International Journal on Engineering Technology and Infrastructure Development (InJET-InDev)
www.lec.edu.np/journal
Volume 2, Issue no. 1, Page 168-183
Received Date: 12/7/2024
Revised Date: 6/7/2025
Accepted Date: 6/15/2025
ISSN: 3021-999X (Print)

Enhancement of Credit Score Prediction Using Artificial Neural Network with Adaptive Particle Swarm Optimization

Purusottam Panta[1]

[1] *Faculty of Science and Technology, Pokhara University Pokhara Metropolitan City-30, Nepal, purusottampanta@gmail.com*

Abstract

Credit scoring is a critical and most important process for every financial and lending institution to make informed lending decisions and effective risk management. Effective credit risk assessment plays a crucial role in minimizing default risk and enhancing recovery assurance, thereby strengthening the confidence of lending institutions in making informed and responsible loan decisions. Historically, traditional statistical methods such as logistic regression, linear discriminant analysis, and decision trees have been used to classify credit applicants. Still, these methods failed to handle non-linear data of the current economic landscape. Machine learning algorithms have improved credit scoring by handling non-linear data more effectively. However, they often struggle with tuning hyper-parameters, highlighting the need for more advanced and efficient models to manage complex data. This study aims to enhance credit score prediction using ANN with APSO. Three models namely: ANN without enhancement, ANN enhanced with the PSO, and ANN enhanced with the APSO were developed and evaluated using a popular real-world German credit dataset which was further split into training, validation, and testing datasets to assess the performance of the developed model with unseen dataset. Performances of the models were assessed using accuracy, precision, recall, F_1-score, and $F_{0.5}$-score. Additionally, a confusion matrix, loss curve, and precision-recall (PR) curve were generated for detailed analysis. The ANN enhanced with the APSO model achieved an accuracy of 80.00%, precision of 84.11%, F1-score of 85.71%, and F0.5-score of 84.74%, demonstrating superior performance over both the baseline ANN and ANN enhanced with the PSO. Specifically, ANN enhanced with the APSO improved the F1-score by 5.7% over the baseline ANN, and by 3.4% over ANN enhanced with the PSO, highlighting its enhanced effectiveness. Compared to optimization methods like Grid Search (GS) and Genetic Algorithm (GA), ANN-APSO also showed more consistent and higher performance metrics. These results suggest that the proposed ANN enhanced with the APSO model offers a more accurate and robust approach to credit scoring, with the potential to significantly improve risk assessment and support more informed, lower-risk lending decisions for financial institutions.

Keywords**:** Credit Scoring, ANN, PSO, APSO, Lending, Machine Learning, Loan Risk Assessment

1. Introduction

Among the various sources of income of financial institutions credit Lending is one of the main businesses that financial and lending institutions make large sums of revenue and profit from. Various types and purposes of loans are provided to individual and corporate institutions for different terms. However, every business poses a risk of failure or non-performance. Likewise, in the credit industry, there is a huge risk of loan defaulting resulting in a decrease in income and profit as well as the reputation of the institution in the market. Therefore, it is important for financial institutions and lending institutions to analyze credit histories and information of credit applicants comprehensively to get insights into the possibility of loan default and the likelihood of loan repayment and decide whether to approve an application for granting credit.

The credit industry including banks and other financial institutions has experienced growth for many decades with a significant increase in the number of credit customers with a variety of backgrounds, credit types, and service areas. This has increased business related to credit lending, which has also increased the risk of default. Since the global recession of 2007-08, many financial institutions have started using credit scoring models widely to improve informed decision-making for sanctioning credit to customers, balancing cash flow, and improving credit collections (Lappas & Yannacopoulos, 2021). Credit scoring model plays a pivotal role in assessing the creditworthiness of individuals and businesses thereby helping the organizations for financial decision-making. Credit scoring models used traditionally are effective to some extent but have difficulties in addressing the complexities of today's dynamic economic landscapes (Baseggio Corrêa & Filomena, 2020). Even credit scoring models based on statistical methods such as Logistic Regression (LR), Linear discriminant analysis (LDA), etc., have been questioned on their appropriateness due to the categorical nature of data. The advent of machine learning techniques such as Artificial Neural Networks (ANN), Decision Tree (DT), and Support Vector Machine (SVM) have presented their robustness and opportunities to improve the adaptability and accuracy of credit scoring even with the dynamic nonlinear data (Baseggio Corrêa & Filomena, 2020) (Louzada & et al., 2016). However, the challenge lies in optimizing the neural network architecture and parameters to maximize predictive performance.

Particle Swarm Optimization (PSO) is a nature-inspired optimization algorithm that simulates the collective behavior of a swarm of particles to find optimal solutions in complex search spaces (Dastile & et al., 2020) (Clerc, 2006) (Kennedy & Eberhart, 1995) (Kennedy, et al., 2001). This algorithm has demonstrated success in various optimization tasks, making it a compelling candidate for enhancing the performance of ANN models in credit-scoring applications (Zhang & Qiu, 2020) (Ye & et al., 2022). Adaptive Particle Swarm Optimization (APSO), is a modified form of standard classical PSO and is more efficient than standard PSO (Zhan & et al., 2009) (Hu & et al., 2013). APSO performs with faster convergence speed while finding global optimal solutions in the entire search space. APSO has the ability to control inertia weight, acceleration coefficients, and other algorithmic parameters at the run time automatically to improve search efficiency and convergence speed (Zhan & et al., 2009). The harmonious collaboration between APSO and ANN holds the potential to address the limitations of traditional credit scoring models, providing a more nuanced and adaptive approach to credit risk assessment.

The existing literature mostly focuses on the evolution of credit scoring methodologies, from rule-based systems to statistical models, and now to machine learning-based approaches (Louzada & et al., 2016). While machine learning models, including ANN, have shown promise in improving predictive accuracy, their effective application requires careful consideration of model complexity, interpretability, and adaptability to changing economic conditions. This background underscores the need for advanced techniques that can optimize neural network architectures and parameters to strike a balance between accuracy and interpretability in credit scoring. Recent studies have explored the application of APSO in various optimization tasks (Kennedy & Eberhart, 1995) (Kennedy, et al., 2001), ranging from engineering design to financial modeling (Zhang & Qiu, 2020) (Qin & et al., 2021) (Marini & Walczak, 2015) (Jain & et al., 2022) (Wihartiko & et al., 2018). However, its specific application to enhance ANN models for credit scoring remains an underexplored area. This research aims to fill this gap by investigating the potential of an ANN enhanced with APSO model for credit scoring, with a focus on optimizing architecture, feature selection, and weight assignment along with adaptability to changing economic conditions. The research is motivated by the aspiration to contribute to the advancement of credit scoring methodologies, providing financial institutions with more accurate, adaptable, and efficient tools for risk assessment. By leveraging the collective intelligence of APSO to enhance ANN models, this research seeks to offer a novel and comprehensive framework that addresses the multifaceted challenges posed by contemporary credit landscapes.

2. Literature Review

In financial decision-making, credit scoring plays a very important role but it has been revolutionized with the introduction of AI. Among AI systems, Artificial Neural Networks have shown signs of being able to analyze intricate trends in credit data (Louzada & et al., 2016). However, the optimization of neural network architectures and parameters remains a difficult task. This literature review explores the integration of ANN with Adaptive Particle Swarm Optimization (APSO) for credit scoring, addressing the need for improved model accuracy and precision. The XGBoost with APSO was proposed for credit scoring and achieved a predictive score of 77.48% which was 0.8% higher than the default XGBoost model on German and Australian Datasets, where a maximum number of iterations n is 200 (Qin & et al., 2021).

A Hybrid data mining model of feature selection and ensemble learning classifier for feature selection and data mining was proposed, where ANN and AdaBoost are used as the main model, compared the model's result with several other statistical and machine learning models and found that the ensemble learning classifier using multiple algorithms outperformed the models built using single algorithm alone (Koutanaei & et al., 2015). In addition, the multi-stage self-adaptive classifier method was introduced to construct a credit scoring model, where individual models were developed and evaluated and in the next step more than one best model was combined to yield better performance and accuracy (Guo & et al., 2019). Hybrid Neural Network (NN) model for credit risk predictions, by combining Decision Trees (DT), Logistic Regression (LR), and Discriminant Analysis (DA) with four types of different NN: Deep Neural Networks (DNN), Multilayer Perceptron (MLP), Adaptive Neuro-fuzzy Inference Systems (ANFIS) and Radial Basis Function (RBF). The capacity of developed approaches was evaluated by ten different performance measure techniques and it found that LR and MLP-based hybrid models always performed best with an accuracy rate of 98.08% (Chi & et al., 2019). The survey on several research papers on credit scoring concluded that in general, an ensemble of classifiers outperforms single classifiers (Dastile & et al., 2020).

While most of the research focused on increasing the performance of the model, a new model was introduced combining expert knowledge with Genetic Algorithm (GA) for feature selection, where the focus was on identifying the model's behavior while injecting expert knowledge into the model parameters instead of enhancement of the model (Lappas & Yannacopoulos, 2021). The result was surprising while comparing Neural Networks (NN), Support Vector Machine (SVM), and Decision Trees (DT) to forecast corporate credit ratings, that DT outperformed SVM and Multilayer Perceptron (MLP) when the Notches Distance concept was used to quantify how far the predictions from the true value (Golbayani & et al., 2020). ANN and SVM are the most widely researched topics in credit scoring as an individual model in various comparative studies. Comparison of ANN and SVM model was done, where ANN (Accuracy 87.39%) outperformed SVM (Accuracy 86.81%) model (Nwulu, et al., 2012) whereas Radial Basis Function (RBF) kernel SVM (AUC 91.58%) outperformed ANN (AUC 87.56%) on (Khemakhem & Boujelbene, 2017).

It is necessary that the result obtained from the model developed should be understandable by the human i.e. the result and the input parameters must be interpreted and valid reasons for using different hyper parameters, features, and variables should be extracted from the model's result, to achieve this to the extent possible an interpretable credit scoring model was developed (Trinkle & Baldwin, 2007). A deep learning model for behavioral credit rating for future prediction of the bank was proposed in which performance, predictability, and accuracy with other traditional models (Merćep & et al., 2020).

Most of the research on the development of credit-scoring models is mainly focused on using ANN for model constructions, (Koutanaei & et al., 2015) (Khemakhem & Boujelbene, 2017) (Trinkle & Baldwin, 2007) (Boguslauskas & Mileris, 2009) (Khashman, 2010) (Teles & et al., 2020) (A.H. & et al., 2022) some tried to implement ensemble models, an Artificial Intelligent model for Loan recovery predictions was proposed in (Adewusi & et al., 2016). Only a few researchers tried to Optimize credit scoring models based on different machine learning algorithms using different techniques such as Adaptive Particle Swarm Optimization (APSO), Genetic Algorithms (GA), etc., (Zhang & Qiu, 2020) (Qin & et al., 2021) (Bahnsen & Gonzalez, 2011) (Oreski & Oreski, 2014) (Kazemi & et al., 2021). Back Propagation Artificial Neural Network (BP-ANN) model was optimized with several Swarm Intelligence algorithms such as Bat Algorithm (BA), Cuckoo Search Optimization (CSO), Firefly Algorithm (FA), Gravitational Search Algorithm (GS), Gray Wolf Optimization Algorithm (GWO), Particle Swarm Optimization (PSO), Social Spider Algorithm (SSA), Whale Swarm Algorithm (WSA). The training and test results showed that PSO-BP-ANN with an AUC of 80.04% outperformed most of the other models (Zhang & Qiu, 2020). An approach to optimize XGBoost using Bayesian hyper parameter optimization to improve predictability was formulated in (Xia & et al., 2017).

Feature selection is the most important task/step of every machine learning model. It is used to make the process more optimized and accurate. In feature selection, the most critical variables are selected, and irrelevant and redundant features are eliminated to increase the predictive power of the Machine Learning Algorithms. The effect of feature selection on financial distress prediction was studied in (Liang & et al., 2015). APSO for feature selection on leadership learning models was used in (Ye & et al., 2022). Genetic-based heuristics for feature selection in credit risk assessments is used in (Oreski & Oreski, 2014).

3. Research Problem

Traditional credit scoring models, rooted in statistical methodologies, often face challenges in adapting to the dynamic and complex nature of modern financial environments. As financial transactions become increasingly digitized and economic conditions evolve, there is a growing need for credit scoring models that are not only accurate but also adaptive, interpretable, and capable of handling intricate patterns within vast datasets. Artificial Neural Networks (ANN) have demonstrated considerable promise in capturing non-linear relationships and patterns, presenting an opportunity to enhance the predictive power of credit scoring models. However, the effective utilization of ANN is contingent upon addressing critical challenges, including the optimal configuration of network architecture, feature selection, and weight assignment. The complexity of these tasks poses a significant obstacle to realizing the full potential of ANN in credit risk assessment. Manual Tuning of hyper parameters such as the number of hidden layers, number of neurons per layer, and learning rates on ANN model can be time-consuming and suboptimal, in addition ANN models are prone to Overfitting, especially when dealing with small and/or noisy datasets. Addressing these issues will contribute to the development of a sophisticated and adaptive credit scoring model that leverages the strengths of APSO and ANN. The outcomes of this research are anticipated to offer financial institutions enhanced tools for accurate and timely credit risk assessment in the face of evolving economic complexities.

Although there have been various research and models in the field of credit scoring using both traditional and Machine learning methods, they are still struggling to predict the dynamic economic environments hence this research is motivated by the overarching goal of advancing credit scoring methodologies to meet the demands of a rapidly changing financial landscape, leveraging the synergies between ANN and APSO for enhanced accuracy, interpretability, and robustness in credit risk assessment.

4. Methodology

The study aims to enhance the predictability of the ANN model that is used for credit scoring using APSO. The methodology employed in the research study is divided into data collection and preprocessing, model development and training, model evaluation, and final model as shown in Figure 1.

4.1. Data Collection and Dataset Description

Data required for the research are obtained from the UCI Machine Learning Repository (Hofmann, 1994). The data set is a popular real-world dataset and is used by several researchers, this way the developed model's performance can be compared with other already developed models. The dataset contains 1000 instances and 20 features (as shown in Table 2) and a target binary variable. Out of 1000 instances, 700 instances are good credit and 300 instances are bad credit as shown in Table 1. The data have no missing values with three types of features namely Categorical, integer, and binary data type.

Table 1 Distribution of Credit Score Dataset

Dataset	Total	Good	Bad
German Dataset	1000	700	300

Table 2 Features Available in the Dataset

Variable Name	Role	Type	Demographic	Description	Units
Attribute1	Feature	Categorical		Status of existing checking account	
Attribute2	Feature	Integer		Duration	months
Attribute3	Feature	Categorical		Credit history	
Attribute4	Feature	Categorical		Purpose	

Variable Name	Role	Type	Demographic	Description	Units
Attribute5	Feature	Integer		Credit amount	
Attribute6	Feature	Categorical		Savings account/bonds	
Attribute7	Feature	Categorical	Other	Present employment since	
Attribute8	Feature	Integer		Installment rate in percentage of disposable income	
Attribute9	Feature	Categorical	Marital Status	Personal status and sex	
Attribute10	Feature	Categorical		Other debtors/guarantors	
Attribute11	Feature	Integer		Present residence since	
Attribute12	Feature	Categorical		Property	
Attribute13	Feature	Integer	Age	Age	years
Attribute14	Feature	Categorical		Other installment plans	
Attribute15	Feature	Categorical	Other	Housing	
Attribute16	Feature	Integer		Number of existing credits at this bank	
Attribute17	Feature	Categorical	Occupation	Job	
Attribute18	Feature	Integer		Number of people being liable to provide maintenance for	
Attribute19	Feature	Binary		Telephone	
Attribute20	Feature	Binary	Other	foreign worker	
Class	Target	Binary		1 = Good, 2 = Bad	

4.2. Data Preprocessing

Collected data contains several categorical features, since most ANN algorithms and models do not support training and testing using categorical values, the data is converted into numerical values using One-Hot encoding. Encoded data was then normalized using z-score normalization using the formula as shown in equation 1.

$$z = \frac{(x - \mu)}{\sigma} \quad \text{(Equation 1)}$$

Where z is the normalized value, x is the original value of the feature, μ is the mean value of the feature, and σ is the standard deviation of the feature.

After normalization, the data is then further split into Train, Test, and Validation sets.

4.3. Model development, training and validation

During the research experiment a base ANN model was developed as shown in Figure 2, and was trained multiple times with one and two hidden layers to find an optimal and effective credit scoring model. Each hidden layer consists of 32 neurons which are chosen from 2 to 256 neurons after several iterations and experiments. Results obtained from training the model using the training dataset were then validated using 5-fold cross-validation. While doing experiments for this research, an ANN model is developed, since the model was developed using randomly chosen hyper-parameters it may not be optimal, hence an APSO is used to optimize those hyper-parameters. Inertia weights w, cognitive coefficient c_1, and social coefficient c_2 are adjusted throughout the process and velocities and position of each particle are updated accordingly.

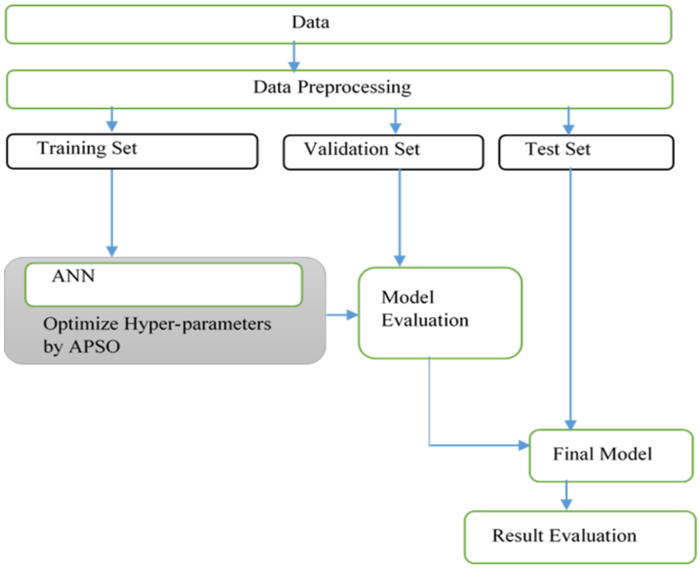

Figure 1 Block Diagram of Proposed Methodology

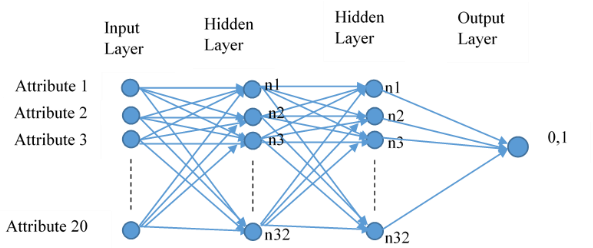

Figure 2 A simplified block diagram of the ANN model used to train the Dataset

The model was run for 200 iterations for each of 10, 20,30,40, and 50 particles at a time. A diversity threshold and stagnation threshold are set to measure the diversity of the mean position of each particle for each iteration. When several improvements are seen up to the stagnation threshold or diversity is less than the diversity threshold, the improvement counter is reset and exploration is increased.

Since credit scoring deals directly with the financial well-being of lending and financial institutions, sanctioning credit to a customer who cannot repay or default the loan is highly risky. Throughout each iteration, each particle's fitness values namely accuracy and false positive value are measured. The goal of this research is to decrease the number of false positive instances and at the same time increase the accuracy of the model.

Algorithm 1 Adaptive Particle Swarm Optimization (APSO)

Input: *Population size N, maximum iterations MaxIter*

Output: *Best solution found*

1: *procedure APSO (N, MaxIter)*

2:    *Initialize population P with N particles*

3:    *for each particle i in P do*

4:       *Initialize velocity* v_i *and position* p_i

5:       *Evaluate the fitness* f_i *of particle i*

6:       *Set personal-best* $pBest_i \leftarrow$ *position of particle i*

7:    *end for*

8:    *Set global best gBest* $\leftarrow$ *best* $pBest_i$ *in P*

9:    *for t* $\leftarrow$ *1 to MaxIter do*

10:       *for each particle i in P do*

```
        Update velocity vi using adaptive parameters
        Update the position of particle i
        Evaluate the fitness fi of particle i
        if fi < fitness of pBesti then
            Update pBesti ← position of particle i
        end if
        if fi < fitness of gBest then
            Update gBest ← position of particle i
        end if
    end for
  end for
  return gBest
end procedure
```

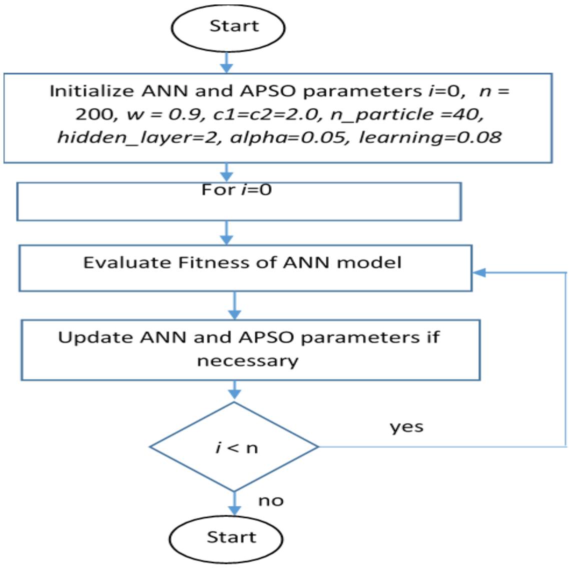

Figure 3 Flow chart of APSO Implementation

The flow chart of APSO implementation is shown in Figure 3 and the pseudocode for the implementation is presented in Algorithm 1. A detailed description of every step involved in the APSO implementation is described in the following paragraphs.

Initialization: Randomly initialize the position x_i and velocity v_i of each particle in the swarm. Evaluate the fitness of each particle's initial position x_i using the objective function $f(x_i)$.

Update velocity and Position: For each particle i, update its velocity v_i and position x_i based on the particle's previous velocity, position, and the global best position ***gBest*** found so far. The velocity update is guided by both cognitive and social components, controlling the exploration and exploitation aspects of the search. The mathematical equation to update the velocity of the particle is shown in equation (2), and the position update equation is shown in equation (3) (Jain & et al., 2022).

$$v_i(t+1) = w.v_i(t) + c_1.r_1.\left(p^i_{best,t} - x_i(t)\right) + c_2.r_2.(g_{best,t} - x_i(t)) \quad \text{(Equation 2)}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad \text{(Equation 3)}$$

Here $v_i(t)$ and $x_i(t)$ are the velocity and position of particle *i* at time *t*, *w* is inertia weight c_1 and c_2 are cognitive and social coefficients, r_1 and r_2 are random uniform numbers $p_{best,t}^{i}$ is the personal best position of particle *i* at time *t* and $g_{best,t}$ is the global best position found by the swarm at time *t*.

Update Inertia Weight and Parameters: Adaptive mechanism of APSO is used to update *w*, c_1, and c_2 adaptively along with other parameters of the algorithm to balance exploration and exploitation throughout the optimization process. Inertia weight *w* can be adapted using equation (4) and c_1 and c_2 can also be adapted using equations (5) and (6) respectively (Jain & et al., 2022).

$$w(t) = w_{max} - \left(\frac{w_{max} - w_{min}}{T_{max}}\right).t \qquad \text{(Equation 4)}$$

$$c_1(t) = c_{1_{max}} - \left(\frac{c_{1_{max}} - c_{1_{min}}}{T_{max}}\right).t \qquad \text{(Equation 5)}$$

$$c_2(t) = c_{2_{max}} + \left(\frac{c_{2_{max}} - c_{2_{min}}}{T_{max}}\right).t \qquad \text{(Equation 6)}$$

Here T_{max} is the maximum number of iterations, *t* is the current iteration, w_{max} and w_{min} are maximum and minimum values of inertia weight, $c_{1_{max}}$ and $c_{1_{min}}$ and $c_{2_{max}}$ and $c_{2_{min}}$ are maximum and minimum values for cognitive and social coefficient respectively. A strong relationship between *w* with acceleration coefficients c_1 and c_2 is considered if the relation satisfies equation (7).

$$w > \frac{1}{2}(c_1 + c_2) - 1 \qquad \text{(Equation 7)}$$

Evaluate Fitness: Fitness of new position $x_i(t+1)$ is to be evaluated after updating the position of the particle using objective function $f(x_i(t+1))$.

Update Personal and Global Bests: Update personal best position $p_{best,t}^{i}$ and global best position $g_{best,t}$ if the new position $x_i(t+1)$ leads to better fitness values.

Iteration: Steps update velocity and position, update inertia weights and parameters, evaluate fitness, and update personal and global best must be repeated for the predefined number of iterations or until the termination criteria are met, such as achieving a satisfactory solution.

Return Global Best: At the end of the procedure after the loop is terminated, return the global best $g_{best,t}$ and its corresponding fitness value $f(g_{best,t})$, as an optimized solution obtained from the APSO.

Cost Function: The cost function associated with APSO is usually a cost function that the algorithm tries to optimize. The cost function associated with training neural networks using APSO are Binary Entropy-Cross loss for classification problems as shown in equation (8) and Mean Squared Error for Regression as shown in equation (9) (Clerc, 2006) (Jain & et al., 2022).

$$J_{accuracy} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \qquad \text{(Equation 8)}$$

$$J_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad \text{(Equation 9)}$$

Model training and testing were conducted on a 12th Gen Intel i5 CPU (10 cores, 1.3 GHz) with 16 GB RAM. The base ANN model completed training in approximately 1 minute, while the ANN optimized using PSO required around 1.17 hours. In comparison, the proposed ANN-APSO model achieved a balance between performance and computational efficiency, completing training in approximately 46 minutes.

This demonstrates that while the optimization process increases training time compared to the base ANN, APSO significantly reduces the time required compared to standard PSO—by nearly 31%. The reduced training time, combined with improved predictive performance, highlights the ANN-APSO model as a more computationally efficient alternative to traditional PSO-based optimization for credit scoring applications.

5. Result And Discussion

After conducting 200 iterations for each model using a range of initialization parameters, the results from all models were compared to evaluate optimization effectiveness. Both APSO and PSO were applied with different particle sizes and varying bounds for ANN hyper-parameters. The ANN model optimized using APSO consistently achieved higher training and validation accuracy relative to the standard ANN and the PSO-optimized version. The specific hyper-parameters corresponding to this improved performance are presented in Table 3.

Table 3 Best Hyper-parameters Obtained from the Experiments

num_particles	40
max_iteration	200
number of neurons	32
Number of hidden layers	2
L2 Regularization (alpha)	0.05
learning_rate_init	0.08
momentum	0.7

5.1. Confusion Matrix

The confusion matrix obtained from the credit scoring model using ANN with APSO is shown in Figure 4, it is seen that out of 150 test data ANN model enhanced with APSO predicted 120 records correctly, out of all correct predictions 90 records were predicted as good credit, and 30 records predicted as bad credit which aligns with the true value of the datasets. Even though the model has great predictive accuracy it falsely predicted 13 records as bad credit while they fell under good credit likewise 17 records were predicted as good credit while they truly fell under bad credit.

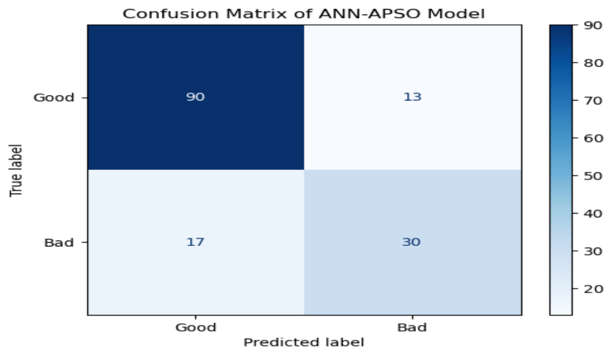

Figure 4 Confusion Matrix of ANN optimized with APSO

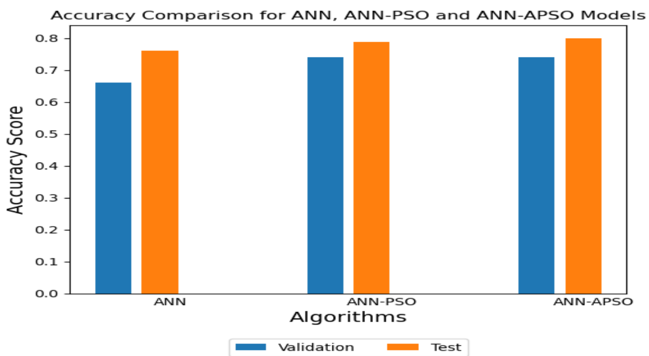

Figure 5 Comparison of Validation and Test Accuracies for Different Models

5.2. Comparison of Different Evaluation Matrices

Three models ANN without APSO, ANN enhanced with PSO, and ANN enhanced with APSO were developed and compared during the research. Test and validation accuracies obtained from the experiment are presented in Table 4 and Figure 5. Table 5 and Figure 6 present the matrices comparison of ANN without APSO, ANN with PSO, and ANN with APSO. The validation accuracy of ANN without PSO and APSO model is 65.99%, which increased around 10% during the test and resulted in an accuracy of 76.00%.

The accuracy of PSO optimized model was 74.00% during validation and 78.66% during testing. The proposed ANN enhanced with APSO model showed significantly higher accuracy during validation and testing with accuracies of 74.67% during validation and 80.00% on test datasets, which is significantly higher than NN model enhanced with model and base ANN model. This means that APSO enhanced model can predict good and bad loans more accurately than any of the PSO and ANN models.

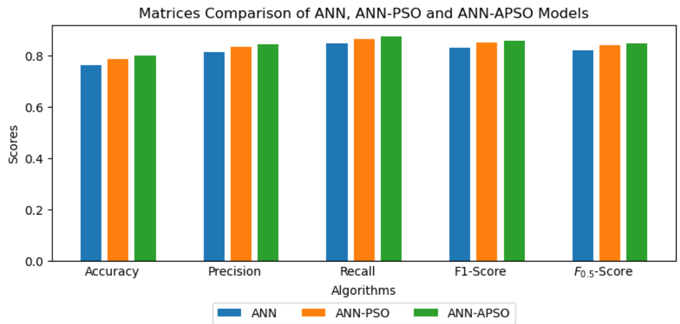

Figure 6 Comparison of Accuracy, Precision, Recall, f1-Score, and F0.5-Score for Different Models

Table 4 Comparison of Test and Validation Accuracies of ANN, ANN with PSO, and ANN with APSO

Algorithm	Validation	Test
ANN	0.6599	0.76
ANN-PSO	0.74	0.7866
ANN-APSO	**0.7467**	**0.8**

Table 5 Comparison of Accuracy, Precision, Recall, f1-Score, and F0.5-Score of ANN, ANN with PSO, and ANN with APSO

Algorithm	Accuracy	Precision	Recall	F1-Score	$F_{0.5}$-Score
ANN	0.76	0.813	0.8446	0.8285	0.8192
ANN-PSO	0.7866	0.8317	0.864	0.8476	0.838
ANN-APSO	**0.8**	**0.8411**	**0.8737**	**0.8571**	**0.8474**

During the experiment, it was found that the precision of ANN, ANN with PSO, and ANN with APSO models resulted in 81.30%, 83.17%, and 84.11% respectively. Likewise, bas ANN model has a recall value of 84.46%, ANN enhanced with PSO has a recall value of 86.40% and ANN enhanced with APSO has a recall of 87.37%. F1-Score obtained as a result of the experiment in this research are 82.85%, 84.76%, and 85.71% for ANN, ANN enhanced with PSO, and ANN enhanced with APSO models respectively. F1-Score equals 1 means perfect precision and recall and 0 means worst combination of precision and recall. Generalization of F1-Score that allows to weigh precision more heavily than precision or vice-versa is called F_{β}-Score. From the experiment when more emphasis is given to precision $F_{0.5}$-Score obtained are 81.92%, 83.80%, and 84.74% for base ANN, ANN enhanced with PSO, and ANN enhanced with APSO models respectively. From Table 5, Figure 5, and Figure 6, it can be summarized that the ANN enhanced with APSO model has relatively high accuracy along with high precision and recall as well as F1-Score and $F_{0.5}$-score than other models, this indicates that the model is better at predicting class good loans.

Table 6 Classification Report Obtained Using Various Train-Test Split Ratio for the ANN with APSO model

Data Split Ratio	Accuracy	Precision	Recall	F1-Score	$F_{0.5}$-Score
50:25:25	0.708	0.7593	0.8352	0.7955	0.7734
60:20:20	0.74	0.7592	0.9044	0.8255	0.7844
70:15:15	**0.8**	**0.8411**	**0.8737**	**0.8571**	**0.8474**
80:10:10	0.78	0.8701	0.8481	0.8589	0.8656

Table 6 presents classification reports obtained with similar conditions but with different data split ratios for training, validation, and testing. Looking at the matrices present in Table 6, the accuracy, recall, and F_1-score of the model when data are split in the ratio of 70:15:15 is higher than other split ratios.

5.3. Accuracy Curve

The accuracy curve obtained from the ANN model without using any optimization method is shown in Figure 7. From the graph, it is seen that both training and testing accuracy increase as the number of epochs in the training increases. The green solid line in the curve represents the testing accuracy and the dotted blue line represents training accuracy. The smooth increase in the accuracy over a number of epochs suggests that the model is learning effectively without any large fluctuations, indicating stable training. The close alignment between training and testing accuracy curves suggests the model generalizes well with new unseen data with minimal overfitting.

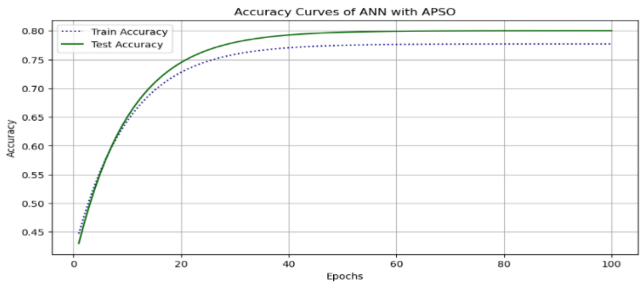

Figure 7 Accuracy Curve of ANN with APSO

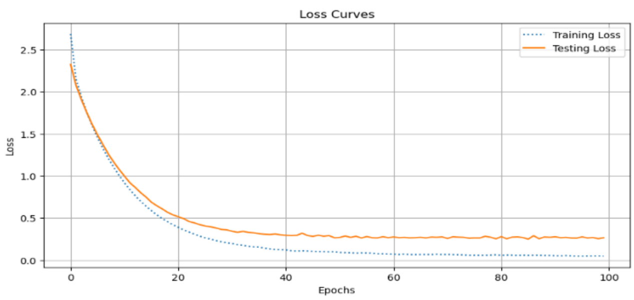

Figure 8 Loss Curve Obtained from Experiment

5.4. Loss Curve

The mean square error loss curve obtained from the model without using APSO is shown in Figure 8. From the graph, it is seen that the training loss decreases sharply with the increase in the number of epochs and levels off when it reaches close to zero. This indicates that the model becomes better at minimizing the loss when training progresses. Looking at the testing loss, it starts slightly lower than the training loss with a steep decline less than that of training loss. Both losses in the graph decrease and level off near zero which suggests that the model is good at maintaining balance between learning and generalization.

5.5. Precision-Recall (PR) Curve

A graphical representation used to illustrate the trade-offs between precision and recall which is used to evaluate the performance obtained from the binary classification model is called the Precision-Recall curve. In Figure 9 dashed line represents the base ANN, the dotted line represents ANN enhanced with PSO, and the solid line represents ANN enhanced with APSO models. From Figure 9 it appears that the ANN with APSO model performs better than the other two models across most thresholds, as its curve is closer to the top right corner of the plot and has maintained high precision across all levels of recall. This indicates that ANN enhanced with the APSO model has a better balance of precision and recall, suggesting that the ANN enhanced with the APSO model is likely the most suitable model among the three for the given task.

5.6. Mean Loss Curve

As shown in Figure 10, the dotted line represents the base ANN model. The loss for the base ANN model was high at the beginning and started to decrease over the increasing number of iterations showing learning and performance increase over a number of iterations. Loss for the ANN enhanced with the PSO model is represented by the dashed line in Figure 10. The loss value for ANN enhanced with the PSO model starts slightly above the base ANN model and decreases with about similar rates over a number of iterations. The solid line in Figure 10 represents a plot of the mean loss function for ANN enhanced with the APSO model. This also starts with similar values like base ANN and ANN enhanced with PSO model, but decreased sharply over a number of iterations suggesting its faster convergence than the other two models.

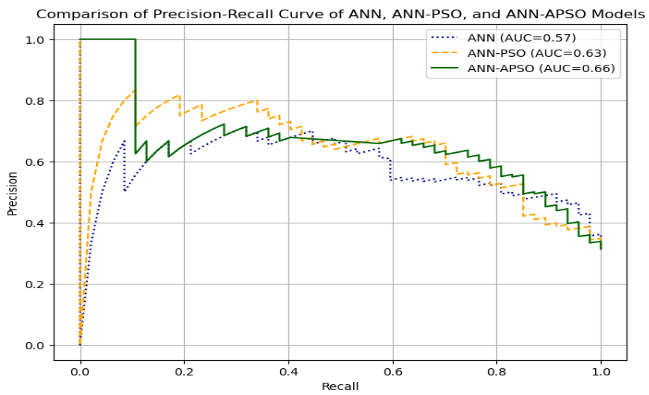

Figure 9 Comparison of PR-Curve of ANN, ANN with PSO and ANN with APSO

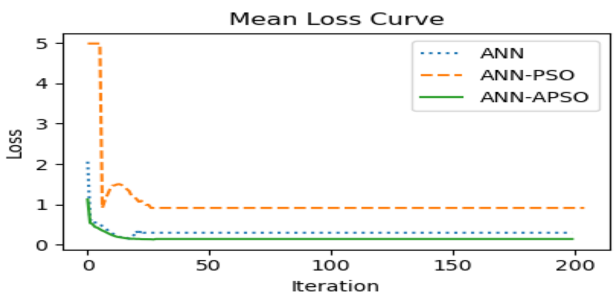

Figure 10 Plot of the mean value of loss function for iteration n=200

5.7. Comparison with other Optimization Models

Table 7 Comparison of Accuracy, Precision, Recall, f1-Score, and F0.5-Score of Proposed Model with Other Optimization Models

Algorithm	Accuracy	Precision	Recall	F1-Score	$F_{0.5}$-Score
ANN	0.76	0.813	0.8446	0.8285	0.8192
ANN-GS	0.76	0.7913	0.8835	0.8349	0.8083
ANN-GA	0.78	0.807	0.8932	0.8478	0.8232
ANN-PSO	0.7866	0.8317	0.864	0.8476	0.838
ANN-APSO	**0.8**	**0.8411**	**0.8737**	**0.8571**	**0.8474**

Looking at the classification comparison table 7, it is observed that the ANN enhanced with the APSO has the highest accuracy of 80.00% followed by ANN enhanced with the PSO and ANN-GA with accuracy of 78.66% and 78.00% respectively. ANN enhanced with APSO also leads precision, F_1-Score, and $F_{0.5}$-Score with a score of 84.11%, 85.71%, and 84.74% respectively. While the recall score of ANN with the GA model is higher with a score of 89.32% than other models, the recall score of the ANN enhanced with APSO model is also quite high at 87.37%. Taking into consideration all these scores ANN enhanced with the model is performing well with the highest score of precision to minimize costly false positives and overall best performance with high accuracy and F_1-score.

5.8. Comparison with State-of-Art Models

Table 8 Accuracy Comparison of Proposed Model with Other Research

Method	Algorithm used	Accuracy
(Xia & et al., 2017)	XGBoost-TPE	0.7734
(Guo & et al., 2019)	MLP-Bays-PSO	0.783
(Qin & et al., 2021)	APSO-XGBoost	0.7748
(Guo & et al., 2019)	BP-ANN-PSO	0.761
(Boguslauskas & Mileris, 2009)	HGA-NN	0.789
Proposed Model	**ANN-APSO**	**0.8**

When comparing the performance of the proposed model with other state-of-the-art models as shown in Table 8, the proposed model has shown the highest accuracy of 80.00%, when compared to other models proposed model achieved enhancements ranging from 2.10% to 3.40% which is a significant improvement over existing method. The superior performance of the ANN enhanced with APSO model is attributed to an effective combination of ANN with APSO to yield better performance than other optimization techniques such as Hybrid Genetic Algorithm (HGA) or Tree-structured Parzen Estimator (TPE). Consistent outperformance by ANN enhanced with APSO across different algorithms suggests that the ANN enhanced with APSO model is a particularly effective approach to credit classification offering a new benchmark for research in this field in the future.

6. Limitations

Despite its improved performance, the ANN enhanced with the APSO model has some limitations. It requires high computational resources and may be prone to overfitting without careful validation. The model also lacks interpretability, which can be a concern in regulated financial settings. Additionally, its effectiveness depends on the quality of the training data and it takes slightly more time for training and testing compared to the base ANN model.

7. Conclusion And Future Recommendations

The research is focused on optimizing the predictive ability of the ANN model using APSO. To validate the performance increase, the research compared three credit scoring models base ANN without optimization, ANN optimized with PSO, and ANN optimized with APSO. The proposed ANN-APSO model is constantly outperforming other models across various matrices such as the highest accuracy (80.00%), precision (84.11%), F_1-Score (85.71%), and $F_{0.5}$-Score (84.74%). As evidenced by the mean loss curve the ANN enhanced with the APSO model showed faster convergence with better learning rates. When comparing ANN enhanced with the APSO model with other optimization techniques such as GS and GA, the proposed ANN enhanced with the APSO model showed superior performance against GS and GA. The proposed method also outperformed other several models from previous studies with improvements of 2.10% to 3.40%. Hence, looking at the results and analysis and comparison report with other models and methods, it can be concluded that the proposed ANN enhanced with the APSO model is suitable for credit score classification compared with other models.

Since the research is conducted to increase the performance of the ANN model by using APSO and is trained, validated, and tested on a single dataset, the method might have several limitations and these can be improved in the future. Interested research in the field of credit scoring can use this model to test and validate other several datasets, investigation of the impact of APSO's parameter can be another research area. Another research area could be an ensemble of the ANN enhanced with the APSO model with other several machine learning algorithms to potentially improve and yield better performance, feature engineering can also be employed on this model to identify the most influential feature. In addition, future researchers can also explore the adaptation of the model to a dynamically changing economic landscape and compare the model's performance with other deep learning approaches.

Acknowledgements

I would like to extend my sincere thanks to professors, friends, and colleagues for providing me with all the essential co-operation, and valuable suggestions for choosing this research topic.

I am grateful to my wonderful supervisor Dr. Ashim Khadka for providing me with valuable information, knowledge, and guidance regarding this research. He consistently allowed this paper to be my work but steered me in the right direction whenever and wherever he thought I needed it.

I am grateful to the program coordinator of Nepal College of Information Technology (NCIT) college Assoc. Prof. Saroj Shakya for providing coordination and all the support related to this research.

References

Lappas, P. Z. & Yannacopoulos, A. N., 2021. A Machine Learning Approach Combining Expert Knowledge with Genetic Algorithms in Feature Selection for Credit Risk Assessment. *Applied Soft Computing*, August, Volume 107, p. 107391.

Baseggio Corrêa, R. & Filomena, T. P., 2020. Traditional and Machine Learning Methods for Credit Scoring: An Introduction. s.l.: s.n.

Louzada, F. & et al., 2016. Classification Methods Applied to Credit Scoring: Systematic Review and Overall Comparison. *Surveys in Operations Research and Management Science*, December, 21(2), p. 117–134.

Dastile, X. & et al., 2020. Statistical and Machine Learning Models in Credit Scoring: A Systematic Literature Survey. *Applied Soft Computing*, June, Volume 91, p. 106263.

Clerc, M., 2006. *Particle swarm optimization*. s.l.:Wiley-ISTE.

Kennedy, J. & Eberhart, R., 1995. Particle Swarm Optimization. s.l., IEEE, p. 1942–1948.

Kennedy, J., Eberhart, R. C. & Shi, Y., 2001. *Swarm Intelligence*. s.l.:Morgan Kaufmann.

Zhang, R. & Qiu, Z., 2020. Optimizing Hyper-Parameters of Neural Networks with Swarm Intelligence: A Novel Framework for Credit Scoring. *PLOS ONE*, June, 15(6), p. e0234254.

Ye, Z. & et al., 2022. Feature Selection Based on Adaptive Particle Swarm Optimization with Leadership Learning. *Computational Intelligence and Neuroscience*, August, Volume 2022, p. 1–18.

Zhan, Z.-H. & et al., 2009. Adaptive Particle Swarm Optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, December, 39(6), p. 1362–81.

Hu, M. & et al., 2013. An Adaptive Particle Swarm Optimization with Multiple Adaptive Methods. *IEEE Transactions on Evolutionary Computation*, October, 17(5), p. 705–720.

Qin, C. & et al., 2021. XGBoost Optimized by Adaptive Particle Swarm Optimization for Credit Scoring. *Mathematical Problems in Engineering*, March, Volume 2021, p. 1–18.

Marini, F. & Walczak, B., 2015. Particle Swarm Optimization (PSO). A Tutorial. *Chemometrics and Intelligent Laboratory Systems*, December, Volume 149, p. 153–165.

Jain, M. & et al., 2022. An Overview of Variants and Advancements of PSO Algorithm. *Applied Sciences*, August, 12(17), p. 8392.

Wihartiko, F. D. & et al., 2018. Performance Comparison of Genetic Algorithms and Particle Swarm Optimization for Model Integer Programming Bus Timetabling Problem. s.l., s.n., p. 012020.

Bishop, C. M., 2006. *Pattern Recognition and Machine Learning*. s.l.:Springer.

Goodfellow, I. & et al., 2016. *Deep Learning*. s.l. The MIT Press.

Moscato, V. & et al., 2021. A Benchmark of Machine Learning Approaches for Credit Score Prediction. *Expert Systems with Applications*, March, Volume 165, p. 113986.

Koutanaei, F. N. & et al., 2015. A Hybrid Data Mining Model of Feature Selection Algorithms and Ensemble Learning Classifiers for Credit Scoring. *Journal of Retailing and Consumer Services*, November, Volume 27, p. 11–23.

Guo, S. & et al., 2019. A Multi-Stage Self-Adaptive Classifier Ensemble Model With Application in Credit Scoring. *IEEE Access*, Volume 7, p. 78549–59.

Chi, G. & et al., 2019. Hybrid Model for Credit Risk Prediction: An Application of Neural Network Approaches. *International Journal on Artificial Intelligence Tools*, August, 28(05), p. 1950017.

Golbayani, P. & et al., 2020. A Comparative Study of Forecasting Corporate Credit Ratings Using Neural Networks, Support Vector Machines, and Decision Trees. *The North American Journal of Economics and Finance*, November, Volume 54, p. 101251.

Nwulu, N., Oroja, S. & İlkan, M., 2012. A comparative analysis of machine learning techniques for credit scoring. *International Information Institute*, October, 15(10), p. 4129–4145.

Khemakhem, S. & Boujelbene, Y., 2017. Artificial intelligence for credit risk assessment: artificial neural network and support vector machines. *ACRN Oxford Journal of Finance and Risk Perspectives*, November, 6(2), p. 1–17.

Trinkle, B. S. & Baldwin, A. A., 2007. Interpretable Credit Model Development via Artificial Neural Networks. *Intelligent Systems in Accounting, Finance and Management*, July, 15(3–4), p. 123–47.

Merćep, A. & et al., 2020. Deep Neural Networks for Behavioral Credit Rating. *Entropy*, December, 23(1), p. 27.

Boguslauskas, V. & Mileris, R., 2009. Estimation of credit risk by artificial neural networks models. *Inžinerinė Ekonomika*, October, 64(4), p. 7–14.

Khashman, A., 2010. Neural Networks for Credit Risk Evaluation: Investigation of Different Neural Models and Learning Schemes. *Expert Systems with Applications*, September, 37(9), p. 6233–39.

Teles, G. & et al., 2020. Artificial Neural Network and Bayesian Network Models for Credit Risk Prediction. *Journal of Artificial Intelligence and Systems*, 2(1), p. 118–32.

A.H., C. & et al., 2022. Machine learning and artificial neural networks to construct P2P lending credit-scoring model: A case using Lending Club data. *Quantitative Finance and Economics*, January, 6(2), p. 303–325.

Adewusi, A. O. & et al., 2016. Application of Artificial Neural Network to Loan Recovery Prediction. *International Journal of Housing Markets and Analysis*, June, 9(2), p. 222–38.

Bahnsen, A. C. & Gonzalez, A. M., 2011. Evolutionary Algorithms for Selecting the Architecture of a MLP Neural Network: A Credit Scoring Case. s.l., s.n., pp. 725-732.

Oreski, S. & Oreski, G., 2014. Genetic Algorithm-Based Heuristic for Feature Selection in Credit Risk Assessment. *Expert Systems with Applications*, March, 41(4), p. 2052–64.

Kazemi, H. R. & et al., 2021. Tuning Structural Parameters of Neural Networks Using Genetic Algorithm: A Credit Scoring Application. *Expert Systems*, November, 38(7), p. e12733.

Xia, Y. & et al., 2017. A Boosted Decision Tree Approach Using Bayesian Hyper-Parameter Optimization for Credit Scoring. *Expert Systems with Applications*, July, Volume 78, p. 225–41.

Liang, D. & et al., 2015. The Effect of Feature Selection on Financial Distress Prediction. *Knowledge-Based Systems*, January, Volume 73, p. 289–97.

Hofmann, H., 1994. *Statlog (German Credit Data)*. s.l.: s.n.

Shami, T. M. & et al., 2022. Particle Swarm Optimization: A Comprehensive Survey. *IEEE Access*, Volume 10, p. 10031–61.