

# Innovative credit card fraud detection: a hybrid model combining artificial neural networks and support vector machines

Oussama Ndama, Ismail Bensassi, El Mokhtar En-Naimi

Computer Science and Smart Systems Laboratory, Faculty of Science and Technology Tangier, Abdelmalek Essaâdi University, Tetouan, Morocco

## Article Info

### Article history:

Received Nov 26, 2023

Revised Jan 24, 2024

Accepted Feb 10, 2024

### Keywords:

Artificial neural networks  
Credit card fraud detection  
Hybrid models  
Support vector machines  
Synthetic minority over-sampling technique

## ABSTRACT

In recent years, escalating fraudulent activities have led to significant financial losses across industries, intensifying the critical challenge of fraud detection. This study introduces a novel hybrid model that combines artificial neural networks (ANN) with support vector machines (SVM) to construct a robust additive model for fraud detection. Emphasizing the synthetic minority over-sampling technique (SMOTE), our investigation addresses the imbalanced nature of fraud versus non-fraud transactions. The clear novelty of our research lies in the seamless integration of these two powerful tools, offering a comprehensive and effective solution to the challenges posed by credit card fraud detection. Furthermore, our study stands out by emphasizing the collaborative synergy between ANN and SVM, particularly through the integration of multiple kernels, which improves the adaptability and accuracy of the proposed hybrid model. We conducted a thorough examination of 284,807 anonymized transactions, placing special emphasis on comparing the hybrid approach's performance and showcasing its superiority over traditional methodologies in the realm of fraud detection.

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



## Corresponding Author:

Oussama Ndama

Computer Science and Smart Systems Laboratory, Faculty of Science and Technology Tangier

Abdelmalek Essaâdi University

Tetouan, Morocco

Email: oussama.ndama@gmail.com

## 1. INTRODUCTION

The realm of modern commerce has experienced an unprecedented shift towards electronic transactions, significantly reshaping global financial ecosystems. However, concomitant with the digital revolution, the prevalence of fraudulent activities, particularly within credit card transactions, has become a ubiquitous challenge. Swift and robust detection and prevention of fraudulent transactions stand as imperative elements in preserving the trust and integrity of financial systems worldwide.

In recent years, the advent of machine learning (ML) techniques has presented a promising frontier in combating the escalating threat of credit card fraud. Notably, artificial neural networks (ANN) and support vector machines (SVM) have emerged as prominent contenders, each offering distinct advantages in predictive modeling and classification tasks [1]–[5]. This research embarks on an innovative exploration, seeking to harness the collective strengths of ANN and SVM through a sophisticated hybrid approach for advanced credit card fraud detection. The amalgamation of these methodologies is poised to fortify the

detection process, enhance predictive accuracy, and discern complex patterns inherent in fraudulent transactions.

The primary aim of this study is to present a novel framework that integrates the power of ANN and SVM in a symbiotic manner, harnessing their complementary attributes. By combining the flexibility and ability to learn on their own of ANN with the structural robustness and optimal margin separation of SVM [6], [7], the proposed hybrid model aims to make credit card fraud detection systems much more accurate and efficient. Through a rigorous investigation and comparative analysis, this paper aims to demonstrate the heightened capabilities of the hybrid ANN-SVM model in identifying fraudulent behaviors within credit card transactions. Leveraging this fusion not only promises to strengthen fraud detection accuracy but also contributes to the ongoing evolution of ML methodologies in combating financial fraud.

In the upcoming sections, we will begin with a review of related works in the field, providing context and highlighting advancements made by other researchers in credit card fraud detection methodologies. Following this, we delve into our research method, outlining the experimental setup. Finally, we conduct a thorough evaluation of the hybrid model's ability to detect fraudulent patterns in credit card transactions.

## 2. RELATED WORKS

This section is dedicated to the examination of credit card fraud detection, with a pronounced focus on the utilization of ML and deep learning (DL) methodologies. Notably, our investigation accentuates the importance of hybrid methodologies, wherein SVM and ANN assume pivotal roles. In addition, we are interested in academic research that aims to improve hybrid methods by using synthetic minority over-sampling technique (SMOTE).

In their comprehensive analysis spanning from 2009 to 2019, Al-Hashedi and Magalingam [8] classified 75 articles, revealing SVM as the dominant method, constituting 23% of research. Notably, 81.33% focused on bank and insurance fraud. For credit card fraud detection, Mienye and Sun [9] proposed a DL ensemble achieving exceptional sensitivity (1.000) and specificity (0.997), surpassing traditional ML classifiers, using long short-term memory (LSTM) and gated recurrent unit (GRU) networks, complemented by a multilayer perceptron (MLP) meta-learner and synthetic minority oversampling technique with edited nearest neighbor (SMOTE-ENN). Verma and Tyagi [10] explored credit card fraud intricacies in e-commerce and online banking, favoring supervised vector classifiers and logistic regression on uneven datasets. Jayanthi *et al.* [11] introduced innovative strategies, employing cluster and classifier-based decision trees, logistic regression, and random forest methodologies, outperforming other methods. Ahmed and Saini [12] emphasized artificial intelligence and ML for fraud detection, with SVM emerging as the most reliable, achieving high accuracy. Karthik *et al.* [13] proposed an effective hybrid ensemble model for credit card fraud detection, surpassing existing methods in real-time detection and addressing data imbalance challenges. Rtayli and Enneya [14] contributed a novel hybrid technique using GridSearchCV, recursive feature elimination, and SMOTE, surpassing previous efforts in speed and efficiency. Sadgali *et al.* [15] explored cardholder behavior patterns, employing a hybrid technique with a scoring mechanism for efficient fraud detection, achieving a per-transaction processing time of 6 milliseconds. Shahapurkar and Patil [16] navigated real-time fraud detection, addressing concept drift using XGBoost as the primary model and four auxiliary algorithms, proving superior performance in accuracy, precision, and recall across diverse industries.

In summary, these studies underscore the pivotal role of SVM and DL in credit card fraud detection, with innovative methodologies such as hybrid techniques, ensemble models, and advanced algorithms contributing to improved accuracy and efficiency. The landscape reflects a significant focus on SVM, emphasizing its reliability, while DL techniques, especially those incorporating LSTM, GRU, and SMOTE, showcase promising results. The exploration of diverse strategies highlights the evolving nature of fraud detection methodologies, adapting to challenges posed by diverse datasets and real-time processing requirements.

## 3. RESEARCH METHOD

This section provides a full explanation of the architectural framework and algorithms that support our hybrid model, which has been specifically developed for credit card fraud detection. This section examines the integration of ANN and SVM in a scientific manner, emphasizing the significance of algorithmic selection, comprehensive data preprocessing, and strategic resampling methods for attaining optimal accuracy. Each phase of our fraud detection methodology has been carefully designed with the aim of enhancing the dependability and effectiveness of the whole approach.

### 3.1. Architectural framework

Our credit card fraud detection system has a hybrid design that integrates two powerful ML techniques, namely ANN and SVM. The uniqueness of our approach lies in the meticulous orchestration of a multifaceted methodology, commencing with a rigorous data preprocessing phase encompassing normalization, feature engineering, and data balancing techniques. This comprehensive approach not only upholds data integrity but also strives for an optimal representation of information.

Innovatively, our study strategically deploys ANN as the primary tool for feature extraction, leveraging its capabilities [17], [18]. The extracted features undergo optional hyperparameter adjustment before contributing to the training of the SVM classifier, which is positioned as the principal model for fraud classification. The architectural design orchestrates the harmonious integration of ANN and SVM [19]–[21], presenting a promising solution to the intricate challenges associated with credit card fraud detection.

A key differentiator is the adeptness of SVM in precise categorization, thereby enhancing ANN's capacity to discern complex patterns. This synergy culminates in the development of a robust hybrid model capable of proficiently predicting fraud and fortifying security in financial transactions. The architectural design extends its innovation by incorporating diverse kernels in SVM, coupled with hyperparameter fine-tuning [22]. This dynamic adaptation mechanism intrinsic to the SVM model amplifies its flexibility in collecting intricate patterns from diverse data sources.

The predictive prowess of our hybrid model is further amplified through the use of multiple kernels, such as linear, radial basis function (RBF), and polynomial. This iterative modification of hyperparameters [23], [24] substantiates the model's adaptability, aiming to optimize the balance between accurate categorization and the reduction of false positives and false negatives. The architectural design, as shown in Figure 1, emphasizes the seamless integration of ANN and SVM, a critical aspect of our novel hybrid model that will be discussed further in subsequent sections. This innovative amalgamation sets the stage for an advanced credit card fraud detection system, improving overall effectiveness and dependability.

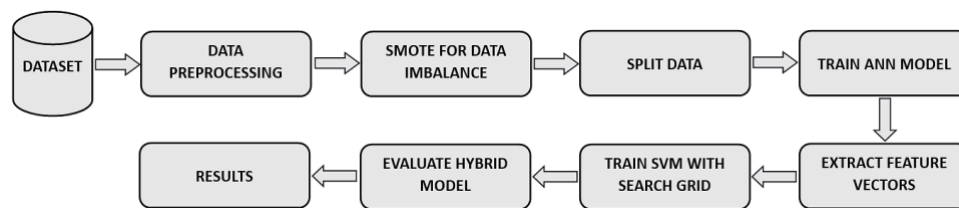


Figure 1. The architectural design of the proposed framework for credit card fraud detection

## 3.2. Algorithms

### 3.2.1. Artificial neural network

The core of our fraud detection framework is centered on the ANN, a highly advanced model that draws inspiration from the complex architecture of the neural network found in the human brain, shown in Figure 2. The ANN is composed of connected layers of nodes, where each node represents an artificial neuron. The primary function of the ANN is to analyze information and identify complex patterns within the given dataset. The ANN has outstanding capability in extracting essential information for the purpose of fraud detection due to its unique layers comprising input, hidden, and output [25]. By fine-tuning the weights of the connections between layers over and over again, the backpropagation technique helps the system learn and make predictions better. The ANN plays a crucial role in uncovering complex connections and enhancing the effectiveness of the framework [26].

### 3.2.2. Support vector machine

The SVM is a widely used discriminative model that is well acknowledged for its robustness and efficacy in performing binary classification tasks. The SVM algorithm does this by determining the ideal hyperplane that effectively divides data points into separate categories [27]. The primary advantage of it is its ability to effectively use many kernels, including linear, RBF, and polynomial kernels. The use of several kernels allows for the establishment of distinct decision limits in the feature space, which allows SVM to effectively handle intricate and nonlinear interactions present within the dataset. The linear kernel is suitable for datasets that are linearly separable, while the RBF kernel is particularly effective in capturing patterns in datasets that are nonlinear. On the other hand, the polynomial kernel is able to tolerate complexity by adjusting the degree parameter [28]. The unique operations of each kernel in the SVM contribute to its

overall resilience and flexibility in the field of fraud detection. This is achieved by allowing the SVM to adapt to varied dataset formats.

– Linear kernel:

$$k(x_i, x_j) = x_i^T x_j \quad (1)$$

– RBF kernel:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (2)$$

– Polynomial kernel:

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d \quad (3)$$

Parameters include  $\gamma$  (scale),  $r$  (coefficient), and  $d$  (degree)

Here are some mathematical formulas that show how SVM kernel functions work. These functions describe how the algorithm changes and measures the connections between data points in the feature space. SVM can easily adapt to different datasets and find complex patterns because each kernel function has its own set of properties. This makes it easier to spot fraud.

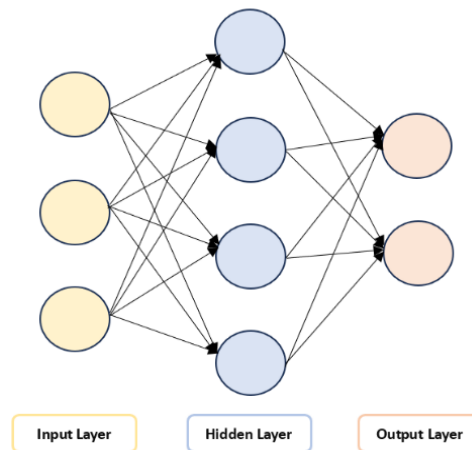


Figure 2. ANN architecture

### 3.3. Data preprocessing

The hybrid credit card fraud detection framework is built upon a strong basis that involves thorough data preparation to guarantee the dataset is properly prepared for future model training and assessment. This phase comprises fundamental procedures in the cleaning and preparation of data. The preprocessing stage is deliberately designed to convert the raw data into a form that is more suitable for analytical purposes [29]. The first step is the normalization of the 'Amount' column, which is a crucial attribute in the realm of credit card fraud detection. By using the StandardScaler capability provided by the scikit-learn library, the 'Amount' values are transformed to conform to a standardized range of [-1, +1]. This normalized feature is introduced as a new column, 'NormalizedAmount,' ensuring standardized data representation. Furthermore, in order to improve the model's capacity to detect patterns and prioritize the most significant characteristics, some columns that may have little value in detecting fraudulent actions, such as 'Amount' and 'Time', are carefully eliminated from the dataset. The maintenance of data integrity is ensured by doing a thorough analysis of the first rows in the updated dataset and then eliminating any possible duplicate entries. This process strengthens the dependability of the dataset for future analytical methods. Furthermore, the dataset is partitioned into two primary components: the feature set, represented as 'X', and the target variable, indicated as 'y', which includes the 'Class' column. The variable 'X' includes all attributes with the exception of the 'Class' column, while 'y' simply represents the 'Class' column, which is essential for identifying between fraudulent and non-fraudulent transactions. These meticulously orchestrated steps in data preparation set the cornerstone for the subsequent deployment of the hybrid framework, laying the groundwork for accurate and robust credit card fraud detection.

### 3.4. Data resampling and segmentation

In the next phase of our hybrid framework for credit card fraud detection, we address the challenge of unbalanced data by using SMOTE from the imbalanced-learn package. This methodology entails the generation of synthetic samples for the underrepresented class by producing instances along the linear trajectories linking pre-existing samples of the underrepresented class. To enhance the balance of our dataset, we are including more instances of the minority class, hence increasing the number of cases available [30], [31]. Following the use of SMOTE, the feature set ( $X$  resample) and the target variable ( $y$  resample) that have undergone resampling are converted into Pandas DataFrames. This conversion is carried out to establish a distinct separation between instances classified as fraudulent and those classified as non-fraudulent. The dataset is prepared in order to facilitate further model training and testing within our hybrid framework.

Subsequently, the resampled dataset is partitioned into training and testing subsets with the 'train\_test\_split' function provided by the scikit-learn module. The inclusion of this stage is of utmost importance in the process of assessing and verifying the efficacy of our methodology. The division is conducted in a stratified way in order to preserve the proportionate representation of class distributions inside the subsets. The comprehensive technique used ensures the dependability and inclusiveness of the subsets within the dataset, hence enhancing the resilience of our overarching hybrid model designed for the detection of credit card fraud.

### 3.5. Model architecture and evaluation highlights

This section gives you a peek into how we built and tested our hybrid model for credit card fraud detection. We kick things off by training our ANN model on carefully processed data. Picture it like a team with 16 players in the input layer, followed by two hidden layers with 24 and 1 players, respectively. This team is coached with the Adam optimizer and binary cross-entropy loss function, making sure they're sharp at catching any fishy transactions. Once our ANN team is trained, they step up to the plate to create feature vectors from the training data. These vectors then become the training material for our SVM model in the next round. We fine-tune the SVM with hyperparameters like 'C' (our regularization parameter), 'gamma' (the kernel coefficients), kernel type, and polynomial degree. It's like giving our SVM player the perfect gear for the game, making sure they're in top form. After our SVM player is all set, we put the hybrid model to the test using a separate set of data. We use the ANN model to grab feature vectors from the test set and let the optimized SVM model predict if there's anything shady going on. The evaluation comes with a bunch of performance scores: accuracy, precision, recall, and the F1 score. Plus, we throw in some confusion matrices to give you a visual on how well our model is distinguishing between fishy and non-fishy transactions.

In the end, we gather up all the results and put them into a fancy dataframe. This not only helps us see how our hybrid model did but also lets us compare its performance with other models. It's like the final scorecard after a game giving you the lowdown on how our model stacks up in the world of credit card fraud detection. We prioritize transparency in decision-making and continual model refinement to stay ahead of emerging fraud strategies and ensure consistent accuracy in fraud detection systems [32]–[38].

## 4. RESULTS AND DISCUSSION

The findings from the assessment of the model indicate the effectiveness of several hybrid configurations of ANN and SVM in identifying instances of credit card fraud. Every configuration provides valuable insights into the compromises between various hyperparameters, resulting in diverse measures of accuracy, precision, recall, and F1 scores. The following are various ANN and SVM hybrid configurations.

### 4.1. ANN-SVM with a linear kernel

The performance of the ANN-SVM model with a linear kernel is evaluated using both the test and the full datasets, as shown in Tables 1 and 2, respectively. Across varied regularization parameter values (C), ANN-SVM linear C=0.1 emerges as the top performer, exhibiting the highest accuracy, precision, and recall. Notably, on the test dataset, the model consistently achieves accuracy above 99.8% with a minimal false negative rate. While precision varies on the full dataset, emphasizing the sensitivity to model tuning, ANN-SVM Linear C=0.1 maintains a robust balance.

Table 1. Performance metrics of ANN-SVM with a linear kernel on the test dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM Linear C 0.1	0.998916	0.000387	0.999613	0.998223	0.998917
ANN-SVM Linear C 1.0	0.998456	0.000509	0.999491	0.997426	0.998458
ANN-SVM Linear C 10.0	0.998813	0.000327	0.999673	0.997957	0.998814

Table 2. Performance metrics of ANN-SVM with a linear kernel on the full dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM Linear C 0.1	0.998596	0.002114	0.997886	0.550117	0.709241
ANN-SVM Linear C 1.0	0.997856	0.004228	0.995772	0.444340	0.614481
ANN-SVM Linear C 10.0	0.998436	0.004228	0.995772	0.523333	0.686089

#### 4.2. ANN-SVM with an RBF kernel

The performance of the ANN-SVM with an RBF model in fraud detection is thoroughly analyzed in Tables 3 and 4, which provide extensive insights into its effectiveness on both the test and complete datasets. In the test dataset, the model demonstrates exceptional performance, especially at gamma = 10.0, reaching a remarkable accuracy of 99.92%, immaculate recall, and outstanding precision. When considering the whole dataset, the model continues to demonstrate its dominance, as seen by gamma = 10.0 consistently achieving high accuracy, no false positives, and perfect recall, precision, and F1 score. The findings underscore the efficacy of the ANN-SVM with an RBF model in both controlled and real-world settings, rendering it highly suitable for practical applications in credit card fraud detection. Moreover, the observed fluctuations in performance across various gamma values highlight the significance of parameter optimization, as elaborated in the full evaluation shown in Tables 3 and 4.

Table 3. Performance metrics of ANN-SVM with a RBF kernel on the test dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM RBF Gamma 0.1	0.998970	0.000206	0.999794	0.998151	0.998972
ANN-SVM RBF Gamma 1.0	0.999055	0.000097	0.999903	0.998211	0.999056
ANN-SVM RBF Gamma 10.0	0.999194	0.000109	0.999891	0.998501	0.999195

Table 4. Performance metrics of ANN-SVM with a RBF kernel on the full dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM RBF Gamma 0.1	0.998560	0.002114	0.997886	0.543779	0.703952
ANN-SVM RBF Gamma 1.0	0.998585	0.000000	1.000000	0.548088	0.708084
ANN-SVM RBF Gamma 10.0	0.998883	0.000000	1.000000	0.605634	0.754386

#### 4.3. ANN-SVM with a polynomial kernel

Tables 5 and 6 showcase the performance of the ANN-SVM model with a polynomial kernel on the test and full datasets, revealing its efficacy for credit card fraud detection. In Table 5, ANN-SVM poly degree 4 emerges as the standout performer, achieving 99.89% accuracy and demonstrating a fine balance between precision and recall. This superior performance extends to Table 6, where the model maintains robust accuracy, ranging from 99.84% to 99.86%. Particularly, ANN-SVM poly degree 4 exhibits noteworthy precision and recall, reinforcing its suitability for handling the intricacies of a larger dataset. These findings underscore the potential of the ANN-SVM model with a polynomial kernel, especially at degree 4, in enhancing the accuracy and reliability of credit card fraud detection systems.

Table 5. Performance metrics of ANN-SVM with a polynomial kernel on the test dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM Poly degree 2	0.998740	0.000944	0.999056	0.998427	0.998741
ANN-SVM Poly degree 3	0.998510	0.001187	0.998813	0.998209	0.998511
ANN-SVM poly degree 4	0.998904	0.000460	0.999540	0.998271	0.998905

Table 6. Performance metrics of ANN-SVM with a polynomial kernel on the full dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM Poly degree 2	0.998836	0.008457	0.991543	0.596692	0.745036
ANN-SVM Poly degree 3	0.998440	0.008457	0.991543	0.524022	0.685673
ANN-SVM Poly degree 4	0.998647	0.002114	0.997886	0.559242	0.716781

#### 4.4. The optimal models across different kernels

In the context of credit card fraud detection, Tables 7 and 8 provide an in-depth evaluation of three high-performing SVM kernel models. These models include ANN-SVM Linear with a regularization parameter (C) set to 0.1, ANN-SVM RBF with a gamma value of 10.0, and ANN-SVM Poly with a polynomial degree of 4. The ANN-SVM linear model demonstrated outstanding recall 99.96% in the test dataset, accompanied by a negligible rate of false negatives 0.038% and a high level of accuracy 99.89%. In

the experiment, it was observed that the ANN-SVM RBF model with a gamma value of 10.0 exhibited flawless accuracy and recall. The ANN-SVM poly 4 model exhibited a remarkable recall rate of 99.95%, a precision rate of 99.83%, and a very high accuracy rate of 99.89%.

Table 7. Performance metrics of the optimal models across different kernels on the test dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM Linear C 0.1	0.998916	0.000387	0.999613	0.998223	0.998917
ANN-SVM RBF Gamma 10.0	0.999194	0.000109	0.999891	0.998501	0.999195
ANN-SVM poly degree 4	0.998904	0.000460	0.999540	0.998271	0.998905

Table 8. Performance metrics of the optimal across from different kernels on the full dataset

Model	Accuracy	FalseNegRate	Recall	Precision	F1 Score
ANN-SVM Linear C 0.1	0.998596	0.002114	0.997886	0.550117	0.709241
ANN-SVM RBF Gamma 10.0	0.998883	0.000000	1.000000	0.605634	0.754386
ANN-SVM poly degree 4	0.998904	0.000460	0.999540	0.998271	0.998905

Upon extending the evaluation to include the whole dataset, it was seen that these models consistently showed high levels of performance in terms of recall, precision, false negative rates, and accuracy. The ANN-SVM RBF model with a Gamma value of 10.0 demonstrated exceptional performance in terms of recall, achieving a perfect score of 100% and no false negatives. This highlights the robustness of the model, with a remarkable accuracy of 99.89%. The ANN-SVM linear and ANN-SVM poly 4 models exhibited notable recall rates of 99.79%, correspondingly, along with commendable levels of accuracy, measuring tied at 99.86%. The ANN-SVM RBF model with a Gamma value of 10.0 demonstrated exceptional performance across all evaluation measures, positioning it as a very favorable option for the practical implementation of credit card fraud detection systems. The results of this study highlight the significant importance of SVM kernel selection, since these models provide superior performance across key criteria.

## 5. CONCLUSION

After studying and practically applying various hybrid models for credit card fraud detection, our research has shown the effectiveness of combining ANN and SVM. Exploring different kernels within this framework provided valuable insights into their impact on precision and recall. The results highlighted trade-offs across various kernels in the ANN-SVM model. RBF kernels with higher gamma values were adept at identifying fraudulent activities, although with a slight compromise on precision. Linear and polynomial kernels, while not performing as well as RBF, still offered flexibility in tailored detection strategies. It's crucial to acknowledge limitations that accompany the strengths of this model. One significant constraint is the sensitivity to parameter tuning, especially in the RBF kernels of SVMs. Achieving optimal performance requires meticulous tuning, which may present challenges in real-world, dynamic scenarios. Furthermore, our focus on ANN and SVM didn't explore advanced DL structures like CNNs or RNNs that could better capture complex transaction patterns. This limitation means our model may not fully adapt to evolving fraud strategies. Although we addressed the issue of imbalanced datasets, no model can entirely overcome the constantly changing nature of fraud. Despite these limitations, applying hybrid models and evaluating them with diverse metrics has provided crucial insights into the intricate landscape of credit card fraud detection. These findings pave the way for customized fraud detection models, reinforcing transaction security and offering a robust defense against fraudulent activities. Combining ANN and SVMs, particularly with different kernel configurations, allows us to build a more advanced and flexible solution for real-world fraud detection. In our upcoming work, we plan to enhance hybrid models using advanced DL structures, which could better capture complex patterns and refine fraud detection strategies. Additionally, developing real-time detection systems will enable rapid responses to emerging fraud while maintaining transparency and continuous refinement of these models.

## REFERENCES

- [1] C. Li, N. Ding, Y. Zhai, and H. Dong, "Comparative study on credit card fraud detection based on different support vector machines," *Intelligent Data Analysis*, vol. 25, no. 1, pp. 105–119, Jan. 2021, doi: 10.3233/IDA-195011.
- [2] S. C. Dubey, K. S. Mundhe, and A. A. Kadam, "Credit card fraud detection using artificial neural network and backpropagation," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, May 2020, pp. 268–273, doi: 10.1109/ICICCS48265.2020.9120957.
- [3] S. Mittal and S. Tyagi, "Performance evaluation of machine learning algorithms for credit card fraud detection," in *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, Jan. 2019, pp. 320–324, doi:




- 10.1109/CONFLUENCE.2019.8776925.
- [4] B. Gedela and P. R. Karthikeyan, "Credit card fraud detection using support vector machine algorithm in comparison with various machine learning algorithms to measure accuracy, sensitivity, specificity, precision and f-score," in *AIP Conference Proceedings*, AIP Publishing, vol. 2587, no. 1, 2023, doi: 10.1063/5.0150792.
- [5] T. K. Behera and S. Panigrahi, "Credit card fraud detection: a hybrid approach using fuzzy clustering & neural network," in *2015 Second International Conference on Advances in Computing and Communication Engineering*, IEEE, May 2015, pp. 494–499, doi: 10.1109/ICACCE.2015.33.
- [6] G. I. Diaz, A. F. -Nkoutche, G. Nannicini, and H. Samulowitz, "An effective algorithm for hyperparameter optimization of neural networks," *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 1-11, Jul. 2017, doi: 10.1147/JRD.2017.2709578.
- [7] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998, doi: 10.1023/A:1009715923555.
- [8] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Computer Science Review*, vol. 40, May 2021, doi: 10.1016/j.cosrev.2021.100402.
- [9] I. D. Mienye and Y. Sun, "A deep learning ensemble with data resampling for credit card fraud detection," *IEEE Access*, vol. 11, pp. 30628–30638, 2023, doi: 10.1109/ACCESS.2023.3262020.
- [10] P. Verma and P. Tyagi, "Analysis of supervised machine learning algorithms in the context of fraud detection," *ECS Transactions*, vol. 107, no. 1, pp. 7189–7200, Apr. 2022, doi: 10.1149/10701.7189ecst.
- [11] E. Jayanthi *et al.*, "Cybersecurity enhancement to detect credit card frauds in health care using new machine learning strategies," *Soft Computing*, vol. 27, no. 11, pp. 7555–7565, 2023, doi: 10.1007/s00500-023-07954-y.
- [12] A. N. Ahmed and R. Saini, "Detection of credit card fraudulent transactions utilizing machine learning algorithms," in *2023 2nd International Conference for Innovation in Technology (INOCON)*, 2023, pp. 1–5, doi: 10.1109/INOCON57975.2023.10101137.
- [13] V. S. S. Karthik, A. Mishra, and U. S. Reddy, "Credit card fraud detection by modelling behaviour pattern using hybrid ensemble model," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 1987–1997, 2022, doi: 10.1007/s13369-021-06147-9.
- [14] N. Rtayli and N. Enneya, "Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization," *Journal of Information Security and Applications*, vol. 55, 2020, doi: 10.1016/j.jjsa.2020.102596.
- [15] I. Sadgali, N. Sael, and F. Benabbou, "Human behavior scoring in credit card fraud detection," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 3, pp. 698–706, 2021, doi: 10.11591/IJAI.V10.I3.PP698-706.
- [16] A. Shahapurkar and R. Patil, "Concept drift and machine learning model for detecting fraudulent transactions in streaming environment," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 5, pp. 5560–5568, 2023, doi: 10.11591/ijece.v13i5.pp5560-5568.
- [17] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "A package-SFERCB-‘Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors,'" *Applied Soft Computing*, vol. 47, pp. 151–167, Oct. 2016, doi: 10.1016/j.asoc.2016.05.020.
- [18] D. A. Bashar, "Survey on evolving deep learning neural network architectures," *Journal of Artificial Intelligence and Capsule Networks*, vol. 2019, no. 2, pp. 73–82, Dec. 2019, doi: 10.36548/jaicn.2019.2.003.
- [19] K. Sabanci, M. F. Aslan, E. Ropelewska, and M. F. Unleren, "A convolutional neural network-based comparative study for pepper seed classification: Analysis of selected deep features with support vector machine," *Journal of Food Process Engineering*, vol. 45, no. 6, 2022, doi: 10.1111/jfpe.13955.
- [20] F. B. Banadkooki, M. Ehteram, F. Panahi, S. S. Sammen, F. B. Othman, and A. E. -Shafie, "Estimation of total dissolved solids (TDS) using new hybrid machine learning models," *Journal of Hydrology*, vol. 587, 2020, doi: 10.1016/j.jhydrol.2020.124989.
- [21] A. Selakov *et al.*, "A comparative analysis of SVM and ANN based hybrid model for short term load forecasting," in *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*, IEEE, May 2012, pp. 1–5, doi: 10.1109/TDC.2012.6281502.
- [22] B. Schölkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*, Massachusetts, USA: The MIT Press, 2018, doi: 10.7551/mitpress/4175.001.0001.
- [23] A. Tharwat, "Parameter investigation of support vector machine classifier with kernel functions," *Knowledge and Information Systems*, vol. 61, no. 3, pp. 1269–1302, 2019, doi: 10.1007/s10115-019-01335-4.
- [24] L. -F. Hu, W. Gong, L. -X. Qi, and P. Wang, "A method for feature selection based on the optimal hyperplane of SVM and independent analysis," in *2013 International Conference on Machine Learning and Cybernetics*, IEEE, Jul. 2013, pp. 114–117, doi: 10.1109/ICMLC.2013.6890454.
- [25] C. Mishra and D. L. Gupta, "Deep machine learning and neural networks: an overview," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 6, no. 2, pp. 66–73, 2017, doi: 10.11591/ijai.v6.i2.pp66-73.
- [26] J. Zhou *et al.*, "Graph neural networks: a review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020, doi: 10.1016/j.aiopen.2021.01.001.
- [27] Y. Wang and Y. Xu, "A non-convex robust small sphere and large margin support vector machine for imbalanced data classification," *Neural Computing and Applications*, vol. 35, no. 4, pp. 3245–3261, 2023, doi: 10.1007/s00521-022-07882-2.
- [28] L. Cao and Y. Zhai, "Imbalanced data classification based on a hybrid resampling SVM method," in *2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*, IEEE, Aug. 2015, pp. 1533–1536, doi: 10.1109/UIC-ATC-ScalCom-CBDCCom-IoP.2015.275.
- [29] C. V. G. Zelaya, "Towards explaining the effects of data preprocessing on machine learning," in *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, IEEE, Apr. 2019, pp. 2086–2090, doi: 10.1109/ICDE.2019.00245.
- [30] R. Blagus and L. Lusa, "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinformatics*, vol. 14, Mar. 2013, doi: 10.1186/1471-2105-14-106.
- [31] A. Desiani, S. Yahdin, A. Kartikasari, and Irmeilyana, "Handling the imbalanced data with missing value elimination smote in the classification of the relevance education background with graduates employment," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 2, pp. 346–354, 2021, doi: 10.11591/ijai.v10.i2.pp346-354.
- [32] S. Bhaskaran, R. Marappan, and B. Santhi, "Design and comparative analysis of new personalized recommender algorithms with specific features for large scale datasets," *Mathematics*, vol. 8, no. 7, 2020, doi: 10.3390/math8071106.
- [33] S. Bhaskaran, R. Marappan, and B. Santhi, "Design and analysis of a cluster-based intelligent hybrid recommendation system for e-learning applications," *Mathematics*, vol. 9, no. 2, pp. 1–23, 2021, doi: 10.3390/math9020197.
- [34] R. Marappan *et al.*, "Efficient evolutionary modeling in solving maximization of lifetime of wireless sensor healthcare networks," *Soft Computing*, vol. 27, no. 16, pp. 11853–11867, 2023, doi: 10.1007/s00500-023-08623-w.
- [35] S. Balakrishnan, T. Suresh, R. Marappan, R. Venkatesan, and A. Sabri, "New hybrid decentralized evolutionary approach for DIMACS challenge graph coloring & wireless network instances," *International Journal of Cognitive Computing in Engineering*,






- vol. 4, pp. 259–265, Jun. 2023, doi: 10.1016/j.ijcce.2023.07.002.
- [36] S. Bhaskaran and R. Marappan, “Enhanced personalized recommendation system for machine learning public datasets: generalized modeling, simulation, significant results and analysis,” *International Journal of Information Technology*, vol. 15, no. 3, pp. 1583–1595, 2023, doi: 10.1007/s41870-023-01165-2.
- [37] R. Marappan and G. Sethumadhavan, “Solution to graph coloring using genetic and tabu search procedures,” *Arabian Journal for Science and Engineering*, vol. 43, no. 2, pp. 525–542, 2018, doi: 10.1007/s13369-017-2686-9.
- [38] S. Bhaskaran, N. Bharathiraja, K. Pradeepa, M. V. Kumar, N. V. Ravindhar, and R. Marappan, “New recommender system for online courses using knowledge graph modeling,” in *2023 International Conference on Computer Communication and Informatics, ICCCI 2023*, IEEE, 2023, pp. 1-6, doi: 10.1109/ICCCI56745.2023.10128262.

## BIOGRAPHIES OF AUTHORS






**Oussama Ndama**    is a Ph.D. student in data science, artificial intelligence and smart systems (DSAI2S) research team, Computer Science and Smart Systems (C3S) Laboratory, Faculty of Sciences and Technologies (FST), Tangier, Morocco. He had his Master in Computer Science and Big Data, Laureate of FST of Tangier. He is also a business intelligence engineer with more than 5 years of experience in different multinational companies. The research topics of interest are smart systems, machine learning, deep learning, NLP, ANN, sentiment analysis, and smart cities. He can be contacted at email: oussama.ndama@etu.uae.ac.ma.



**Ismail Bensassi**    is a Ph.D. student in DSAI2S, C3S Laboratory, Faculty of Sciences and Technologies (FST), Tangier, Morocco. He is an engineer in computer Science, Laureate of FST of Tangier. The research topics of interest are smart connection of user profiles in a big data context, multi-agent systems (MAS), case-based reasoning (CBR), ontology, machine learning, smart cities, and eLearning/MOOC/SPOC. He can be contacted at email: bensassi.ismail@gmail.com.



**Dr. El Mokhtar En-Naimi**    is a full professor in the University of Abdelmalek Essaâdi (UAE), Faculty of Sciences and Technologies of Tangier (FSTT), Department of Computer Sciences. He was temporary professor from 1999 to 2003 and permanent professor since 2003/2004 until Now. Actually, he is a full professor in UAE, FST of Tangier. He was a Head at Department of Computer Sciences, since October 2016 until the end of December 2020. He was responsible for a Licence of Science and Technology, LST Computer Engineering (“Licence LST-GI”), from January 2012 to October 2016. He is a chief of data science, artificial intelligence and smart systems (DSAI2S) research team since the academic year 2022/2023. He is also a founding member of the both laboratories: Laboratoire d’Informatique, Systèmes et Télécommunications (LIST) Laboratory (from 2008 to 2022) and C3S Laboratory since the academic year 2022/2023 until now, the University of Abdelmalek Essaâdi, FST of Tangier, Morocco. He is also an expert evaluator with the ANEAQ, since the academic year 2016/2017 until now, that an expert of the private establishments belonging to the territory of the UAE and also an expert of the initial or fundamental formations and formations continuous at the Ministry of Higher Education, Scientific Research and Executive Training and also at the UAE University and the FST Tangier since 2012/2013 until now. He is an author/co-authors of several articles, published in the international journals in computer sciences, in particular, in multi-agent systems (MAS), CBR, artificial intelligent (AI), machine learning (ML), deep learning (DL), eLearning, MOOC/SPOC, big data, data-mining, wireless sensor network, VANet, MANet, and smart city. He was/is also director of several doctoral theses in computer sciences. He has too served as a general chair, technical program chair, technical program committee member, organizing committee member, session chair, and reviewer for many international conferences and workshops. In addition, he is an associate member of the ISCN - Institute of Complex Systems in Normandy, the University of the Havre, France, since 2009 until now. He can be contacted at email: en-naimi@uae.ac.ma.