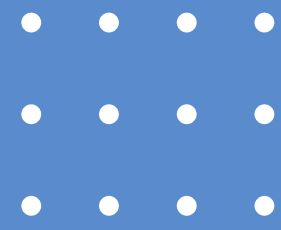


**GROUP 4**



# HOME LOAN APPROVAL PREDICTION

**Presented To :  
Dr. Vinay Singh**



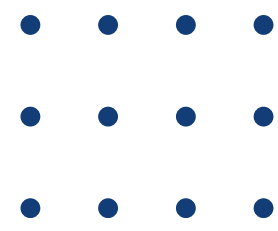
# Our Team

Gourav Waskle (2020IMG-024)  
Sourabh Singh Jindal (2020IMG-062)

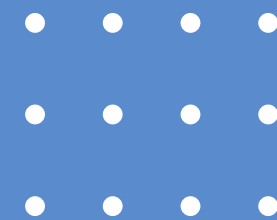


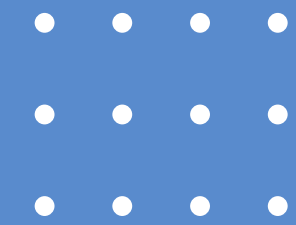
# Outline

## Table of Content



- **Introduction**
- **Motivation and Problem statement**
- **About Dataset**
- **Proposed Methodology**
- **Result**
- **Conclusion**
- **Novelty and Future scope**

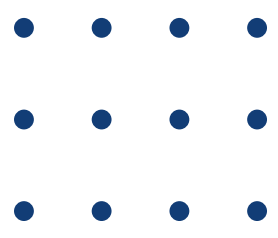




# Introduction

Loans are essential for various needs, like education and buying homes or cars. However, deciding if someone qualifies for a loan can be a complex process. Machine Learning and Python are used to simplify this task. It aids in predicting if an applicant's profile aligns with loan approval criteria. Typically, bank personnel make these decisions, but the process can be slow and sometimes biased. This project's goal is to improve it. The aim is to enhance the loan approval process, making it more efficient and accurate. Ultimately, the objective is to create a fair and helpful financial system accessible to all.

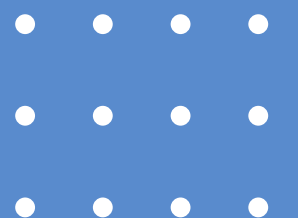


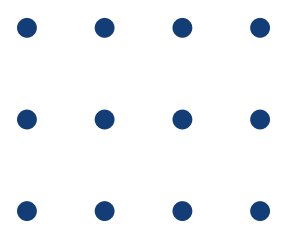


# Motivation and Problem statement



- **Enhance efficiency** : Automating the loan approval process saves time and effort, benefiting both bank staff and customers.
- **Speed up service** : Automation can significantly improve the speed of service provided to customers, reducing processing time.
- **Improve accuracy** : Machine learning algorithms can provide more accurate predictions for loan approval based on applicant information.
- **Complex decision-making** : Loan approval involves evaluating various aspects of an applicant's profile, making it a challenging process for banks.

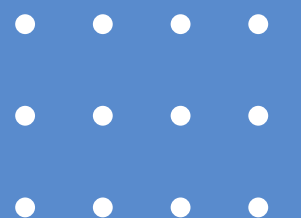




# Motivation and Problem statement



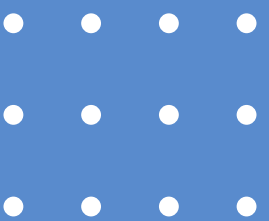
- **Digital transformation** : The financial sector is undergoing a digital transformation, with AI and ML leading this revolution.
- **Subjectivity in old process** : The old loan approval process relies on subjective human judgment, leading to potential unfairness and inconsistency.
- **Create a transparent system** : The project aims to establish a transparent, efficient, and inclusive financial system, benefiting both lenders and borrowers.



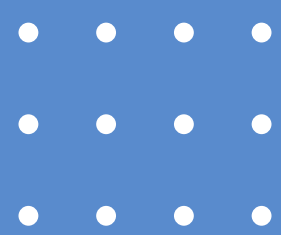
# About Dataset

• • • •  
• • • • The dataset contains 12 features  
• • • •

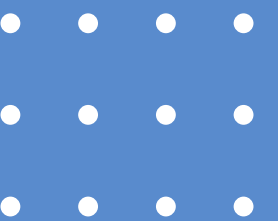
1. **Loan\_id** : A unique id
2. **Gender** : Gender of the applicant Male/female
3. **Married** : Marital Status of the applicant, values will be Yes/ No
4. **Dependents** : It tells whether the applicant has any dependents or not.
5. **Education** : It will tell us whether the applicant is Graduated or not.
6. **Self\_Employed** : This defines that the applicant is self-employed i.e. Yes/ No
7. **ApplicantIncome** : Applicant income



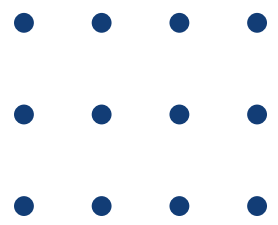
# About Dataset



- 8. **CoapplicantIncome** : Co-applicant income
- 9. **LoanAmount** : Loan amount (in thousands)
- 10. **Loan\_Amount\_Term** : Terms of loan (in months)
- 11. **Credit\_History** : Credit history of individual's repayment of their debts
- 12. **Property\_Area** : Area of property i.e. Rural/Urban/Semi-urban

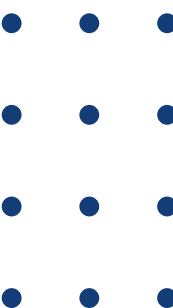
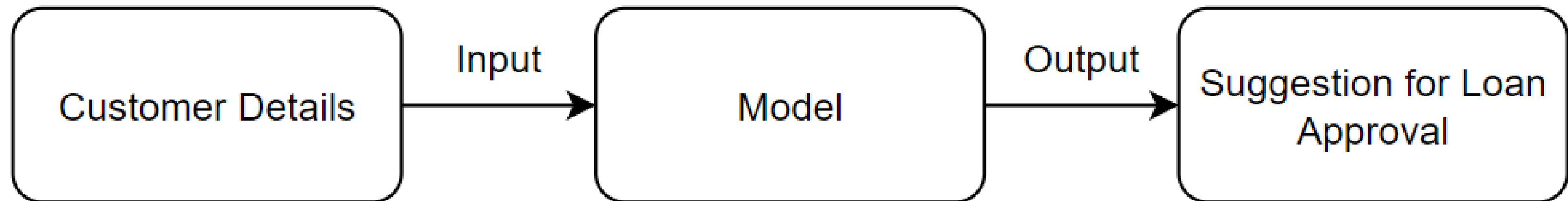






# Methodology

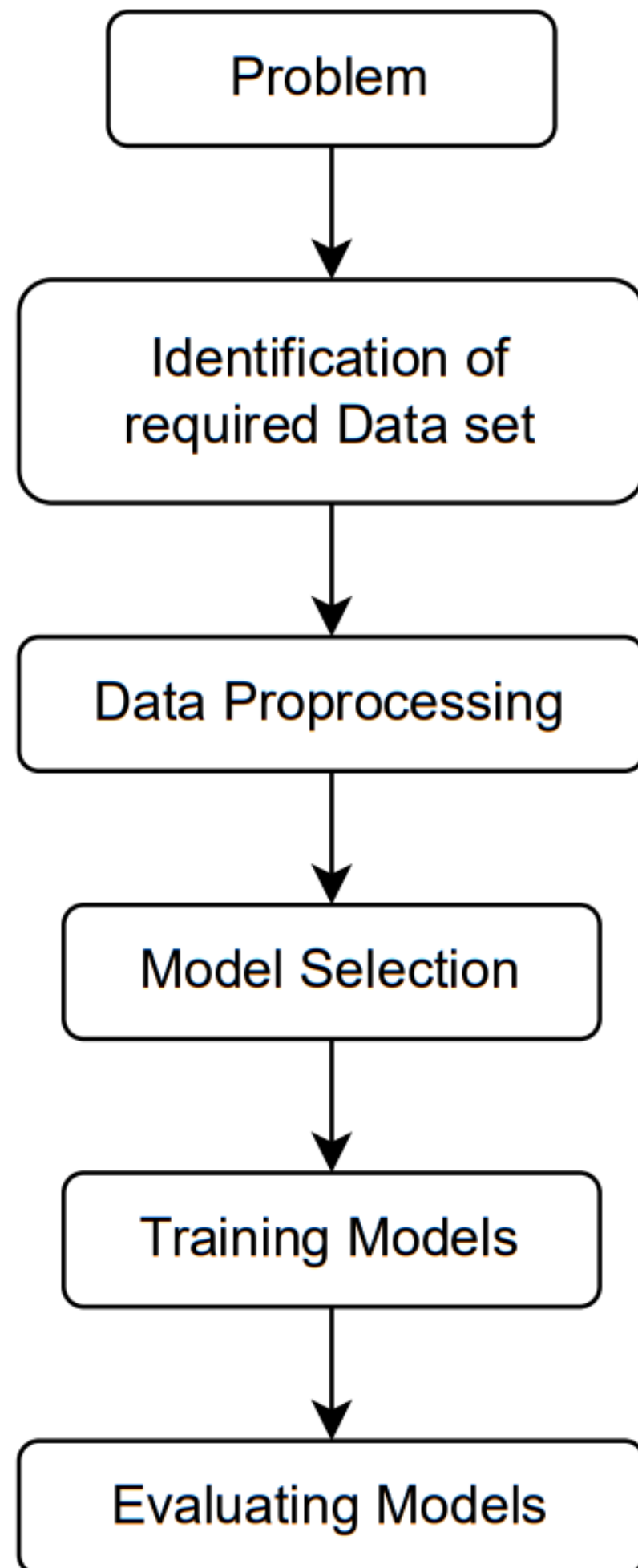
## Process Flow Diagram



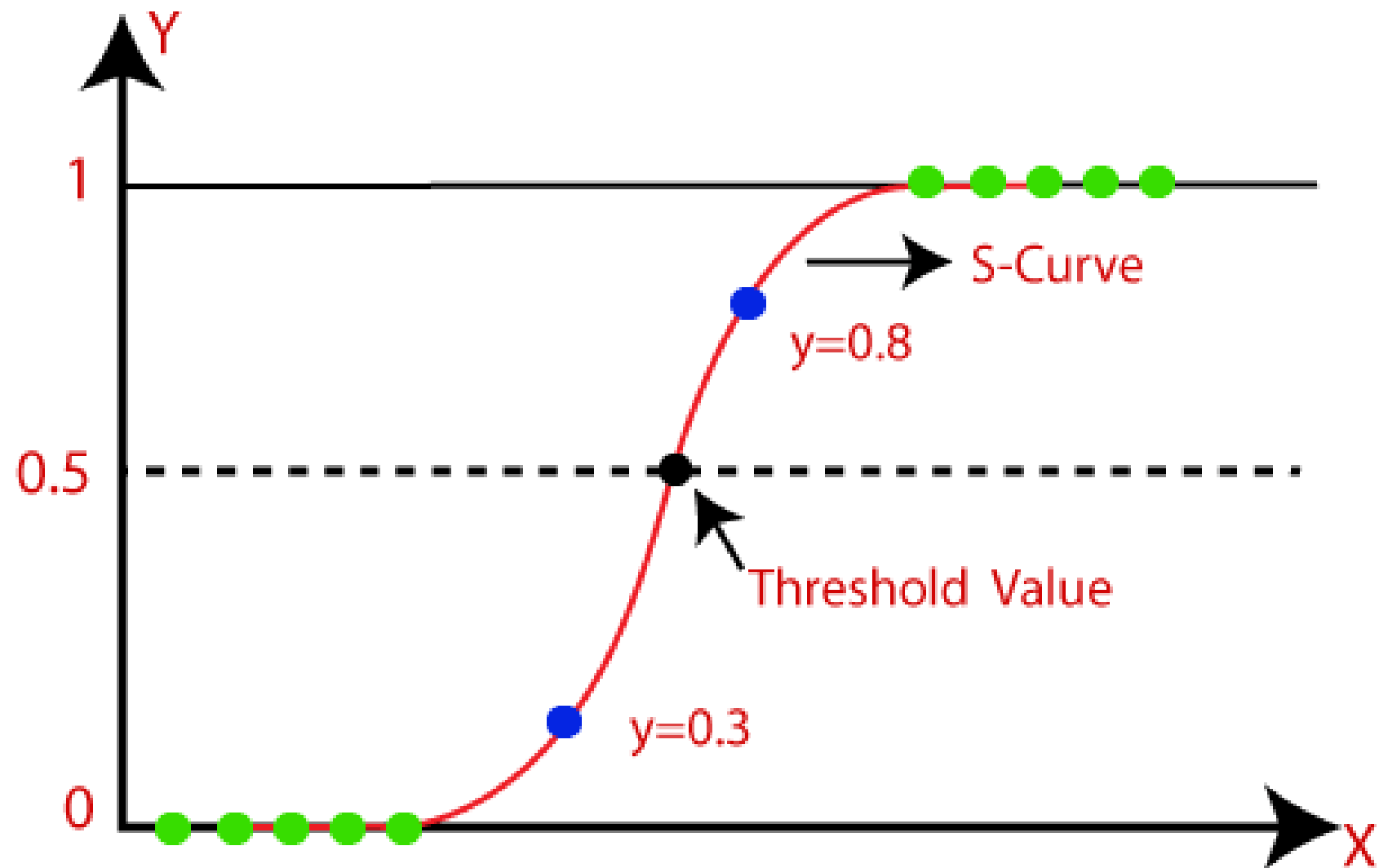
# Methodology

## Frame work

Training a machine learning model for binary classification involves the following steps: data collection, data preprocessing, splitting data into training and testing sets, selecting an appropriate algorithm (e.g., logistic regression or random forest), model training, and evaluating the model's performance using accuracy score.



# Methodology



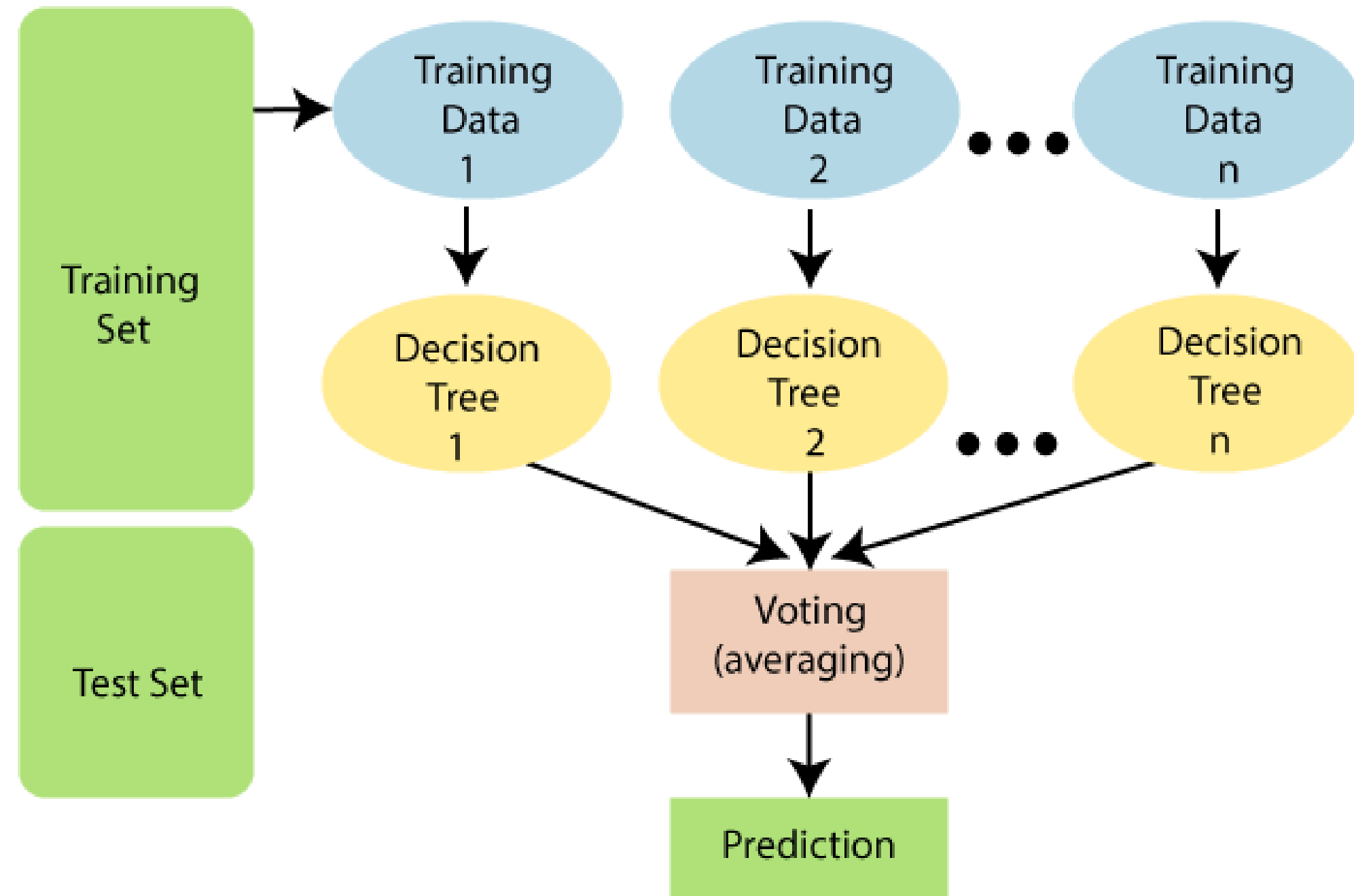
## Logistic Regression

- Logistic Regression models the probability of an instance belonging to a particular class using the sigmoid (logistic) function.
- It employs a cost function (often the cross-entropy loss) to measure the difference between predicted probabilities and actual labels. The model aims to minimize this cost during training.

# Methodology

## Random Forest Classifier

- Random Forest is an ensemble method that aggregates the predictions of multiple decision trees using majority voting.
- Creates subsets of the training data with replacement. Each tree is trained on one of these subsets.
- Selects a random subset of features for each tree.



# Results

- The Dataset consists of 614 entries out of which 60% of the data is used for training and 40% is used testing.
- We are getting different accuracies through the models used.

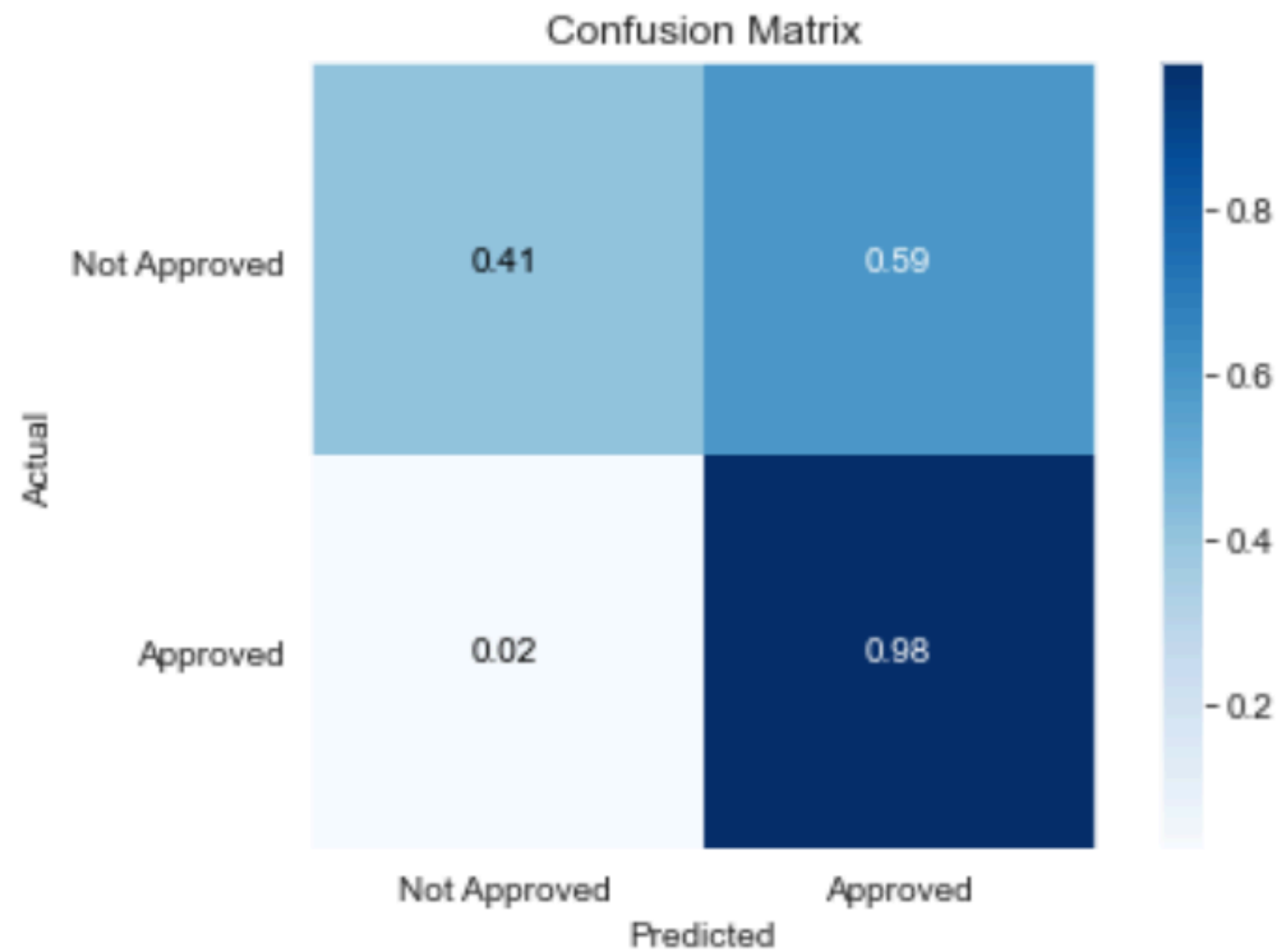
Models used and their accuracy score :

- Logistic Regression- 80.13%
- Random Forest Classifier- 76.01%

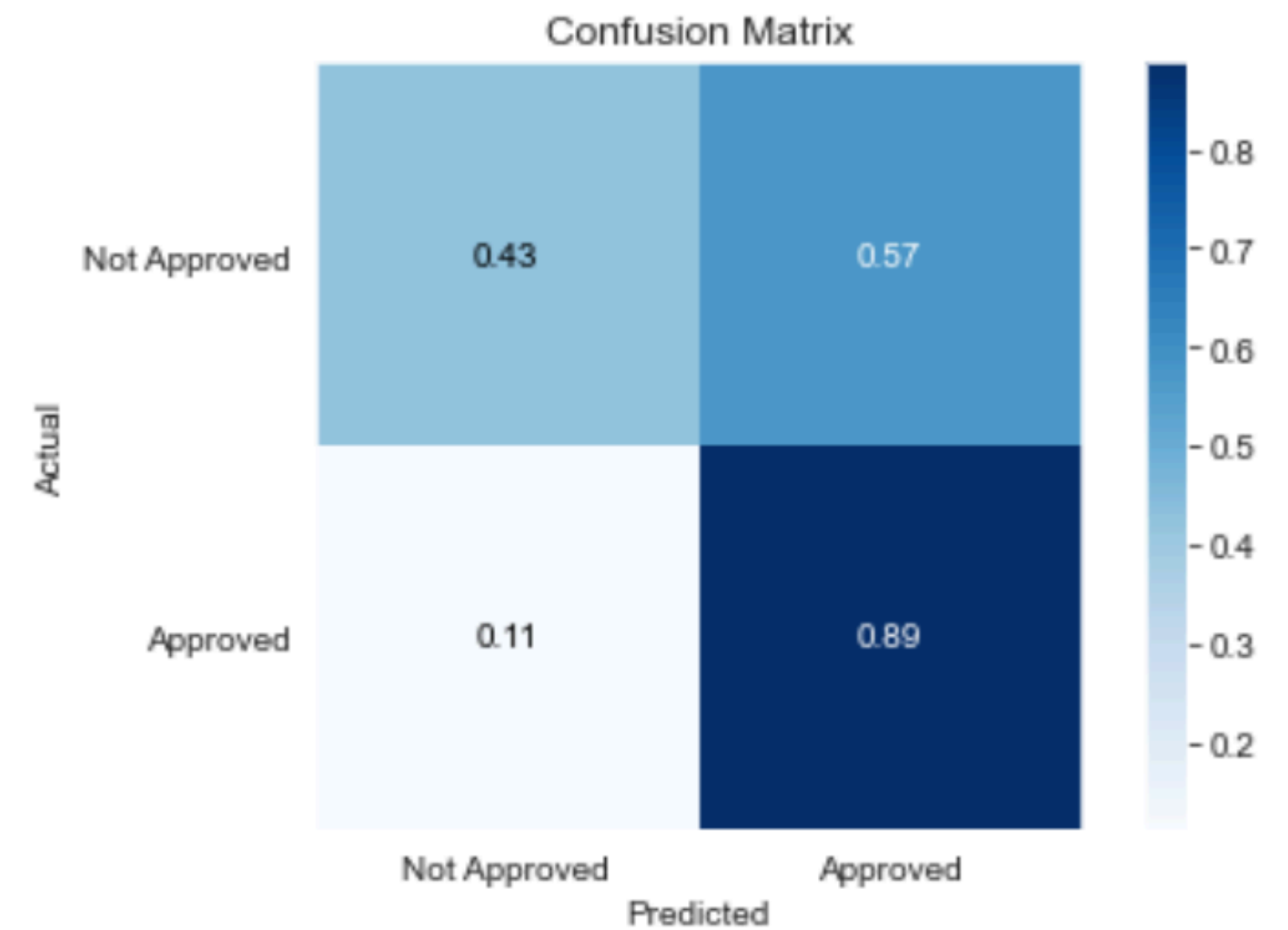


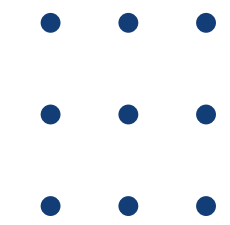
# Results

## Logistic Regression



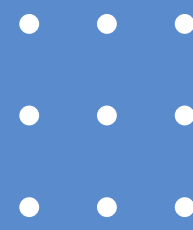
## Random Forest Classifier





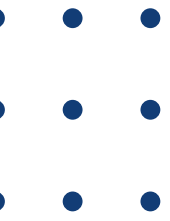
# Conclusion

- This project has practical implications for financial institutions and banks, enabling them to automate and improve loan approval processes.
- The predictor can help institutions better manage risk by making data-driven decisions in the loan approval process, thus reducing the chances of default.
- The predictor enhances the financial institutions and banks' experience by providing quick and transparent loan approval decisions.





# Future Scope and Novelty



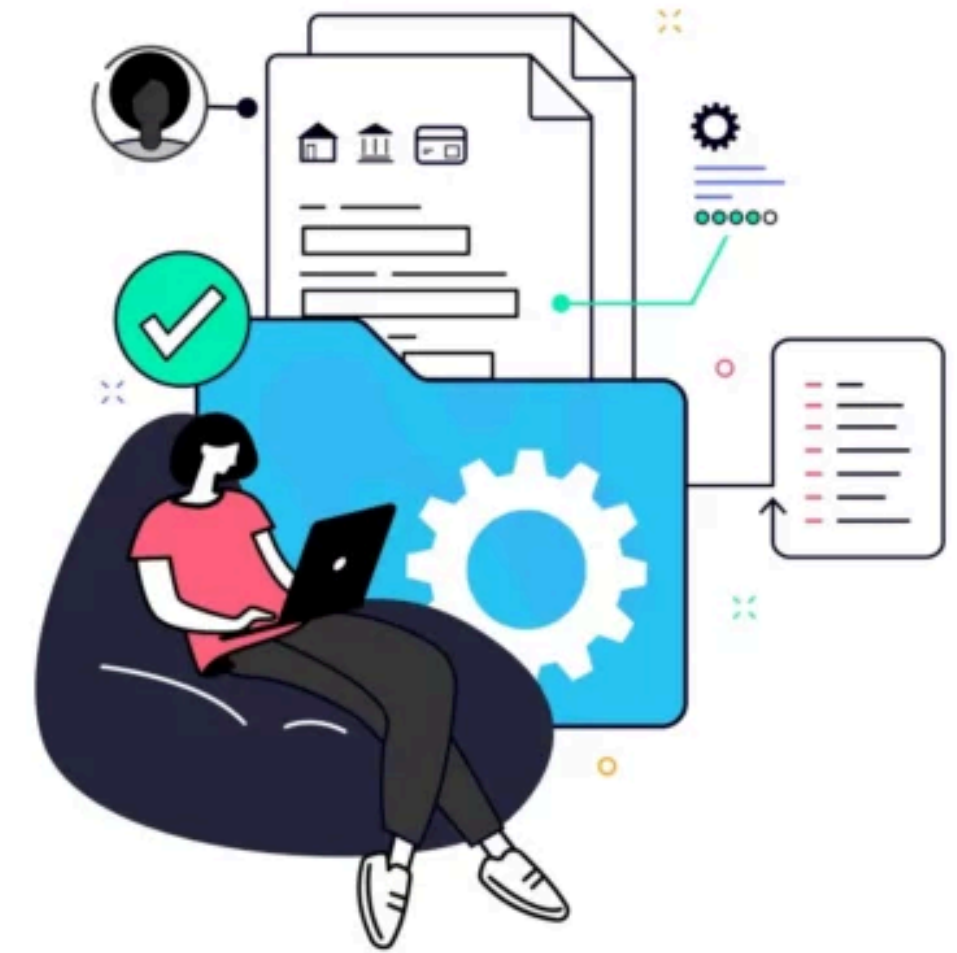
- **Integration with Financial Data:** To integrate with external financial data sources, such as customer's credit history, debt-to-income ratio, and other financial indicators, to enhance the accuracy of loan eligibility predictions.
- **Real-Time Processing:** The system could be enhanced to handle real-time loan applications. Customers could receive instant feedback on their eligibility as soon as they submit their application, improving the overall customer experience.
- **Mobile Application:** Developing a mobile application for loan eligibility checks would be a novel approach. Customers could use their smartphones to input their information and receive instant loan eligibility results.





# Future Scope and Novelty

- **Risk Assessment:** Expanding the model to assess the risk associated with a loan application is another valuable extension. This could help to determine interest rates and terms based on the customer's risk profile.
- **Personalized Loan Offers:** Tailoring loan offers to individual customers based on their eligibility, financial history, and preferences can be a unique selling point. Providing personalized offers can increase customer satisfaction and loyalty.
- **Customer Feedback Loop:** Creating a feedback loop with customers to collect their input on the loan eligibility process and model can lead to continuous improvement and innovation.



**THANK YOU FOR  
YOUR ATTENTION**

