AI6101: AI ETHICS ASSIGNMENT

Group Name: AKEZA

Joseph UWINEZA
G2303477F
JOSEPH005@e.ntu.edu.sg

Etienne NTAMBARA
G2304253K
NTAM0001@e.ntu.edu.sg

Shahrul AI-NIZAM LING
G2304209B
SLING005@e.ntu.edu.sg

MUKANYANDWI Joselyne
G2304244J
joselyne001@e.ntu.edu.sg

Qiu JUNPING
G2304210E
jqiu007@e.ntu.edu.sg

Program: MSAI

November 2023

Mini-project: Loan Approval AI System

Abstract

The project proposal called Loan Approval by Group AKEZA focuses on developing an AI-based loan approval system and application to improve efficiency, fairness, and scalability in financial institutions such as banks. The proposed AI application automates the loan assessment process, utilizing advanced algorithms and machine learning to minimize processing times and reduce biases from the existing system used by financial institutions. The system will ensure data privacy through techniques like federated learning, differential privacy, and secure multi-party computation. Key aspects include responsible AI practices, emphasizing explainability, fairness, privacy preservation, and ethical standards in data usage. This innovative approach is expected to transform financial operations, providing significant benefits to both financial institutions and customers.

1 Introduction

The loan approval process is important for accessing funds like housing loans and credit cards in today's modern financial system. Nevertheless, the existing system, which is entirely dependent on a manual examination conducted by the officers within banks, may not be efficient, objective, and extendable enough[5]. The traditional loan approval process is exhaustive and resource-consuming. Documents such as government-issued IDs, bank statements, employment proofs, salary slips, property paperwork, and other necessary documents are collected by the loan officer [7]. AI is not intended to minimize human workers or to completely replace the human function in credit decision-making. No matter how advanced technology is, some human touch is always required [4].

2 Background

As the number of loan applications keeps increasing, these shortcomings become more obvious, and the paradigm shift in the approval workflow becomes imperative. The relevant documents are processed at scale by AI-based document processing software to gather particular information. It also verifies submitted documents, and forwards applications to the concerned departments for approval. It also helps with faster loan approval and enhanced customer experience.

2.1 Problem Statement

The existing manual loan approval system has many challenges, such as high processing times, inconsistent decision-making, and potential biases. Given that this is an old model that is becoming more apparent in its weaknesses as demand for loans diversifies and increases, The situation is complicated by different loan types, which result in poor performance on the part of financial institutions as well as an unsatisfactory experience for applicants with suboptimal results.

The loan approval process, a cornerstone of facilitating access to financial resources such as housing loans and credit cards, is currently encumbered by inefficiencies inherent in its predominantly manual methodology. The existing system, largely dependent on bank officers conducting manual reviews, suffers from a lack of efficiency, objectivity, and scalability. This traditional process necessitates the labor-intensive gathering and analysis of a plethora of documents, including government-issued IDs, bank statements, employment proofs, salary slips, and property papers. While the adoption of artificial intelligence (AI) in this domain is not intended to diminish the human workforce or entirely replace human judgment in credit decision-making, the escalating volume of loan applications accentuates the limitations of the current approach and underscores the urgency for a transformative shift in the loan approval workflow.

One potential solution is the integration of AI, especially through AI-based document processing software. It can expedite the loan approval process overall, process and verify a large volume of documents efficiently, and route applications to the appropriate departments in an efficient manner. This greatly improves the client experience while also guaranteeing speedier loan approvals. In order to ensure that technological advancements complement human judgement rather than replace it, it is crucial to maintain a balanced synergy between AI automation and the essential human element."

2.2 Current System

In the current paradigm, loan applicants submit personal and financial information, and bank officers meticulously assess eligibility based on factors such as credit scores and income. This process is inherently limited by its reliance on manual scrutiny. Following the compilation of all documents, critical information is personally validated. After verification, final approval from the manager could take days, if not weeks. The assessment process lacks a standardized, scalable framework, often leading to delays, inconsistencies, and the inadvertent introduction of subjective biases. Also, the growing interconnectivity of consumer data across various applications poses challenges in holistic decision-making within the current system.

Currently, most financial institutions have non-updated customer data, which also leads to inconsistency while they are making manual decisions for approving customer's loan requests and it takes a long time to track the real accurate information. In this case, the way the loan is approved tends to be likely unfair because the manager would tend to approve a loan to someone with a unique relationship with relaxation on the rules and regulations of the bank.

2.3 Proposed Solutions:

We proposed a new AI-based software platform that will help change the current process of approving loans. This novel method seeks to inject efficiency, objectivity, and scalability into the decision-making workflow, thereby creating a fairer system. This AI-driven application will automate the assessment by using advanced algorithms and machine learning models to process large quantities of data in order to make informed decisions. The automation will greatly cut down on processing times, which will afford applicants faster access to financial resources.

In addition, AI can process data with a possible minimum bias, thereby reducing the possibility of bias in the evaluation and assuring a fair evaluation for all the clients. Objectivity will also help to ensure non-discrimination based on race, gender, social-class affiliation, etc. Accordingly, the proposed AI-enabled application will have a flexible design to accommodate the special criteria of different loan types, such as housing loans and credit card applications, among other financial instruments. Besides, data privacy and integrity will be considered in the application's core design. The commitment to responsible AI adoption will mean that there will be no violations or unnoticed processing of user data, as well as no hidden automated decisions on behalf of the applicants.

The applicant will submit his or her request via a web system or mobile app for pre-processing. The information submitted will be analyzed automatically by the system based on the current information of the applicant. If the loan criteria are met, the system will alert the manager for further analysis. The applicant will receive notification about the loan decision.

2.4 Purpose

We proposed this AI application to make the loan approval procedure more efficient for financial institution officers and clients through Machine learning algorithms. This project aims to automate a loan application process in a fair and effective manner consequently providing valuable information needed by bank officers for decision making.

2.5 Objectives

- Ensuring fair approval of loans by considering the eligibilities kept in place to determine who is likely to get their loan approved.
- Save effort and time for both the bank officers to go through the whole bunch of documents and the customers to go back and forth asking if their loan is approved or not.
- Using AI applications to minimize the biases in the process of loan approval.

3 Application Description

Our innovative AI system is set to automate the loan approval process, supporting bank officers in making well-informed decisions and increasing their productivity. We will develop advanced AI models based on credit history, income, employment, and market trends to assess creditworthiness and evaluate risk levels. This will enable bank officers to make informed decisions with more confidence and accuracy. The interface we will use is simple and will make it easy for borrowers. Also, we will integrate our AI system with other algorithms such as federated learning (FL) and zero-knowledge proofs (ZKP) to make sure that the system has tight security of personal financial details, and that the people are who they claim to be since they will be login page on the system.

Here is the process of our proposed system will be working at the end:

Data collection: Data is a vital part of each model or system, due to the fact each model decides or produces results based on the data it is trained on, we will begin by collecting the loan applications. Through this AI system, the client will be allowed to upload and fill in the relevant documents from various sources, including loan application forms, financial files, credit score rankings, personal history, educational background, and other needed information that can prove the creditworthy of the applicant.

Data processing and analysis: Usually, the customers submit their loan application documents to the bank and wait for the bank officers to go through the documents, and check if there is any missing information and it takes days for humans to do that considering the number of applicants. However, our AI system will leverage machine learning and AI technologies like Natural language processing (NLP) and Optical character recognition (OCR) to process and analyze data promptly, flagging any missing information and notifying customers for immediate rectification. Upon data completion, the system will assess the applicant's eligibility based on predefined criteria and generate a comprehensive report.

Decision Making: Bank officers will review the AI-generated report alongside bank policies and regulations to make informed loan approval or rejection decisions.

Expected outcomes:

- Fair loan approval: By using this AI application, there will be a fair selection process. The AI system does not have human emotions, it makes decisions based on the given criteria which means that if the customer meets the eligibility criteria required their loan will be guaranteed.

- Prompt customer service: the AI system will enhance customer satisfaction by reducing the loan processing time, hence enabling faster loan approvals.

System Flowchart

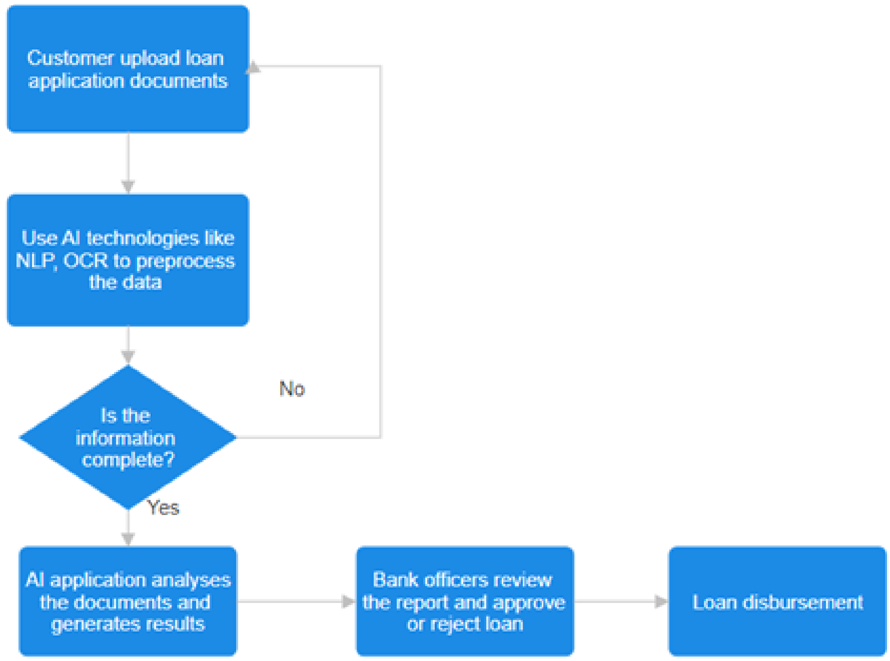

Figure 1: This figure shows the flow of information in the implemented AI system

4 Use of AI Technologies

4.1 Federated learning technique

Financial institutions hold the most confidential information for their clients. If that information leaks it can severely affect not only the institution's reputation but also the clients whose information is leaked. We adopted the federated learning algorithm where each bank will hold its local model responsible for assessing loan applications to avoid the sharing of confidential information and upload the local model updates to the central model through the encrypted channel to maintain the privacy of clients' data. The central model will aggregate local models to develop a unified model applicable across all the banks. This federated learning will play an inevitable role in maintaining the privacy of clients' data.

Federated learning will provide a secure and robust infrastructure that enables each bank sub-branch to collaboratively learn a shared prediction model while keeping all the training data at the same branch. This negates the need to store the data in the cloud or on physical servers on the main branch exclusively.

We take advantage of the ability to conduct model training at the sub-branches rather than make predictions alone. Although training models at the sub-branches help preserve user privacy as data remains at the location collected, a single branch with limited data might not be able to train a high-quality model. To collaboratively train a model without sharing private data across branches, each branch can share their local models instead of their data with a central server. These models will be aggregated, give an average, and produce a global model that can then be redeployed to all branches.

The figure 2 shows the simple structure of how the Federated Learning algorithm will be implemented in the banks while preserving the privacy of the bank data.

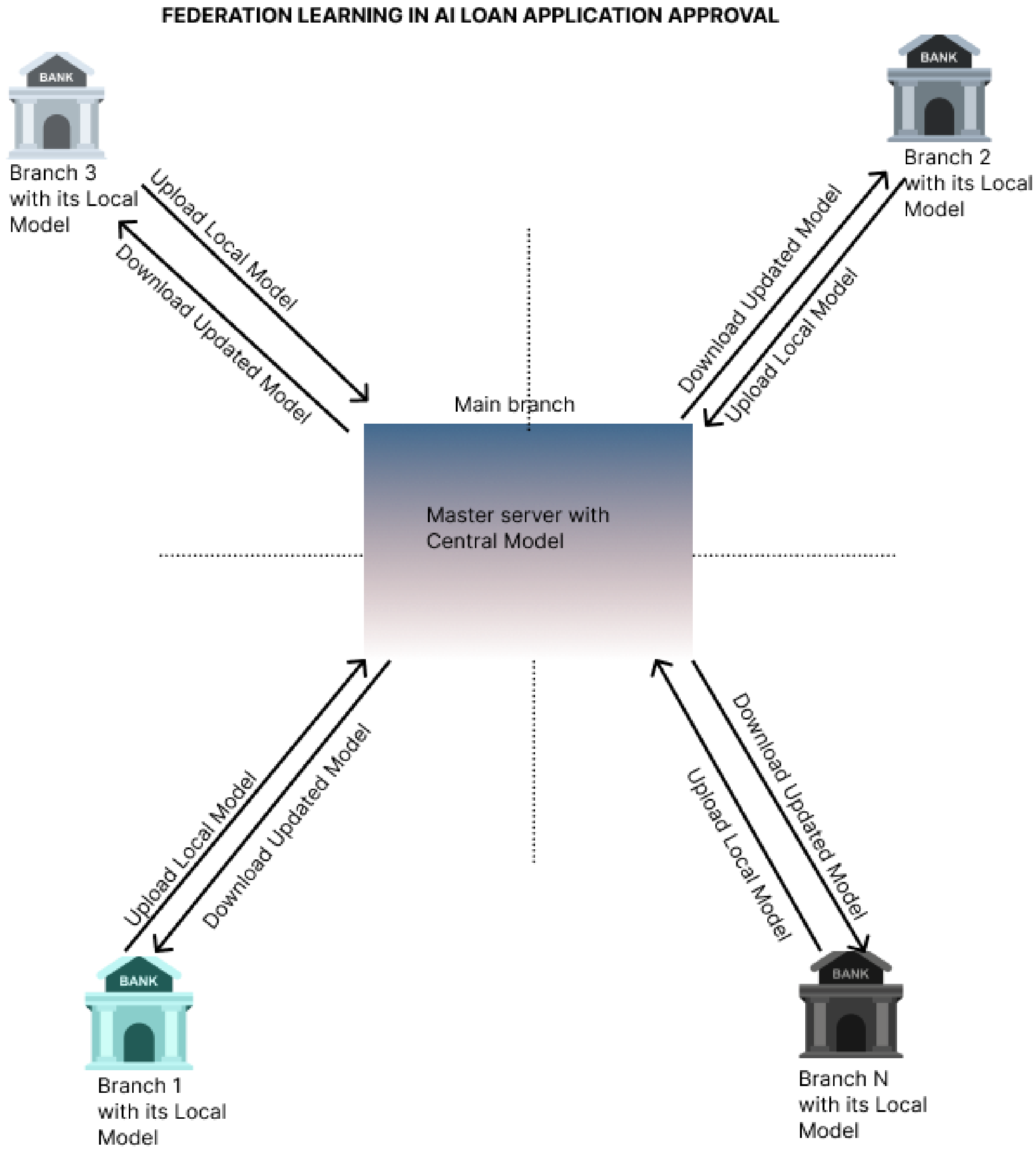

Figure 2: The figure shows the simple structure of how Federated Learning will be implemented in the banks

4.2 Differential Privacy

Differential privacy This technique introduces noise into the model parameters during federated learning to make it difficult for unauthorized users to access personal data from the model while improving data privacy. With the help of the privacy budget, differential privacy allows users to manage the trade-off between data privacy and model accuracy.[2]

As data sharing is required between different branches, we will adopt the differential privacy methodology as another useful method that preserves user privacy. It works by analyzing a dataset and computing statistics about it, such as the mean, variance, median, and mode of the data. It is called differentially private because it makes it impossible to discern if user data was included in the original dataset when reviewing the output. A differentially private algorithm does not change its behavior when a single data point joins or leaves the dataset, offering the guarantee that individual-level information about people in the database remains confidential and is not leaked.

4.3 Secure Multi-Party Computation

Secure Multi-Party Computation (SMPC): This is a cryptographic protocol through which different entities work together without exposing their individual inputs. In our system, this protocol will help us calculate the loan risk score for each lender without revealing their personal financial data, and this will allow the collaboration of different bank branches on loan risk assessment without revealing the client's personal data [1].

4.4 Data Minimization

This is an AI accountability tool that limits data exposure through secure aggregation (a class of secure, multi-party computation algorithms that hold private and aggregate values) and secure enclaves (a secret memory sector whose contents are safeguarded by hardware-grade encryption and hardware-isolation mechanisms), which provides further user privacy protection.

In secure aggregation, bank branches agree on shared random numbers and mask their local models in a way that preserves the aggregated result. The main server does not have any knowledge of how each sub-branch modified its model. It enables all branches to collaboratively merge models without revealing each individual contribution to the main server.

To further protect the confidentiality of personal information, we will adopt federated learning together with differential privacy, secure multi-party computation, and data minimization. In this way, we can achieve a high level of data privacy while maintaining our model's accuracy.

5 Responsible AI Practice Addressed

5.1 Explainable AI(XAI)

Explainable AI (or XAI) is a new field that can assist banks in managing concerns of transparency and trust, as well as giving greater clarity on their AI governance[3]. XAI strives to make AI models more explainable, intuitive, and intelligible to humans while maintaining performance and forecast accuracy. Explainability is also a growing concern for banking authorities, who want to ensure that AI procedures and outputs are "reasonably understood" by bank personnel. XAI may also assist banks in bringing more pilot projects to light, as a lack of explainability can be a major barrier to using AI models. Explainability tools can reveal various forms of information about a model based on the type of answers sought and the type of modeling used. The explainability of AI comes with a cost; the following figure3 depicts how maintaining high explainability goes with model complexity. In the bank, we should consider partnerships with **think tanks, universities and re-**

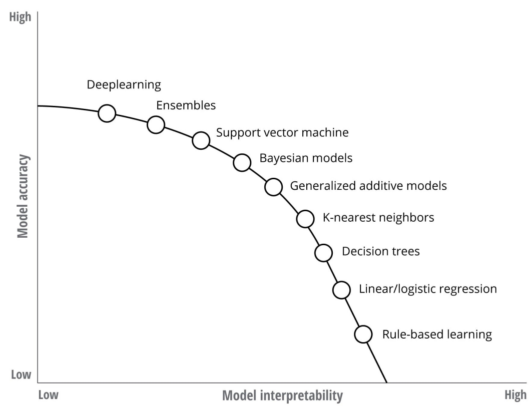

Figure 3: Show explainability vs. trade-off in model performance. Explainability has become more important within banks as a result of the increasing complexity of AI algorithms, which is occurring at a breakneck rate due to increased computer power, the explosion of large data, and breakthroughs in modeling approaches like neural networks and deep learning [3].

search institutions, and groups such as Singapore's Veritas Consortium, which brings together significant financial institutions from around the world to set industry-wide rules.

In addition, the bank should be an active participant in conferences and workshops that cover emerging XAI topics and collaborate on research that can drive the field and its practical applications forward. We will also push vendors to continue making prepackaged models more explainable so they can more easily adopt third-party solutions. Banks must share their experiences with regulators and collaborate with government agencies and other oversight entities to set norms that allow AI research and development.

Moreover, we will explore explainable AI (XAI) methodologies to ensure fairness. AI models have progressed such that the calculation processes of the algorithm may be a “black box." Even the engineers, developers, or research and data scientists who created the algorithm cannot understand or explain the reasons behind the algorithm’s decision or output. We improve fairness by being able to ensure the system is working as intended and by being able to provide an explanation to the end user who was affected by the decision. XAI also helps debug and improve the model development process while promoting user trust and auditability. To this end, we can make use of readily available toolkits, such as:

- SHAP (Shapley Additive exPlanations) A unified framework based on game theory (a theoretical framework built for developing social situations among competing players) is designed to interpret the predictions and outputs of ML models by assigning a value of importance for a particular prediction to each “feature” of the model. SHAP estimates and demonstrates how each feature in the model influences the model. It calculates values for each feature and value and approximates the contribution of the given output using a particular data point.
- LIME (Local Interpretable Model-Agnostic Explanations)

  LIME provides “local explanations” of the classifier for a single instance. It works by creating a series of artificial data from the input data and assigning these to different categories or classifiers, which provides further insights into the model. LIME can be used across various classifiers, including text, image, and tabular data.
- Explainer Dashboard

  An open-source library that helps developers and data scientists build interactive dashboards to analyze and predict ML models It investigates SHAP values, permutation importance, interaction effects, partial dependence plots, all kinds of performance plots, and even individual decision trees inside a random forest.

We will integrate these explainable AI methods into our system to improve fairness, transparency, and user trust, which are important parts of our proposed AI model, as they would allow us to detect and resolve any bias issues that may have occurred in our system.

5.2 Fairness

As banks increasingly use AI systems to make credit decisions, they must confront an inconvenient fact about lending: It has a history of biases against sensitive data such as race, gender, and location [6]. Such biases are available in financial institutions’ decisions about who allowed credit and on what terms. In this case, relying on algorithms to make credit decisions rather than human judgment appears to be an obvious solution. The problem is that humans are frequently incapable of detecting unfairness in the large number of data sets that machine-learning systems study. As a result, banks are increasingly relying on artificial intelligence to detect, predict, and reduce biases against protected classes that are accidentally built into algorithms.

Remove bias from data before building a model.
To minimize bias in a credit decision, remove discrimination from the data before generating the model. However, this needs more adjustment than simply minimizing data variables that clearly indicate gender or race, because earlier bias has ripple effects throughout. For instance, loan data samples for female customers are typically smaller than those for male customers. As a result, underrepresented and unevenly treated female applicants experience greater errors and erroneous assumptions. Manual interventions to correct data bias can result in self-fulfilling prophecies when errors or assumptions are repeated and magnified.

To prevent this, we can now use AI to detect and correct patterns of historical bias against female customers in raw data, accounting for changes over time by purposefully modifying this data to provide an artificially higher, more equitable chance of approval. For instance, one bank revealed that in the past, female customers had to earn 30% more than male customers on average in order for equivalent-sized loans to be granted. We will use this approach to balance the data that went into developing and testing the AI-driven credit decision model by modifying the gender distribution and the fraction of past loans given to female customers to be closer to the same amount as for male customers with a similar risk profile while retaining the relative ranking, which will improve fairness and adhere to the line of how our bank wishes to extend credit more equitably in the future.

5.3 Privacy preservation

This element focuses on securing individuals' personal and sensitive information. Privacy preservation in loan approvals would imply that any data utilized by the AI system is handled securely, with access limits and encryption, and is compatible with data protection requirements such as federated learning and differential privacy to promote data privacy. It would also imply that the system only uses the data required to make a judgment and does not use personal information for any other reason. This implies a commitment to safeguarding applicants' sensitive information throughout the process, from data collection to storage and processing, in compliance with applicable privacy laws and regulations.

In this case, AI developers should consider several processes while developing a system for loan approval. They have the task of securing customers' sensitive information, such as personal identification documents, account information, etc...

6 Data Ethics

Since our system deals with highly sensitive information, we prioritized data ethics as one of the key components during system implementation. We'll make sure that our AI system is dedicated to upholding ethical standards in the use of customer data. We took into consideration several fundamental concepts, including ownership and control, which provide the client authority over their personal data through express consent and unambiguous opt-out choices. All processing operations will prioritize transparency, and users will receive explicit communication of all decisions. By utilizing strong security and AI privacy, our suggested method also encourages fair access to data and privacy. Following the characterization of the problem and the assessment of the data gathering aims, a cooperative, moral strategy including experts and stakeholders will be used prior to execution. Continuous monitoring and analysis of results is done to identify potential risks, promoting inclusiveness and justice.

Conclusion

In conclusion, our Loan Approval project proposal presents a transformative AI-based solution to the inefficiencies and biases plaguing the current manual loan approval system. By integrating advanced AI technologies like federated learning and differential privacy, our system not only improves efficiency and scalability but also ensures fairness and privacy in the loan approval process. This modern approach promises significant benefits for both financial institutions and their customers, leading to speed-up loan processing, enhanced client satisfaction, and more equitable loan decisions. The adoption of such AI solutions is not just an advancement in technology but a necessary step towards a more efficient, fair, and secure financial future. We strongly advocate for financial institutions to practice this innovative shift, ensuring a more reliable and customer-centric approach to loan approvals. The era of AI in finance is upon AI developers and users, and it is imperative that we harness its potential to reshape and improve essential financial processes like loan approvals.

Recommendations to the Future work

Artificial Intelligence continues to transform industries around the world, and financial institutions need to use the potential of AI to improve performance optimization and reduce employee workloads. We recommend other researchers focus more on developing AI-based solutions that can automate different routine tasks, simplify processes, and provide decision-making support in various areas of the financial industry.

Acknowledgment

We have used QuillBot for paraphrasing and ChatGPT to make our content follow academic formatting.

References

[1] M. Abadi and D. Lindell. Secure multi-party computation: An introductory survey. *Communications of the ACM*, 57(4):54–64, 2013.

[2] Cynthia Dwork. *Differential privacy: From theory to practice*. Cambridge University Press, 2008.

[3] Alexey Surkov Val Srinivas Jill Gregorie. Unleashing the power of machine learning models in banking through explainable artificial intelligence (xai). https://www2.deloitte.com/xe/en/insights/industry/financial-services/explainable-ai-in-banking.html, 2023. Accessed: 2023-10-25.

[4] Garrett Laird. Traditional ai can help banks compete for better loans. https://thefinancialbrand.com/news/data-analytics-banking/artificial-intelligence-banking/traditional-ai-can-help-banks-compete-for-better-loans-171418/, 2023. Accessed: 2023-11-24.

[5] Michael Tetrick. Six essential steps of the loan process. Rely services Reinvent, 2023. Accessed: 2023-11-28.

[6] Sian Townson. Ai can make bank loans more fair. *Harvard Business Review*, 2020. Accessed: 2023-10-25.

[7] Pankaj Tripathi. Artificial intelligence-based loan management. https://www.docsumo.com/blog/artificial-intelligence-based-loan-management, 2023. Accessed: 2023-11-28.

-End-