TEMPLATE-BASED DYNAMIC TIME WARPING CREDIT CARDS' FRAUD PREDICTION MODEL

Olayinka O. Ogundile[1] – H.S. Ogunmade[1] – A. A. Owoade[1] – Oluwaseyi P. Babalola[2*] – Vipin Balyan[2]

[1]Department of Computer Science, Tai Solarin University of Education, Ogun State, Nigeria.
[2]Department of Electrical, Electronics, and Computer Engineering, Cape Peninsula University of Technology, Bellville, South Africa

ARTICLE INFO

***Article history*:**
Received: 22.04.2024.
Received in revised form: 15.03.2025.
Accepted: 25.03.2025.

Keywords:
Credit card fraud
DTW
DT
FDR
LR
RF

DOI: https://doi.org/10.30765/er.2512

Abstract:

The use of credit cards for electronic commerce purposes has been on the increase in recent time. Credit card being the most acceptable and popular mode of payment, the number of fraud cases associated with it is also on the rise. As a result of the extensive nature of credit card transaction details, it is challenging to identify fraud in a credit card system in recent years, which implies that the identification of credit cards' fraud accurately, quickly, and effectively is a gray area of research. In this article, an automated template-based credit cards' fraud prediction (CCFP) model is developed using dynamic time warping (DTW) technique. This proposed CCFP technique is novel, as it has not been previously used to predict fraud in credit cards' systems. The performance of this proposed DTW-CCFP model is verified using the dataset that contains the credit card transactions of European cardholders. In addition, the performance of this proposed DTW-CCFP model is compared with three different machine learning (ML) prediction models: logistic regression (LR), decision tree (DT) and random forest (RF). The results were documented using two different performance metrics: sensitivity and false discovery rate (FDR). The proposed DTW-CCFP model outperforms the LR, DT and RF techniques. Of interest, the proposed DTW-CCFP model achieves this feat using a small portion (less than 5%) of the dataset to build the template in comparison to the LR, DT and RF techniques, which uses between 20%-30% of the dataset for training to achieve a fairly reasonable result.

1 Introduction

Credit cards are increasingly used for several purchases due to the exponential growth of the internet and the popularity of services like e-commerce, tap, and pay systems, online bill payment systems, and digital transfers [1]. Banks and other financial institutions offer a variety of payment alternatives, including credit card transactions, internet payments, point-of-sale payments, automated teller machines (ATM), cash deposit machines (CDM), and kiosk transactions, to properly service consumers [2]. Due to the digitalization of banking activities, the banking industry generates a sizeable volume of sensitive and secret data regarding various financial transactions. For online credit card transactions, some information from the cardholders' is required, including the card verification value (CVV) number, cardholder's name, credit card number, cardholder's address, expiration date, personal identification number (PIN), security question, and so on.

Yet, this wonderful and effective method of trading comes with a high level of risk. As a result, credit card thieves are working harder than ever to steal money from transactions. Identity theft and fraud are on the rise, as more people use credit cards globally. The card's details, such as its number, expiration date, security code,

* Corresponding author
E-mail address:ogundileoo@tasued.edu.ng

and so on, are all that are needed to complete a virtual card transaction. Usually, these purchases are made on the phone or online. All that is necessary to engage in fraud on these types of purchases is knowledge of the card information. Credit card information should be kept confidential. Just two of the security procedures used to safeguard credit card transactions are tokenization and data encryption [3]. Although these techniques mainly work, they do not completely guard against fraud when using credit cards. Information about credit card privacy should not be compromised. Phishing websites lost or stolen credit cards, cloned credit cards, the theft of card information, and intercepted cards are a few examples of ways that credit card information can be taken [4].

For safety concerns, avoid engaging in the aforementioned activities. Simple card information is all that is needed for online fraud, which occurs remotely. The moment of the transaction does not call for a manual signature, PIN, or card imprint. Frequently, genuine cardholders are unaware that someone else has accessed or stolen their card information [5]. Using the data that is currently available, the simplest method to identify this kind of fraud is to look at each card's spending patterns and check for any deviations from typical spending patterns. However, a credit card system's ability to predict fraud has become increasingly difficult in recent years due to the extensive quantity of transactional information. There seems to be no workable mechanism to guarantee the security of the card while it is used by a third party or concealed in the owner's pocket. Accordingly, the precise, quick, and effective recognition of credit cards has been the subject of extensive research recently.

The use of machine learning (ML) techniques is a crucial strategy for reducing or eliminating credit card fraud. Without being particularly designed to do so, machine learning enables computers to learn from past data and enhance their forecasting skills [6, 7]. Many ML techniques such as logistic regression (LR), decision trees (DT), and random forest (RF) have previously been used to predict fraudulent credit card transactions [8-12]. However, research is still ongoing to improve the prediction ability of these ML techniques. In the same vein, this article proposes a template-based credit card fraud prediction (CCFP) model based on the principle of dynamic time warping (DTW). This proposed CCFP model builds a template based on previous non-fraudulent credit card transactions using DTW. Thus, the DTW-CCPF model matches all credit card transactions against this template while setting a threshold to determine if the transaction is fraudulent or otherwise. The performance of this proposed DTW-CCFP model is documented using the dataset that contains the credit card transactions of European cardholders [13]. Besides, this article verified the performance of the proposed DTW-CCFP model in comparison to three other ML models; LR, DT, and RF based on their prediction sensitivity and false discovery rate (FDR: prediction reliability).

The contribution and relevance of this article are as follows. This article develops a reliable and efficient DTW-CCFP model that can quickly identify fraudulent card transactions. To the best of the authors' knowledge, this proposed DTW-CCFP model has not been previously used to predict fraud in credit card systems. Likewise, the article reviews the performances of three different ML tools that have been used in the literature for the prediction of credit card fraud, while also comparing their performances with the proposed DTW-CCFP model. Comparing this proposed work's results to the aforementioned techniques, it is innovative and produces good sensitivity while keeping a low FDR. Therefore, financial institutions can reliably decide on the ML tool to deploy based on their requirement. In addition, it is envisaged that this article will save customers and financial institutions from losing money daily and boost customers' confidence in online banking that requires credit card details.

The remainder of this article is structured as follows. Section 2 presents a review on the related works on credit card fraud. In Section 3, the ML models used for performance comparison in this article are briefly discussed. Section 4 explains the proposed template-based DTW-CCFP model in detail. Firstly, this section explains the principle of the DTW techniques before elaborating on how it is employed to predict credit card fraud. Additionally, the Section explains the proposed DTW-CCFP model with a numerical example. The results and discussion are presented in Section 5. Lastly, the article is concluded in Section 6 with discernible comments.

2 Related Works

In this section, the related works on credit card fraud prediction are discussed. [7] developed a genetic algorithm (GA) based feature selection combined with artificial neural network (ANN), DT, Naive Bayes (NB), LR, and RF classifiers to detect fraudulent credit card transactions. They proved that their work outperforms the traditional approach without the GA-optimised feature selection. Nevertheless, their results

portray the GA-RF algorithm as the most effective tool in comparison to the GA-ANN, GA-DT, GA-NB, and GA-LR algorithms using various performance metrics such as the operating characteristic curve (AUC), accuracy, recall, precision, and F1-score (F-measure).

The author in [14] proposed an optimized support vector machine (SVM) employing the cuckoo search algorithm. The CS-SVM algorithm is applied to a European cardholder's dataset to identify fraudulent credit card transactions. They verified the performance of their proposed model with traditional classifiers such as SVM, LR, RF DT, and NB using various performance metrics, which include AUC, accuracy, recall, precision, and F1-score. Their documented results showcased the proposed CS-SVM as the most efficient approach in terms of these metrics; yet it exhibited more computational complexity in comparison to these traditional classifiers.

A framework for credit card fraud detection that combines the potentials of meta-learning ensemble approaches with a cost-sensitive learning paradigm was proposed in [15]. The proposed framework took the approach of letting base classifiers fit conventionally while incorporating cost-sensitive learning into the ensemble learning process to fit the cost-sensitive meta-classifier without requiring cost-sensitive learning to be imposed on each of the base classifiers. The prediction accuracy of this trained cost-sensitive meta-classifier and base classifiers such as multilayer perceptron (MLP), KNN, and DT were demonstrated in terms of the AUC. The classification of unknown data yielded results that the cost-sensitive ensemble classifier retains a superb AUC value, demonstrating constant performance across various fraud rates in the dataset. Accordingly, these findings show that, in comparison to the results of conventional ensemble classifiers, the cost-sensitive ensemble framework is effective in producing cost-sensitive ensemble classifiers that are capable of accurately detecting fraudulent transactions in various databases of payment systems regardless of the proportion of fraud cases.

The author in [9] investigated the performance of three ML tools; RF, SVM, LR, and DT on a credit card fraud dataset. Their results were documented in terms of accuracy, specificity, sensitivity, and precision. It was observed from their work that the RF tool offered the best performance, closely followed by the LR tool, while the DT tool exhibited the worst performance.

In [10], a performance comparison of two RF classifiers is carried out on a credit card transaction dataset obtained from an e-commerce company in China. The two types of RF classifiers are employed to model the behavioural features of fraudulent and non-fraudulent transactions. The authors experimentally demonstrated that their modified RF classifier outperforms the traditional RF approach. However, while their proposed modified RF classifier produced good results on small datasets, there are still some issues, such as skewed data, that limit its effectiveness.

The performance of NB, KNN, and LR data mining tools was investigated on a highly imbalanced credit card dataset in [16]. The authors carried out a hybrid method of under-sampling and oversampling on the imbalanced dataset. The performance of these data mining techniques was documented in terms of precision, sensitivity, specificity, balanced classification rate, and the Matthews correlation coefficient (MCC). Their comparative results show that the KNN algorithm offered the best performance while the LR algorithm exhibited the worst performance in terms of the aforementioned performance metrics.

In this article, credit card fraud is predicted using a different approach to conventional methods, such as DT, LR, RF, and extreme gradient boost (XGBoost). The article proposed the DTW-CCFP model as a more efficient and reliable approach in comparison to these conventional methods. Aside from offering superior sensitivity and FDR performances, the DTW-CCFP model exhibits a low computational complexity and can be deployed for real-time credit card systems applications.

3 Prediction Techniques

3.1 Logistic Regression

Logistic regression is an ML classification algorithm used to predict the probability of certain classes based on some dependent variables. LR is a statistical and data mining technique that statisticians and researchers use to analyze and classify binary response datasets [17]. The most common application of logistic regression is to handle classification problems, with the outcome being binary (yes or no). In the real world, logistic regression is used in a variety of domains and fields, including health care, finance, and marketing, making it appropriate for this application. Mathematically, in LR, a linear function, w defined by Equation 1 is fed into a logit function to estimate the probability P that decides the prediction of a particular class [7,18].

$$w = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_u X_u, \tag{1}$$

$$P = \frac{1}{1 + e^{-w}}. \tag{2}$$

Note that the value of P must be between 0 and 1. The closer the value of P to 1, the more reliably the LR algorithm predicts the class.

3.2 Decision Tree (DT)

Decision tree is a supervised learning approach that can be used to solve classification and regression issues. However, DT is commonly employed to solve classification problems. DT is a tree-structured classifier, it has internal nodes that reflect dataset attributes, branches that represent decision rules, and each leaf node that represents the final decision [19]. The algorithm begins at the core node. The branch is the point where the decision is reached, which splits the tree while the leaf node denotes the results. DT is utilized in a wide range of applications, including information extraction, machine learning, categorization in scientific research, and biomedical ones [20]. The goal of employing a decision tree is to develop a training model that can be used to forecast the category or value of the target variable by studying straightforward decision rules deduced from previous data.

3.3 Random Forest (RF)

Random forest is a type of supervised ML algorithm that is commonly used in classification and regression problems. It constructs decision trees from several samples and uses their majority vote for classification and average for regression [21]. One of the most important features of the RF algorithm is that it can handle the dataset containing continuous variables as in the case of regression and categorical variables as in the case of classification. The RF algorithm is composed of different decision trees, each with the same nodes, but uses different data that leads to different leaves. It merges the decisions of multiple decision trees to find an answer, which represents the average of all these decision trees [21]. Hence, the working principle of the RF algorithm can be summarized in these five steps [20]:

1. Given a training data set, select i number of data points randomly.
2. Create the decision trees correlated to the selected i points.
3. Determine j number of decision trees to be created.
4. Repeat steps 1 and 2.
5. From each decision tree, identify the prediction for fresh data points, and then group them in the category of the highest votes.

4 Methodology: Proposed DTW Approach

4.1 Dataset

In this article, secondary dataset containing credit card fraud information was collected from an online repository provided by Kaggle.com. The dataset contains credit card transactions made by European cardholders in September 2013. This dataset contains 492 fraud cases out of 284,807 transactions in two days; that is, 0.172% of the card transactions are fraudulent. There are 31 numerical features (31 columns) in each transaction, 28 ($\mathcal{V}_1 - \mathcal{V}_{28}$), which are transformed using principal component analysis (PCA) due to privacy concerns [7,22]. The remaining 3 columns are time (shows the difference in time between the dataset's first transaction and every subsequent one), amount, and class of transaction ('0' indicating a non-fraudulent transaction and '1' indicating a fraudulent transaction). A summary of some key statistical properties of the dataset are presented in Table 1 [23].

Table 1: Statistical properties (S_p) of the dataset.

S_p	$Time$	V_1	V_2	V_3	V_4	V_5
Mean	94814	0.0000	0.0000	0.0000	0.0000	0.0000
Median	84692	0.0181	0.0655	0.1799	-0.0199	-0.5434
Highest	172792	2.4549	22.0577	9.3826	16.8753	34.8017
Lowest	0	-56.4075	-72.7157	-48.3256	-5.6832	-113.7433
1st Quartile	54202	-0.9204	-5.5986	-0.8904	-0.8486	-0.6916
3rd Quartile	13932	1.3156	0.8037	1.0272	0.7433	0.6119
S_p	V_6	V_7	V_8	V_9	V_{10}	V_{11}
Mean	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Median	-0.2742	0.0401	0.0224	-0.0514	-0.0929	-0.0328
Highest	73.3016	120.5895	20.0072	15.5950	23.7451	12.0189
Lowest	-26.1605	-43.5572	-73.2167	-13.4341	-24.5883	-4.7975
1st Quartile	-0.7683	-0.5541	-0.2086	-0.5354	-0.6431	-0.7625
3rd Quartile	0.3986	0.5704	0.32735	0.5971	0.45392	0.7396
S_p	V_{12}	V_{13}	V_{14}	V_{15}	V_{16}	V_{17}
Mean	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Median	0.1400	-0.0136	0.0506	0.0481	0.0664	0.6568
Highest	7.8484	7.1268	10.5268	8.8777	17.3151	9.2535
Lowest	-12.6837	-5.7919	-19.2143	-4.4989	-14.1295	-25.1628
1st Quartile	-0.4056	-0.6485	-0.4256	-0.5829	-0.4680	-0.4838
3rd Quartile	0.6182	0.6625	0.4931	0.6488	0.5233	0.3997
S_p	V_{18}	V_{19}	V_{20}	V_{21}	V_{22}	V_{23}
Mean	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Median	-0.0036	0.0037	-0.0625	-0.0295	0.0068	-0.0112
Highest	5.0411	5.5920	39.4209	27.2028	10.5030	22.5284
Lowest	-9.4987	-7.2135	-54.4977	-34.8304	-10.9331	-44.8077
1st Quartile	-0.4989	-0.4563	-0.2117	-0.2284	-0.5424	-0.1619
3rd Quartile	0.5008	0.4589	0.1330	0.1864	0.5286	0.1476
S_p	V_{24}	V_{25}	V_{26}	V_{27}	V_{28}	$Amount$
Mean	0.0000	0.0000	0.0000	0.0000	0.0000	88.3500
Median	0.0410	0.0166	-0.0521	0.0013	0.0112	22.0000
Highest	4.5846	7.5196	3.5173	31.6122	33.8478	25691.1600
Lowest	-2.8366	-10.2954	-2.6046	-22.5657	-15.4301	0.0000
1st Quartile	-0.3546	-0.3172	-0.3270	-0.0708	-0.0530	5.6000
3rd Quartile	0.4395	0.3507	0.2410	0.0910	0.0783	77.1700
S_p	$Class$					
Mean	0.0017					
Median	0.0000					
Highest	1.0000					
Lowest	0.0000					
1st Quartile	0.0000					
3rd Quartile	0.0000					

4.2 Dynamic Time Warping (DTW)

DTW which was first introduced in [24], is frequently used in manufacturing, data mining, gesture recognition, medical, speech recognition, animal sound recognition, and so on [25,26]. Meanwhile, to the best of our knowledge, DTW has not been deployed in credit card fraud prediction as proposed in this article. DTW algorithm seeks the optimal alignment between two-time sequences using the temporal distortions between these sequences. The time sequences are warped effectively in a non-linear way to fit one another. For instance,

consider the two-time sequences $\mathcal{C}_1$ and $\mathcal{C}_2$ of lengths ϕ and ρ respectively, the two-time sequences are aligned by the DTW algorithm to form a $\phi \times \rho$ difference matrix is defined as [24]:

$$D = min\begin{pmatrix} D[\phi-1,\rho-1] \\ D[\phi-1,\rho] \\ D[\phi,\rho-1] \end{pmatrix} + \left|\mathcal{C}_{1_\phi} - \mathcal{C}_{2_\rho}\right|, \tag{3}$$

where the similarities between the two times sequences $\mathcal{C}_1$ and $\mathcal{C}_2$ at positions ϕ and ρ, respectively, are represented by each element in D. In this article, the difference between the two times sequences is defined as $D_{\phi^{th},\rho^{th}}$. This indicates the value of element at the ϕ^{th} and ρ^{th} position of D. Refer to [24, 25, 27] for more information on the principle of DTW.

4.3 Proposed Template-based DTW-CCFP Model

In developing the prediction model, some of the non-fraudulent transactions are manually identified from the credit card holder transaction dataset, $\mathcal{D}$ ($\mathcal{D}$ contains only the PCA transformed 28 columns ($\mathcal{V}_1 - \mathcal{V}_{28}$), where the other three columns have been removed) and used to derive the template. As mentioned earlier, any transaction in the dataset with class variable '0' is non-fraudulent. Therefore, the chosen non-fraudulent transactions are formulated in a matrix form given as:

$$\mathcal{F} = \begin{bmatrix} \mathcal{V}_{1,1} & \mathcal{V}_{1,2} & \dots & \mathcal{V}_{1,k} \\ \mathcal{V}_{2,1} & \mathcal{V}_{2,2} & \dots & \mathcal{V}_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{V}_{\psi,1} & \mathcal{V}_{\psi,2} & \dots & \mathcal{V}_{\psi,k} \end{bmatrix}, \tag{4}$$

where k is the last element in each transaction; in this case, k is constant, k=28. The number of transactions in the template is denoted as ψ. The value of ψ is varied in this article to be 10, 20, and 30. This means the sizes of the formulated matrix, $\mathcal{F}$ of non-fraudulent transactions $\psi \times k$ are 10×28, 20×28 and 30×28. Note that the chosen values of ψ are less than 5% ($< 5\%$) of the entire dataset.

Subsequently, the DTW algorithm is employed in the model to compute the similarities between the selected non-fraudulent credit card transactions. Therefore, the formulated matrix, $\mathcal{F}$ with ψ several manually identified non-fraudulent transactions are warped with each other using Equation 3 to generate a $\psi \times \psi$ dissimilarities template matrix defined as:

$$\mathcal{T} = \begin{bmatrix} 0 & \mathcal{T}_{1,2} & \dots & \mathcal{T}_{1,\psi} \\ \mathcal{T}_{2,1} & 0 & \dots & \mathcal{T}_{2,\psi} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{T}_{\psi,1} & \mathcal{T}_{\psi,2} & \dots & 0 \end{bmatrix}, \tag{5}$$

where each element in $\mathcal{T}$ represents the $D_{\phi^{th},\rho^{th}}$ similarities between the transactions in the formulated matrix $\mathcal{F}$. Next, the model computes the maximum value of each column of the developed template $\mathcal{T}$ to generate a $1 \times \psi$ row vector is given as:

$$\mathcal{T}_{max} = [\mathcal{T}_{max,1} \quad \mathcal{T}_{max,2} \quad \mathcal{T}_{max,3} \quad \dots \quad \mathcal{T}_{max,\psi}]. \tag{6}$$

Then, the prediction process begins by computing the similarities between each transaction in the dataset, $\mathcal{D}$ and all the ψ non-fraudulent transactions in the formulated matrix $\mathcal{F}$ to obtain a $1 \times \psi$ row vector is given as:

$$\mathcal{Y}_n = [\mathcal{Y}_{n_1} \quad \mathcal{Y}_{n_2} \quad \mathcal{Y}_{n_3} \quad \dots \quad \mathcal{Y}_{n_\psi}], \qquad n = 1,2,\dots,N, \tag{7}$$

where N is total number of credit card transactions in the dataset, N=284,807. Similarly, each element in $\mathcal{Y}_n$ represents the $D_{\phi^{th},\rho^{th}}$ similarities between each transaction in the dataset and the transactions in formulated matrix $\mathcal{F}$. Afterwards, the model computes the match value, χ by matching $\mathcal{T}_{max}$ and $\mathcal{Y}_n$. The matching value, χ, is calculated by counting the instances in which each column value of $\mathcal{Y}_n$ is less than or equal ($\leq$) to the corresponding column value of $\mathcal{T}_{max}$. Accordingly, χ is compared with an empirically determined reliability value, λ, which spans between $0-1$. If $\chi \geq \lfloor \lambda * \psi \rfloor$ the input is taken to be a non-fraudulent transaction, μ; otherwise, it is taken as a fraudulent transaction, ν. This predetermined value, λ specifies the performance of this proposed DTW-CCFP model. That is, the smaller the value of λ, the higher the sensitivity performance of the DTW-CCFP model; but, at the expense of increased FDR.

4.4 DTW-CCFP: Numerical Example

Given that 10 non-fraudulent transactions were manually identified from the credit card cardholder's transaction dataset, to form a 10×28 matrix $\mathcal{F}$, defined by:

$$\mathcal{F} = \begin{pmatrix} -1.35981 & -0.07278 & 2.536347 & \dots & -0.02105 \\ 1.191857 & 0.266151 & 0.16648 & \dots & 0.014724 \\ -1.35835 & -1.34016 & 1.773209 & \dots & -0.05975 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -0.89429 & 0.286157 & -0.11319 & \dots & 0.142404 \\ -0.33826 & 1.119593 & 1.044367 & \dots & 0.083076 \end{pmatrix}. \tag{8}$$

Afterwards, the columns of $\mathcal{F}$ are warped with one another to generate a 10×10 dissimilarities template matrix $\mathcal{T}$, defined as:

$$\mathcal{T} = \begin{bmatrix} 0 & 11.9227 & 18.5188 & \dots & 9.4573 \\ 11.9227 & 0 & 22.2071 & \dots & 8.5707 \\ 18.5188 & 22.2071 & 0 & \dots & 19.3182 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 10.7496 & 12.0118 & 20.6518 & \dots & 13.3191 \\ 9.4573 & 8.5707 & 19.3182 & \dots & 0 \end{bmatrix} \tag{9}$$

The maximum value of each column of the developed template $\mathcal{T}$ is then computed to generate a 1×10 row vector $\mathcal{T}_{max}$, given as:

$$\mathcal{T}_{max} = [20.0452 \quad 22.2071 \quad 22.2071 \quad \dots \quad 19.3182]. \tag{10}$$

Therefore, to predict if a transaction is fraudulent or not, the proposed model computes the similarities between that transaction and each transaction in $\mathcal{F}$. Thus, it selects the maximum from each warping (that is, the output of the similarities computation) to form a 1×10 row vector given as:

$$\mathcal{Y}_1 = [9.7806 \quad 12.3320 \quad 18.1407 \quad \dots \quad 11.6820]. \tag{11}$$

Afterwards, the proposed model computes the matching value of χ by matching $\mathcal{T}_{max}$ (Eqn. 10) and $\mathcal{Y}_n$ (Eqn. 11). This proposed template-based DTW-CCFP model is summarised in Algorithm 1.

Algorithm 1: DTW-CCFP Model.
Input: $\mathcal{D}$, ψ, λ
Output: μ *or* ν
1: *Formulate the* $\psi \times k$ *non-fraudulent matrix* $\mathcal{F}$
2: *Compute the* $\psi \times \psi$ $\mathcal{T}$ *matrix from Eqn. (3)*
3: *Find* $\mathcal{T}_{max}$
4: *Compute* $\mathcal{Y}_n$ *from Equation 3*
5: *Match* $\mathcal{T}_{max}$ *in Eqn. (6) and* $\mathcal{Y}_n$ *in Eqn. (7) to obtain* χ
6: ***if*** $\chi \geq \lfloor \lambda * \psi \rfloor$
return μ
7: ***else***
return ν

5 Results and discussion

5.1 Performance Metric

In this article, the performance of the developed DTW-CCFP model is demonstrated based on two standard metrics: sensitivity, $\mathcal{S}$ and false discovery rate (FDR) [26, 28].

1. Sensitivity, ($\mathcal{S}$): The $\mathcal{S}$ measures the prediction accuracy of the proposed DTW-CCFP model. That is, the capability of automated predictor to correctly differentiate between non-fraudulent and fraudulent transactions. Mathematically, it is given as:

$$\mathcal{S} = \frac{TP}{TP + FN}, \qquad (12)$$

where TP is called the true positives, which indicates the number of credit card transactions that were accurately predicted. The false negative (FN) indicates the number of times the output of the automated predictor does not match the manually identified credit card transactions. A high value of the sensitivity is preferable as it determines the performance of the proposed model.

2. False Discovery Rate, (FDR): The FDR measures the reliability of the proposed DTW-CCFP model. Mathematically, it is given as:

$$FDR = \frac{FP}{TP + FP}, \qquad (13)$$

where FP is called the false positive, which indicates the number of incorrect non-fraudulent credit card transaction prediction. The smaller the value of the FDR, the more reliable the prediction model.

5.2 DTW-CCFP Performance as a Function of ψ *and* λ

Table 2. DTW-CCFP Performance as a Function of ψ and λ

ψ	$\mathcal{S}$ (%)	FDR (%)	$\mathcal{S}$ (%)	FDR (%)	$\mathcal{S}$ (%)	FDR (%)	$\mathcal{S}$ (%)	FDR (%)
	$\lambda=\frac{5}{6}$		$\lambda=\frac{4}{6}$		$\lambda=\frac{3}{6}$		$\lambda=\frac{2}{6}$	
10	91.11	1.33	92.03	2.02	92.22	2.99	93.01	3.33
20	91.63	1.21	92.37	1.76	92.68	2.79	93.55	3.01
30	92.18	0.99	93.88	1.20	94.23	1.88	95.77	2.61

Table 2 depicts the performance of the proposed DTW-CCFP model as a function of ψ for various empirically selected value of λ. From the table, firstly observe that the performance of the proposed DTW-CCFP improves as the template size, ψ increases from 10 to 30. For instance, the $\mathcal{S}$ performance improves by 0.54% as the template size, ψ increases from 10 to 20 for λ=$\frac{2}{6}$. A further increase in ψ (from 20 to 30) at the same λ improves the performance of $\mathcal{S}$ by 2.22%. On the other hand, the FDR value reduces by 0.32% and 0.4% as ψ increases from 10 to 20 and 20 to 30 for λ=$\frac{2}{6}$ respectively, which signifies a performance improvement. Note that however small, this performance improvement in $\mathcal{S}$ and FDR cuts across all values of λ. Therefore, it can be concluded that the performance of the DTW-CCFP improves with an increase in the template size, ψ. Nonetheless, it is emphasised that as the value of ψ continues to increase, there will be a saturation point where there is no further significant increase in the performance of $\mathcal{S}$ and FDR; rather, it will only increase the computational complexity of the model.

In addition, observe from the table as the performance of the $\mathcal{S}$ improves with a reduction in the value of λ while the performance of FDR worsens. As mentioned earlier, the smaller the value of λ, the higher the sensitivity performance of the DTW-CCFP model but at the expense of an increased FDR. This is logical mathematically because the FDR determines the reliability of the model's outputs; thus, as λ goes closer to zero, the value of the FDR diminishes. For example, for $\psi = 30$, the $\mathcal{S}$ performance improves by 3.59% as λ reduces from $\frac{5}{6}$ to $\frac{2}{6}$ while the FDR performance plummets by 1.62%. Therefore, λ should be selected carefully, as this indicates that there is a trade-off in the $\mathcal{S}$ and FDR performances when choosing the value of λ.

To buttress, Figs. 1 and 2 depict the sensitivity and FDR performances of the DTW-CCFP model respectively, as a function of ψ and λ. As shown in Fig. 1, the sensitivity performance of the DTW-CCFP model improves as λ grows. Moreover, increasing the template size, ψ, improves the sensitivity performance of the DTW-CCFP model. On the other hand, as shown in Fig. 2, the FDR performance of the DTW-CCFP model decreases with increasing λ. Yet, the FDR performance improves with an increase in the template size, ψ.

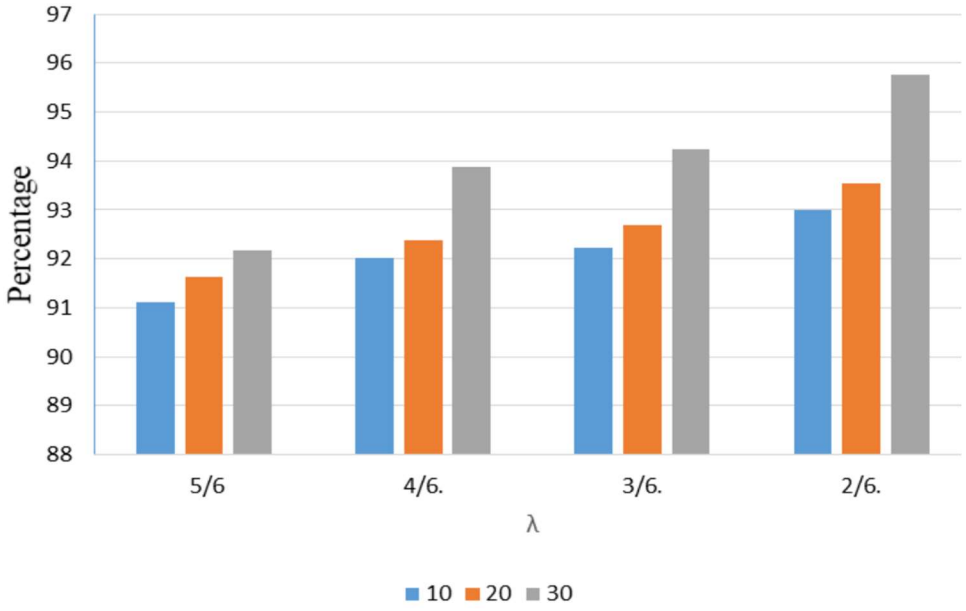

Figure 1: A graphical representation of DTW-CCFP model sensitivity performance as a function of ψ and λ.

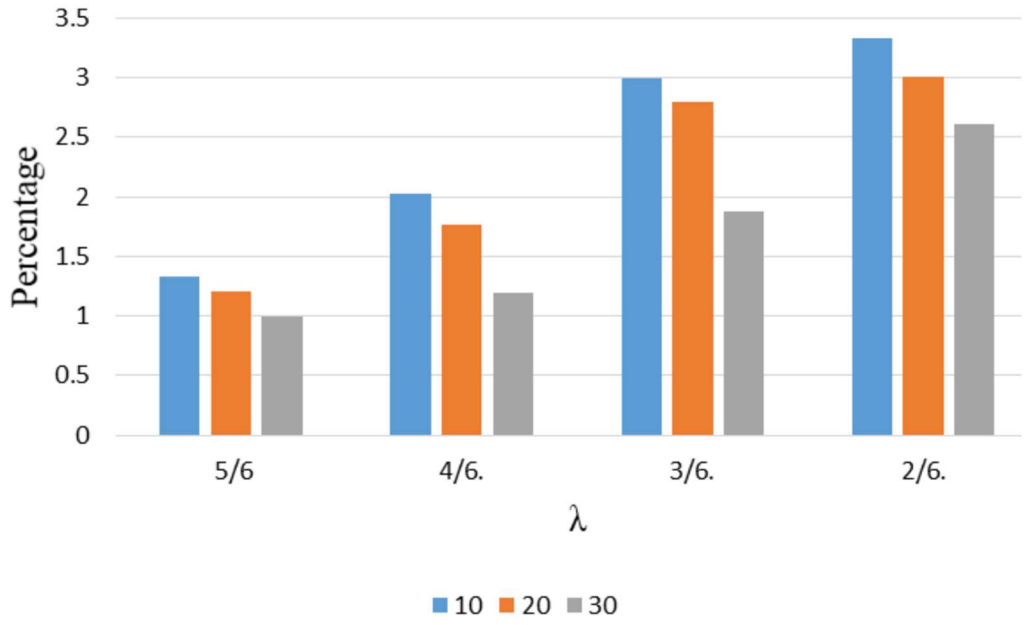

Figure 2: A graphical representation of DTW-CCFP model FDR performance as a function of ψ and λ.

5.3 DTW-CCFP Performance with Other ML Models

In this section, the performance of different conventional ML models is firstly demonstrated as a function of the training size, T_s as shown in Table 3. Here, the dataset is divided into two portions: small and large portion. The small portion is used to train the models, and the large portion is used to test the model. The small portion used to train the models is referred to in this article as the training size, T_s. The T_s is selected to be between 10%-30% to allow a relative fair comparison with the DTW-CCFP model that uses between ψ=10-30 non-fraudulent transactions for training. This implies that training to testing ratios used for these ML models are 1:9, 2:8 and 3:7. From Table 3, notice as the performances of the $\mathcal{S}$ and FDR improve with increase in T_s. Consider the case of the RF model, its $\mathcal{S}$ performance improves by 0.52% as T_s increases from 10% to 30%. Also, there is a gain of 1.35% in FDR performance as T_s increases from 10% to 30%. These performance gains are visible with the other ML models as presented in Table 3. Nevertheless, the RF model offered the best performance in comparison to the LR and DT models.

Table 3: Performance comparison of different ML models as a function of T_s

T_s(%)	$\mathcal{S}$ (%)	FDR (%)	$\mathcal{S}$ (%)	FDR (%)	$\mathcal{S}$ (%)	FDR (%)
	RF		*LR*		*DT*	
10	89.39	5.31	86.18	10.39	86.16	10.42
20	89.52	5.04	88.24	10.29	88.22	10.15
30	89.91	3.96	88.66	9.57	88.88	9.66

Table 4 presents a performance comparison of the proposed DTW-CCFP model with the ML models presented in Table 3. In the table, the performance of proposed DTW-CCFP is given for $\psi = 30$ and λ=$\frac{2}{6}$ while a training size of $T_s = 30\%$ is assumed for the other ML models. Note the $\mathcal{S}$ and FDR performance gain achieved by the proposed DTW-CCFP in comparison with the other ML models. To buttress, the DTW-CCFP exhibits a $\mathcal{S}$ performance gain of 5.86% in comparison to the RF model, which offered the best performance as compared to the LR and DT models. Likewise, it achieves a FDR performance gain of 1.35% in comparison to the RF model. Another key advantage of the proposed DTW-CCFP model is the size of the training set. While these conventional ML models use 30% (which amounts to about 85,442 credit card transactions) of the dataset to train the model to achieve reasonable $\mathcal{S}$ and FDR results, the proposed DTW-CCFP only require

30 non-fraudulent transactions (less than 5%) to build the template used to validate any credit card transaction. This implies that the DTW-CCFP is less computationally intensive in comparison to these other ML models. Accordingly, it is established in this article that this proposed DTW-CCFP model is a performance-efficient alternative to these ML models for credit card fraud prediction.

Table 4. Performance comparison of the DTW-CCFP and other ML models: $T_s = 30\%$, $\psi = 30$, $\lambda = \frac{2}{6}$

Prediction	$\mathcal{S}$ (%)	FDR (%)
DTW-CCFP	95.77	2.61
RF	89.91	3.96
LR	88.66	9.57
DT	88.88	9.66

For emphasis, Figs. 3 and 4 illustrates the sensitivity and FDR performances of the DTW-CCFP model in comparison to the other ML models. It can be clearly seen that the DTW-CCFP model offered the best performance as compared to the RF, LR and DT models. Accordingly, the developed DTW-CCFP model is a performance-efficient alternative to these ML models for the prediction of credit card frauds.

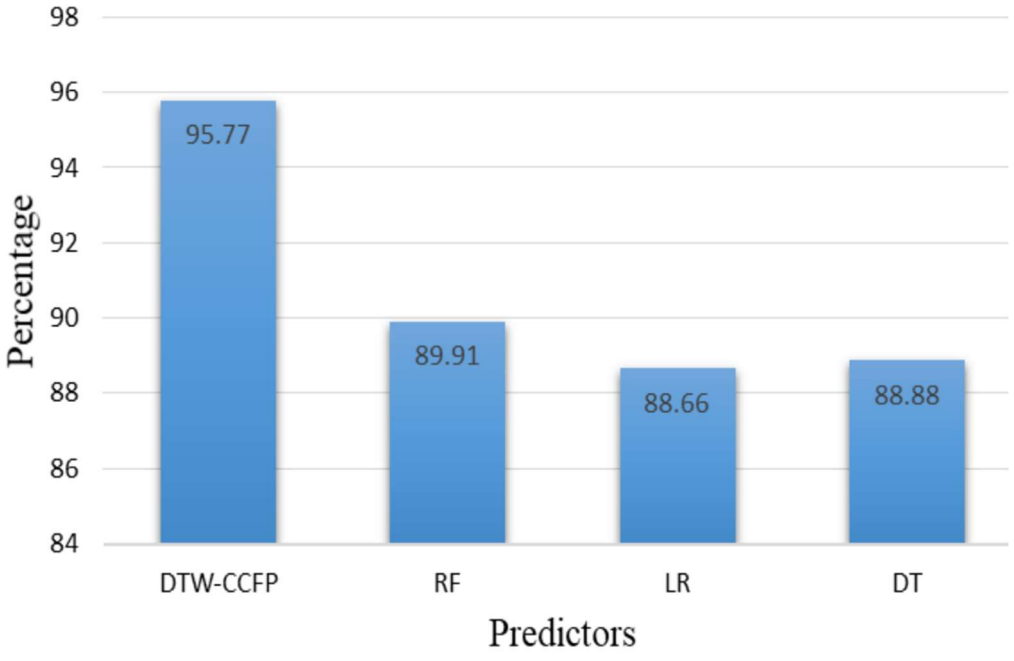

Figure 3: A graphical representation showing the sensitivity performance comparison of DTW-CCFP model and other ML models

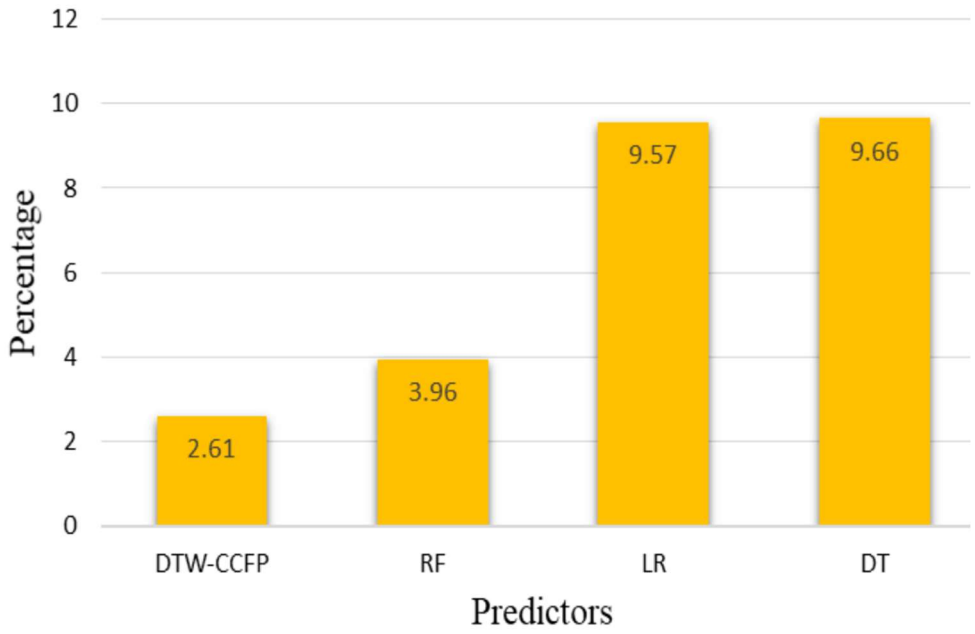

Figure 4: A graphical representation showing the FDR performance comparison of DTW-CCFP model and other ML models

6 Conclusion

In this article, a novel credit card fraud prediction model called DTW-CCFP was proposed based on the principles of dynamic time warping. This proposed model utilized a few non-fraudulent credit card transactions to develop a template that was used to screen credit card transactions. The DTW-CCFP model was demonstrated to exhibit a high sensitivity at a reduced false discovery rate. Besides, the article showcased that the proposed DTW-CCFP model offered good performance in comparison to other ML models such as RF, LR, and DT. More so, the proposed DTW-CCFP model can be deployed in real-time to validate credit card transactions because it offers reasonable computational time complexity; even, in comparison to these other ML models. Future research could compare the DTW-CCFP model with other ML and deep learning models, highlighting its benefits and drawbacks.

References

[1] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques," *International Journal of System Assurance Engineering and Management,* vol. 8, pp. 937–953, 2017.

[2] A. Roy, J. Sun, R. Mahoney, L. Alozi, S. Adam, and P. Beling, "Deep learning detecting fraud in credit card transactions," in 2018 Systems and Information Engineering Design Symposium (SIEDS), 2018, pp. 129–134.

[3] G. B. Iwasokun, T. G. Omomule, and R. O. Akinyede, "Encryption and tokenization-based system for credit card information security," *International Journal of Cyber Security and Digital Forensics*, vol. 7, no. 3, pp. 283–293, 2018.

[4] G. Rushin, C. Stancil, S. Sun, S. Adam, and P. Beling, "Horse race analysis in credit card fraud—deep learning, logistic regression, and Gradient Boosted Tree," in 2017 Systems and Information Engineering Design Symposium (SIEDS), 2017, pp. 117–121.

[5] M. L. Bhasin, "The role of technology in combating bank frauds: perspectives and prospects," *Ecoforum Journal*, vol. 5, no. 2, pp. 200–212, 2016.

[6] A. Burkov, The hundred-page machine learning book. Quebec City, QC, Canada, 2019.

[7] R. B. Sulaiman, V. Schetinin, and P. Sant, "Review of machine learning approach on credit card fraud detection," *Human-Centric Intelligent Systems*, vol. 2, pp. 55–68, 2022.

[8] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," *J Big Data*, vol. 9, no. 24, pp. 1-17, 2022.
[9] N. Khare and S. Y. Sait, "Credit card fraud detection using machine learning models and collating machine learning models," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 20, pp. 825–837, 2018.
[10] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random forest for credit card fraud detection," in 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, China, 2018, pp. 1–6.
[11] Y. Lucas, P.-E Portier, L. Laporte, L. He-Guelton, O. Caelen, M. Granitzer, and S. Calabretto, "Towards Automated Feature Engineering for Credit Card Fraud Detection Using Multi-Perspective HMMs," *Future generation computer systems,* vol. 102, pp. 393-402, 2020.
[12] W. N. Robinson and A. Aria, "Sequential Fraud Detection for Prepaid Cards Using Hidden Markov Model Divergence," *Expert Syst. Appl.,* vol. 91, pp. 235-251, 2018.
[13] https://www.kaggle.com/mlg-ulb/creditcardfraud.
[14] C. Li, N. Ding, H. Dong, and Y. Zhai, "Application of credit card fraud detection based on CS-SVM," Int J Mach Learn Comput, vol. 11, no. 1, pp. 34-39, 2021.
[15] T. A. Olowookere and O. S. Adewale, "A framework for detecting credit card fraud with cost-sensitive meta-learning ensemble approach," *Scientific Africa*, vol. 8, no. 2020, pp. 1–15, 2020.
[16] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques," in 2017 International Conference on Computing Networking and Informatics (ICCNI), Lagos, Nigeria, 2017, pp. 1-9.
[17] P. Kulkarni and R. Ade, "Logistic regression learning model for handling concept drift with unbalanced data in credit card fraud detection system," in Proceedings of the Second International Conference on Computer and Communication Technologies: IC3T, 2016, pp. 681–689.
[18] A. Robles-Velasco, P. Corte´s, J. Mun˜uzuri, and L. Onieva, "Prediction of pipe failures in water supply networks using logistic regression and support vector classification," Reliability Engineering & System Safety, vol. 196, p. 106754, 2020.
[19] P. Save, P. Tiwarekar, N. Ketan, and N. Mahyavanshi, "A novel idea for credit card fraud detection using decision tree," *International Journal of Computer Applications*, vol. 161, pp. 6–9, 2017.
[20] O. O. Ogundile, A. A. Owoade, and P. B. Emeka, "Animals' classification: A review of different machine learning classifiers," *Journal of Science and Logics in ICT Research*, vol. 9, no. 1, pp. 106–114, 2023.
[21] M. M. Ghiasi and S. Zendehboudi, "Application of decision tree-based ensemble learning in the classification of breast cancer," *Computer in Biology and Medicine*, vol. 128, p. 104089, 2020.
[22] V. N. Dornadula and S. Geetha, "Credit Card Fraud Detection using Machine Learning Algorithms," Proc Comput Sci., vol. 165, pp. 631–641, 2019.
[23] O. O. Ogundile, O. P. Babalola, A. S. Ogunbanwo, O. M. Ogundile and V. Balyan, "Credit card fraud: Analysis of feature extraction techniques for ensemble hidden Markov model prediction approach," *MDPI, appliced science,* vol. 14 7389, pp. 1-18, 2024.
[24] F. Itakura, "Minimum prediction residual principle applied to speech recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing,* vol. 23 no. 1, pp. 67-72, 1975.
[25] E. J. Keogh and M. J. Pazzani, "Derivative dynamic time warping," 2001, pp. 1–11. [Online]. Available: IEEE Xplore. [Accessed: June 06, 2023].
[26] O. O. Ogundile and D. J. J. Versfeld, "Analysis of template-based detection algorithms for inshore Bryde's whale short pulse calls," *IEEE Access*, vol. 8, pp. 14377-14385, 2019.
[27] C. Myers, L. Rabiner, and A. Rosenberg, "Performance tradeoffs in dynamic time warping algorithms for isolated word recognition," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 28, no. 6, pp. 623–635, 1981.
[28] O. O. Ogundile, A. M. Usman, and O. P. Babalola, "Dynamic mode decomposition: A feature extraction technique based hidden Markov model for detection of Mysticetes' vocalisations," *Ecological Informatics*, vol. 63, p. 101306, 2021.