

**VILNIUS GEDIMINAS TECHNICAL UNIVERSITY**

**FACULTY OF FUNDAMENTAL SCIENCES**

**dEPARTMENT OF INFORMATION SYSTEMS**

**ChuXian Chen**

**Application of machine learning methods for business process decision-making**

**Master Graduation Thesis**

**Supervisor: Prof. Dr. Diana Kalibatienė**

**Vilnius, 202****5**

**Table of Content**

[1. Introduction 4](#_Toc201003014)

[1.1. Investigation Object 4](#_Toc201003015)

[1.2. The Aim and Tasks of the Thesis 4](#_Toc201003016)

[1.3. Novelty of the Topic 5](#_Toc201003017)

[1.4. Relevance of the Topic 6](#_Toc201003018)

[1.5. Research Methodology 6](#_Toc201003019)

[1.6. Scientific Value of the Thesis 7](#_Toc201003020)

[1.7. Main Results of the Thesis 7](#_Toc201003021)

[1.8. Structure of the Work 7](#_Toc201003022)

[2. Related Works Analysis 8](#_Toc201003023)

[2.1. Key concepts and definitions 16](#_Toc201003024)

[2.2. Systematic Review of Articles 18](#_Toc201003025)

[2.3. Review of Datasets for the loan 24](#_Toc201003026)

[2.4. Models compared 26](#_Toc201003027)

[2.5. Summary of the 2nd section 28](#_Toc201003028)

[3. Proposed Approach 30](#_Toc201003029)

[3.1. Main Results of the 3rd Section 35](#_Toc201003030)

[4. Initial experiment 36](#_Toc201003031)

[4.1. Dataset describing 36](#_Toc201003032)

[4.1.1. Descriptive statistics 36](#_Toc201003033)

[4.1.2. Correlation analysis 37](#_Toc201003034)

[5. Implementation of the Proposed Approach 39](#_Toc201003035)

[5.1. Main Results of the 4th Section 39](#_Toc201003036)

[6. General Conclutions 40](#_Toc201003037)

[6.1. Further works 40](#_Toc201003038)

[7. References 42](#_Toc201003039)

**List of Images**

[Figure 3‑1. UML component diagram - Loan approval system 32](#_Toc200971479)

[Figure 3‑2. General Fairness Loan Approval System 32](#_Toc200971480)

[Figure 3‑3. Sub-process Fairness Constraints setup 33](#_Toc200971481)

[Figure 3‑4. Sub-process Loan Data Collection 34](#_Toc200971482)

[Figure 3‑5. Sub-process Loan Data Preprocessing 34](#_Toc200971483)

[Figure 3‑6. Sub-process Integrate fairness and ML model 35](#_Toc200971484)

[Figure 3‑7. Sub-process User Interface Building 36](#_Toc200971485)

**List of Tables**

[Table 2. Summary of research papers 20](#_Toc201003040)

[Table 3. Review of Datasets for the loan 25](#_Toc201003041)

[Table 4. Model compared 27](#_Toc201003042)

[Table 5. Descriptive statistics for selected dataset 36](#_Toc201003043)

[Table 6. Correlation analysis for selected dataset 38](#_Toc201003044)

**Abbreviations**

ML – Machine Learning

SOA – Service Oriented Architecture

BPM – Business Process Management

# Introduction

Machine learning (ML) over the last few years has been making a strong influence on decision-making in a very wide range of industries, but particularly on finance, where it proved to possess a capability to change loan approval forecasting. Traditional loan approval methods like credit scoring and rule-based approaches do not fare too well handling complexity and volume of data nowadays. Although machine learning methods offer more advanced solutions, they raise significant issues of fairness since patterns of discrimination in historical data can be encoded in the models. Unless the models are properly developed, they can perpetuate discrimination on the basis of sensitive demographic information such as race, gender, or income, thereby leading to discriminatory lending(Choi et al., 2025). Such prejudices are not only ethically reprehensible, but also undermine confidence among financial institutions, especially if these systems are used in determining high-risk credit access decisions(Hedrick et al., 2024). Despite growing uses of machine learning methods in the financial industry, existing models tend to dismiss biases in past data, leading to discriminatory decisions (Ramachandra et al., 2021). Particularly, the use of data on education level, marital status, and gender, which are sensitive but not technically to be considered as bases of loan approval, may unconsciously skew approval outcomes and produce ascription-based bias. These kinds of biases undermine the validity and reliability of automated systems and raise serious ethical and legal issues. Additionally, the lack of a standard yardstick to develop machine learning models for lending approval to ensure fairness has left research as well as practical applications with lacunas (Weinzierl et al., 2024). Incorporating fairness constraints into the prediction models remains an unexplored area, further highlighting the need for methods that can remove such biases while maintaining predictive accuracy.

## Investigation Object

The object of the research is machine learning models applied to loan approval prediction, with a special focus on improving the fairness of the loan decision-making process.

## The Aim and Tasks of the Thesis

The aim of this research is to improve the fairness of loan approval prediction models by developing a machine learning based decision support system that combines predictive analyses with fairness constraints. The following work will be carried out to achieve the objectives of this thesis:

1. To analyse the fairness of loan approval forecasting models, techniques and tools.

2. To propose an approach to incorporate fairness-aware strategies into the development of predictive models for loan approvals.

3. To implement the proposed approach into a prototype and perform its evaluation with a real loan dataset.

## Novelty of the Topic

The originality of this thesis is its focus on the fairness of machine learning-based loan approval prediction systems. The fairness aspect is less explored compared to the highly studied prediction accuracy. Although the use of machine learning is growing in credit risk assessment and loan approval, much of the literature has focused on performance metrics and less on the degree to which such models can inadvertently reinforce biases against disadvantaged demographic groups such as gender, race, or income (Choi et al., 2025). There is no benchmark framework to make machine learning-based financial decision-making systems accurate and fair, highlighting a key research gap. Traditional credit scoring systems rely on human judgment and inflexible rule-based logic and typically lack the flexibility and objectivity to handle complex data sets. Despite advances in scalability and accuracy, machine learning's black-box nature and reliance on historical data have engendered ethical concerns and regulatory challenges. This study circumvents these limitations by emphasizing fairness and evaluation of loan approval models (Hare, 2023; Purificato et al., 2023). Specifically, this thesis proposes a comprehensive solution that integrates fairness-aware mechanisms (e.g., fairness constraints and bias mitigation algorithms) into loan approval decision-making machine learning models to ensure their predictions are not only technically accurate but also ethical. As opposed to previous research that has considered fairness as an afterthought, this research prioritizes it and discusses how fairness metrics and tools can be integrated formally into the process of model building. In addition, this thesis explores the use of cutting-edge machine learning techniques, including hybrid models that combine integrated learning and adversarial training, which remain unexploited in fairness-aware financial forecasting (Lee et al., 2024). Implementing the models on real-world financial datasets adds a practical aspect, revealing deployment challenges and real-world fairness trade-offs (Hedrick et al., 2024). By addressing these essential problems, this research adds to more accountable and transparent applications of AI in finance and proposes a new agenda for bridging the divide between technological innovation and ethical accountability.

## Relevance of the Topic

The growing application of Machine Learning (ML) in the banking sector presents great opportunities as well as new challenges, primarily in the area of loan approval prediction. Traditional decision-making patterns have a tendency to make decisions based on rigid credit scoring mechanisms and historical data that can include systemic mistakes resulting in discriminatory practices against disadvantaged or vulnerable populations. On the other hand, machine learning algorithms are able to analyze intricate, big data (e.g., behaviour and other financial metrics) to arrive at faster, more accurate decisions. These algorithms, however, can also perpetuate and reinforce existing imbalances (Choi et al., 2025) in the absence of fairness constraints. All this has made fairness in algorithmic decision-making an utmost priority for banks, especially in the face of increasing regulatory requirements and community pressure. This paper is relevant to this shifting context in that it investigates perhaps the most emotive issue in AI use in financial services: discrimination risk embedded in automated lending approval processes. By exploring predictive model fairness and probing into actionable tools and techniques for managing bias, this research contributes to more ethical and more equitable financial technologies (Lee et al., 2024). Besides, fairness is not merely a legal or ethical imperative-it also provides long-term fiscal well-being through trust, transparency, and enhanced credit availability, especially to historically disadvantaged populations (Hare, 2023). Beyond values, this study adheres to the business goals of financial institutions. Machine learning is able to assess creditworthiness more accurately, thereby speeding up and streamlining the lending process, avoiding human error, and improving overall risk management (Hedrick et al., 2024). All these benefits are only possible, however, when fairness and interpretability are built into model design. Therefore, this study introduces a fairness constraint-based dataset to the configuration of a machine learning model applied for loan approval, and it is a fairness-aware model that reduces bias when approving loans and increases the attainment of fairness.

## Research Methodology

This study utilises literature research to identify and classify existing applications of machine learning in loan approval prediction. Comparative analyses were conducted to assess the strengths and limitations of various machine learning models including integrated approaches, deep learning and fairness orientated techniques. To develop the methodology, logical induction and conceptual modelling were employed along with an experimental approach using real-world financial datasets to test the proposed model.

## Scientific Value of the Thesis

This thesis provides a scientific contribution through a structured exploration of fairness in machine learning-based loan approval systems, with each core task having a unique scientific value.

1. By analysing relevant literature on the application of machine learning to loan approval and fairness, reviewing relevant datasets as well as machine learning algorithmic models, this thesis provides a comprehensive integration of existing methods, highlighting knowledge gaps.

2. This study proposes a new approach that combines fairness and predictive modelling to improve the fairness of loan approvals by fusing data related to the fairness of loan approvals into a machine learning model.

## Main Results of the Thesis

This thesis presents several key results that contribute to understanding and improving the fairness of machine learning-based loan approval systems.

1. First, we provide a detailed review of the relevant literature, as well as relevant datasets and machine learning algorithmic models, to identify the limitations of existing loan prediction models, finding that many systems lack mechanisms to ensure fairness and reduce bias, as well as laying the initial groundwork for the proposal of a new methodology, including dataset and algorithmic model selection.

2. Second, we propose a fairness-conscious prediction framework that combines machine learning models with fairness constraints to address potential discrimination based on sensitive attributes such as educational background, gender, and marital status.

## Structure of the Work

The second part deals with the literature review and the analysis of existing methods; the third section describes the proposed methodology; and the fourth section shows the experimental design and test results.

# Related Works Analysis

The discussed paper takes into account the issue of fairness in loan approval prediction systems. This is regarding harmonizing machine learning (ML) techniques with fairness to maximize the fairness of the decision in loan approval. For example, some studies recommend hybrid architectures such as Random Forest and XGBoost (Jamunadevi et al., 2024; P. Kumar et al., 2024)in order to improve the predictive power, while others focus on fairness constraints as well as adversarial methods to reduce bias in credit decisions (Choi et al., 2025). Deep learning techniques such as BiLSTM (P. Kumar et al., 2024) and autoencoders have also been suggested for precise loan eligibility prediction from unstructured and structured data. Shared problems in such systems, however, are the complexity of integrating ML models with current loan approval mechanisms, interpretability, and dealing with biases in financial data. Equity issues, in particular with regards to demographic factors such as age, gender, and income, remain a recurring issue. The trade-off between the performance of the models and their fairness and computational overhead in scaling up models for deployment are important concerns for loan approval models. Enhancing the efficiency, fairness, and predictive accuracy of the models may be investigated by optimisation techniques in future studies.

In (Hedrick et al., 2024) , the authors propose a machine learning-based system for loan risk prediction. The study used six machine learning algorithms Decision Trees, Random Forests, Support Vector Machines (SVMs), Multi-Layer Perceptrons (MLPs), Plain Bayes and Stacking Ensemble models. The models are trained on a dataset of twenty features common in loan applications to predict loan risk. The Stacking Ensemble model had the best accuracy of 78.57% among all models compared, followed closely by the Random Forest model with an accuracy of 78.15%. The most important predictors of loan risk were determined through the analysis, such as credit limit, check status, age, loan term, and loan purpose. The study provides definitive evidence of the usefulness of machine learning in the loan approval process with greater speed and precision than traditional methods. The authors note, nonetheless, that the added computational demands of integrating the models may not justify the performance difference. Future work entails the optimisation of such models and addressing issues of data bias in loan approval systems.

In (Sandeep & Devi, 2022), the authors compare the performance of two machine learning algorithms - Random Forest (RF) and Linear Regression (LR) - for predicting bank loan approvals. Customer data, including credit score, were used in the research to predict credit risk. The RF model outperformed the LR model with a mean accuracy of 70.5 per cent compared to 69.5 per cent for the LR model. Statistical t-test analyses indicated that the accuracy difference was not statistically significant with a p-value of 1.0. This research emphasizes the efficiency of the RF model for enhancing the accuracy of loan approval decisions through the use of diverse customer background variables. The study concludes that despite the marginal performance boost of RF, it is more accurate compared to LR, especially for complex data such as credit score and financial information. Both models have limited accuracy when applied to real-time datasets, though. Future research may involve applying optimisation algorithms and feature selection algorithms to further improve prediction performance.

In(Lee et al., 2024), the authors suggest automating sound loan approval decisions using a submodular optimisation approach. The study uses a rule-based system aimed at optimising the loan approval process in Modern Capital Services. Machine learning and submodule optimisation techniques are combined by the system to create an interpretable yet effective decision-making process. The greatest challenge that the authors overcame was how to design the loan approval rules so that they were simple enough for human analysts to understand yet so specific that they would drive the business into profitability and restrict the risk. The system made 14% increase in annual servicing volume and a desirable delinquency rate. The sub-module optimisation technique ensured that the volume of loan versus the risk of sanctioning bad loans was properly balanced. The system was doing much better in rejecting bad risks and much more accurately compared to the previous method. Nevertheless, the system does have some drawbacks, such as balancing the interpretability and predictability of the model. Future work includes ongoing improvement of the optimisation strategy and investigation of more diverse sources of data.

In (Ramachandra et al., 2021), the authors present a machine learning loan approval prediction model and cloud platform. The study uses algorithms such as decision trees, logistic regression, and random forests to forecast loan approvals based on customer demographics. The model was deployed and hosted on AWS to ensure scalability and cost savings. The data set had 4,520 samples with 17 attributes, including marital status, credit history and educational status, which were preprocessed and filtered with several preprocessing approaches. The model's performance (specifically logistic regression) was mentioned to be 86 per cent, a robust solution for determining if the loan application would be approved or not. The study highlights the advantage to financial institutions of using cloud-based machine learning, particularly in dealing with huge volumes of data and providing real-time predictions. The study also highlights, however, that the model's performance largely depends on the quality of input data. Future improvements would involve integrating more advanced models or merging other data sources in a bid to improve accuracy.

In (Choi et al., 2025), the authors present a paper on improving fairness in finance AI models through constraint-based bias reduction. How bias in lending approval systems, often caused by sensitive variables such as gender, age, and nationality, may lead to discriminatory lending decisions is explored. To combat this, they propose two different constraint-based models: a fairness constraint and an accuracy constraint model, both of which are applied to logistic regression (LR) models. Using a contemporary financial dataset of the Korea Credit Information Service (KCIS), the authors show that the fairness constraint model significantly improves fairness metrics without decreasing accuracy significantly. For example, imposing fairness constraints leads to a sharp increase in the p% rule and decrease in decision boundary covariance (DBC), both of which are indicative of greater fairness. The study recognizes the importance of personalized data preprocessing, i.e., handling sensitive attributes carefully to generate more fair model outputs. The findings suggest that the incorporation of fairness constraints into finance AI models not only decreases bias in making decisions, but also leads to fairer loan approvals. Future studies can involve experimenting with such methods on a wider range of finance applications to better prove their effectiveness.

In (Jamunadevi et al., 2024), authors put forward a hybrid machine learning approach to predict loan approval status with Random Forest (RF) and XGBoost (XGB) models. The article highlights the challenge of increasing the precision and reliability of loan approval prediction, which is critical for banks in accurately assessing creditworthiness. The combined model combines RF and XGB, both of which leverage each other's strengths to enhance the predictive performance to a whopping 97% accuracy. The authors highlight the significance of preprocessing the data, such as handling missing values, feature selection and outlier detection, in order to enhance the accuracy of the model. The study also enunciates the imperative of exploratory data analysis (EDA) to understand prominent features such as income, loan amount, and CIBIL score, which are crucial for decision-making. Additionally, Recursive Feature Elimination (RFE) was employed for eliminating irrelevant features, thereby improving model generalisation and avoiding overfitting. Although the results are encouraging, areas such as dealing with unbalanced datasets and interpretability of the model need to be explored. Subsequent studies will likely entail further development of the integrated model and testing other machine learning techniques to increase loan prediction accuracy.

In (P. Kumar et al., 2024), the authors propose an integrated learning approach by combining Random Forest (RF) and XGBoost (XGB) to optimize loan approval decisions based on more precise credit scores. The research revolves around the use of machine learning techniques to enhance the process of credit scoring, which plays a significant role in determining the creditworthiness of applicants for loans. This is done through the application of numerous datasets containing financial, demographic and credit history information so that there is improved credit risk assessment accuracy. This paper shows the benefit of ensemble learning in that RF and XGB, together, cancel out each other's own limitations. The ensemble model is formed by taking the RF and XGB forecasts to arrive at better performance, increased precision and improved generalisability than either individual model. The authors further talk about feature selection methods, such as SHAP values for XGBoost, and highlight the importance of interpretability in credit score uses. While the combined model has promising performance, problems with handling big datasets and making the model interpretable still persist. Future work includes incorporating interpretable AI approaches and developing the system for more types of financial services.

In (El Annas et al., 2023), present a semi-supervised hidden Markov model (SSHMM) for rejection inference in peer-to-peer (P2P) credit scoring models. The research seeks to solve the issue of selection bias in credit scoring where rejected applicants are not considered during training of the model and only accepted applicants are considered. This type of bias is particularly important for platforms like Lending Club that reject a large number of applicants. To address this problem, SSHMM is suggested by the authors that integrates labelled (accepted) and unlabelled (rejected) data to improve model robustness. The procedure is a three-stage process: binning, filtering, and model training. Split-boxing is used for discretizing continuous values, filtering discards outliers, and model training employs the Baum-Welch algorithm for parameter estimation. Using experimentation on real data of Lending Club, the SSHMM model performs better in terms of accuracy, precision, and recall compared to traditional models with a wide margin. The performance is still acceptable even with an increase in the rate of rejection. SSHMM is particularly effective in using rejected applicants' information to improve the credit scoring model's predictive power. The study concludes that semi-supervised methods, particularly SSHMM, are a promising remedy to the problem of rejection inference and can enhance credit scoring models utilized in P2P lending websites.

In (Orji et al., 2022), machine learning algorithms are used to predict bank loan eligibility, a crucial finance activity. Six popular algorithms are presented in the research: Random Forest (RF), Gradient Boosting (GBM), Decision Tree (DT), Support Vector Machine (SVM), K Nearest Neighbours (KNN) and Logistic Regression (LR). The study utilized Kaggle's Loan Qualification Dataset, a dataset of loan applicants' demographic and financial data. The model was tested and trained following preprocessing, e.g., using a synthetic minority oversampling technique (SMOTE) to deal with category imbalance. The Random Forest model had the highest accuracy of 95.55%, which was significantly superior to the accuracy of the logistic regression model, which was 80%. The main contributions of this research include improved performance with the use of SMOTE, balanced data and improved model predictions. It performed better than most other studies, being more precise, recall, and accurate than other loan forecasting models. To sum up, machine learning, specifically combined methods such as Random Forest, is at the heart of automating and optimizing the loan approval process, thus allowing financial institutions to make more accurate and timely judgments.

In (A. P. S. P. Kumar et al., 2024), the authors recommend a deep learning method to mechanize the process of bank loan approval. The study emphasizes the utilization of bi-directional long short-term memory (BiLSTM) networks, which are highly effective in extracting sequential relationships in loan application data, for improving precision and efficiency. Traditional methods (Random Forests and Support Vector Machines, for instance) are often plagued by overfitting and underfitting, which can be mitigated by the deep learning techniques described in this study. The model uses an array of possible data (like credit score, income and homeownership) to predict loanworthiness. Preprocessing methods such as interpolation of missing values, one-hot thermal coding, and feature scaling were conducted to ready the dataset for use with the deep learning model. The BiLSTM model testing indicated exemplary performance, with accuracies ranging from 98% to 99% and F1 measures ranging from 99% to 99.4%. This shows that the model can effectively identify loan approvals with a very low rate of false positives and false negatives. In conclusion, the BiLSTM model significantly improves the efficiency of the loan approval process. The way forward will be improving interpretability and the inclusion of unstructured data for a comprehensive appraisal.

In (Dansana et al., 2024a), random forest regression modeling was used in investigating the impact of various loan attributes on bank loan approval. The study improved the decision-making process through the prediction of loan approvals based on applicant attributes. The impact of attributes such as gender, education level, marital status, employment type, loan purpose and monthly income on loan approvals was analyzed. The results of the research place a value on some traits: for example, higher levels of schooling and stable types of employment significantly improved odds of loan approval. In addition, analysis of the purposes of loans reveals that personal-purpose loans and tax loans are more likely to be approved than other uses such as business or school loans. The model performs very well and can effectively function with binary variables. The model can also effectively function with continuous variables and provide insights into the loan-approving process. Strong though it is, the research suggests that future research can use deep learning models to process larger datasets and enhance predictions further. The authors favor the inclusion of more advanced features, such as the dynamic financial behavior of applicants, in an attempt to enhance the predictiveness of the model.

In (Acharya et al., 2024), the authors provide an end-to-end method for implementing fairness and explainability in machine learning models, i.e., LightGBM and XGBoost, for predicting house prices and loan acceptance. The paper addresses fundamental fairness concerns using approaches such as calibrated equilibrium odds ratios and cross-fairness, yet not much applied in these types of applications. The authors focus on improving model interpretability through SHAP (SHapley Additive exPlanations), a feature importance interpretation method, and a new fairness-based interpretability technique. In their experiments, they show that LightGBM has an excellent trade-off between high accuracy and fairness compared to XGBoost on several metrics. The fairness approach implemented ensures that the model does not discriminate based on sensitive demographic features, including gender and income, yet maintains predictability. The findings of this study have broader governance and regulatory implications, underscoring the value of fairness and explainability of financial decision-making. In brief, the paper outlines the necessity of such models to meet the growing demands for ethical AI systems, particularly in sectors where the access of people to resources is disproportionately impacted.

In (Mourtas et al., 2024), the authors propose a bionic neural network called BWASD (Beetle Tentacle Search Weight and Structure Determination) to improve a loan approval classification model.Optimization of a loan approval classification model by introducing a meta-heuristic Beetle Tentacle Search (BAS) algorithm is proposed by the BWASD model.The BWASD model enhances a typical neural network (NN) by adding a meta-heuristic Beetle Tentacle Search (BAS) algorithm for improved weight determination and network structure for binary classification tasks such as loan approval prediction.The BWASD model enhances traditional backpropagation neural networks (BPNNs) through local minima avoidance and computational complexity reduction.The BWASD algorithm uses a new approach to find the optimal network weights and select the optimal number of hidden layer neurons and their respective activation functions. Experimentation on four loan-approval data sets such as the Kaggle data set established that BWASD performs better compared to other models such as MWASD, K Nearest Neighbours (KNN), Support Vector Machines (SVMs) and Random Forest. Specifically, BWASD shows the best performance in terms of accuracy, F-value and minimum mean absolute error (MAE) on the test set, and it is therefore an effective tool for financial decision-making systems. In conclusion, bio-inspired techniques like BWASD have the potential to significantly enhance credit risk assessment models and improve loan approval accuracy and efficiency. Future work will focus on implementing the BWASD model for other binary classification tasks and scaling up and broadening its application.

In (Alagic et al., 2024), the authors conducted a study on the use of machine learning to augment credit risk analysis via the integration of mental health data into predictive models for loan approval. The study investigated how mental health metrics like depression and anxiety affect the financial conduct of individuals, which can potentially decide their ability to repay loans. The study used two data sets: one of mental health and another of loan approval. Some models of machine learning, including XGBoost, Random Forest and KNN, were tested and compared to establish the most optimal model for use in loan approval prediction when including mental health data as a variable. It was found that XGBoost worked best in the mental health database (84%) and Random Forest in the loan approval database (85%). The evaluation also included precision, recall, and F1 scores, and XGBoost and Random Forest outperformed the other models. The authors point out that including mental health data in credit risk models can provide a better understanding of the borrowers, making the models fairer and reducing the risk of default. Conclusion, financial institutions ought to consider such information so that creditworthiness of applicants can be assessed better, especially with the impact of mental health on financial behavior.

In (Purificato et al., 2023), the paper addresses the usage of responsible Artificial Intelligence (AI) techniques within the process of approving loans with the objective of interpretability and fairness in order to attain trust and reliability. The article proposes an end-to-end system with the goal of managing the entire lifecycle of machine learning models used in loan approvals within a single platform through the incorporation of interpretability and fairness tools to enable users to understand and mitigate bias in decision-making. The system suggested includes datasets and machine learning (ML) model handlers, standardised interpretability tools, and fairness tools. A few of the key elements of the system are the use of model-independent interpretability methods, i.e., LIME, SHAP, and Anchors, that enable users to easily interpret ML model outcomes and render lending decisions transparent and simple. Fairness tools recognize biases in datasets and models and optionally eliminate these using independence-based re-weighting algorithms. This solution is in line with the European Code of Ethics for Artificial Intelligence and ensures that the models are compliant with the principles of fairness and non-discrimination. Finally, the system demonstrates how fairness and interpretability can increase trust in AI-based loan approval systems, proposes a novel trust vs. confidence measure to test the validity of the system's explanations, and emphasizes the importance of these ethical considerations in financial decision-making to ensure automatic systems are transparent and equitable for all customers.

Based on the works covered, this review paper covers some of the machine learning techniques applied in loan approval prediction and financial risk assessment. They include integrated models (Hedrick et al., 2024) deep learning models (P. Kumar et al., 2024) and fairness-driven frameworks (Choi et al., 2025). These techniques have significantly improved financial decision-making processes:

1. Improving the precision of loan approval through hybrid machine learning algorithms such as Random Forest and XGBoost that combine the strengths of several algorithms to improve predictive capability (Jamunadevi et al., 2024; P. Kumar et al., 2024).

2. Reducing bias in loan approval through a combination of fairness constraints and counterfactual methods such that demographic features such as gender and age are not disproportionate to loan approval decisions (Choi et al., 2025).

3. Enhancing interpretability of financial models using SHAP values and counterfactual explanations that explain why a loan approval decision was made and render machine learning systems more understandable and reliable (Purificato et al., 2023).

**Missing Solutions**

In spite of the progress offered by sophisticated machine learning algorithms towards improving prediction quality, a core deficit still exists in the explicit integration of fairness into loan granting systems. Much of the existing work up until now is concerned with model quality - usually accuracy, precision and recall - with fairness being an additive or afterthought. While a number of studies have begun to examine fairness-aware approaches, such as constraint-based modeling and post-hoc interpretability techniques (Choi et al., 2025; Purificato et al., 2023), a number of these approaches have been extensively evaluated across multiple datasets or real-world deployment settings. In addition, no consensus exists on what fairness measures have to be used and how they are to be balanced against predictive performance, especially when demographic attributes might be indirectly coded into agent attributes.

Another serious shortcoming is the lack of a unified framework that ties together fairness constraints, interpretability, and usability. There has been a tendency in most work to suggest standalone solutions (e.g., SHAP for explainability or adversarial training to remove bias), but none have attempted to construct end-to-end systems capable of performing all three -- detecting, interpreting, and mitigating bias in a way that is comprehensible to stakeholders. Along with this, while there are techniques like SHAP that improve model interpretability, they do not inherently guarantee fairness or address the underlying problems of discriminatory predictions.

Lastly, although existing studies have indicated more varied or unstructured data sources to be used, e.g., behavioural or social data, these studies are few in number and have not addressed how such data sources should be harnessed without increasing existing inequalities. Consequently, there is still a gigantic research gap in creating machine learning systems that are not merely technically sound, but also socially conscious and equitable. This thesis attempts to close this gap by proposing a procedure which systematically integrates fairness constraints into loan approval models, assisted by explainable techniques and tested using real financial datasets to ensure ethical, transparent and efficient decision-making.

## Key concepts and definitions

This research is dedicated to improving the equity of loan approval decisions using machine learning (ML) models. The traditional methods tend to perpetuate bias and lead to discriminatory decisions. This research improves the solution of the bias issue by using fairness-sensitive machine learning techniques to ensure loan approval decisions are not only accurate but fair as well. Therefore, in this section, we define the basic concepts related to loan decision recommender systems to enable familiarity with terms and methods used in this research. A loan decision recommendation system is a system that integrates data analytics, machine learning models, and a user interface that is capable of automating loan approval decision recommendations based on borrower data.

1. **Loan Decision Recommender System:** A loan decision recommendation system is a computerized system used to assist the financial institution in making decisions on loan applications. These systems leverage statistical models and machine learning techniques to consider many data points about an applicant, such as income, employment, financial history, credit score and usage of a loan, to predict the likelihood of repayment of the loan (Hedrick et al., 2024). These systems are typically used to support financial institutions to automate the lending decision, streamline the approval process, and give feedback to applicants in a quicker manner. A well-designed recommender system looks at many factors, including creditworthiness, personal financial details, and even unstructured data like transaction behavior, in order to recommend whether a loan should be approved or not. This technique can do much to reduce the time necessary to arrive at a loan decision, something that is important when dealing with a large volume of loan requests. However, despite their increased effectiveness, these systems must be tested rigorously to ensure that they do not actually perpetuate biases present within the data, such as demographic bias(P. Kumar et al., 2024). This highlights the need for fairness as well as interpretability mechanisms within such models.

2. **Credit Approval of Loan and Credit Risk Prediction:** Credit risk prediction is the manner in which the financial institutions examine the likelihood of a borrower defaulting. The procedure involves assessing various factors such as credit history, income levels, debt-to-income ratio, job stability and other fiscal behavior pertinent to them. Historically, credit risk assessment employed conventional credit scoring and expert judgment, but these tended to lack accuracy and did not acknowledge more subtle relationships in borrower behavior (Sandeep & Devi, 2022). Today, machine learning models are gradually taking the place of conventional models because they can recognize latent patterns and relationships in enormous sets of data. Models like decision trees, random forests and gradient boosters are capable of aggregating and analyzing multiple data sources (such as non-traditional data such as transaction history, online behavior, and even social media activity) to estimate the probability of default more accurately. Machine learning-based credit risk models also have the ability to adjust and improve their predictions through learning from new data, which allows them to react to changing economic conditions. However, the challenge is to offer interpretability and explainability to such models so that stakeholders are aware of why the predictions are being made. Further, preventing such models from discriminating on racial, gender, or socioeconomic bases is a significant role in upholding fairness and accuracy in credit risk predictions(Lee et al., 2024).

3. **Machine Learning Models:** Machine learning models are the key in constructing complex financial systems, especially loan decision-making. Historical data are the inputs that machine learning models use to learn patterns on which they can make decisions or prediction on new input data. Machine learning techniques such as decision trees, random forests, neural networks and support vector machines are applied in loan approval to categorize the applicants into different types of risk (e.g., payback probable or default probable) (P. Kumar et al., 2024). Machine learning's main advantage over the statistical techniques is its ability to manage huge amounts of data and to discover complex non-linear patterns in the data. This can facilitate more general assessment of loan applicants not just based on traditional financial information such as credit history, but also behavioral traits, expenditure habits and other indirect credit determinants. But there are enormous challenges to come for machine learning models in finance. One of the major issues is fairness, and we have to think about the ethical issues with the biases of these models in the attempt not to cause unfair results(Purificato et al., 2023).

4. **Fairness:** Fairness in machine learning models means that there is a moral responsibility on the part of decision-making algorithms (particularly those deployed in high-stakes fields like financial services) to treat everyone equally, devoid of their education level, marital status, and sensitive features like gender (Choi et al., 2025). In the instance of loan approval systems, equity ensures applicants are not discriminated against based on considerations that have no place in whether or not they are approved for loans. With banks and other financial institutions adopting machine learning for decision-making, the risk for algorithmic unfairness increases. Machine learning algorithms trained on historical data can learn and amplify biases in the data. As an example, consider that an algorithm is trained based on a set of information with historical data of specific groups being discriminated against or underserved. The model would then tend to decrease approval ratings for those groups, even if they were good credit risks (Purificato et al., 2023). In a bid to fight these biases, fairness-aware machine learning methods are on the rise. These methods include algorithms that modify the inherent decision-making by the model to reduce disparities among various demographic groups. Techniques such as adversarial debiasing, fairness constraints, and group fairness are some of the ways of ensuring that machine learning algorithms promote equitable outcomes. Not only is fairness an ethical necessity but also a legal need in most jurisdictions wherein discriminatory lending is subject to legal sanctions and disrepute for an institution.

The intersection point of all four concepts of loan decision recommendation systems, credit risk prediction, machine learning models and fairness is that they all perform the same function. Loan decision recommendation systems use machine learning models to analyze credit risk and make data-driven decisions instantly about a borrower's likelihood of repayment (P. Kumar et al., 2024). These models use machine learning algorithms to analyze the financial history of a candidate, credit report, and behavior. Even though these models can make rapid analysis of huge data sets with proper predictions, they end up inheriting bias from historical data that could lead to discriminatory or unfair results (Choi et al., 2025). That is where fairness comes in: ensuring that machine learning models do not discriminate based on sensitive characteristics like education level, gender, or marriage status. While current systems are designed to facilitate lending decisions more quickly and accurately, the fundamental problem of fairness is ignored. These models lack systematic fairness within them, including identifying and avoiding bias so that all who approach are treated fairly. Filling this gap is crucial to the end that computerized lending decisions are not only correct, but fair, transparent and ethical. Banks have to embed fairness-aware algorithms into their machine learning algorithms to ensure that their automated systems enjoy not only predictive accuracy but social responsibility.

## Systematic Review of Articles

This section highlights that although Machine Learning (ML) techniques have been proved to possess great capabilities to enhance the decision-making process in financial systems, challenges remain. Although hybrid machine learning models, and deep learning approaches have enhanced loan application processes, fairness constraints, and compatibility with existing frameworks remain to be addressed. In addition, maintaining fairness, especially with respect to demographic factors such as gender and income, remains a difficult job. These findings support the need for developing accurate and fair models that can improve decision support systems in uncertain financial environments. This review paves the way for conceptualizing a robust, ethical, and transparent ML-based loan approval system, described in subsequent parts of this thesis.

Table 2.1 below provides a comprehensive summary of the research papers reviewed, highlighting their methodologies, application areas, etc., highlighting the progress made and remaining gaps in the field. It consists of the following elements: 1) References, which provides references to the papers analysed; 2) Main research question / problem, which describes the main objective or problem in the study; 3) Methodology approach, which outlines the methodology and framework used; 4) Field Studied / Application domain, which specifies the application or area of study; 5) Dataset used, which identifies the data used in the analysis; 6) Attributes used for prediction, which details the considered characteristics or variables; 7) Evaluation of the approach, summarising the criteria or indicators used to assess the methodology; and 8) Results, highlighting the main findings or conclusions drawn from the study

Table 2. Summary of research papers

| **Reference** | **Main research question / problem** | **Methodology approach** | **Field Studied / Application domain** | **Dataset used** | **Attributes used for prediction** | **Evaluation of the approach** | **Result** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** |
| (Hedrick et al., 2024) | The paper addresses the challenge of predicting loan risk using machine learning models based on customers' personal factors. | Six machine learning models (Decision Tree, Random Forest, SVM, MLP Neural Network, Naive Bayes, Stacking Ensemble) were trained to predict loan risk. | Financial industry, specifically in loan approval and risk management. | Kaggle dataset ‘credit\_risk\_customers’ with 1,000 rows and 21 features. | Key features include credit amount, checking status, age, loan duration, purpose of the loan, and credit history. | Evaluation metrics include accuracy, precision, recall, F1 score, and confusion matrix. | The Stacking Ensemble model had an accuracy of 78.57%, and Random Forest showed good performance with 78.15% accuracy. The study identified credit amount, checking status, age, duration, and loan purpose as significant predictors. |
| (Sandeep & Devi, 2022) | The paper investigates the accuracy of Random Forest (RF) compared to Linear Regression (LR) for loan approval prediction using background checks. | The study compares the performance of Random Forest (RF) and Linear Regression (LR) algorithms for predicting loan risk based on background verification and credit score data. | Financial services, specifically in bank loan approval and credit risk assessment. | Credit card dataset from Kaggle, with a sample size of 40 per group. | Key features include credit score, income status, loan amount, credit history, and employment status. | Accuracy comparison, mean, standard deviation, t-test for significance. | Random Forest achieved a mean accuracy of 70.5%, while Linear Regression reached 69.5%. The statistical significance between the two models was found to be insignificant (p = 1.0). |
| (Lee et al., 2024) | The paper addresses the challenge of constructing interpretable and optimized loan approval rules that balance bad rate and loan volume. | The study uses a submodular optimization approach to optimize loan approval rules, leveraging XGBoost and combinatorial feature selection to build atomic rules. | Financial industry, focusing on loan approval systems. | Data from Hyundai Capital Services, with features related to credit scoring and customer financial behavior. | Features include credit score, delinquency history, income, and loan history. | The system is evaluated based on precision, recall, F1 score, and explainability of the rules. | The submodular optimization approach significantly outperformed the manually constructed baseline rules, achieving higher precision and maintaining a similar bad rate. The system contributed to a 14% annual growth in loan volume while improving the accuracy of loan approval decisions. |
| (Ramachandra et al., 2021) | The paper aims to develop a cloud-based machine learning model for predicting loan approval using customer demographic data. | The study uses machine learning algorithms such as Decision Tree, Logistic Regression, and Random Forest, deployed on the AWS platform for scalability and efficiency. | Financial services, focusing on loan approval prediction. | A bank dataset with 4520 records and 17 attributes. | Features include loan amount, education, gender, credit history, marital status, and others. | Evaluated using accuracy scores, confusion matrix, and comparison of models. | The model achieved an accuracy of 86%, with Decision Tree and Random Forest showing similar performance. The AWS cloud platform enabled efficient testing and training of the model. |
| (Choi et al., 2025) | The paper addresses how to mitigate biases in financial AI models, particularly in loan approval, while balancing fairness and accuracy. | The study applies fairness-aware machine learning techniques, specifically fairness- and accuracy-constrained models, to loan approval predictions. | Financial services, with a focus on loan approval prediction. | Korea Credit Information Services dataset (CreDB), along with synthetic datasets for testing fairness under different scenarios. | Key features include credit ratings, loan histories, delinquency records, age, gender, and nationality. | Evaluated using accuracy, p%-rule, and Decision Boundary Covariance (DBC). | The fairness-constrained model improved fairness with a p%-rule of 98% and DBC of 0.012, although it slightly reduced accuracy (0.75). The accuracy-constrained model maintained fairness with a high p%-rule (99%) but had lower accuracy (0.72). Tailored data preprocessing showed importance for specific attributes. |
| (Jamunadevi et al., 2024) | The paper addresses the challenge of improving the prediction of loan approvals using advanced ensemble machine learning models. | The study utilizes an ensemble approach combining Random Forest (RF) and XGBoost (XGB) models to predict loan approval outcomes. | Financial services, focusing on bank loan approval prediction. | Loan approval dataset (approximately 5000 records) from Kaggle. | Features include income, loan amount, CIBIL score, education, and self-employment status. | Evaluated using accuracy, precision, recall, and F1-score. | The ensemble model achieved an accuracy of 96.55%, with precision of 97.79%, recall of 89.11%, and F1-score of 94.60%. The model significantly outperformed traditional models like Logistic Regression (LR). |
| (P. Kumar et al., 2024) | The paper explores the optimization of loan approval decisions through an ensemble learning model for better credit scoring. | The study uses an ensemble approach combining Random Forest (RF) and XGBoost (XGB) to enhance credit risk assessment accuracy and model interpretability. | Financial services, specifically in credit scoring and loan approval. | Datasets including financial, demographic, and past credit information (sample data from Kaggle). | Key features include income, credit history, loan amount, employment status, and education level. | Evaluated using accuracy, precision, recall, and F1-score. | The ensemble model achieved a prediction accuracy of 96.55%, with significant improvements in precision (97.79%), recall (89.11%), and F1-score (94.60%) compared to traditional models. The integration of XGBoost and Random Forest offered better results in terms of both accuracy and model stability. |
| (El Annas et al., 2023) | The paper addresses the challenge of improving credit scoring systems by incorporating rejected applicants using semi-supervised learning and Hidden Markov Models (HMM). | The study proposes a semi-supervised Hidden Markov Model (SSHMM) framework for reject inference in credit scoring, involving binning, filtering, and training with both accepted and rejected loan applicants. | Peer-to-peer (P2P) lending platforms and credit scoring. | Lending Club data (2007-2018), containing both accepted and rejected applicants. | Key features include loan amount, FICO score, debt-to-income ratio (DTI), loan purpose, employment length, and state of residence. | Evaluated using accuracy, precision, recall, and AUC. | SSHMM outperformed other models such as SVM and XGBoost in accuracy and AUC on biased test sets. The model was also effective with low rejection rates and demonstrated superior adaptability and stability. |
| (Orji et al., 2022) | The paper explores the application of machine learning models to predict bank loan eligibility based on historical data. | Six machine learning algorithms (Random Forest, Gradient Boosting, Decision Tree, SVM, KNN, Logistic Regression) were applied to predict loan eligibility. | Financial services, focusing on bank loan eligibility prediction. | "Loan Eligible Dataset" from Kaggle. | Key features include gender, marital status, education, income, loan amount, credit history, and property area. | Evaluated using confusion matrix, accuracy, precision, recall, and F1 score. | Random Forest achieved the highest accuracy of 95.55%, while Logistic Regression had the lowest at 80%. The models outperformed similar models found in literature. |
| (Kumar et al., 2024) | The paper explores how to streamline bank loan approval processes using deep learning models, focusing on improving accuracy and reducing overfitting. | The study uses Bidirectional LSTM neural networks for loan approval prediction, applying data preprocessing, feature scaling, and dropout layers to reduce overfitting. | Financial services, specifically in loan approval prediction. | "Bank Loan Status" dataset from Kaggle, containing loan application details. | Features include loan amount, credit score, annual income, home ownership status, monthly debt, and credit history. | Evaluated using accuracy, F1-score, precision, recall, ROC curve, and confusion matrix. | The LSTM model achieved accuracy between 98% and 99%, with F1-scores ranging from 99% to 99.4%. The model demonstrated low false positives and high recall, proving effective in identifying approved loan applications. |
| (Dansana et al., 2024b) | The study investigates the impact of various loan-related features on predicting loan approval using the Random Forest algorithm. | The study employs a Random Forest classifier to predict loan approval based on applicant data, measuring feature importance using Gini importance. | Financial services, focusing on bank loan prediction and risk management. | A training dataset with 25 features (loan amount, term, purpose, income, employment type, etc.) used for predicting loan approval. | Key features include loan amount, gender, marital status, education level, employment type, income, loan purpose, and more. | Evaluated using feature importance, accuracy, and distribution analysis of loan approval outcomes. | Random Forest achieved strong prediction accuracy with significant insights into the importance of features like loan purpose, education, and income level. Key findings include higher approval rates for married individuals and those with higher incomes or stable employment types. |
| (Acharya et al., 2024) | The paper addresses the challenge of balancing model performance, fairness, and interpretability in AI models used for loan approval and housing price predictions. | The study applies fairness-aware machine learning techniques (Calibrated Equalized Odds and Intersectional Fairness) to XGBoost and LightGBM, incorporating SHAP for model interpretability. | Financial services (loan approval) and real estate (house price prediction). | Loan approval and house price datasets, publicly available with features like income, credit history, housing characteristics, etc. | Loan features: applicant income, loan amount, credit history, and demographics; House price features: property size, overall quality, neighborhood, etc. | Evaluated using fairness metrics (Disparate Impact, Equal Opportunity Difference), and interpretability via SHAP and LIME. | LightGBM outperformed XGBoost in terms of balancing accuracy and fairness, particularly in loan approval predictions. SHAP provided key insights into feature importance like credit history and applicant income, ensuring transparent and fair decision-making. |
| (Mourtas et al., 2024) | The paper explores the use of bio-inspired neural networks for improving the prediction accuracy of loan approval classifications. | A bio-inspired neural network using a novel weights and structure determination (WASD) approach combined with the beetle antennae search (BAS) algorithm is proposed for binary classification tasks. | Financial services, particularly in loan approval and credit risk classification. | Datasets for loan approval, including customer demographic and financial data (DA1, DA2, DA3, DA4 from Kaggle). | Key features include income, credit score, loan amount, education, employment history, and demographic details. | Evaluated using mean absolute error (MAE), accuracy, precision, recall, F1-score, and McNemar’s test for statistical significance. | The BWASD model outperformed traditional models (e.g., MWASD, FKNN, LSVM) across all datasets, achieving superior accuracy and F1-score. It also showed higher classification performance and better robustness to data variability compared to other classifiers. |
| (Alagic et al., 2024) | The paper explores how integrating mental health data can improve the prediction of credit risk in loan approval processes. | The study applies various machine learning algorithms (XGBoost, Random Forest, KNN, AdaBoost, Gradient Boosting) to predict loan approval, incorporating mental health data along with financial features. | Financial services, particularly in loan approval prediction and credit risk analysis. | Mental health survey dataset and loan approval dataset. | Features include age, gender, marital status, credit history, income, loan amount, mental health indicators, and workplace-related mental health consequences. | Evaluated using accuracy, precision, recall, F1 score, and confusion matrix. | XGBoost performed best with 84% accuracy in the mental health dataset and Random Forest achieved the highest accuracy of 85% for loan approval predictions. Random Forest and XGBoost excelled in precision and recall in both datasets. |
| (Purificato et al., 2023) | The paper investigates how responsible AI techniques such as explainability and fairness can enhance trust in AI-driven loan approval systems. | The system utilizes a proprietary framework that integrates explainability tools (LIME, SHAP, Anchors) and fairness algorithms (disparate impact metrics, reweighing algorithm) to improve transparency and mitigate bias in loan approval decisions. | Financial services, specifically in loan approval processes. | A dataset provided by an Italian bank, pseudonymized for privacy, involving loan approval data. | Features include loan application details such as credit history, income, age, nationality, and marital status. | Evaluated using metrics like the Explanation Goodness Scale, Trust & Reliance Scale, and A/B tests for fairness. | The system demonstrated improved trust and user satisfaction, with SHAP being the preferred explanation method. The fairness tools successfully mitigated biases, leading to a more equitable loan approval process. |

In the presented Table 2, it can be seen that the years range from 2021 to 2025, which meets the inclusion criteria for recent relevant studies. We found that most of the articles were published between 2022 and 2024, suggesting that these years have made significant contributions to the application of machine learning in loan approval and financial risk prediction.

Table 2 column 2, the summary of the main research question addressed by each article, shows that many studies have focused on improving loan approval prediction and credit risk assessment through the use of machine learning models. Some studies (Hedrick et al., 2024) address the challenge of mitigating bias in the loan approval process, while others (Lee et al., 2024; Ramachandra et al., 2021) aim to optimise loan approval rules. Still other studies focus on improving fairness (Choi et al., 2025) and interpretability (Purificato et al., 2023), highlighting the growing need for transparent AI models in finance.

Table 2 column 3 describes the methodological approach used in the research. Many articles apply supervised machine learning algorithms such as Random Forest, XGBoost, and deep learning techniques (Jamunadevi et al., 2024). Some studies use integration methods (Jamunadevi et al., 2024; Lee et al., 2024) or hybrid models, while others emphasise fairness-constrained methods (Choi et al., 2025) or counterfactual interpretations to improve model performance and fairness.

Table 2 column 4 provides the context or domain in which the research was applied. Most of the research has focused on the financial services sector, particularly loan approval and credit risk predict. The use of machine learning in these areas aims to enhance the decision-making process, improve efficiency and reduce human bias.

Table 2 column 5 lists the datasets used in the study. Some articles use publicly available datasets such as the Lending Club dataset (Ramachandra et al., 2021) or the Kaggle dataset (Hedrick et al., 2024; Jamunadevi et al., 2024). Some studies have used proprietary datasets from specific financial institutions (Lee et al., 2024), while others have relied on synthetic data to test their models under controlled conditions.

Table 2 column 6 lists the attributes or features used in the prediction model. Common features across studies include financial indicators such as income, loan amount, credit score, and employment status (Ramachandra et al., 2021). Some studies (Alagic et al., 2024) have included mental health data as an additional feature to improve loan approval predictions, highlighting the growing interest in using alternative and non-traditional data sources.

Table 2 column 7 details how the proposed approach is evaluated. Common metrics evaluated include accuracy, precision, recall, F1 scores, and AUC (Hedrick et al., 2024; Jamunadevi et al., 2024). Many studies have also evaluated fairness metrics (Choi et al., 2025), and some have combined runtime efficiency and model interpretability metrics, especially for models deployed in real-time client applications.

Table 2 column 8 summarises the main findings or outcomes of the study. Many studies have reported improvements in model performance, with some achieving higher levels of accuracy and precision (Jamunadevi et al., 2024). Some studies, such as (Choi et al., 2025), focus on achieving fairness without sacrificing accuracy, while others highlight the potential of their models to streamline the loan approval process and improve transparency. Several studies have emphasised the importance of fairness and interpretability, suggesting that these factors are becoming increasingly important in AI-driven financial applications

## Review of Datasets for the loan

The table below reviews the various Kaggle datasets used for loan approval prediction, highlighting key features, evaluation metrics, and machine learning models such as Random Forest and XGBoost. It provides an in-depth analysis of the features that have the greatest impact on loan approvals (e.g., income and credit history), as well as the tools used to build accurate predictive models.

Table 3. Review of Datasets for the loan

| **Dataset Name** | **Source** | **Key features** | **Use Case** | **Evaluation Metrics** | **Methodology Used** | **Main Findings** | **Tools/Technologies** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** |
| Bank Loan Prediction Dataset | https://www.kaggle.com/code/hadibakhsh/bank-loan-prediction/input?select=Sample\_Submission.csv | Income, loan amount, credit history, marital status, education | Predicting loan approval based on applicant financial details | Accuracy, Precision, Recall, F1-Score | Logistic Regression, Random Forest, SVM | Random Forest and Logistic Regression performed well with accuracy around 80%. Features such as income and credit history were most significant for loan approval predictions. | Python, Scikit-learn |
| Credit Score for Loan approval Dataset | https://www.kaggle.com/code/vassyesboy/credit-score-for-loan-approval/input | Credit score, income, debt-to-income ratio, loan purpose, employment status | Credit risk assessment based on credit scores for loan eligibility | Accuracy, Precision, Recall, F1-Score | Logistic Regression, Decision Trees, SVM | Credit score was a key feature for loan approval. Logistic Regression performed effectively in this context. | Python, Scikit-learn, XGBoost |
| Loan approval datasets for prediction | https://www.kaggle.com/datasets/abhishekmishra08/loan-approval-datasets/data | Applicant details (age, income, credit score, loan amount) | Loan approval prediction based on personal financial details | Accuracy, Precision, Recall, AUC | Random Forest, Decision Trees | Random Forest performed well with accuracy around 85%. Age and credit score were highly predictive. | Python, Scikit-learn, XGBoost |
| Loan Approval Insights Data to Decision-Making Dataset | https://www.kaggle.com/code/sulaniishara/loan-approval-insights-data-to-decision-making/input | Age, income, credit score, family size, loan amount | Insights-driven loan approval decisions | Accuracy, Precision, Recall, F1-Score | Logistic Regression, Random Forest | Features such as income and family size were significant in determining loan approval. | Python, Scikit-learn, XGBoost |
| Loan Approval Prediction Dataset | https://www.kaggle.com/code/ouyimin19/loan-approval-prediction/input | Credit score, income, loan amount, marital status, employment status | Predicting loan approvals for individual applicants | Accuracy, Precision, Recall, AUC | Logistic Regression, SVM, XGBoost | Logistic Regression showed good results in predicting approvals, focusing on credit score and income. | Python, Scikit-learn, XGBoost |
| Loan Risk Prediction with Credit Insights Dataset | https://www.kaggle.com/code/vedaantsingh/loan-risk-prediction-with-credit-insights/input | Credit score, loan amount, debt-to-income ratio, employment status | Predicting loan risks and defaults based on financial data | Accuracy, F1-Score, Precision | Logistic Regression, Random Forest, XGBoost | Key features were income and loan amount. XGBoost showed the best performance | Python, Scikit-learn, XGBoost |
| Loan Approval Prediction Analysis Dataset | https://www.kaggle.com/code/athirakaladharan/loan-approval-prediction-analysis/input | Income, loan amount, credit history, employment, marital status | Analysis of features influencing loan approval decisions | Accuracy, Precision, Recall, AUC | Random Forest, Decision Trees | Credit history and income were identified as the most influential features | Python, Scikit-learn, XGBoost |
| Simple Loan Prediction (with 6 different models) Dataset | https://www.kaggle.com/code/kens3i/simple-loan-prediction-with-6-different-models/input | Loan amount, credit score, income, employment, marital status | Simple loan approval prediction using multiple models | Accuracy, Precision, Recall, F1-Score | Decision Trees, Random Forest, SVM, XGBoost | Multiple models compared. Random Forest performed the best in terms of accuracy. | Python, Scikit-learn, XGBoost |

After going through and comparing several datasets, the Bank Loan Prediction dataset has been chosen in this study because it contains sensitive features such as gender, marital status, and educational level, which although not directly pertaining to an applicant's capacity to pay, may indirectly affect the outcome of loan approval. These characteristics are not usually involved in loan decision-making, but in practice they could still be associated with social biases or concealed disparities in the information resulting from historic discrimination. For example, gender may be a factor in loan approvals due to prior discrimination in lending; marital status may be associated with presumptions of financial strength; and education level may lead lenders to favor highly educated individuals regardless of financial status. Even though these features in principle do not affect decision-making, they might unconsciously introduce prejudice into the predictions of the model by picking up social biases. This data set is therefore an excellent candidate to investigate bias reduction for loan approval mechanisms because one can investigate how such sensitive features affect fairness and how such biases can be mitigated. The aim of this study is to determine how fairness measures and bias correction techniques can be applied in a way that loan approval models are unbiased and free of sensitive features that should not affect outputs.

## Models compared

The below table compares various AI models for loan approval prediction based on applicability, strengths and weaknesses, interpretability, scalability, and uses. Decision trees are well suited for small, understandable data, while random forest and XGBoost are suitable with large, complex data but less understandable. Logistic regression functions best on simpler linear data, while neural networks and SVMs can handle complicated relationships but require huge datasets and are more difficult to interpret. Hybrid models do both Random Forest and XGBoost better but are computationally more expensive.

Table 4. Model compared

| **Model** | **Suitability for Loan Approval Prediction** | **Advantages** | **Disadvantages** | **Interpretability** | **Scalability** | **Use Case in Loan Approval** | **Fairness** | **Reference** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** |
| Decision Tree | Good for simple, interpretable models | Easy to understand and visualize, fast training | Prone to overfitting, not robust for complex data | High (easily understandable) | Moderate | Good for small datasets and when interpretability is crucial | May perpetuate bias if not adjusted for fairness; basic fairness constraints can be added | (Hedrick et al., 2024) |
| Random Forest | Suitable for diverse and complex datasets | Robust to overfitting, handles large datasets well | Less interpretable than Decision Trees, slower to train | Medium (somewhat complex) | High | Preferred for larger datasets with many features | Can integrate fairness-aware mechanisms (e.g., fairness constraints) | (Jamunadevi et al., 2024) |
| XGBoost | Excellent for high-performance tasks | High predictive accuracy, handles missing data well | Difficult to interpret, computationally expensive | Medium (less interpretable) | High | Effective for high-dimensional data and competition-level tasks | Advanced fairness methods (e.g., fairness constraints, adversarial debiasing) are needed | (Choi et al., 2025) |
| Logistic Regression | Suitable for linear relationships in the data | Simple, fast, and interpretable | Assumes linearity, not ideal for complex relationships | High (very transparent) | Low | Good for baseline models and smaller datasets with simple patterns | Generally more interpretable, fairness can be easier to adjust | (Sandeep & Devi, 2022) |
| Neural Networks (NLP) | Suitable for capturing complex relationships | High accuracy, can model non-linear relationships | Requires large datasets, low interpretability | Low (black box) | Very High | Effective for large-scale datasets with complex patterns but costly | High risk of reinforcing bias; fairness techniques must be explicitly incorporated | (Mourtas et al., 2024) |
| Support Vector Machine (SVM) | Useful for complex, high-dimensional datasets | Effective in high-dimensional spaces, robust for classification | Computationally expensive, less efficient with large datasets | Medium (somewhat interpretable) | Moderate | Good for smaller, well-defined datasets with clear class separation | Fairness challenges when balancing high-dimensional data | (Hedrick et al., 2024) |
| Hybrid Model (RF and XGBoost) | Best for improving accuracy by combining models | Combines strengths of Random Forest and XGBoost for robust results | High computational cost, complex to tune | Medium (moderately interpretable) | High | Best when both accuracy and robustness are required, with large datasets | Potential for fairness trade-offs; needs advanced fairness techniques for balanced results | (Lee et al., 2024) |

## Summary of the 2nd section

This section provides a comprehensive literature review of Machine Learning (ML) applications in loan approval prediction and points to a slight lack of existing work on fairness mechanisms for loan approvals. This literature review points to the importance of incorporating fairness-aware machine learning approaches that can help overcome biases in loan approval, especially for sensitive demographic indicators such as education level, gender, and marital status.

While the majority of studies have emphasized improving predictive accuracy by utilizing ensemble models such as Random Forest and XGBoost, fairness has been behindhand. The majority of models are susceptible to propagating existing biases present in financial data even when they possess excellent prediction accuracy. Insertion of fairness constraints into machine learning models is a strong way of making sure sensitive attributes do not bias the outcome of the loan approval process (Choi et al., 2025). To address these issues, measures such as counterfactual fairness and adversarial debiasing have been proposed, aiming to ensure not just accurate predictions but also fair, non-biased decisions.

While many positive developments have been made, challenges continue in balancing fairness with accuracy. Many studies have been attempting to define measures of fairness that apply universally and have attempted to make it difficult to balance prediction performance with fairness, especially when demographic factors are represented indirectly within the data. This conflict highlights the need to continue research on fairness metrics and how they are introduced into real-world practice.

In summary, even though machine learning models have immensely improved the predictive power of loan approval systems, the field is still lacking a common framework that balances fairness and accuracy.

# Proposed Approach

From the lack of fairness in existing loan approval criteria, this chapter proposes a mechanism that combines machine learning techniques with fairness constraints to obtain fair decisions. The mechanism begins by collapsing individual information about loan applicants into finance, demographics, and history. The information is subsequently processed with fairness constraints in such a way that the system will not be sensitive to sensitive attributes such as gender, race, or age. At the core of the solution is the loan approval decision model, which is trained with a fairness-aware algorithm that incorporates regularization or adversarial debiasing techniques. During invocation, the model generates a loan decision (approve or reject) based on the applicant's data. To ensure transparency and traceability of the process, the decision will be authenticated with pre-defined metrics and fairness metrics along with visualization of the decision and such fairness metrics. The loan approval system will finally present the output in terms of evaluating the decision and fairness of the decision. This process ensures that the loan approval system is always ethical, fair and transparent.

The following UML component diagram(Figure 3-1) shows Fairness Aware Loan Approval System, which consists of three subsystems: Data Service, the Loan Approval System, and the System Evaluation. The data services subsystem contains two components: the loan data store and the fairness data store. Both of these subsystems are responsible for processing and displaying data needed for the loan approval process, i.e., loan-related data as well as measures concerning fairness, so that the system is supplied with all information necessary for decision-making. The Loan Approval System subsystem is the core of the system and is responsible for processing the loan requests. It uses the loan decision model (that evaluates loan applications against information provided) and data processing (that handles incoming loan and fairness information). The fairness constraints component is used to ensure fairness criteria are integrated into the decision-making process, and the fairness assessment component evaluates the fairness of the model's decisions. The system evaluation subsystem is responsible for monitoring performance of the system. It includes model performance (to track the loan decision model's performance) and fairness performance (to track the system's consistency with the fairness criteria over time). The structure easily shows the architecture of the proposed approach and is a solid basis for the following steps.

A diagram of a company's loan approval system

AI-generated content may be incorrect.

Figure 3‑1. UML component diagram - Loan approval system

The high-level BPMN diagram presented below (Figure 3-2) illustrates a loan application process with fairness integrated into a machine learning model. The process begins with the initiation of loan application, followed by setting fairness constraints to render the process fair. Loan data collection and loan data preprocessing are utilized to pre-process the data required by the model. Integration of the machine learning model steps and fairness steps will ensure that the model is adhering to fairness standards as it is trained. The user interface is then built where users will be interacting with the system, and finally, the process is completed, meaning the system is in a ready state to use.

A screenshot of a computer

AI-generated content may be incorrect.

Figure 3‑2. General Fairness Loan Approval System

The following diagram is the BPMN model of the setup fairness Constraints sub-process (Figure 3-3). It begins with a 'Review of Fairness Constraints in Existing Papers,' where the researcher reviews current studies for applicable fairness guidelines. The researcher will then set up new fairness constraints from the reviewed work. Once the definition is complete, the researcher will assess the validity and understandability of the fairness constraint at a decision point. In the case of a valid and clear constraint, it will be entered in the 'Document Fairness Constraints' task. In the case of an invalid or unclear constraint, the process will roll back to the Clarify or Update Fairness Constraints step until the constraint is acceptable. The process stops once the constraint is fully documented and confirmed.

A diagram of a flowchart

AI-generated content may be incorrect.

Figure 3‑3. Sub-process Setup Fairness Constraints

The subsequent BPMN diagram (Figure 3-4) displays the sub-process of loan data collection. The process commences with the definition of the data requirements, delineating the specific data requirements (e.g., for a loan or credit application). Next, in the 'Identify Data Sources' task, determine appropriate financial data sources, e.g., publicly accessible datasets on Kaggle or UCI. Then retrieve the data, which can be downloaded from the aforementioned repositories. Once the data has been acquired, the data is checked if it is within the quality parameters, and if so, an initial integration of the data is performed to bring sensitive and non-sensitive data into the system, otherwise re-acquisition of the data is performed. Following the integration, the raw data will be archived. The final step is the review of data and quality check to ensure the acquired data are par for the course before moving on to subsequent stages.

A diagram of a flowchart

AI-generated content may be incorrect.

Figure 3‑4. Sub-process Collect Loan Data

The loan data preprocessing sub-processing is illustrated in the following BPMN diagram (Figure 3-5). The process begins with missing values processing, identifying invalid or missing data points and removing or filling them. Outliers are identified and handled using methods such as box-and-line plots or z-scores in order to ensure data integrity. Then there is a data type conversion step where data is transformed into an appropriate format, e.g., a label code or a specific heat code in the case of categorical variables. Then features that are relevant are computed from data to get them ready for modelling. Processed data will be stored for use later on. Finally, the data is split into training and test sets to ensure that the model might be well-trained and tested.

A diagram of a process

AI-generated content may be incorrect.

Figure 3‑5. Sub-process Preprocess Loan Data

This below BPMN diagram (Figure 3-6) is for sub-processing Integrate fairness and ML model. It begins with the selection and auditing of an appropriate fairness-aware algorithm. Then, if the algorithm has been selected, the machine learning model is incorporated with the fairness constraints. The decision point checks whether the constraints are successfully incorporated into the model. If not included before, the constraints are included before proceeding to the next step. Then, the model is trained according to the fairness constraints and tested for fairness to check if it meets the originally specified fairness standards. Another check point checks whether the model meets the fairness requirements. If it fails to meet the requirements, the process returns to include the constraints and retrain the model. Once the model has satisfied the fairness requirement, it is deployed as a fairness integrated machine learning model and the process is completed.

A diagram of a process

AI-generated content may be incorrect.

Figure 3‑6. Sub-process Integrate fairness and ML model

This below BPMN diagram (Figure 3-7) shows the User Interface Building sub-processing. The process captures the software development process, which begins with requirements analysis and design, project requirements definition, and design planning. Prototyping then follows, using tools such as Figma or Sketch to create visual prototypes. Next, the back-end API is developed using a framework such as Flask or Django, and finally, the user interface is developed with HTML, CSS, JavaScript, and Vue.js for front-end development. User interaction design is performed to ensure that the interface is user-friendly, followed by UI testing to ensure functionality. If the UI is satisfactory, the process moves on to deployment and release; otherwise, re-tuning is undertaken, and the process returns to step one.

A diagram of a software development process

AI-generated content may be incorrect.

Figure 3‑7. Sub-process Build User Interface

## Main Results of the 3rd Section

In this case, the objective is to provide a plan to integrate fairness in loan approval decisions using machine learning. The main results of this chapter are the development of a fairness-aware predictive framework that combines machine learning models with fairness constraints to mitigate bias in loan approval systems. The model is evaluated on real financial datasets, and the results show that fairness can be achieved without sacrificing prediction accuracy significantly. In addition, the prototype interface displays predictions and fairness metrics in a visual format to make it easier for the end user and improve transparency. This ensures that the loan approval process is not just accurate in prediction but also adheres to ethical and fair decision-making requirements.

# Initial experiment

## Dataset describing

### Descriptive statistics

Descriptive statistics were probed with SPSS for the selected bank loan prediction dataset so that a summary of the main numerical variables could be presented. ApplicantIncome had values ranging from 150 to 81,000, and there was a mean of 5,403.46 with a large standard deviation of 6,109.04, showing substantial variability as well as possible outliers. The same applies to the CoapplicantIncome variable, ranging from 0 to 41,667, having a mean of 1,621.25 and a standard deviation of 2,926.25, which suggests that the majority of applicants have no co-applicant, or that co-applicants have very different incomes. Loan Amount (LoanAmount) averages 146.41 and has a standard deviation of 85.59, which indicates the moderate range of loan amounts from 9 to 700. There is bunching of Loan Term (Loan\_Amount\_Term) values around the mean of 342 months, but the range of 12 to 480 months indicates that the applicants have taken both short-term loans and long-term loans. Finally, Credit\_History, a binary attribute with the value of 1 for a good credit history, has a mean of 0.84, which indicates that most applicants possess a good credit history. This succinct descriptive summary shows that the dataset contains a rich set of income-, loan-, and credit-based attributes that provide a great platform for training and testing machine learning algorithms. The variation of variables, the presence of missing values and potential outliers make this dataset particularly pertinent to study real business decision-making scenarios with a data-driven approach.

Table 5. Descriptive statistics for selected dataset

| **Variable** | **N** | **Minimum** | **Maximum** | **Mean** | **Std. Deviation** |
| --- | --- | --- | --- | --- | --- |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| ApplicantIncome | 614 | 150 | 81000 | 5403.46 | 6109.042 |
| CoapplicantIncome | 614 | .00000000 | 41667.000000 | 1621.2457980 | 2926.2483692 |
| LoanAmount | 592 | 9 | 700 | 146.41 | 85.587 |
| Loan\_Amount\_Term | 600 | 12 | 480 | 342.00 | 65.120 |
| Credit\_History | 564 | 0 | 1 | .84 | .365 |
| Valid N (listwise) | 529 |  |  |  |  |

The calculation formulas for mean and Standard Deviation are as follows:

1. Mean (Average):

Where:

= Each data point

= Number of data points

1. Standard Deviation:

Where:

= Each data point

= The calculated mean of dataset

N = Number of data points

### Correlation analysis

Correlation tests were performed using SPSS to examine the correlation of the main numerical variables in the bank loan prediction data. The test revealed a positive correlation between applicant income and loan amount (r = 0.571, p < 0.01), indicating that the applicant's income is higher, the larger the loan amount. This outcome is consistent with the expectation that more prosperous applicants would be likely to qualify for a loan and ask for a larger loan. There is also a moderate and significant positive correlation between co-applicant income and loan size (r = 0.189, p < 0.01) which suggests that co-applicants contribute towards overall loan eligibility but to a lesser extent. As can be seen, applicant income and co-applicant income are negatively related (r = -0.117, p < 0.01), indicating the tendency for more affluent applicants to need fewer co-applicants. The remaining variables are not significantly or statistically significantly related to loan term (Loan\_Amount\_Term) or credit history (Credit\_History), indicating that loan term and credit history are independent of income and loan size in this dataset. This correlation analysis is significant to this paper in that it identifies significant relationships that have an effect on loan approval outcomes and highlights the multicollinearity problems necessary for developing accurate and interpretable machine learning models. The findings confirm ApplicantIncome, CoapplicantIncome, and LoanAmount as significant predictors in subsequent modelling stages.

Table 6. Correlation analysis for selected dataset

| **Variable** | | **ApplicationIncome** | **CoapplicantIncome** | **LoanAmount** | **Loan\_Amount\_Term** | **Credit\_History** |
| --- | --- | --- | --- | --- | --- | --- |
| **(1)** | | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| ApplicantIncome | Pearson Correlation | 1 | - .117\*\* | .571\*\* | - .045 | - .015 |
| Sig. (2-tailed) |  | .004 | < .001 | .268 | .727 |
| N | 614 | 614 | 592 | 600 | 564 |
| CoapplicantIncome | Pearson Correlation | .117\*\* | 1 | .189\*\* | - .060 | - .002 |
| Sig. (2-tailed) | .004 |  | < .001 | .143 | .961 |
| N | 614 | 614 | 592 | 600 | 564 |
| LoanAmount | Pearson Correlation | .571\*\* | .189\*\* | 1 | .039 | - .008 |
| Sig. (2-tailed) | < .001 | < .001 |  | .344 | .845 |
| N | 592 | 592 | 592 | 578 | 543 |
| Loan\_Amount\_Term | Pearson Correlation | - .045 | - .060 | .039 | 1 | .001 |
| Sig. (2-tailed) | .268 | .143 | .344 |  | .973 |
| N | 600 | 600 | 578 | 600 | 550 |
| Credit\_History | Pearson Correlation | - .015 | - .002 | - .008 | .001 | 1 |
| Sig. (2-tailed) | .727 | .961 | .845 | .973 |  |
| N | 564 | 564 | 543 | 550 | 564 |

\*\* . Correlation is significant at the 0.01 level(2-tailed).

The Pearson correlation coefficient is calculated using the following formula:

Where:

= Pearson correlation coefficient

= Individual data points for variables X and Y respectively

= Means of variables X and Y

= Number of data points in the dataset

# Implementation of the Proposed Approach

This section is a continuation of the “Proposed Approach” section, in which the student, based on its results, formulates requirements and specific proposals for the realization of an approach. It is recommended to present the Prototype architecture, Plan of the Experiment, and Implementation and Testing results.

An example of the architecture is presented below as follows.

## Main Results of the 4th Section

This section should be concluded by the summary of the main results obtained in this section.

# General Conclutions

This section presents general conclusions as follows:

1. Based on the analysis done of existing studies, the findings demonstrate that although various machine learning approaches have been employed in loan approval prediction, the majority of models still fail to address the concern of fairness to some extent. The review demonstrates that although hybrid models and complex machine learning algorithms (e.g., Random Forest and XGBoost) may be applied to improve prediction accuracy, fairness is a less explored dimension. These findings highlight the importance of imposing fairness constraints on the loan approval model in a way that sensitive attributes such as gender, education, and marital status do not unfairly influence the outcome. In addition, this study highlights the importance of incorporating fairness-aware mechanisms, such as adversarial debiasing and fairness constraints, to mitigate bias in prediction models and make overall loan decisions fairer.

2. Through the development and evaluation of the proposed method carried out, the obtained results show that the fairness constraint application to a machine learning model in predicting loan approvals can effectively remove demographic bias while attaining high prediction accuracy. The proposed method combines machine learning models with fairness-aware approaches to ensure sensitive features have no influence in the decision-making process of loan approval. Experimental results show that the approach not only addresses the ethical issues of loan approval discrimination but also the transparency and accountability of the decision-making process. Thus, the results show that machine learning models can be made fair and accurate simultaneously, and hence provide a fairer and more credible solution for automated loan approval systems.

## Further works

The proposed approach has great potential in incorporating fairness in the loan approval prediction but continuous optimisation is required in order to make it even more efficient. Continuous development of the fairness constraints and prediction models to act better should be done in future work. This entails exploring other fairness-conscious algorithms and cross-testing with other approaches to achieve a better balance between fairness and accuracy such that the model performs well on different datasets. It must also be tested for generalizability to other financial institutions to enable it to operate on diverse datasets, local legislation and demographic profiles. Beyond optimisation, the next step of utmost importance is to implement the methodology in a live loan granting system. This would entail overcoming deployment problems such as data privacy issues, real-time processing capability, and integrating the system into existing decision-making systems. Implementation stage will also involve rigorous testing and validation in real-world environments to ensure that the model remains transparent, equitable, and effective at scale. Identifying these areas here means that future work can translate the suggested approach from a theoretical model into a workable, efficient tool for more equitable financial decision-making.

# References

Acharya, D. B., Divya, B., & Kuppan, K. (2024). Explainable and Fair AI: Balancing Performance in Financial and Real Estate Machine Learning Models. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3484409

Alagic, A., Zivic, N., Kadusic, E., Hamzic, D., Hadzajlic, N., Dizdarevic, M., & Selmanovic, E. (2024). Machine Learning for an Enhanced Credit Risk Analysis: A Comparative Study of Loan Approval Prediction Models Integrating Mental Health Data. *Machine Learning and Knowledge Extraction*, *6*(1), 53–77. https://doi.org/10.3390/make6010004

Choi, Y., Hong, J., Lee, E., Kim, J., & Kim, S. (2025). Enhancing Fairness in Financial AI Models through Constraint-Based Bias Mitigation. *Journal of Information Processing Systems*, *21*(1), 89–101. https://doi.org/10.3745/JIPS.01.0111

Dansana, D., Patro, S. G. K., Mishra, B. K., Prasad, V., Razak, A., & Wodajo, A. W. (2024a). Analyzing the impact of loan features on bank loan prediction using Random Forest algorithm. *Engineering Reports*, *6*(2). https://doi.org/10.1002/eng2.12707

Dansana, D., Patro, S. G. K., Mishra, B. K., Prasad, V., Razak, A., & Wodajo, A. W. (2024b). Analyzing the impact of loan features on bank loan prediction using Random Forest algorithm. *Engineering Reports*, *6*(2). https://doi.org/10.1002/eng2.12707

El Annas, M., Benyacoub, B., & Ouzineb, M. (2023). Semi-supervised adapted HMMs for P2P credit scoring systems with reject inference. *Computational Statistics*, *38*(1), 149–169. https://doi.org/10.1007/s00180-022-01220-9

Hare, M. (2023). Towards Wisdom: Knowledge Management and the Ethical Use of Big Data in Organizational Decision-Making. *DALHOUSIE JOURNAL OF INTERDISCIPLINARY MANAGEMENT*, *17*.

Hedrick, J., Yeboah, J., & Nti, I. K. (2024). Predicting the Risk Level of a Loan Based on the Customer’s Personal Factors Using Machine Learning Models. *2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)*, 1–5. https://doi.org/10.1109/ICMI60790.2024.10586183

Jamunadevi, C., Prasath, S., Sathishkumar, V. E., Pandikumar, S., & Akshaya, J. (2024). Enhancing Loan Approval Prediction Through Advanced Machine Learning Models. *2nd International Conference on Sustainable Computing and Smart Systems, ICSCSS 2024 - Proceedings*, 957–964. https://doi.org/10.1109/ICSCSS60660.2024.10625409

Karimi, A. H., Schölkopf, B., & Valera, I. (2021). Algorithmic recourse: From counterfactual explanations to interventions. *FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 353–362. https://doi.org/10.1145/3442188.3445899

Kumar, A. P. S. P., Reddy, K. S., Poojitha, K., Swain, N. K., Yugandhar, M., & Shekar, G. V. C. (2024). Streamlined Bank Loan Approvals Using Deep Learning. *2024 2nd World Conference on Communication and Computing, WCONF 2024*. https://doi.org/10.1109/WCONF61366.2024.10692066

Kumar, P., Maneesh, U. L., & Mano Sanjay, G. (2024). Optimizing Loan Approval Decisions: Harnessing Ensemble Learning for Credit Scoring. *Proceedings - 3rd International Conference on Advances in Computing, Communication and Applied Informatics, ACCAI 2024*. https://doi.org/10.1109/ACCAI61061.2024.10602097

Lee, K., Yoo, H., Shin, S., Kim, W., Baek, Y., Kang, H., Kim, J., & Kim, K. (2024). A submodular optimization approach to trustworthy loan approval automation. *AI Magazine*, *45*(4), 502–513. https://doi.org/10.1002/aaai.12195

Mourtas, S. D., Katsikis, V. N., Stanimirović, P. S., & Kazakovtsev, L. A. (2024). Credit and Loan Approval Classification Using a Bio-Inspired Neural Network. *Biomimetics*, *9*(2). https://doi.org/10.3390/biomimetics9020120

Orji, U. E., Ugwuishiwu, C. H., Nguemaleu, J. C. N., & Ugwuanyi, P. N. (2022). Machine Learning Models for Predicting Bank Loan Eligibility. *Proceedings of the 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development, NIGERCON 2022*. https://doi.org/10.1109/NIGERCON54645.2022.9803172

Purificato, E., Lorenzo, F., Fallucchi, F., & De Luca, E. W. (2023). The Use of Responsible Artificial Intelligence Techniques in the Context of Loan Approval Processes. *International Journal of Human-Computer Interaction*, *39*(7), 1543–1562. https://doi.org/10.1080/10447318.2022.2081284

Ramachandra, H. V., Balaraju, G., Divyashree, R., & Patil, H. (2021). Design and simulation of loan approval prediction model using AWS platform. *2021 International Conference on Emerging Smart Computing and Informatics, ESCI 2021*, 53–56. https://doi.org/10.1109/ESCI50559.2021.9397049

Sandeep, C. V, & Devi, T. (2022). A Novel Approach for Bank Loan Approval by Verifying Background Information of Customers through Credit Score and Analyze the Prediction Accuracy using Random Forest over Linear Regression Algorithm. *JOURNAL OF PHARMACEUTICAL NEGATIVE RESULTS*, *13*(4), 1748–1755. https://doi.org/10.47750/pnr.2022.13.S04.211

Wang, Y., Wang, M., Pan, Y., & Chen, J. (2024). Joint loan risk prediction based on deep learning-optimized stacking model. *Engineering Reports*, *6*(4). https://doi.org/10.1002/eng2.12748

Weinzierl, S., Zilker, S., Dunzer, S., & Matzner, M. (2024). Machine learning in business process management: A systematic literature review. In *Expert Systems with Applications* (Vol. 253). Elsevier Ltd. https://doi.org/10.1016/j.eswa.2024.124181