2019.

[28] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *Proc. 23rd Int. Conf. Mach. Learn.*, Philadelphia, PA, USA, 2006, pp. 233–240.
[29] A. D. Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi, "Learned lessons in credit card fraud detection from a practitioner perspective," *Expert Syst. Appl.*, vol. 41, no. 10, pp. 4915–4928, Aug. 2014.
[30] H. A. El Bour, Y. Oubrahim, M. Y. Ghoumari, and M. Azzouazi, "Using isolation forest in anomaly detection: The case of credit card transactions," *Periodicals Eng. Natural Sci.*, vol. 6, no. 2, pp. 394–400, 2018.
[31] A. Jog and A. A. Chandavale, "Implementation of credit card fraud detection system with concept drifts adaptation," in *Proc. ICICC*, Singapore, 2018, pp. 467–477.
[32] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, "Credit card fraud detection using AdaBoost and majority voting," *IEEE Access*, vol. 6, pp. 14277–14284, 2018.
[33] F. Carcillo, Y.-A. Le Borgne, O. Caelen, Y. Kessaci, F. Oblé, and G. Bontempi, "Combining unsupervised and supervised learning in credit card fraud detection," *Inf. Sci.*, to be published.
[34] F. Carcillo, A. D. Pozzolo, Y.-A. Le Borgne, O. Caelen, Y. Mazzer, and G. Bontempi, "SCARFF: A scalable framework for streaming credit card fraud detection with spark," *Inf. Fusion*, vol. 41, pp. 182–194, May 2018.
[35] R. Saia and S. Carta, "A frequency-domain-based pattern mining for credit card fraud detection," in *Proc. 2nd Int. Conf. Internet Things, Big Data Secur.*, 2017, pp. 386–391.
[36] S. Yuan, X. Wu, J. Li, and A. Lu, "Spectrum-based deep neural networks for fraud detection," in *Proc. ACM Conf. Inf. Knowl. Manage. (CIKM)*, 2017, pp. 2419–2422.
[37] R. Saia, "A discrete wavelet transform approach to fraud detection," in *Proc. Int. Conf. Netw. Syst. Secur.* Cham, Switzerland: Springer, 2017, pp. 464–474.
[38] J. West and M. Bhattacharya, "Intelligent financial fraud detection: A comprehensive review," *Comput. Secur.*, vol. 57, pp. 47–66, Mar. 2016.
[39] *UCSD: University of California, San Diego Data Mining Contest 2009*. Accessed: Jan. 14, 2019. [Online]. Available: https://www.cs.purdue.edu/commugrate/data/credit_card/
[40] S. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. London, U.K.: Pearson, 2016.
[41] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3146–3154.
[42] S. Dhankhad, E. Mohammed, and B. Far, "Supervised machine learning algorithms for credit card fraudulent transaction detection: A comparative study," in *Proc. IEEE Int. Conf. Inf. Reuse Integr. (IRI)*, Jul. 2018, pp. 122–125.
[43] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi, and G. Bontempi, "Credit card fraud detection and concept-drift adaptation with delayed supervised information," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2015, pp. 1–8.
[44] A. Jović, K. Brkić, and N. Bogunović, "A review of feature selection methods with applications," in *Proc. 38th Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO)*, 2015, pp. 1200–1205.
[45] S. V. S. S. Lakshmi and S. D. Kavilla, "Machine learning for credit card fraud detection system," *Int. J. Appl. Eng. Res.*, vol. 13, no. 24, pp. 16819–16824, 2018.
[46] A. Rohilla, "Comparative analysis of various classification algorithms in the case of fraud detection," *Int. J. Eng. Res. Technol.*, vol. 6, no. 9, pp. 1–6, 2017.
[47] S. Kamaruddin and V. Ravi, "Credit card fraud detection using big data analytics: Use of PSOAANN based one-class classification," in *Proc. Int. Conf. Informat. Anal.*, 2016, pp. 1–8.

**ALTYEB ALTAHER TAHA** received the Ph.D. degree in computer science from the University of Khartoum, Sudan.

From 2010 to 2012, he was a Postdoctoral Research Fellow with the National Advanced IPv6 Center, University Science Malaysia, Malaysia. From 2016 to 2017, he was the Head of the IT Department, Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia. He is currently a Professor of computer science with the Faculty of Computing and Information Technology, King Abdulaziz University. His research interests include information security, machine learning, data mining, and artificial intelligence.

**SHARAF JAMEEL MALEBARY** received the Ph.D. degree in computer science and engineering from the University of South Carolina, USA. He is very enthusiastic individual and has shown and proven strong leadership potentials. As a result, he serves as the Head of the Information Technology Department. Moreover, he was nominated to serve as the Vice-Dean for Graduate Studies and Scientific Research at the Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh Branch. He is currently an Assistant Professor with King Abdulaziz University, Rabigh. His research interests are autonomous systems, wireless communications, artificial intelligence, and machine learning.