International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Scoring of Borrowers Solvability by SVM and MLP hybridized to GA: Evidence from Banking Dataset

Azzeddine Idhmad[1*,] Mohammed Kaicer[2], JihaneAlami Chentoufi[3], Zie Mamadou Coulibaly[4]

Submitted: 16/11/2022 **Revised**: 13/01/2023 **Accepted**: 11/02/2023

Abstracts: In this paper, we treat the problem of credit default risk or risk of non-repayment in banks using credit scoring models. As the methods currently used have some gaps in predicting the solvency of loan applicants, which could cause in losses for the banks, our contribution is to propose a new credit scoring method based on Machine Learning algorithms. We adopt two strategies: first, we hybridize Genetic Algorithms (GA) with Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) to evaluate impact of GA on prediction performance; and second, we test SVM and MLP with their hyperparameters, and then hybridize MLP with Artificial Neural Network (ANN). To compare our method with the methods proposed in [28], we realized simulations using Python on our banking dataset. The generated results show that the hybridization with GA yields less significant results compared to the strategy of SVM, MLP with their hyperparameters, and MLP-ANN that generate the improved values of AUC, Accuracy, confusion matrix and F1-Score compared to [28]. Furthermore, even for our database the same metrics is also significant with best values.

Keywords: *Credit scoring, Machine Learning, Optimization, Solvability, Banking*

1. Introduction

The credit market represents a fundamental lever for the world economy with trillions of US dollars outstanding. This is why the Bank system's stakeholders attach great importance to ensuring that this market is optimal and meets the requirements of the markets. Among the major challenges of the sector is the prediction of creditworthiness or credit risk of borrowers, be they individuals, companies, government agencies, countries or even banks.

Whenever a financial institution such as a bank, a credit card company, or any other type of business that grants credit, it has to transmit information to the rating agencies about whether repayments are made on time. The rating agencies collect information on the borrower's profile and the length of time to repay the credit.

This information constitutes the so-called credit history. Processing of the history generates a score that makes it possible to decide whether to grant or reject the client's credit application. This is credit scoring, which is one of the tools used by banks or credit agencies to analyse the risk of default or non-repayment of clients [33] and to respect credit contracts.

Credit scoring is a decision support engineering based on statistical models as several works have addressed credit scoring models[30].However, the current economic and financial environment, characterised by chaos, requires the application of robust methods to properly adjust decisions;In many fields such as medical diagnosis, predictive justice, facial recognition, fraud detection, job search or education, data science is becoming necessary to improve decisions; thus artificial intelligence (AI) and Machine Learning (ML) techniques are particularly useful in decision support and customer knowledge, [31].

There are many machine learning methods see [32], but in general one can distinguish two categories: supervised classification methods, which are designed to partition the predictor space to predict events; the most commonly used in the credit risk domain is the support vector machine SVM. The most commonly used in the field of credit risk is the support vector machine (SVM), a neural network. The second method is the ensemble method, which uses rules to combine predictions from a set of several base models to arrive at the final classification. In our paper, we rely on the first category which is the use of SVM to predict credit scores and hyper parameters to optimise the Score.

In this paper we present the different credit scoring techniques used in the literature, then we explain the material and the methodology adopted by describing the problem, the data and the proposed ML models for the

[1]*Laboratory of Analysis,Geometry and Applications Faculty of sciences Ibn TofailUniversityKenitra , Morocco*
Email: azzeddine.idhmad@uit.ac.ma
ORCID ID : 0000-0001-7389-0970
[2] *Laboratory of Analysis,Geometry and Applications Faculty of sciences Ibn Tofail University Kenitra , Morocco*
Email: mohammed.kaicer@uit.ac.ma
ORCID ID : 0000-0003-1782-6872
[3]*Laboratory of research in Informatics Faculty of sciences Ibn Tofail University kenitra, morroco Email: j.alami@uit.ac.ma*
ORCID ID :0000-0002-7167-7620
[4]*Agriculture and food data management Unilasalle polytechnic institute Rouen, France*
Email: zie-mamadou.coulibaly@etu.unilasalle.fr
ORCID ID : 0000-0003-1174-0250
** Corresponding Author Email:azzeddine.idhmad@uit.ac*

simple and hybrid case and defining the associated metrics; then we move on to the simulation to validate the results and finally we conclude and discuss the results.

2. Literature Review

With the exponential growth of the banking and financial sectors, the credit industry has been thriving and its influence is far-reaching. A credit card is a convenient financial tool that has become an integral part of our daily lives. It is an easier alternative payment tool instead of cash [26]. Thusly, it became the number one accepted method of payment. it also offers safety which many consumers appreciate since they don't have to carry much money if they has a credit card Although all the aforementioned benefits are only a scratch on the surface of the tremendous advantages of using credit cards, the credit card market faces a very serious problem that is, credit card risk [9]. And since credit-cards have become such a popular trend in paying consumer expenses, after each purchase the amount the consumer charges are supplementary to the total, commonly referred as the balance in the credit card [4]. The money owed to the bank needs to be returned monthly or in instalments. In this case, when credit service providers offer credit cards to consumers, there is a risk that a customer may not repay the money to the bank [5]. This major risk may propose huge losses to banks and consequently to the financial sector [20]. Consistent payments every month on a credit card are a solid way to develop a credit history and high credit scoring, which enables the banking sector to categorize credit card applicants as having either good or negative credit history [16]. This method allows credit card service providers to mitigate risk since approving applicants with a bad credit history will retrieve great loss for the bank. The scoring of credit is a statistical tool to forecast the probability of a loan applicant, incumbent borrower, or adversary defaulting or becoming overdue. It estimates the likelihood of defaults or delinquency, which is commonly employed for commercial credit cards [11].

The primary benefit of a credit score is that it enables Credit Service Providers to assess an applicant's potential for a credit card in a promptly, reliable, and effective manner [3]. This is why bankers and researchers view the credit scoring system as a vital method in determining the creditworthiness of a credit card applicant. To properly manage and reduce the risks associated with information such as mobile phone data, data from social media, Coarse Transactional Data, Bank transactional data, and so on are now essential elements of credit risk assessment. Banks and credit card companies evaluate prospective borrowers using an underwriting procedure that mainly relies on their credit histories and credit ratings. Furthermore, data are critical in the construction, surveillance, and ongoing upkeep of models for credit scoring that will be the focus of the current investigation.

When issuing credit, financial organisations must use credit scoring algorithms to identify defaulters and non-defaulters [19]. In general credit scoring models are categorized into two main groups: Traditional credit scoring techniques, Artificial Intelligence and Machine Learning Credit Scoring determine which is responsible for credit card admission decisions. Linear Discriminant Analysis, K-nearest Neighbour Classifier, and Support Vector Machines (SVM) are some of these scoring methods. To begin, Linear Discriminant Analysis is, as the name implies, a linear framework for categorization and reduction of dimensionality.

Most typically used in pattern classification tasks for feature extraction. Linear discriminant analysis was the initial statistical tool used to explain which companies went bankrupt using ratios from accounting and other financial data. Sir Ronald Fisher devised the initial discriminant analysis linear model in 1936 (Fisher, R. A., 1936). However, this model was first introduced into credit scoring by Reichert, Cho, and Wagner in 1983. That being said, the other statistical credit scoring model is the K-nearest Neighbor Classifier. The k-nearest neighbor classifier serves as an example of the non-parametric statistical approach. This technique assesses the similarities between the pattern identified in the training set and the input pattern. One chooses a metric on the space of applicants and takes the k-nearest neighbor of the input pattern that is nearest in some metric sense. A new applicant will be classified in the class to which the majority of the neighbors belong (in the case when the costs of misclassification are equal) or according to the rule expressed by the equation.

Chatterjee and Barcun were the first to adopt this strategy in 1970 [6]. The non-parametric structure of the method allows for the modelling of anomalies in the possibility function over the space of features, which is clearly one of its merits. The k-NN technique has been shown to outperform other non-parametric methods, and because it is a very intuitive model, it can be readily described and communicated to the managers who will adopt it [8]. Finally, Vapnik Vladimir [27] was the first to introduce (SVM) Support Vector Machines in the setting of statistical learning theory. SVM has numerous successful applications in pattern recognition, suggesting that it is a competitive classifier. SVM is a relatively new methodology in data mining that is a new tool for solving machine learning using optimization methods and is a machine-learning algorithm based on statistical learning theory. The widespread use of credit cards prompted banks to employ scoring. As the number of people applying for credit cards grows on a daily basis, banks must limit the risk associated with this process [28].

3. Material and Methodology

The procedure followed in our approach to propose the best credit scoring model is shown in the following figure 1:

Figure 1 describes the model of our classification system used in this study.

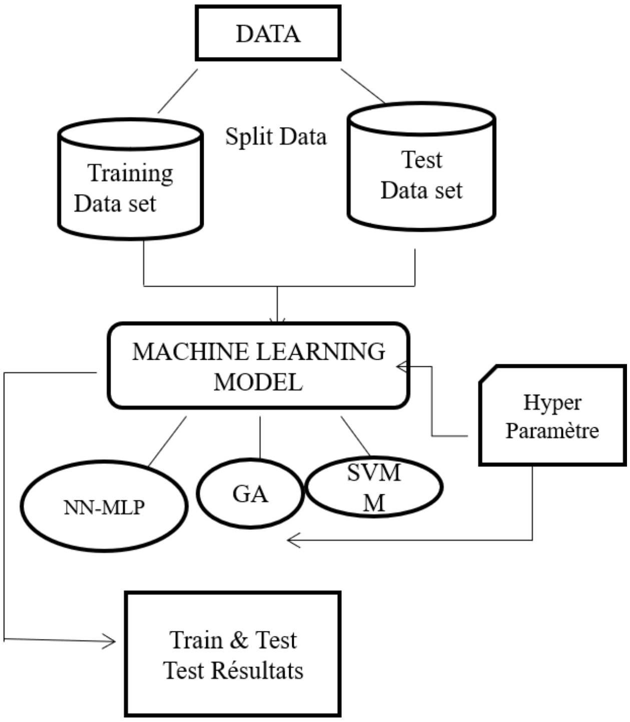

Fig 1. The hybrid SVM method process

The purpose of this research is to propose the best credit scoring model that can best expect if credit- card holders will pay their owed credit balance or not. A good credit scoring model, without a doubt, will be able to properly categorize clients into either default or non-default categories, allowing banks or credit card service providers to save money based on how the model is efficient. And in light of the current issue, proposing an effective model and calculating unpaid risks will allow banks to take the necessary course of action needed in time.

Our proposed method scoring is realised using the modelling and simulation of artificial intelligence (AI) and machine learning (ML) techniques such as the Support Vector Machine (SVM), Multi-layer Perceptron Neural Network (MLP), a form of ANN, and Genetic techniques (GA). The simulation of all the algorithms is through Python that has some libraries of the AI. Therefore, in order to validate and calibrate our methods, we have tested and processed a Banking Dataset.

3.1. Basic concepts of Support Vector Machine

In the process of proposing the most efficient credit scoring method to lower the risk of financial loss that may face banks and credit card service providers due to default in paying back credit card balances, many models were proposed to solve this major issue. One of them is the Support Vector Machines classifier which was introduced by [25], illustrating the concept of SVM and how it is applied, as a two-class model is the main focus of this research. SVM is a supervised learning method that analyses data, recognizes patterns, and is employed for statistical classification and regression analysis.

SVM selects the best hyperplane with the biggest margin as the decision margin among the two classes. The largest margin approach involves selecting the line with the highest margin. Assume you've been handed a dataset of pairings (x_i, y_i), where i=1,2,3,…,n where $x_i \epsilon R^n$ and $y_i \in \{-1, +1\}$, the y_i is used to show which class the point x_i belongs. The value of xiwill help us determine the maximum margin distance between the hyperplanes that segregate the position having di = 1 from those having d_i = −1.

In SVM, the decision boundary that differentiates the two classes is often written as:

$$w \cdot x - b = 0 \qquad (1)$$

$$w \cdot x - b = 1$$

In this case "." expresses the vector and dot product "w" is a normal vector which is perpendicular to the hyperplane. To maximize the margin, the parameter b / ‖w‖is used where the short distance measured perpendicularly and it is possible to determine the normal vector "n" from the origin, by choosing "n" and "b" the margin is maximized. To achieve optimal separating hyperplane, the value of ‖w‖has to be minimized satisfying the constraint:

$$y_i(w.x_i - b) \geq 1 \text{ for } i = \{1,2,\dots,n\} \qquad (2)$$

However, when executing the optimization process by the above constraint the square root calculation is required for determining the norm, and the computation of the square root may sometimes prove difficult. For this reason, the optimization equation has been altered to minimize 1/2 $w^T w$

$$y_i(w.x_i - b) \geq 1 \text{ for } i = \{1,2,\dots,n\} \qquad (3)$$

The above optimization problem is a quadratic one, the saddle point, also known as the minimax point, needs to be found. In this way the previous optimization equation can also be written as:

$$\min_{w,b} \max_{\propto} \{\frac{1}{2} w^T w - \sum_{i=1}^{n} \propto_i (y_i(w.x_i - b) - 1)\} \qquad (4)$$

The SVM optimization problem can also be solved with Lagrange Multipliers. This technique can be used to transform the above-constrained optimization problem into a formulation whose solution is equivalent to the above. Considering αi is the Lagrange multiplier, consequently αi≥ 0. In this scenario, a linear amalgamation of the training vectors is the solution to optimization and is expressed as:

$$w = \sum_{i=1}^{n} \propto_i x_i y_i \qquad (5)$$

This will help ease the computation process, we can further apply a differentiation concerning "w" and "b", and also present the Karush–Kuhn–Tucker (KKT) condition [27]. As a result, of this differentiation, the above expression can be changed into the dual Lagrangian LD (α):

$$\max_{\propto} L_D(\propto) = \sum_{i=1}^{n} \propto_i - \frac{1}{2}\sum_{i.j=1}^{n} \propto_i \propto_j \, y_i y_j < X_i . X_j > \quad (6)$$

Subject to: $\alpha i \geq 0$ for i = 1, 2,…, n and $\sum_{i=1}^{n} \propto_i \, y_i = 0$

When a concern can be converted to another problem whose answer is simpler to calculate and also provides the solution to the initial challenge, both challenges are said to exhibit Duality, and vice versa. In the case of dual optimisation issues, the solution provides the parameters w* and b* of the ideal hyperplane. And after training in dual form, our classification rule for the optimal hyperplane decision function becomes:

$$f(x, \propto^*, b^*) = \mathrm{sgn}(\sum_{i=1}^{n} y_i \propto_i^* < x_i . x_j > + b^*) \quad (7)$$

Nonetheless, only few Lagrange multipliers are usually positive, and these vectors are close to the ideal hyperplane, and the training vectors for these positive Lagrange multipliers serve as support vectors. In non-separable circumstances, however, some level of mistake must be permitted. As a result, for the goal of minimising training errors, the challenge of determining the plane of motion for this inspired hyperplane is shown as:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i \quad (8)$$

Subject to: $y_i(<w.x_i> +b) + \xi_i - 1 \geq 0 \; and \xi_i \geq 0$

Where ξ_i is the positive slack variable, and the variable C is the penalty strength, which specifies how much we care about errors (training points that are on the wrong side). Accordingly, the optimization problem in the above expression is solved using Lagrangian method. Thus the corresponding optimization expression is written as the following:

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i.j=1}^{n} \alpha_i \alpha_j w_i w_j < x_i . x_j > \quad (9)$$

Subject to: $0 \leq \alpha_i \leq C \; for \; \iota = 1,2, \ldots, n \; and \; \sum_{i=1}^{n} \propto_i \, d_i = 0$

While the defined penalty parameter C is supposedly the upper bound of αi, kernel functions transform non-linear spaces into linear spaces. It transforms data into another dimension so that the data can be classified. This happens by using the mapping function ϕ, and the inner product has to be replaced by the kernel function [20] as the following expression:

$$\left(\phi(x_i) . \phi(x_j)\right) = k(x_i, x_j) \quad (10)$$

And this equation will be further altered into:

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad (11)$$

With the limited condition to: $0 \leq \alpha \leq C \; for \; i = 1,2, \ldots, n \; and \; \sum_{i=1}^{n} \alpha_i \, d_i = 0$

But on the other hand, in linear generalized case, the decision function is as illustrated below:

$$f(x, \propto^*, b^*) = sgn(\sum_{i=1}^{n} y_i \propto_i^* \, k(x_i . x_j) + \upsilon \quad (12)$$

3.2. Multi-layer Perceptron Neural Network

Artificial neural networks (ANNs) are a class of artificial intelligence algorithms that are concerned with addressing the different aspects or elements of learning. One of the most useful and effective types is certainly the multi-layer perceptron model. It has been used as the benchmark model by many researchers [21]-[22].. Adding to that, MLP is a class of feedforward artificial neural networks that are considered global approximators and can be trained to implement any given nonlinear input-output mapping [15]-[16]-[19]. Furthermore, as the name implies, MLP is made up of many layers: one input layer that receives the signal, an output layer which generates a decision or forecast about the input, and an arbitrary number of hidden layers that serve as the MLP's true computational engine. [12].

Furthermore, the multilayer Perceptron is utilised to generalise a nonlinear function f for function modelling with one anticipated variable f: X ∈ RD → Y ∈ R1 which is expressed by the following equation : $f(X) = b_2 + W_2 \times (f\mathrm{A}(b_1 + W_1 \times X))$. In this case, W1 and W2 are the weight matrix of the hidden and output layers, respectively. Whereas $b_1 = [b_{11}, b_{12}, \cdots, b_{1N}]$ expresses the bias vector of the hidden layer, and b2 is a bias vector of the output layer. And finally, fA represents an activation function. In the above equation W_1 and W_2 (the weight metrics), as well as b_1 and b_2 (the bias vectors) are randomly initialized and then updated through a training process. That is, the weights are initialised very close to zero, but at random. As a result, symmetry is broken, and each neuron is no longer doing the same computation. The value of each hidden node is then determined by adding the sum of the multiplications of the values of the input neurons and the corresponding connection weights. Ultimately, the output value is transmitted forward from the hidden node and connection weight values in the same manner. By evaluating the output response data to the target values, the discrepancy can be determined and minimised in the backward stage by modifying the connection weights. Simply stated, input-output pairs are fed into the network, and weights are modified to minimize the difference between the network output and the real value.

3.3. Genetic algorithms

Genetic Algorithms are a type of quantitative and combinational circuitry optimizer that is particularly useful for tackling complicated linear and non-convex issues through natural development. Genetic Algorithms were developed by John Holland from the University of Michigan and his collaborators in the 1960s and 1970s. He was potentially the first to use inheritance, mutation, selection, and crossover in the study of adaptive and artificial systems.

3.4. Basic concepts of ROC curve

The ROC curve is a two-dimensional in-nature evaluation of classification performance [8–13]. It can be thought of as a plot of the likelihood of properly classifying positive cases versus the rate of wrongly identifying real negative examples. This curve can thus be interpreted as an evaluation of the classifier throughout the entire spectrum of class distribution and error costs. A decision rule is typically carried out by determining a threshold that distinguishes between the positive and negative classifications. One of the most intriguing aspects of the ROC curve is that even when the error costs and class distributions are unidentified, the performance of classifiers can still be characterised and optimised. The diagonal line represents the ROC curve of the classifier that forecasts the class at random, and the performance improves as the curve approaches the upper left corner of the plot. The measured value of the area under the curve, typically abbreviated as AUC, is the most generally used performance measure retrieved from the ROC curve. If the threshold is correctly determined, the classifier achieves perfect accuracy when AUC is equal to 1, while a classifier that assumes the category at random has an associated AUC of 0.5. Another intriguing aspect of the AUC is that it represents the classifier's overall behaviour because it is distinct of the threshold used to get a class label. In the continuous situation, processing the AUC would necessitate the computation of an integral; however, in the discrete case, this area may be computed with step functions, and the following property holds

$$AUC = \frac{\sum_{i=1}^{n^+}\sum_{j=1}^{n^-} 1_{f(x_i^+)>f(x_j^-)}}{mn} \qquad (13)$$

where f(•) denotes the scoring function (the decision-making function of a classifier is the most common scoring function in machine learning), x+ and x indicate the positive and negative samples, n+ and n are the total quantity of positive and negative examples, and 1 is defined to be 1 if the predicate holds and 0 otherwise. The Wilcoxon-Mann-Whitney statistic has been identified as the above equation. This equation defines whether the classifier f(x) is satisfying the condition: $f(x_i^+) > f(x_j^-), \forall i = 1,\cdots n^+, \forall j = 1,\cdots n^-$ to achieve a maximum AUC. Any negative sample that is rated greater than the positive sample reduces the AUC. Other commonly used indicators in the machine learning algorithm are linked to the ROC curve via a classifier's confusion matrix. These metrics are determined by the true positive (tp) and false positive (fp) rates, as well as the skew parameter c, which is the ratio of negative to positive cases.

4. Empirical Analysis

4.1 Crédit data sets

In this work, we used real data sets for analysis. The French datasets from a banking institution that detected a notable increase in risk over several months. Proven by the statisticians, after the first score was submitted, the bank found that the performance of the score had deteriorated following the latest modifications in credit regulations which completely changed the profile of the population. It has become crucial to create a new score. To create an updated and more efficient score, we put forward a detailed statistical analysis of the existing database, and it showed that it contains 19670 lines (customers), 50 variables, and certainly there are missing values for certain variables. For instance, MONTANT DEM has 129 missing values and SECTEUR_EMP with 28 missing variables. To manage and mitigate these missing values, we decided to replace them with the mean of the variable (for quantitative variables), or by the mode of the statistical series of this variable (for qualitative variables).

Testing models are based on splitting our database into two parts: one training component 14753 attribute (training dataset) 75% of the total data. The other data to do the test component 5007 attribute (test dataset) of a percentage of 25%. We can see later that the examination of this dataset yielded some intriguing results. Table 1 shows the features of the French real dataset.

Table 1. Characteristics of the real dataset

Data set	**Number of classes**	**Number of instances**	**Test Data 25%**	**Training Data 75%**
French	2	19670	5007	14753

Three strategies were used in this study, namely ‘‘SVM’’, ‘‘NN-MLP,’’ and ‘‘GA.’’ Table 2 summarizes the outcomes for the data sets. The accuracy of classification among the three models for the French data set was 90%, 88.11%, and 50%, respectively. For each "treatment," the experiments for the three techniques used similar training and assessment sets. We used a confusion matrix and ROC curve to evaluate the accuracy of classification of the test set (table 2 and images below).

Table 2. Result of credit scoring models (our banking dataset)

Credit scoring model	Credit scoring results (%)			
	Accuracy test (25%)	Accuracy training (75%)	F1-score	AUC
SVM	90	89.32	91	*0.958*
NN-MLP	88.11	87.75	89	*87.7*
SVM + hyperparameter	95		95	*95.88*
NN-MLP + hyperparameter	89	87.76	89	*95*
GA + hyperparameter + NN-MLP	50	48.99	53	*0.60*
GA + hyperparameter + SVM	50	49.83	67	*0.70*

In Figures 3 & 4, we have the confusion matrix in the left corner containing true positives, these are the client who was correctly identified by the algorithms as the client for whom we'll pay back the credit. True negative is in the bottom left corner these are the client that did not pay back their credit and were correctly identified by the algorithms. False negatives are in the bottom left-hand corner containing the client that paid back their credit and the algorithm wrongly identified and finally, the False positives are the client that didn't pay their credit and the algorithm wrongly identified them as they did pay back the credit.

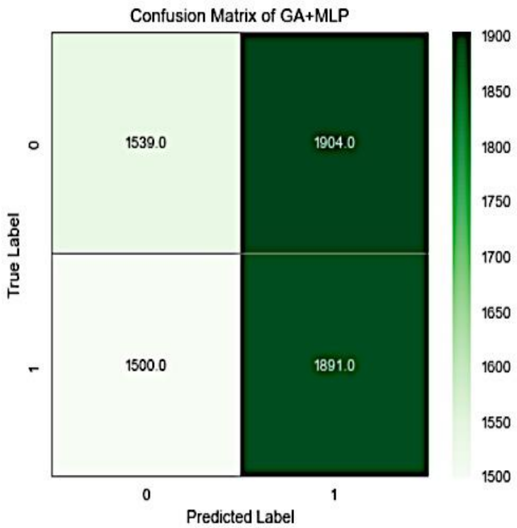

Fig 2. Confusion matrix Of GA+MLP

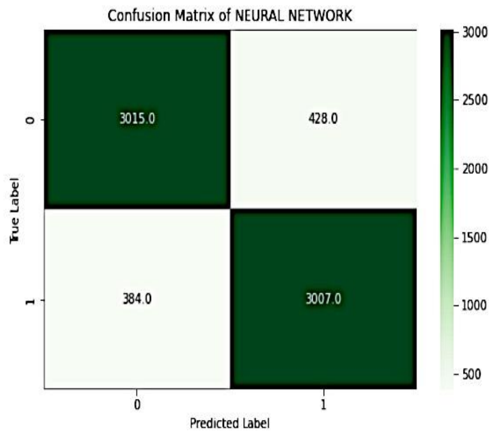

Fig 3. *Confusion matrix of neural network*

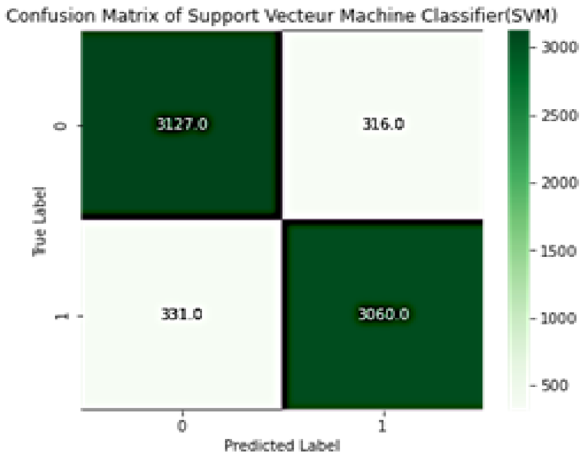

Fig. 4 Confusion matrix of SVM

source / data		German dataset			Australian dataset			Our data Banking dataset		
		AUC	accuracy	F1-score	AUC	accuracy	F1-score	AUC	accuracy	F1-score
Zhang, R. and Qiu, Z., 2020[28]	**SVM**	0.776	0.7550	0.77	0.922	0.8897	0.87	---	----	----
	KNN	0.7624	0.7420	0.751	0.902	0.8435	0.8391	-----	---	----
Our methods	**SVM**	---	---	---	---	---	---	**0.958**	**0.90**	**0.95**
	SVM-GA	---	---	---	---	---	---	0.70	0.50	0.67
	MLP	---	---	---	---	---	---	**0.95003**	0.89	0.888
	MLP-GA	---	---	---	---	---	---	0.60	0.50	0.53
	MLP-ANN	---	---	---	---	---	---	**0.95**	0.881	0.89

Table3. Comparing results of our methods with other different dataset

Our results represent in figures 4 and 3 when we applied the Support Vector Machine classifier (SVM) to the testing data we got 3127.0 as true positive and 3060.0 as True negative however the algorithm miss classified 331.0 as False Negative and the algorithm misclassified 316.0 as False Positive, from these states we can say that NN is worse than the SVM at predicting client that paid their credit back (3015.0 vs 3127.0) And worse at predicting client that did not pay their credit (3007.0 vs 3060.0). So, the SVM is much better, with an accuracy of 90 %, which is close to the results shown by [5] with an accuracy of 81.25%.

Other research conducted by [23] found that the SVM methodology yield a significantly better result 70.00% accuracy.

Our results show in figures 5 & 6 the accuracy of "SVM + hyperparameter" and "NN-MLP + hyperparameter", which allows us to make a comparison between them.

In this case, we use the curve ROC and AUC to determine which of the two models is better (table 2and table 3). we can see that the value of AUC in "SVM + hyperparameter "is 95.88% which is higher than "NN-MLP + hyperparameter" is 95 % so we can say that SVM + hyperparameter is much better at predicting the credit score.

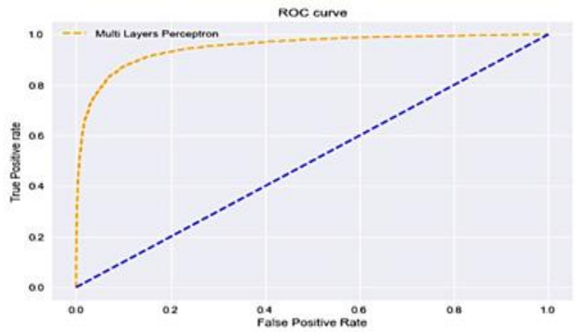

Fig 5. ROC curve of NN-MLP

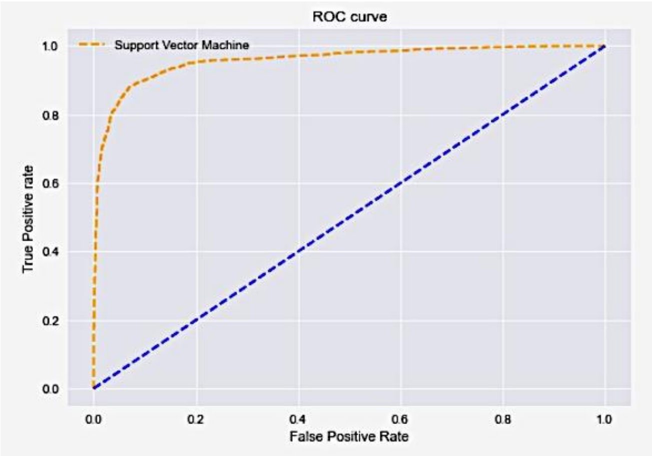

Fig 6. ROC curve of SVM

According to figures 7 & 8, we can see that the support vector machine is much less precise than the neural network, as you can see in (figure 7). The SVM is 87 % while the neural network is 88%, but when we add the hyperparameter to SVM and the NN the SVM increases to 89,5% while the NN remain the same at 88 % (Figure 8).

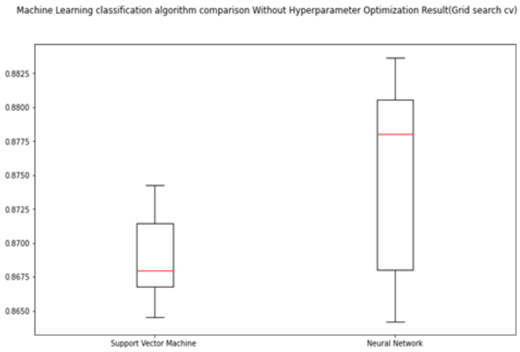

Fig. 7. Machine Learning classification algorithm comparison without Hyperparameter optimization result

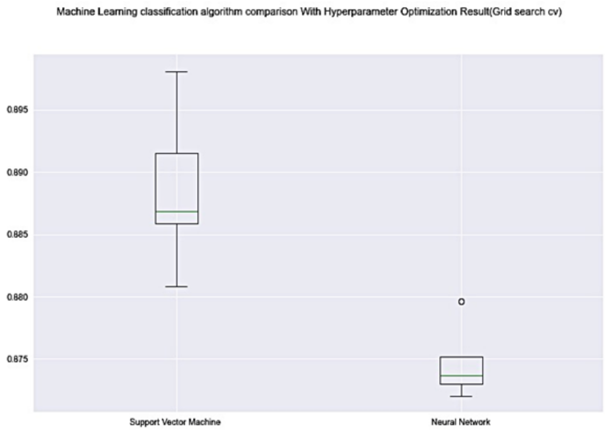

Fig. 8. Machine learning classification algorithm comparison with Hyperparameter optimization result

Research conducted by by [10,11] on two datasets German and Australian, shows the opposite result which represent respectively (86.76%, 86.90%) in the Australian data set and (77.92, 77.26) in German dataset. According to the complete outcomes, the combination of SVM and hyper parameter offers an excellent performance, making it the first classifier in the performance ranking of credit scoring models.

5. Conclusion

The implementation of the SVM classifier for assessing credit score is highlighted and well explained in the present research. SVM provides a practical answer to the constraints of the present crop of classifiers in use. The results suggest that datasets treated with big nonlinear techniques are getting increasingly computationally expensive. As a result, in recent years, data analysts have sought to develop faster computational algorithms in order to uncover patterns in datasets of growing complexity and size without the need of a high-performance classifier. In terms of the model results, SVMs are apt classifiers which can result in accurate classification. Moreover, SVMs with hyper-parameters have better performance than single SVMs. As a result, the credit scoring outcomes evaluated in this study support the notion that the SVM hyper-parameter can be employed in credit scoring applications to improve overall precision from a fraction of a percent to multiple percent. I strongly believe that future studies will focus on examining the hybridization impact of our SVM classifier with another method, such as KNN and random forest for better results.

References:

[1] J. Andersson, L. Baresi, N. Bencomo, R. Lemos, A. Gorla, P. Inverardi and T. Vogel, "Software Engineering Processes for Self-Adaptive Systems," in Software Engineering for Self-Adaptive Systems II, Springer, 2013, pp. 51-75.

[2] K. K. Ganguly and K. Sakib, "Decentralization of Control Loop for Self-Adaptive Software through Reinforcement Learning," in 24th Asia-Pacific Software Engineering Conference Workshops (APSECW), Nanjing, China, 2017.

[3] Abellán J, Mantas CJ. Improving experimental studies about ensembles of classifiers for bankruptcy prediction and credit scoring. Expert Systems with Applications. 2014 Jun 15;41(8):3825-30.

[4] Al-Mudimigh, A. and Anshari, M., 2020. Financial technology and innovative financial inclusion. In Financial technology and disruptive innovation in ASEAN (pp. 119-129). IGI Global.

[5] Birkenmaier, J., Despard, M., Friedline, T. and Huang, J., 2019. Financial inclusion and financial access. In Encyclopedia of social work.

[6] Chatterjee, S. and Barcun, S., 1970. A nonparametric approach to credit screening. Journal of the American statistical Association, 65(329), pp.150-154. Chatterjee, S. and Barcun, S., 1970. A nonparametric approach to credit screening. Journal of the American statistical Association, 65(329), pp.150-154.

[7] Chen, W., Sun, Z., Wang, X.J., Jiang, T., Huang, Z., Fang, D. and Zhang, D.D., 2009. Direct interaction between Nrf2 and p21Cip1/WAF1 upregulates the Nrf2-mediated antioxidant response. Molecular cell, 34(6), pp.663-673.

[8] Chopra, A. and Bhilare, P., 2018. Application of ensemble models in credit scoring models. Business Perspectives and Research, 6(2), pp.129-141.

[9] Durai, T. and Stella, G., 2019. Digital finance and its impact on financial inclusion. Journal of Emerging Technologies and Innovative Research, 6(1), pp.122-127.

[10] Egan, J.P. and Egan, J.P., 1975. Signal detection theory and ROC-analysis. Academic press.

[11] Frolova, E.E., Ermakova, E.P. and Protopopova, O.V., 2020, March. Consumer protection of digital financial services in Russia and abroad. In 13th International Scientific and Practical Conference-Artificial Intelligence Anthropogenic nature Vs. Social Origin (pp. 76-87). Springer, Cham.

[12] Prabakaran. D and S. Ramachandran, "Secure Key Generation from Speech signal Using Enhanced MFCC Algorithm," 2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkur, Karnataka, India, 2021, pp. 1-5, doi: 10.1109/ICMNWC52512.2021.9688398..

[13] Huang, C.L., Chen, M.C. and Wang, C.J., 2007. Credit scoring with a data mining approach based on support vector machines. Expert systems with applications, 33(4), pp.847-856.

[14] Kavzoglu, T. and Mather, P.M., 2003. The use of backpropagating artificial neural networks in land cover classification. International journal of remote sensing, 24(23), pp.4907-4938.

[15] Li, G., Klein, B., Sun, C. and Kou, J., 2020. Applying Receiver-Operating-Characteristic (ROC) to bulk ore sorting using XRF. Minerals Engineering, 146, p.106117.

[16] Nehrebecka, N., 2018. Predicting the default risk of companies. Comparison of credit scoring models: LOGIT vs Support Vector Machines. Ekonometria, 22(2), pp.54-73.

[17] Nguyen, H., Bui, X.N., Tran, Q.H. and Moayedi, H., 2019. Predicting blast-induced peak particle velocity using BGAMs, ANN and SVM: A case study at the Nui Beo open-pit coal mine in Vietnam. Environmental earth sciences, 78(15), pp.1-14.

[18] Shyamala, R., and D. Prabakaran. "A survey on security issues and solutions in virtual private network." International Journal of Pure and Applied Mathematics 119.15 (2018): 3115-3122.

[19] Pławiak, P., Abdar, M. and Acharya, U.R., 2019. Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring. Applied Soft Computing, 84, p.105740.

[20] Qu, Y., Quan, P., Lei, M. and Shi, Y., 2019. Review of bankruptcy prediction using machine learning and deep learning techniques. Procedia Computer Science, 162, pp.895-899.

[21] Sadowski, C., Söderberg, E., Church, L., Sipko, M. and Bacchelli, A., 2018, May. Modern code review: a case study at google. In Proceedings of the 40th International Conference on Software Engineering: Software Engineering in Practice (pp. 181-190).

[22] Schölkopf, B., Smola, A.J., Williamson, R.C. and Bartlett, P.L., 2000. New support vector algorithms. Neural computation, 12(5), pp.1207-1245.

[23] Tien Bui, D., Tuan, T.A., Hoang, N.D., Thanh, N.Q., Nguyen, D.B., Van Liem, N. and Pradhan, B., 2017. Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. Landslides, 14(2), pp.447-458.

[24] Tien Bui, D., Tuan, T.A., Klempe, H., Pradhan, B. and Revhaug, I., 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides, 13(2), pp.361-378.

[25] Van Gestel, T., Suykens, J.A., Baesens, B., Viaene, S., Vanthienen, J., Dedene, G., De Moor, B. and Vandewalle, J., 2004. Benchmarking least squares support vector machine classifiers. Machine learning, 54(1), pp.5-32.

[26] Vanatta, S.H., 2018. Making Credit Convenient: Credit Cards and the Political Economy of Modern America.

[27] Vapnik, V., 1998. The support vector method of function estimation. In Nonlinear modeling (pp. 55-85). Springer, Boston, MA.

[28] D. Prabakaran and S. Ramachandran, "Multi-factor authentication for secured financial transactions in cloud environment," Computers, Materials & Continua, vol. 70, no.1, pp. 1781–1798, 2022.

[29] Zhu, Y.F. and Mao, Z.Y., 2004, December. Online optimal modeling of LS-SVM based on time window. In 2004 IEEE International Conference on Industrial Technology, 2004. IEEE ICIT'04. (Vol. 3, pp. 1325-1330). IEEE.

[30] Michael Bücker Journal of Banking & Finance Volume 37, Issue 3, March 2013, Pages 1040-1045.

[31] ACPR, (2018), Intelligence artifcielle : enjeux pour le secteur financier, Document de réflexion, DÈcembre 2018

[32]] Varian, H.R., (2014), Big data : New tricks for econometrics, Journal of Economic Perspectives, 28, 3-28.

[33] Anderson, R. The Credit Scoring Toolkit: Theory and Practice for Retail Credit RiskManagement and Decision Automation. New York, Oxford University Press, 2007.