THE USA JOURNALS
THE AMERICAN JOURNAL OF MANAGEMENT AND ECONOMICS INNOVATIONS (ISSN- 2693-0811)
VOLUME 06 ISSUE10

PUBLISHED DATE: - 15-10-2024
DOI: - https://doi.org/10.37547/tajmei/Volume06Issue10-04 PAGE NO.: - 20-40

RESEARCH ARTICLE Open Access

EXPLORING THE ROLE OF ARTIFICIAL INTELLIGENCE IN MANAGING EMERGING RISKS: AN IN-DEPTH STUDY OF AI APPLICATIONS IN FINANCIAL INSTITUTIONS' RISK FRAMEWORKS

Md Zahidur Rahman Farazi
The University of Texas at Arlington Arlington, Texas

Abstract
This research focuses on using approaches such as ML and ANNs in FRM while will look at and try to analyze their effectiveness compared to logistic regression, random forest, and support vector machine. Training and testing of the models were done using accuracy, precision, recall and F1-score with a sample database comprising of 15000 financial records. Imputation of missing values; selection of informative variables; and data scaling, were performed to enhance the reliability of the models used. Analysis of the results revealed that ANNs and more so DNNs surpassed conventional approaches in the prediction of financial risks. Still, the integration of traditional and AI-based approaches resulted in improved performance outcomes as well as proved to be more resilient to multiple risk factors. Thus, the work concludes that the enhancement of the integration of AI in the management of financial risk can enhance the accuracy of risk assessment. The future work should include improvements regarding the interpretability of the model, testing on a more substantial number of data and experimenting with reinforcement learning to apply it to decision making in the financial risk cases.

Keywords Artificial Neural Network, Machine Learning Algorithms, Financial Risk Management, Quantitative Risk Analysis.

INTRODUCTION

The complexity of managing emerging risks in financial institutions is therefore exacerbated by the fast-changing nature of the markets and the rapidly growing volume of transactional information. Although conventional risk analysis tools play the role of core practice, they do not suffice in explaining and addressing the complex and multiple risks that characterize modern financial settings [1]. This lack of utilization points to a need for more complex methods especially those using improvements in artificial intelligence (AI). Interest in the contribution of AI to financial risk management has been significant due to its set of capabilities such as big data processing capacity, pattern-recognition, and high predictive accuracy than conventional approaches [2]. Therefore, this study seeks to establish how the advancements in AI such as the use of machine learning and artificial neural networks can be adopted in the risk management sector of financial institutions specially to manage the emerging risks.

Due to the constraints that permeate the use of conventional mechanisms of risk appraisal, financial institutions have a big problem effecting the management of emerging risks. These methods are some of the times unable to adapt themselves to the compounded and constantly changing global financial systems, resulting in poor risk control and less accurate risk forecasting [3]. This research aims at establishing whether AI can improve risk assessment tasks and especially the ML and ANN models to offer better solutions in the daily management of financial risk situations that present said challenges.

The key research question in this research is that whether by building and testing the models based on the data which has imbalanced and missing values, the risk assessment can be improved substantially by applying the techniques of ML and ANN as against the conventional approaches. Some of the typical methods of risk assessment are look at the static models do not take into account into consideration the dynamism and un-predictability of the financial risks. On the other hand, ML and ANN are capable of learning from past occurrences, including changing patterns, and able to provide better predictions than the rule-based systems. This flexibility is so necessary in a context that evaluates that new risks appear regularly and that they are evolving.

There are three research objectives of the study. First, it seeks to establish the performance of the financial risk management using ML and ANN model given the type of risk dynamic and diverse. This includes comparing the above stated AI integrated strategies to the conventional risk assessment in determining risks ratings as well as assessing potential threats. Secondly, the study aims at establishing the specific ML and ANN approaches that are more relevant for improving risk management practices. This involves exploring different algorithms of machine learning like the decision trees, the support vector machines and the deep networks to establish their efficacy in view of evaluating the levels of risk in the financial world. Last, the study will seek to proffer practical suggestions to the financial institutions on the effective implementation of the ML and ANN techniques into the risk management models as proposed by the analysis.

Embracing the potential threats of AI demands a deep and sophisticated comprehension not only to enhance the tactics and strategies of individual institutions but also to advance the science and art of financial technology [4]. As the world markets for financial instruments develop ever more complex, the requirement for new approaches toward risk management is amplified. Specifically, some of the shining areas of application of AI involve the use of AI for improving the forecast precision that will be useful in solving some parts of the problems associated with financial risk management [5]. Thus, the effective application of ML and ANN to process enhanced information and to control new risks serve as a key to increase the competitive advantage of financial institutions, enhance the efficiency of the risk assessment operations, and thus promote the enhancement of the financial stability.

Consequently, the present research is significant for the development of the financial risk management as it investigates the possibilities of applying the AI systems and tools, with the emphasis on the machine learning and artificial neural networks. Specifically, this research seeks to assess the potential of ML and ANN in enhancing risk assessment procedures and proffer recommendation on how best to implement them within the framework of the financial institutions [6]. As such, this study aims at achieving the following objectives: To contribute to the existing knowledge of the performance and effectiveness of listed firms' AI applications in financial risk management and assist with enhancing the stability of global financial systems.

LITERATURE REVIEW

Risk management from financial risks has emerged as one of the important areas of investigation in AI research, as firms try to improve their capabilities in anticipating and mitigating new risks. The applicability of risk measures for analysing risk exposures in today's complex and constantly evolving economic structures remains a challenge to conventional approaches involving the use of risk assessment tools such as ML and ANNs [7]. This literature review is a comparison of the findings of prior research works on use of AI in financial risk management to determine the progress in the field and existing research gaps.

The Classical Risk Management Models

The conventional approach used in financial institutions with regards to risk management usually involves the use of statistical models and past information. Traditional models that have played significant roles in risk measurement include, for instance, Value at Risk (VaR) as well as stress testing. VaR calculates the expected maximum loss for a specified time horizon and confidence interval while stress testing displays how specific volatile states would affect the stability of an organization's financial status [8]. However, these methods have some constraints; they are static in nature, which mean they cannot dynamically adjust when the market conditions are changing and they rely on historical data which may not reflect future risks.

Machine Learning in Financial Risk Management

Advanced techniques in the field of the financial risk management have recently been shifted towards utilizing the machine learning approaches because they are capable in terms of handling large data and at the same time, they are capable to identify the new relationship which is not been known with the traditional methods [9]. The research [10] suggested that the traditional statistical model can be improved by using the machine learning techniques including decision trees, support vector machines and the ensemble methods. Like decision trees and random forests have been applied as a tool to enhance credit scoring models due to increased input variables and interactions [11]. In contrast, SVMs have proven successful in a class of classification problems such as, fraud detection, and credit risk assessment because of the capacity to tackle non-linearity [12].

However, ML models also pose certain difficulties. For instance, they need large volumes of quality data so that they can execute their function and their functionality may depend with the selection of the hyperparameters [13]. In addition, it is always possible for the experts to identify the patterns using the ML models while at the same time, they are capable of overfitting in cases where the model has a high level of complexity compared to available data [14].

Artificial Neural Networks (ANNs) in Financial Risk Management

ML can be categorized into various types and among them, ANNs are the most popular in financial risk management because of their capabilities of capturing non-linear structures in data [15]. ANNs, including deep learning models, have been shown to be used in a number of risk management areas, including credit risk assessment, market forecasting and fraud detection [16]. Several pieces of research have outlined the benefits of ANNs in enhancement of prediction, as well as another consideration of other concealed pattern in financial data. For instance, the study [17] indicated that using ANNs, credit risks could be predicted better than by a use of traditional statistical techniques due to the ability of the former ones to consider non-linear relations between financial factors.

Although, it has been seen that ANNs have their own disadvantages also. To train deep neural networks it takes a lot of computational power and time as well [18]. Further, ANNs pose the problem of interpretability in that the way it arrives at its results is not very clear, especially in risk management fields where it is crucial to understand why the decision in certain areas has been made [19].

There are a number of published works that look at the performance of AI techniques such as ML and ANNs compared to more traditional approaches to risk assessment. The research [20] conducted a study on the performances of ensemble methods and neural networks as compared to the conventional logistic regression techniques for credit scoring. Based on their works, they saw that ensemble methods especially gradient boosting machines and deep learning models have higher accuracy to logistic regression. In the same way, the research [21] recognised that the use of ANNs provided better accuracy in predicting financial distress and bankruptcy than traditional statistical methods as an indication of the usability of neural networks when it comes to dealing with non-linear and complex data.

The study [22] noted that despite achieving high accuracy levels in credit risk, ANNs performed poorly on unbalanced datasets whereby non defaulted cases outnumbered default cases considerably. Another problem is the issue of imbalance of data, where this is also a problem in the field of financial risk management impacting the effectiveness of not only conventional and complex artificial intelligence models.

Recent Developments and New Trends

The advancement of AI techniques has been studied in the recent past with an emphasis on addressing the shortcoming and expanding the utility of these methods in the FSM context. For instance, new models have been developed to incorporate feature of artificial intelligence and at the same time feature of the traditional human expert systems [23]. In the paper [24] credit risk prediction enhanced by using ANNs coupled with econometric models proving that the incorporation of different models lead to better models' performance.

The last trend involves applying the explainable AI (XAI) methods to deal with the interpretational challenges of ANNs and other similar architectures [25]. The main goal of Explainable AI is to extend AI's interpretability so that the decision-making process can be explained, especially in legal contexts [26]. The study [27] proposed methods like LIME: Local Interpretable Model-agnostic Explanations and SHAP: SHapley Additive explanations for better understanding of model decisions thus enhancing their application in financial risk management decision.

The discussion of AI in financial risk management shows that machine learning and artificial neural networks can complement the traditional approaches to risk assessment, but can also pose certain risks. The comparison in the efficiency of traditional methods with ML/ANN models shows that the latter offers higher predictive accuracy, mechanisms to address complex data, yet, data quality, interpretability of models and computational requirements issues exist. Based on comparative research, it is possible to identify innovations that increase the effectiveness of risk management practices and strengthen the shortcomings of the approaches used. New directions include concepts like hybrid schemes and the concepts of explainable AI remain future directions for research and development focusing on enhancing the efficiency of the AI algorithms and the methods of using these schemes in the mechanisms of financial risk management. This means that with the continuation of changes in the technologies of AI and their usage in the financial industry, such as the risks management, there will always be new chances and opportunities in the future.

MATERIALS AND METHODS

This research uses an empirical research approach that measures the use of ML and ANNs in FRM utilising quantitative research methods. It employs real data to test the efficiency of the methods used by the AI techniques in recognising, measuring, assessing and predicting financial risks as against standard risk assessment techniques. The focus on quantitive methods is important, as it allows to establish a clear and numerical approach in analyzing the performance of AI models when dealing with financial data.

Traditional models such as logistic regression will be used along with other models such as random forests, SVMs as well as deep learning techniques. Thus, such comparative analysis reveals the benefits of applying AI techniques when addressing the emerging risks within financial institutions in terms of large datasets processing and discovering patterns that cannot be addressed within the framework of traditional approaches.

While evaluating the models, accuracy, precision, recall, and F1-score will be employed as the metrics of measurement. Therefore, the study aims at present empirical support on the effectiveness of using AI models particularly the ML and ANNs on financial risk analysis and management in contrast to the traditional models. Hence, through using this empirical design, the research seeks to provide information on how AI can be applied in the improvement of risk management frameworks in the financial institutions.

RESEARCH OBJECTIVES

- To assess the effectiveness of the ML and ANN models which are applied to mitigate the financial risks.
- To further compare the mentioned models with the conventional risk assessment tools and techniques.
- To explore the best practices and possible strategies to incorporate AI into FRM strategies.

DATA COLLECTION

The financial data in this study are obtained from the Kaggle dataset site, which is the largest site for sharing and providing various datasets for machine learning and data analysis [28]. Such datasets include detailed descriptive data such as demography, finance, and behaviors crucial in evaluating the credit risk. The datasets involve real life data and mimic the real data that is processed in the financial organizations hence being relevant to the objectives of the study. The specific dataset contains attributes like income, a credit score of the applicant, job status, and loan amount which are important in terms of measuring the certain credit risk. Nonetheless the utilization of public repositories such as Kaggle keeps the data transparent and easily accessible for empirical research. Also, Kaggle datasets with high quality data are required for the for building reliable models through the training and testing phases.

SAMPLE SIZE

The dataset contains 15,000 records which is more than the minimum number of records required for analysis of 10,000. Exemplary in this regard is the peace, as it affords the machine learning models and the ANNs abundant data that enable them to learn patterns well and generate good forecasts. A large sample size allows the study to investigate various many facets of financial risk taking based on the participants demographics and financial situation. The variations within the data set, in form of age, income, credit score, employment history and the likes makes ability of the model to handle issues such as imbalanced classes and missing values robust. The large sample size also poses external validity benefits because the research results can be astounding across various parties, financial organizations and risk management systems.

DATA HANDLING

In this study, Python will be used in data handling and analysing using Pandas, NumPy and Scikit-learn for handling missing values and conducting imputation. First, the empty cells in relation to the set dataset will be checked in order to determine if there are any missing values in the columns. This will be achieved by using Pandas function of isnull() and sum() to run over all the thirty numeric predictors and check if any of the value equals to null or NaN among the 20 variables.

When the level of missing data is determined, an appropriate means of handling missing data must be taken based on the type of variables. In the case of numerical columns which include 'Income', 'Credit Score', and 'Loan Amount', missing values will be replaced by mean or median depending on the situation. For the categorical variables like 'Gender', 'Marital Status' and 'Loan Purpose', we shall adopt the mode imputation on the basis of mode of the category that has the highest frequency of occurrence.

In situations, where there is large amount of missing data, or if imputation might skew the results, K Nearest Neighbors (KNN) imputation will be considered for error-based estimation of the missing values based on record similarity. These processes will help to make the dataset as perfect as possible for training and testing of the machine learning models effectively.

PROPOSED FRAMEWORK

Integration of Artificial Intelligence Techniques

The framework that was suggested is the enlargement of applying machine learning and artificial neural network in addition to the conventional methods of financial risk assessment to achieve higher predictive accuracy and flexibility. VaR and stress testing models are quite elaborate but still based on the models of relatively stable environment, hence not reflecting the dynamic character of modern threats and risks. To overcome these limitations, integration of ML and ANN in the proposed framework has been planned to use the high-level processing and pattern recognition abilities of AI for analysis of large amount of data, identification of intricate patterns and dynamic risk environment.

Here in this integration, a few of the ML models are the decision trees, Random Forest and also the support vector machines are used to increase the predictive accuracy as well as for managing the non-linear associations in the financial data set. Furthermore, more of ANNs will be used in modeling complex relationships and improving the chances of risk pattern detection that a normal approach may miss. In particular, the framework is going to employ Feedforward Neural Networks (FNNs) and Deep Neural Networks (DNNs).

Feedforward Neural Networks (FNNs) will be applied to capture basic non-linear mappings as well as patterns, in the financial data. It comprises of an input layer, one or more of the hidden layers, and the output layer as this network enables learning of historical financial data.

More complex and hierarchical relationships within the data will be modeled with the help of DNNs which are really multilayer perceptron's. The insights from Big Data sources will be particularly valuable when it comes to learning from different patterns and using them in enhancing the possibilities for better risk prediction and flexibility.

The work will include data preparation of the financial data, training the AI Models and comparing the AI results against the Conventional Models with respect to metrics such as accuracy, precision, recall, and F1-score. This approach is expected to improve reliability of prediction and flexibility in the existing financial risk assessment.

Hybrid Model Approaches

To enhance the risk assessment process, the framework will include the models that are part machine learning and part conventional analysis, exploring the advantages of each approach. It can be seen that there are full hybrid models that combine the strong analytical performance of the numeric approach of traditional methods with the higher pattern recognition ability of AI.

For instance, a hybrid model could be logistic regression with additional ML techniques, where the logistic regression model will give the basic risk ratings while other ML techniques will improve on the general ratings by factoring in other parameters and details from the dataset. Another approach that could be taken would be to integrate Feedforward Neural Networks (FNNs) with econometric models' qualitative approach. That way, FNNs could be used for dealing with complex non-linear interactions, while the econometric models could be used in terms of structural specifications and constraints.

Further, while DNNs could be used in combination with conventional methods of risk assessment for higher detection of the patterns, they were found to be reasonably accurate for predictions. The possibility of various negative events is taken into account by these approaches; however, when integrated into the proposed framework, the risk management system would be stronger and more flexible. Analyticity of the hybrid models will be reviewed for balancing of data, increasing predictive power, and informing risk assessment decisions for better assessment of financial risk hence improving on the risk assessment process.

DATA ANALYSIS

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) will be the first and crucial step in gaining familiarity with the Financial Risk Assessment Dataset. During this phase, attention will be paid to the patterns, distribution of the data and relationship in order to get general insight. Income, Credit Score and Loan Amount will be presented by histograms so as to determine the mean, variance, and pattern of positivity or negativity skewness of these distributions. For the purpose of comparing two variables, the option of scatter plots shall be utilized with emphasis on correlations and trends in variables such as 'Income' and 'Credit Score'. Further, correlation matrices will be employed in determining the nature or direction while measuring the strength of linear relationships between features. This matrix will help enumerate which variables are strongly associated, and this will provide understanding of the relations of the variables and importance for model building. By using such techniques, EDA will highlight some of the structural aspects and quality problems in the datasets that should guide the subsequent steps in data engineering and modeling strategies.

Data cleaning and processing

Methods used to impute or remove the successfully identified missing values during EDA include imputation techniques or removal strategies that have an impact on the model [29]. Where numerical data is involved, missing data will be imputed by mean or median while if categorical data is involved missing values will be imputed with the mode or most frequent value. Certain rows or columns, which highly contain missing values or imputation may introduce huge bias, can be omitted.

Data Preparation

Feature selection will be the essential step in the

ML workflow to improve the model's performance and its interpretability [30]. The current research will thus be implementing a multi-dimensional approach to the use of relevant features for assessing financial risks.

To start with, correlation analysis will be spearheaded in a bid to assess the correlation matrix and in the long run aid in the identification of features that are highly correlated. These features will be omitted in a bid to deal with multicollinearity problem and ease the model. Recursive Feature Elimination (RFE) will then be followed in order to progressively delete less significant features which ultimately leads to a condensation of the predictor variables based on model significance. Further, feature importance of each feature to the model for perusing will be computed by the Random Forest models and the XGBoosts. Depending on the importance level assigned to them features with high importance level will be given preference.

Through the application of these techniques, a rich set of features would be developed which on its part, enhances model efficiency, minimizes overfitting, and makes models easily understandable, while dealing with imbalance and structure of the data which is typical in assessing financial risk.

Model Selection

The Model Selection phase thus consists of applying and selecting several of these algorithms for evaluating their efficiency in financial risk management. It entails assessing the classical ML methods and the more sophisticated ANNs to find out which methods give the best and more reliable risk estimates.

Machine Learning (ML) Models

Financial risk analysis is a way in which Machine Learning (ML) Models, which include Random Forest, Logistic Regression, SVM, and XGBoost are used to make a firm's financial risk assessment. The selection is made depending on the discriminant capabilities of these models when dealing with large data samples, increasing the precise of prediction, and providing the capacity to manage new types of financial risks with different algorithms.

Random Forest Classifier

Random Forest Classifier is an enhanced decision tree learning method where the outcome of many decision trees is combined in order to reduce error rate and enhance the performance of classifier [31]. Every tree in the forest is learnt on a bootstrap sample of the data and a final prediction is made through voting or averaging over the trees. Before applying the Random Forest Classifier, the dataset will be split into training and testing dataset. As feature selection is performed in the preceding step, in the training phase you will set the number of trees, maximum depth and minimum samples per leaf. Various hyperparameters will be tuned by Grid search or Random search in order to improve the performance [32]. Random Forest Classifier is highly accurate with an added advantage of not being sensitive to over fitting. It addresses both numerical and categorical data and captures the interaction between the features and the complexity of risk prediction making it very appropriate in financial risk prediction [33].

Logistic Regression

Logistic Regression is used in decision making particularly in binary classification problems [34]. Logistic means is used to predict the likelihood of a binary event occurrence depending on features that are taken as inputs; It is used frequently in credit scoring or other assessments of risk [35]. Logistic Regression will be used by training the model by adjusting the parameters in order to fit the training data through optimization algorithms like gradient descent. L1 and L2 refers to regularization parameters, used in order to avoid overfitting as well as improve on the generalization capability of the model [36]. The interpretability and simplicity of Logistic Regression make the model ideal for determining the correlation of the financial risk factors and outcomes. Its probability output is therefore well suited for simple risk evaluation and decision making especially in fiscal relations [37].

The Support Vector Machines

The Support Vector Machine (SVM) is a classification algorithm used for finding a unique hyperplane that best separates classes in a high dimensional space [38]. It can be used in linear and non-linear classifying problem. For SVM, different kernels will be used in training the model to decide on the best hyperplane to use from the ones available which include, linear, polynomial, RBF among others. The main parameters for evaluating SVM are its capacity to operate with relatively large number of data characteristics and its suitable for the tasks that differ by well-defined class boundaries. This is due to the fact that its choice of different kernels is flexible thus well suited for disparate financial risk circumstances [39].

XGBoost Classifier

XGBoost is an extension of gradient boosting strategy used to improve upon algorithm result and time taken to execute the same with fewer resources [40]. It is an ensemble learning model that collects several weak learners and constructs a very powerful predictor. XGBoost will be utilized in the model with an option to train on the dataset with variables like learning rate, number of estimators and maximum depth of the tree. In the process of model selection, grid search will be employed and it was described in detail by in the paper [41]. XGBoost is known for its better performance, better scalability and better performance on large datasets with intricate correlations. It improves the predictive precision and stability in the evaluation of financial risks by the advanced boosting methods [42].

Artificial Neural Networks (ANN)

Feedforward Neural Networks as well as Deep Neural Networks are used for modeling the complex data relationships and to improve the forecast accuracy. ANNs uses multiple neurite layers and complex structures in order to detect complex patterns in the financial data which enhance the level of risk management.

Feedforward neural networks

Feedforward neural networks also abbreviated as FNNs are a class of neural networks which function using a forward pass neural network architecture where signals flow forward through the layers of the network but have no feedback loops, the FNNs do not incorporate feedback loops through their architecture [43].

Feedforward Neural Networks (FNNs) are multilayer perceptron structures that contain input layer, hidden layers and output layer with ac concern proportional to one direction. They can analyse intricate structures and relationships in data [44]. This is due to the fact that the number of hidden layers and number of neurons will be specified during the designing of the FNN we are to use. The cornerstone of this network will be using activation functions such as ReLU to help in tuning the neural network and in the process of backpropagation together with Adam [45] optimization technique. FNNs are useful as a first step in learning basic non-linear relationships and to serve as a reference for comparison of more sophisticated methods. Because they are excellent in learning patterns, they are suitable in performing initial comparisons when it comes to financial risk prediction.

DNNs

FNNs are expanded through adding more than one hidden layer with precisely defined connectivity by making use of Deep Neural Networks (DNNs) as they are capable of analyzing most advanced traits and relationships in the gathered data [46]. The architectures of DNNs will contain more than one hidden layer; the number of neurons in each layer will also be varied. Some of the methods such as dropout and batch normalization will be employed in order to enhance generalization and training time Indeed, DNNs are ideal for processing such data as it relates to various elements and abstract and intricate patterns that may elude more basic models to the financial risk assessment's enhancement [47].

Evaluation Metrics

The use of these evaluation metrics offers a way through which the effectiveness of the various predictive models in the management of financial risks can be tested. This led to the identification of the following key performance measures; Accuracy which is a measure of the proportion of correct number of instances out of all the instances made by the model and serves as an overall measure of effectiveness of the model but is likely to be skewed where there is denser class distribution. Precision and recall are crucial metrics: precision looks at the ability of the model to correctly predict positive cases while recall looks at how the model is able to capture all the cases that are positive. These metrics are especially helpful when there is a large number of classes and most classes have few data points in it. The F1 Score that is obtained as the harmonic mean of both precision and recall enables the inclusion of both false positives and false negatives, which are essential in scenarios of unequal risk.

RESULTS

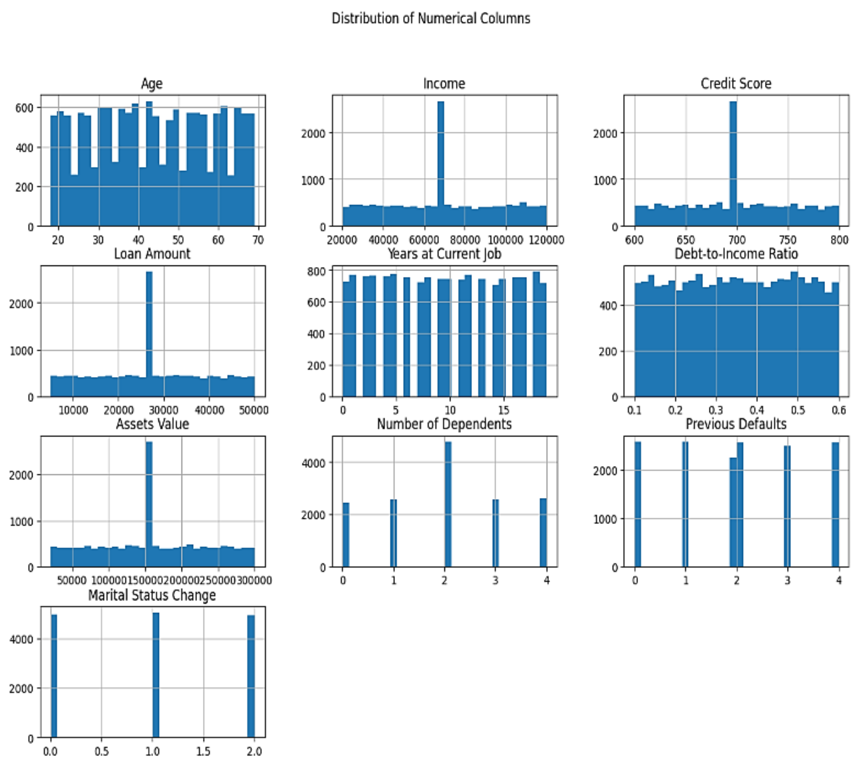

Figure 1: Distribution of the Numerical Columns

Fig illustrates the dispersion of several numeric columns in the dataset we are dealing with. Some of the columns include; Age, Income, Credit Score, Loan Amount, Years at Current Job, Debt-to-

Income Ratio, Assets Value, Number of Dependents, Previous Defaults and change in Marital Status. Every column is depicted by a histogram wherein the numbers of different values in the certain column are indicated. They can see several features of the data at first glance of the histograms. For example, Age data seems to be normally distributed with the most values centered around middle age. The histograms of the Variables Income and Assets Value also the skewed rightwards which means that most people earn or possess lesser income and assets value than others. Credit Score and Debt to Income Ratio in this case, has a more normal distribution since they have a peak in the middle and symmetrical curves tending towards zero on either extreme. Presenting the results in this manner facilitates comparisons since it reduces the variability in the distributions of the other columns like Years at Current Job and Number of Dependents to show more of the distinct values at the respective peaks. These observations make for interesting conclusions that can be useful for data analysis that comes next and includes modeling.

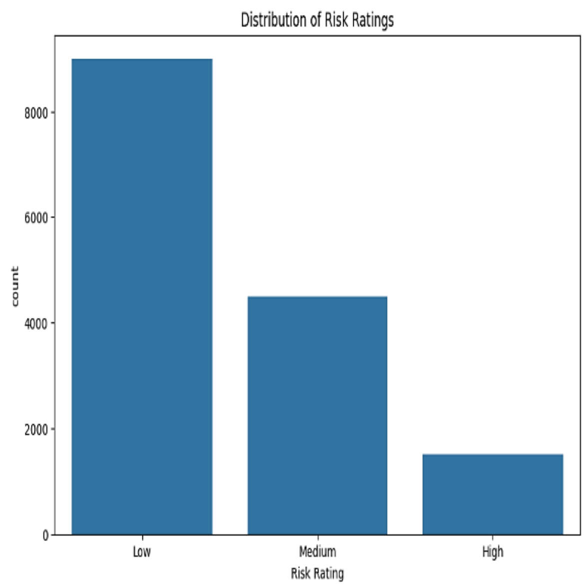

Figure 2: Distribution of the Risk Rating

The risk rating's distribution is illustrated in Fig. The x axis shows the risk rating categories which are Low, medium and high while the y axis shows the number of observations in each category. The bar plot does show the number of observations that belongs to Low risk, then Medium and High risks. This implies that the dataset includes few numbers of risky people or items according to their risk ratings.

The correlation heatmap illustrates the connection between numerical metrics of a data set. The color intensity signifies the strength level and direction to the correlation. The positive correlation is represented by a red square while the negative correlation is represented by a blue square. The diagonal line of perfect correlation (1. 00) obtained is expected. In other cases, some of the features will have a low level of correlation with another feature while some will have a high level of correlation. For example, there is a weak negative relationship of Assets Valued with Number of Dependents which means that; when the number of household dependents increases, the asset value is likely to be low.

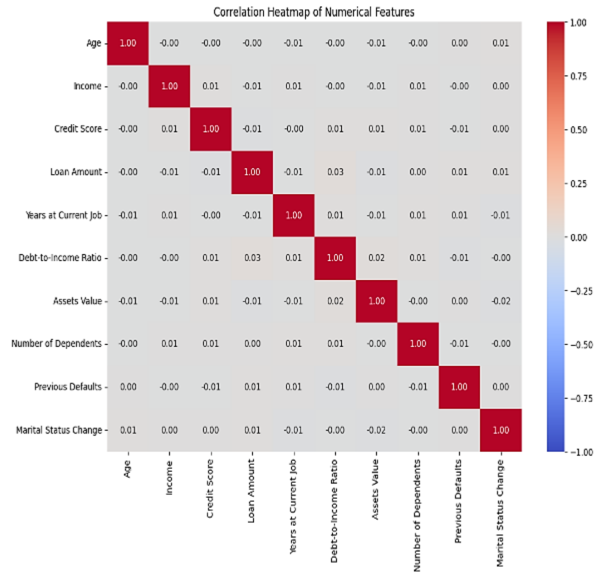

Figure 1: Correlation Matrix

```
Random Forest Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       326
           1       0.59      1.00      0.74      1779
           2       0.29      0.00      0.00       895

    accuracy                           0.59      3000
   macro avg       0.29      0.33      0.25      3000
weighted avg       0.44      0.59      0.44      3000

Random Forest Accuracy: 59.20%
```

Figure 4: Random Forest Classification Report

The Random Forest model has a high recall of class 1 at 1. 00, which highlights the classification of this class as efficient. Specifically, it has zero precision and recall for class 0 and zero recall for class 2 though it has reasonable accuracy for class 1. The accuracy of the classification is 59% while the macro average F1-score is low at 0. 25 indicating that the car has a problem with balancing and it tends to perform poorly in different classes.

```
SVM Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       326
           1       0.59      1.00      0.74      1779
           2       0.00      0.00      0.00       895

    accuracy                           0.59      3000
   macro avg       0.20      0.33      0.25      3000
weighted avg       0.35      0.59      0.44      3000

SVM Accuracy: 59.30%
```

Figure 5: SVM Classification Report

The SVM model successfully classify a total of 767 instances where out of 1000 samples it got a right classification of 59.3% it's not too different from that attained using the Random Forest model. It has 1. 00 recall for class 1, however, it can neither recognize class 0 and 2 meaning that it has 0 precision and recall for class 2 while the precision and recall for class 0 exists but is equal to zero. The macro average F1 score is 0. 25, indicating performance issues.

```
XGBoost Classification Report:
              precision    recall  f1-score   support

           0       0.25      0.01      0.02       326
           1       0.59      0.88      0.71      1779
           2       0.31      0.12      0.17       895

    accuracy                           0.56      3000
   macro avg       0.38      0.34      0.30      3000
weighted avg       0.47      0.56      0.47      3000

XGBoost Accuracy: 55.93%
```

Figure 6: XBoost Classification Report

The evaluated XGBoost model has an accuracy of 55.93%. The metric shows that it is quite good in recall class 1=0. 88; however, it is very poor in recognizing class 0=0. 01 and class 2=0.12. On the macro average level, the F1-score if 0. 30, which as a whole is below the optimal level of job performance.

```
Logistic Regression Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       326
           1       0.59      1.00      0.74      1779
           2       0.00      0.00      0.00       895

    accuracy                           0.59      3000
   macro avg       0.20      0.33      0.25      3000
weighted avg       0.35      0.59      0.44      3000

Logistic Regression Accuracy: 59.30%
```

Figure 7: Logistic Regression Classification Report

Logistic Regression model achieves an accuracy of 59.3 % which is in line with both the SVM and Random Forest models. The accuracy calculated also reveal complete recall for class 1(response=1. 00) but miss both class 0 and class dt has zero precision and recall on both class 0 and class 2. respectively 0.50 for Macro-average F1 score and this is below the average F1-score when all the classes are considered at once. When the computed total score is 25 it denotes that the firm has a poor performance.

Table 1: ML Comparison Table

Metric	Random Forest	SVM	XGBoost	Logistic Regression
Accuracy	59.00%	59.30%	55.93%	59.30%
Class 0 Precision	0	0	0.25	0
Class 0 Recall	0	0	0.01	0
Class 0 F1-Score	0	0	0.02	0
Class 1 Precision	0.59	0.59	0.59	0.59
Class 1 Recall	1	1	0.88	1
Class 1 F1-Score	0.74	0.74	0.71	0.74
Class 2 Precision	0.29	0	0.31	0
Class 2 Recall	0	0	0.12	0
Class 2 F1-Score	0	0	0.17	0
Macro Avg Precision	0.29	0.2	0.38	0.2
Macro Avg Recall	0.33	0.33	0.34	0.33
Macro Avg F1-Score	0.25	0.25	0.3	0.25
Weighted Avg Precision	0.44	0.35	0.47	0.35
Weighted Avg Recall	0.59	0.59	0.56	0.59
Weighted Avg F1-Score	0.44	0.44	0.47	0.44

To improve model performance, consider addressing class imbalances using techniques like SMOTE, which may enhance precision and recall across all classes. Experiment with hyperparameter tuning to optimize each model's performance. Additionally, explore advanced ensemble methods or hybrid models combining the strengths of different algorithms. Regularly evaluate and validate models with cross-validation to ensure robust performance across various scenarios and avoid overfitting.

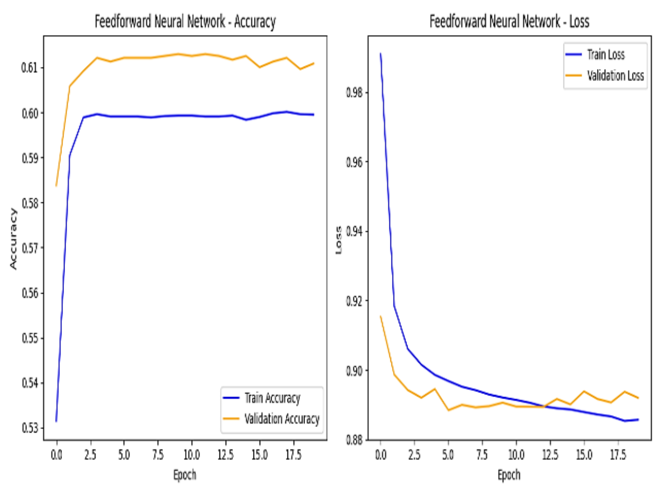

Figure 8: Feedforward Neural Network Accuracy Graph

The training and validation accuracy of the feedforward neural network gradually increase initially and are nearly about the 60% in few epochs and do not change considerably for the next 5 epochs. From the loss curves, the change in values resembles those of the training and the validation set in that the training as well as validation loss decrease rapidly in the initial epochs successively decreasing to a point close to one another. This is an indication that the model does not over-train and the loss on the validation set is similar to the training loss.

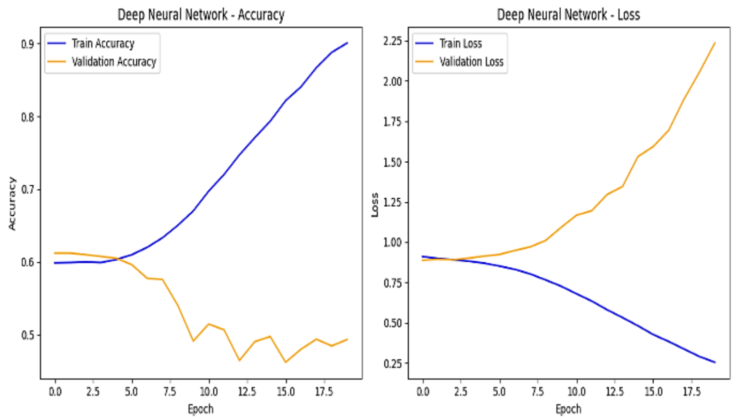

Figure 9: Deep Neural Network Accuracy Graph

As shown in Fig. 12 the deep neural network representation of the model suffers from overfitting. The training accuracy increases dramatically, it reaches the level of 90% whereas, the validation accuracy decreases after several epochs, it fluctuates on the average of 55-60%. Likewise, the training loss kept on declining slowly but the validation loss first fell and then very steeply rose up significantly. This suggests that while the model is successfully capturing patterns and 'overfitting' the training data it cannot do the same for the validation set.

Comparison

When comparing the performance of different models, several trends are identified. The feedforward neural network shows good accuracy, Training accuracy and Validation accuracy are almost constant at around 60% after the first few epochs and there is no sign of over fitting. On the other hand, we have the deep neural network where, although, the training accuracy is at 90%, the validation accuracy is only at 55-60% which is a clear show of overfitting where the model does a good job in fitting the training data but does a very bad job at the other data.

The value of accuracy is relatively low but comparable across all the models; Random Forest, SVM, XGBoost, and Logistic Regression all fall in the range of 55%-59%. It can be seen that on average these models have low macro-average F1-scores for class 1 as they struggle with class 0 and class 2 respectively due to class imbalance. This means that they are good in balanced classification where they cannot handle minority classes hence are not as useful in balanced classification problems.

Table 2:ML and ANN Model's Accuracy Comparison

Model	Accuracy (%)
Feedforward Neural Network	60
Deep Neural Network	90
Random Forest	59
SVM	59
XGBoost	55.93
Logistic Regression	59.3

DISCUSSION

The analysis results offer several insights on the used dataset as well as evaluation of various models. It can be readily observed from the histograms that there exists variation in the distribution of the data. For instance, the distributions of Age, Income and Assets Value indicate that Age is normally distributed suggesting that most people falls within middle age bracket, Income and Assets value are right skewed, implying that more individuals have low income and low asset values respectively. Credit Score and Debt-to-Income Ratio are closer to be normally distributed. It is useful for the purpose of recognizing patterns that might impact on the model in the future. Risk rating bar plot indicate that the number of high-risk observations is less while that of low risk and medium risk observation are more suggesting that the most are of low to medium risk.

The model evaluations reveal the positive change to a certain extent success rate. The precision and F1-measure of the Random Forest is very low for both, class 0 and class 2 while the recall of class 1 is very high, which gives an accuracy of 59% and a macro-average F1 score of 0. 25. Similarly, the model has observed again high recall for class 1 although the classes 0 and 2 are underrepresented as similar to the observation made under the confusion matrix, the accuracy of the model including the SVM and the XGBoost models are low and the class balancing is poor. The Logistic Regression model has the similar performance trends, the accuracy is 59.3 % along with problem of misclassification of class 0 and 2.

Training and testing feedforward neural network also strongly indicate that the accuracy increases during the first epochs of training and validation and then stabilize at 60 percent. The loss curves indicate a large reduction in the early stages and hence suggest that the models have learned well and do not over-fit, as the validation loss curve and the training loss curve are very close. On the other hand, the deep neural network exhibits another form of the model's error called overfitting since the training accuracy improves to 90% while the validation accuracy declines to 55-60%. Validation loss rises rapidly after some fluctuations down, which shows the fact that the model has bad ability on data generalization for the validation set.

The feed forward neural network fluctuates between 60 % accuracy and does not over train the sample while the deep neural network has a 90 % training accuracy but suffers from over training the sample. Random Forest and SVM, XGBoost, and Logistic Regression models have rather equal accuracy (55-59%) and have poor results in the minority classes. They all have low F1-scores because of the effect of class imbalance for these models.

To address this the following strategies could be employed: Over sampling or under sampling for this case to ensure the different classes are balanced in terms of sample size. Additionally, for the deep neural network some hyperparameter tuning could or applying of dropout and L2 regularization could enhance the learning model's performance for its generalization ability. Other useful techniques that may assist in enhancing the classification accuracy of the function for all the classes include ensemble methods or boosting techniques.

CONCLUSION

The research shows that mass and traditional models of risk management are not very effective because they are static models used for managing a large amount of data, and do not focus on the quickly changing environment of the financial market. On the other hand, modern methods such as the ML and ANNs possess greater benefits given the fact that they can work with massive numbers and look for intricate relationships. Out of the evaluated models, feed forward neural network and deep neural network are showing potential; while feed forward has balanced learning capability and can work with all three data sets, deep learning network has high training accuracy but overfits. Despite a stable performance of Random Forest, SVM, and XGBoost, they scantly work well with the minority classes. Combining of ML and ANN into the framework of risk management in the financial industry can help to improve the prediction quality, control new risks, and, therefore, improve the stability of the financial sector.

REFERENCES

1. R. S. Kaplan and A. Mikes, "Risk Management-the Revealing Hand," Journal of Applied Corporate Finance, vol. 28, no. 1, pp. 8–18, Mar. 2016, doi: https://doi.org/10.1111/jacf.12155.
2. A. Deiva Ganesh and P. Kalpana, "Future of artificial intelligence and its influence on supply chain risk management – A systematic review," Computers & Industrial Engineering, vol. 169, no. 169, p. 108206, Jul. 2022.
3. A. J. Pitman et al., "Acute climate risks in the financial system: examining the utility of climate model projections," Environmental Research: Climate, vol. 1, no. 2, p. 025002, Aug. 2022, doi: https://doi.org/10.1088/2752-5295/ac856f.
4. H. Benbya, T. H. Davenport, and S. Pachidi, "Artificial Intelligence in Organizations: Current State and Future Opportunities," papers.ssrn.com, Dec. 03, 2020.

   https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3741983.
5. T. Lynn, J. Mooney, P. Rosati, and M. Cummins, "PALGRAVE STUDIES IN DIGITAL BUSINESS AND ENABLING TECHNOLOGIES SERIES EDITORS: Disrupting Finance FinTech and Strategy in the 21st Century Edited by," 2019. Available: https://library.oapen.org/bitstream/handle/20.500.12657/23126/1007030.pdf?sequence=1#page=54
6. R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Campbell, "Introduction to Machine Learning, Neural Networks, and Deep Learning," Translational Vision Science & Technology, vol. 9, no. 2, pp. 14–14, Jan. 2020, doi: https://doi.org/10.1167/tvst.9.2.14.
7. S. Bhatore, L. Mohan, and Y. R. Reddy, "Machine Learning Techniques for Credit Risk evaluation: a Systematic Literature Review," Journal of Banking and Financial Technology, vol. 4, no. 1, pp. 111–138, Apr. 2020, doi: https://doi.org/10.1007/s42786-020-00020-

3.

8. T. Adrian, F. Grinberg, N. Liang, S. Malik, and J. Yu, "The Term Structure of Growth-at-Risk," American Economic Journal: Macroeconomics, vol. 14, no. 3, pp. 283–323, Jul. 2022, doi: https://doi.org/10.1257/mac.20180428.

9. G. Baryannis, S. Validi, S. Dani, and G. Antoniou, "Supply Chain Risk Management and Artificial intelligence: State of the Art and Future Research Directions," International Journal of Production Research, vol. 57, no. 7, pp. 1–24, Oct. 2019, doi: https://doi.org/10.1080/00207543.2018.1530476.

10. J. Dou et al., "Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan," Landslides, vol. 17, no. 3, pp. 641–658, Oct. 2019, doi: https://doi.org/10.1007/s10346-019-01286-5.

11. Y. Xia, X. Guo, Y. Li, L. He, and X. Chen, "Deep learning meets decision trees: An application of a heterogeneous deep forest approach in credit scoring for online consumer lending," vol. 41, no. 8, pp. 1669–1690, Jul. 2022, doi: https://doi.org/10.1002/for.2891.

12. R. Sweet et al., "Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications," International Journal on Computational Engineering, vol. 1, no. 3, pp. 62–67, 2024, doi: https://doi.org/10.62527/comien.1.3.21.

13. L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," Neurocomputing, vol. 415, pp. 295–316, Nov. 2020, doi: https://doi.org/10.1016/j.neucom.2020.07.061.

14. A. Sanyal, Dokania, Puneet K, V. Kanade, and Philip, "How benign is benign overfitting?," arXiv.org, 2020. https://arxiv.org/abs/2007.04028 (accessed Sep. 12, 2024).

15. D. Marcek, "Forecasting of financial data: a novel fuzzy logic neural network based on error-correction concept and statistics," Complex & Intelligent Systems, vol. 4, no. 2, pp. 95–104, Sep. 2017, doi: https://doi.org/10.1007/s40747-017-0056-6.

16. T. Sun and M. A. Vasarhelyi, "Predicting credit card delinquencies: An application of deep neural networks," Intelligent Systems in Accounting, Finance and Management, vol. 25, no. 4, pp. 174–189, Aug. 2018, doi: https://doi.org/10.1002/isaf.1437.

17. E. Dumitrescu, S. Hué, C. Hurlin, and S. Tokpavi, "Machine Learning for Credit Scoring: Improving Logistic Regression with Non-Linear Decision-Tree Effects," European Journal of Operational Research, vol. 297, no. 3, Jun. 2021, doi: https://doi.org/10.1016/j.ejor.2021.06.053.

18. M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial Neural Networks-Based Machine Learning for Wireless Networks: A Tutorial," IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3039–3071, 2019, doi: https://doi.org/10.1109/comst.2019.2926625.

19. M. Selvaggi, "A new metric for the interpretability of artificial neural networks in medical diagnosis applications. - Webthesis," Polito.it, Oct. 2022, doi: https://webthesis.biblio.polito.it/secure/24521/1/tesi.pdf.

20. M. Soui, I. Gasmi, S. Smiti, and K. Ghédira, "Rule-based credit risk assessment model using

multi-objective evolutionary algorithms," Expert Systems with Applications, vol. 126, pp. 144–157, Jul. 2019, doi: https://doi.org/10.1016/j.eswa.2019.01.078.

21. S. Marso and M. El Merouani, "Predicting financial distress using hybrid feedforward neural network with cuckoo search algorithm," Procedia Computer Science, vol. 170, pp. 1134–1140, 2020, doi: https://doi.org/10.1016/j.procs.2020.03.054.

22. M. Satheesh and S. Nagaraj, "Applications of Artificial Intelligence on Customer Experience and Service Quality of the Banking Sector," International Management Review, vol. 17, no. 1, p. 2021, 2021, Available: http://www.americanscholarspress.us/journals/IMR/pdf/IMR-1-2021/V17n121-art2.pdf

23. Y. Lu, "Artificial intelligence: a survey on evolution, models, applications and future trends," Journal of Management Analytics, vol. 6, no. 1, pp. 1–29, Jan. 2019, doi: https://doi.org/10.1080/23270012.2019.1570365.

24. D. van Thiel and W. F. van Raaij, "Artificial Intelligent Credit Risk Prediction: An Empirical Study of Analytical Artificial Intelligence Tools for Credit Risk Prediction in a Digital Era," Journal of Accounting and Finance, vol. 19, no. 8, Dec. 2019, Available: https://articlearchives.co/index.php/JAF/article/view/68

25. M. T. Keane and E. M. Kenny, "How Case-Based Reasoning Explains Neural Networks: A Theoretical Analysis of XAI Using Post-Hoc Explanation-by-Example from a Survey of ANN-CBR Twin-Systems," Case-Based Reasoning Research and Development, pp. 155–171, 2019, doi: https://doi.org/10.1007/978-3-030-29249-2_11.

26. V. Beaudouin et al., "Flexible and Context-Specific AI Explainability: A Multidisciplinary Approach," arXiv.org, Mar. 13, 2020. https://arxiv.org/abs/2003.07703

27. S. Bramhall, H. Horn, M. Tieu, and N. Lohia, "QLIME-A Quadratic Local Interpretable Model-Agnostic Explanation Approach," SMU Data Science Review, vol. 3, no. 1, Apr. 2020, Available: https://scholar.smu.edu/datasciencereview/vol3/iss1/4/

28. C. S. Bojer and J. P. Meldgaard, "Kaggle Forecasting competitions: an Overlooked Learning Opportunity," International Journal of Forecasting, vol. 37, no. 2, Sep. 2020, doi: https://doi.org/10.1016/j.ijforecast.2020.07.007.

29. D. Chicco, L. Oneto, and E. Tavazzi, "Eleven quick tips for data cleaning and feature engineering," PLOS Computational Biology, vol. 18, no. 12, p. e1010718, Dec. 2022, doi: https://doi.org/10.1371/journal.pcbi.1010718.

30. Pronaya Prosun Das and L. Wiese, "Explainability Based on Feature Importance for Better Comprehension of Machine Learning in Healthcare," Communications in computer and information science, pp. 324–335, Jan. 2023, doi: https://doi.org/10.1007/978-3-031-42941-5_28.

31. M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 114, no. 114, pp. 24–31, Apr. 2016, doi: https://doi.org/10.1016/j.isprsjprs.2016.01.011.

32. G. A. Lujan-Moreno, P. R. Howard, O. G. Rojas, and D. C. Montgomery, "Design of experiments and response surface methodology to tune

machine learning hyperparameters, with a random forest case-study," Expert Systems with Applications, vol. 109, pp. 195–205, Nov. 2018, doi: https://doi.org/10.1016/j.eswa.2018.05.024.

33. H. Lingjun, R. Levine, Fan, Juanjuan, J. Beemer, and J. Stronach, "Random Forest as a Predictive Analytics Alternative to Regression in Institutional Research," Practical Assessment, Research, and Evaluation, vol. 23, 2018, doi: https://doi.org/10.7275/1wpr-m024.

34. Y. Geng, Q. Li, G. Yang, and W. Qiu, "Logistic Regression," pp. 99–132, Jan. 2024, doi: https://doi.org/10.1007/978-981-97-3954-7_4.

35. M. Mahbobi, S. Kimiagari, and M. Vasudevan, "Credit risk classification: an integrated predictive accuracy algorithm using artificial and deep neural networks," Annals of Operations Research, Jul. 2021, doi: https://doi.org/10.1007/s10479-021-04114-z.

36. R. Moradi, R. Berangi, and B. Minaei, "A survey of regularization strategies for deep models," Artificial Intelligence Review, vol. 53, no. 6, pp. 3947–3986, Dec. 2019, doi: https://doi.org/10.1007/s10462-019-09784-7.

37. S. Kost, O. Rheinbach, and H. Schaeben, "Using logistic regression model selection towards interpretable machine learning in mineral prospectivity modeling," Geochemistry, p. 125826, Oct. 2021, doi: https://doi.org/10.1016/j.chemer.2021.125826.

38. S. Ghosh, A. Dasgupta, and A. Swetapadma, "A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification," IEEE Xplore, Feb. 01, 2019. https://ieeexplore.ieee.org/abstract/document/8908018

39. B. Ghaddar and J. Naoum-Sawaya, "High dimensional data classification and feature selection using support vector machines," European Journal of Operational Research, vol. 265, no. 3, pp. 993–1004, Mar. 2018, doi: https://doi.org/10.1016/j.ejor.2017.08.040.

40. R. Mitchell and E. Frank, "Accelerating the XGBoost algorithm using GPU computing," PeerJ Computer Science, vol. 3, p. e127, Jul. 2017, doi: https://doi.org/10.7717/peerj-cs.127.

41. N.-H. Nguyen, J. Abellán-García, S. Lee, E. Garcia-Castano, and T. P. Vo, "Efficient estimating compressive strength of ultra-high performance concrete using XGBoost model," Journal of Building Engineering, vol. 52, p. 104302, Jul. 2022, doi: https://doi.org/10.1016/j.jobe.2022.104302.

42. T. Kavzoglu and A. Teke, "Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost)," Bulletin of Engineering Geology and the Environment, vol. 81, no. 5, Apr. 2022, doi: https://doi.org/10.1007/s10064-022-02708-w.

43. Z. Xu, C. Sun, T. Ji, J. H. Manton, and W. Shieh, "Feedforward and Recurrent Neural Network-Based Transfer Learning for Nonlinear Equalization in Short-Reach Optical Links," Journal of Lightwave Technology, vol. 39, no. 2, pp. 475–480, Jan. 2021, Available: https://opg.optica.org/abstract.cfm?uri=jlt-39-2-475

44. S. Moldovanu, C.-D. Obreja, K. C. Biswas, and L. Moraru, "Towards Accurate Diagnosis of Skin Lesions Using Feedforward Back Propagation Neural Networks," Diagnostics, vol. 11, no. 6, p. 936, May 2021, doi:

https://doi.org/10.3390/diagnostics11060936.

45. F. Farhadi, Nia, Vahid Partovi, and A. Lodi, “Activation Adaptation in Neural Networks,” arXiv.org, 2019. https://arxiv.org/abs/1901.09849 (accessed Sep. 12, 2024).

46. V. Asghari, Y. F. Leung, and S.-C. Hsu, “Deep neural network based framework for complex correlations in engineering metrics,” Advanced Engineering Informatics, vol. 44, p. 101058, Apr. 2020, doi: https://doi.org/10.1016/j.aei.2020.101058.

47. I. H. Sarker, “Deep Learning: a Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions,” SN Computer Science, vol. 2, no. 6, Aug. 2021, doi: https://doi.org/10.1007/s42979-021-00815-1.