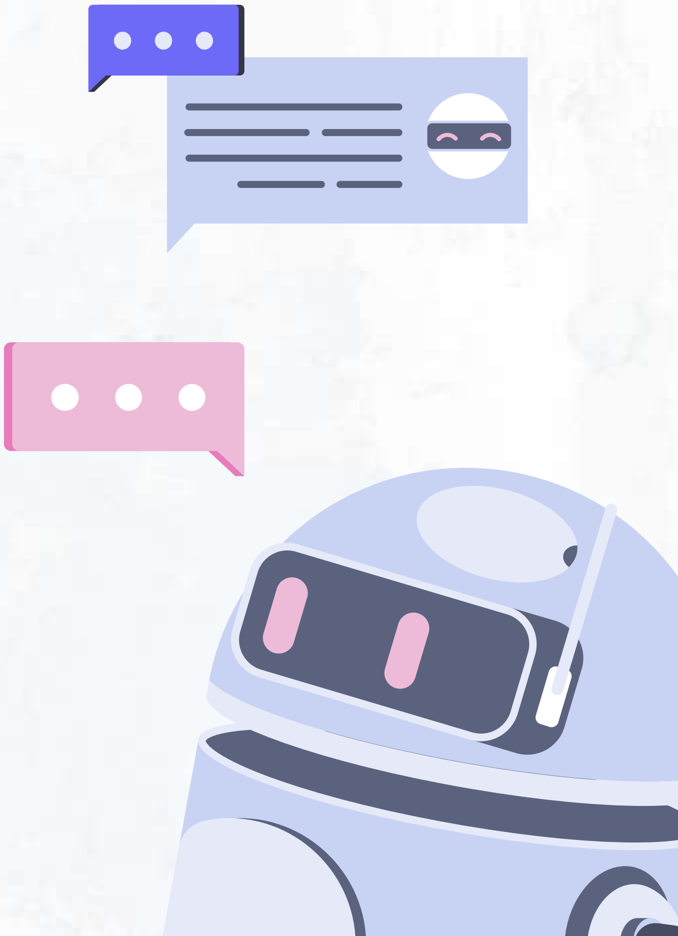
Improving Transparency and Fairness in Loan Approval Algorithms Through Explainable AI

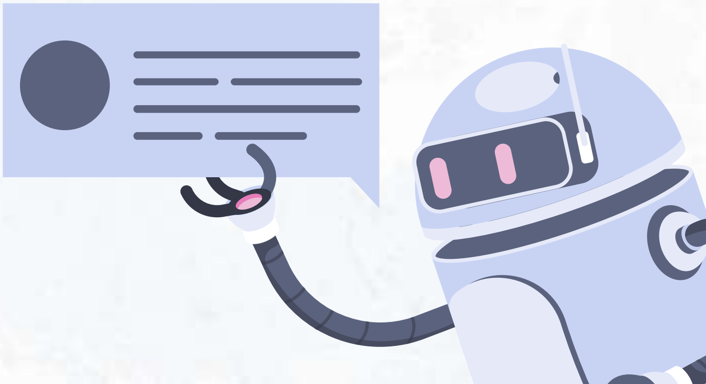


Final Class Project – Two-Person Team

Table of contents

01 →

What is The Problem?



What is The Problem?

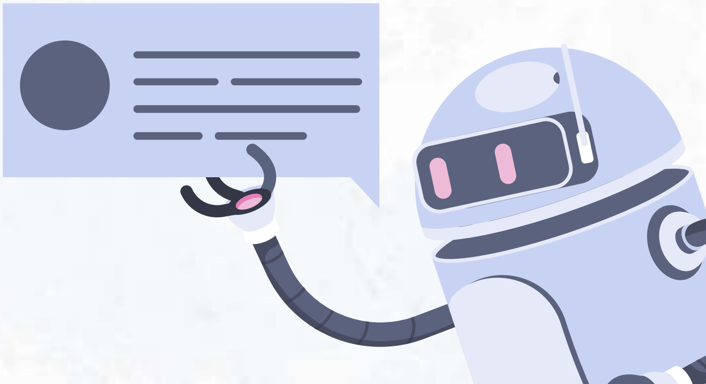
- **Lack of Transparency in Loan Approval**: Loan approval algorithms are often "black boxes," making decisions without clear explanations.

- **Bias in Decision-Making**: These algorithms may inadvertently discriminate based on race, gender, or income, leading to unfair loan denials or approvals.

- **Impact on Marginalized Groups**: Discriminatory decisions disproportionately affect marginalized groups, exacerbating economic inequality.

02 →

Why is This an Interesting Problem?

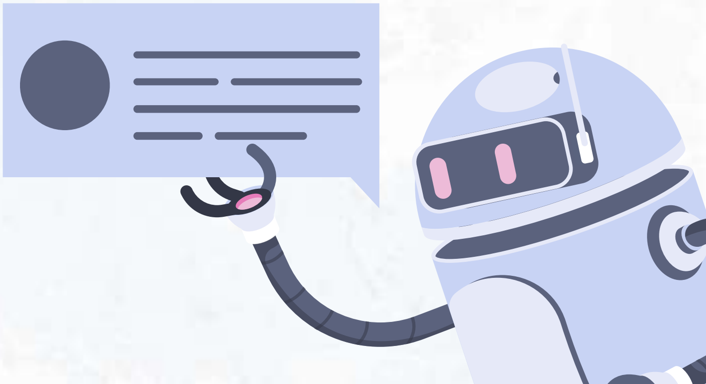


Why is This an Interesting Problem?

- **Ethical and Economic Justice**: Access to fair financial opportunities is a critical aspect of economic justice.
- **Bias in AI**: Machine learning models in financial systems can perpetuate historical biases, making it essential to address these issues.
- **Relevance to Current AI Debates**: There is increasing concern over AI's role in societal outcomes, especially in high-stakes areas like finance.
- **Innovation in XAI**: Few studies have applied Explainable AI (XAI) techniques to loan approval systems, making this research particularly relevant.

03

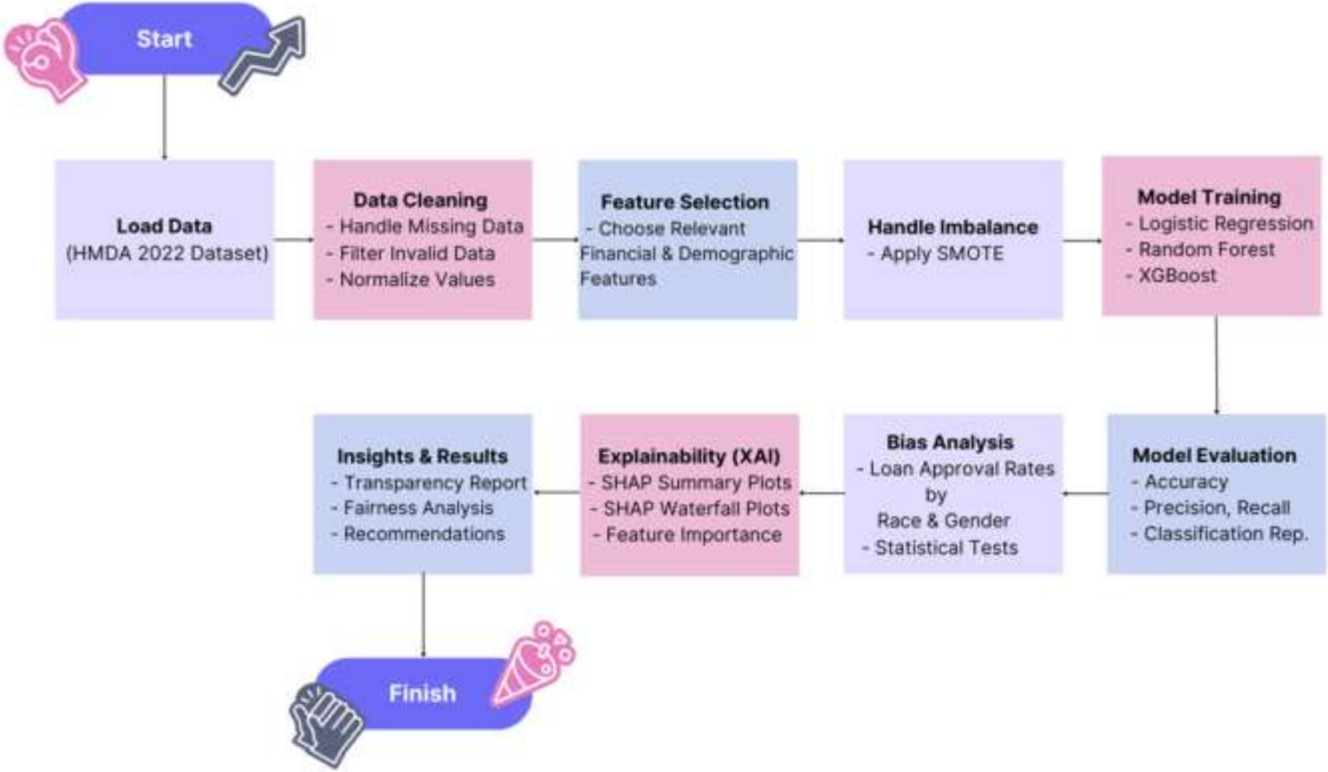
How Did We Tackle The Problem?



How Did We Tackle The Problem?

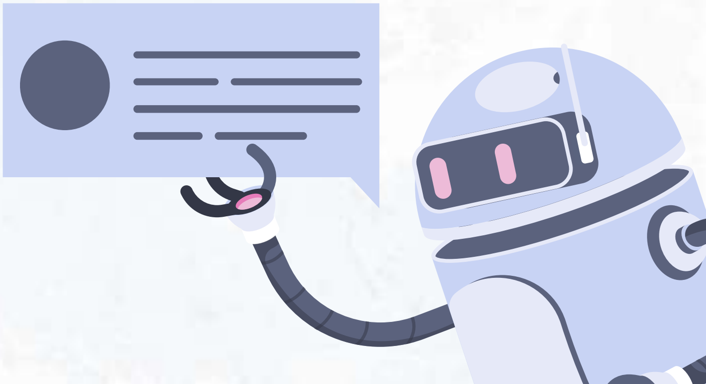
- **Dataset**: We analyzed the *Home Mortgage Disclosure Act (HMDA) 2022* dataset for New Jersey, focusing on factors like race, gender, income, and loan outcomes.
- **Data Cleaning & Preprocessing**: We handled outliers, missing values, and ensured consistency in variables like race and income.
- **Use of XAI**: We applied fairness algorithms and XAI techniques (e.g., SHAP) to enhance the transparency and interpretability of loan decisions.
- **Modeling**: We tested multiple machine learning models, including Logistic Regression, Random Forest, and XGBoost, to assess fairness and accuracy.

Flowchart: Project Workflow



04

Findings



Findings

Biases Identified:

- **Race**: Significant disparities in loan approval rates between racial groups, with White and Asian applicants having higher approval rates.
- **Income**: Higher income levels correlated with higher approval rates.
- **Gender**: Minor gender differences, but statistically significant.

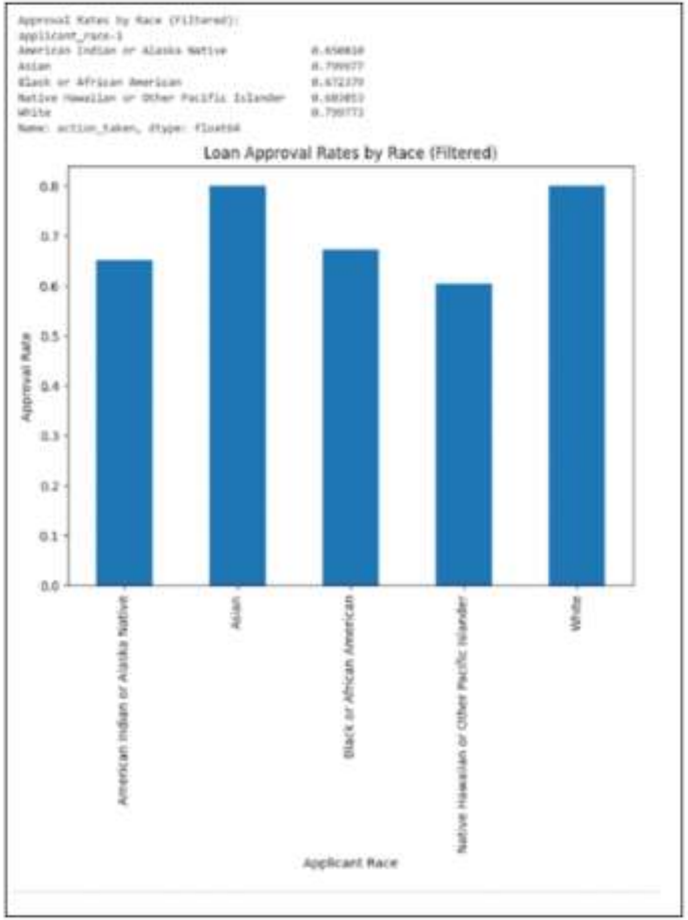


Figure 1: Loan Approval Rates by Race

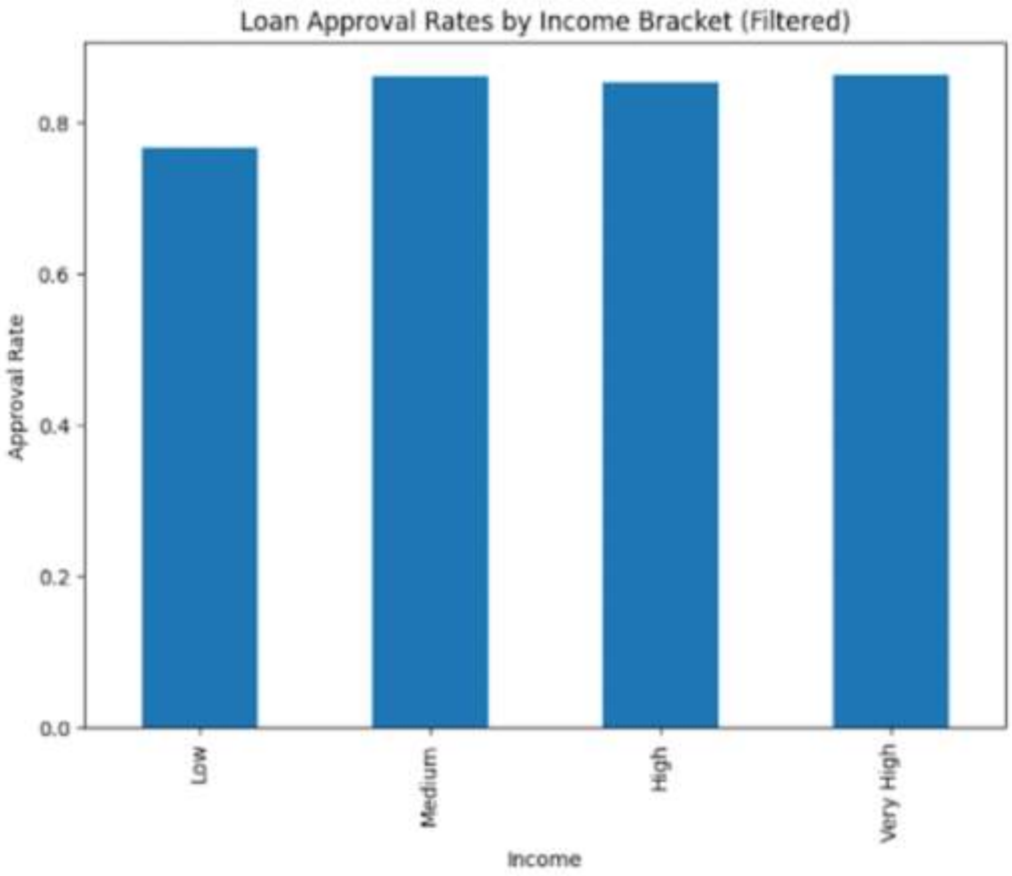


Figure 2: Loan Approval Rates by Income Bracket

```
Approval Rates by Gender (Filtered):
applicant_sex
Female     0.779973
Male       0.786238
Name: action_taken, dtype: float64
```

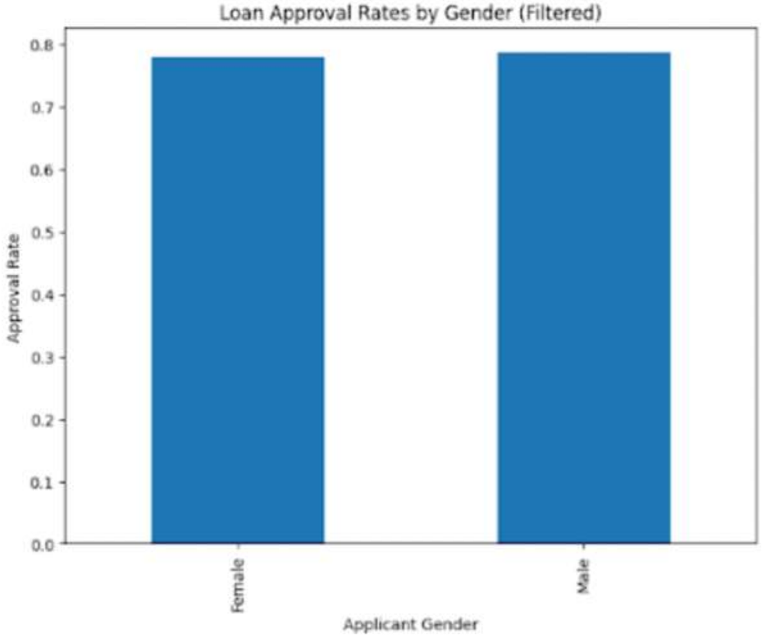
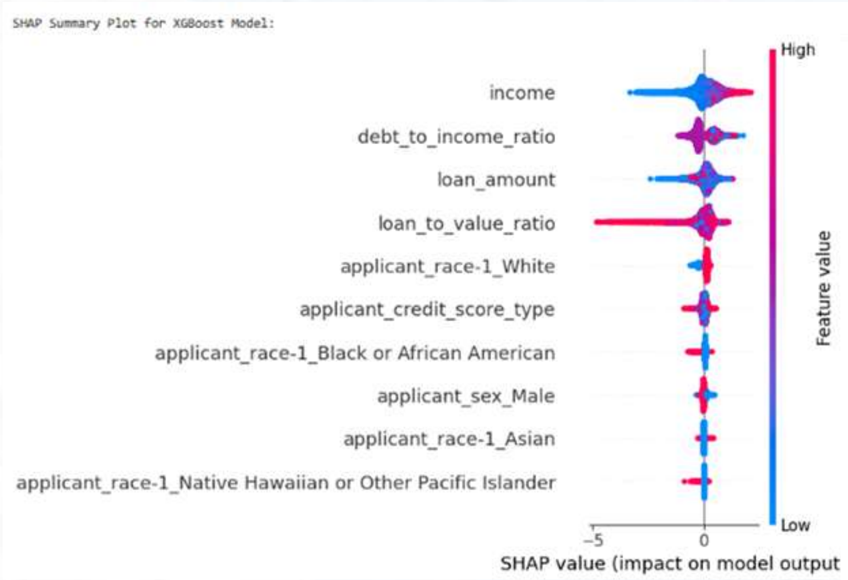
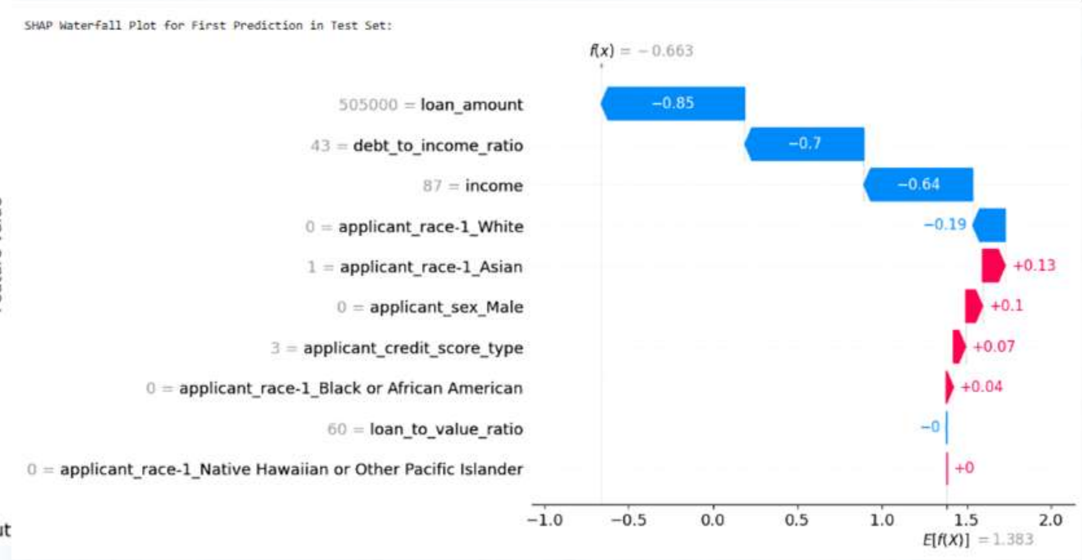


Figure 3: Loan Approval Rates by Gender





Figures 4-5: SHAP Waterfall Plot Analysis

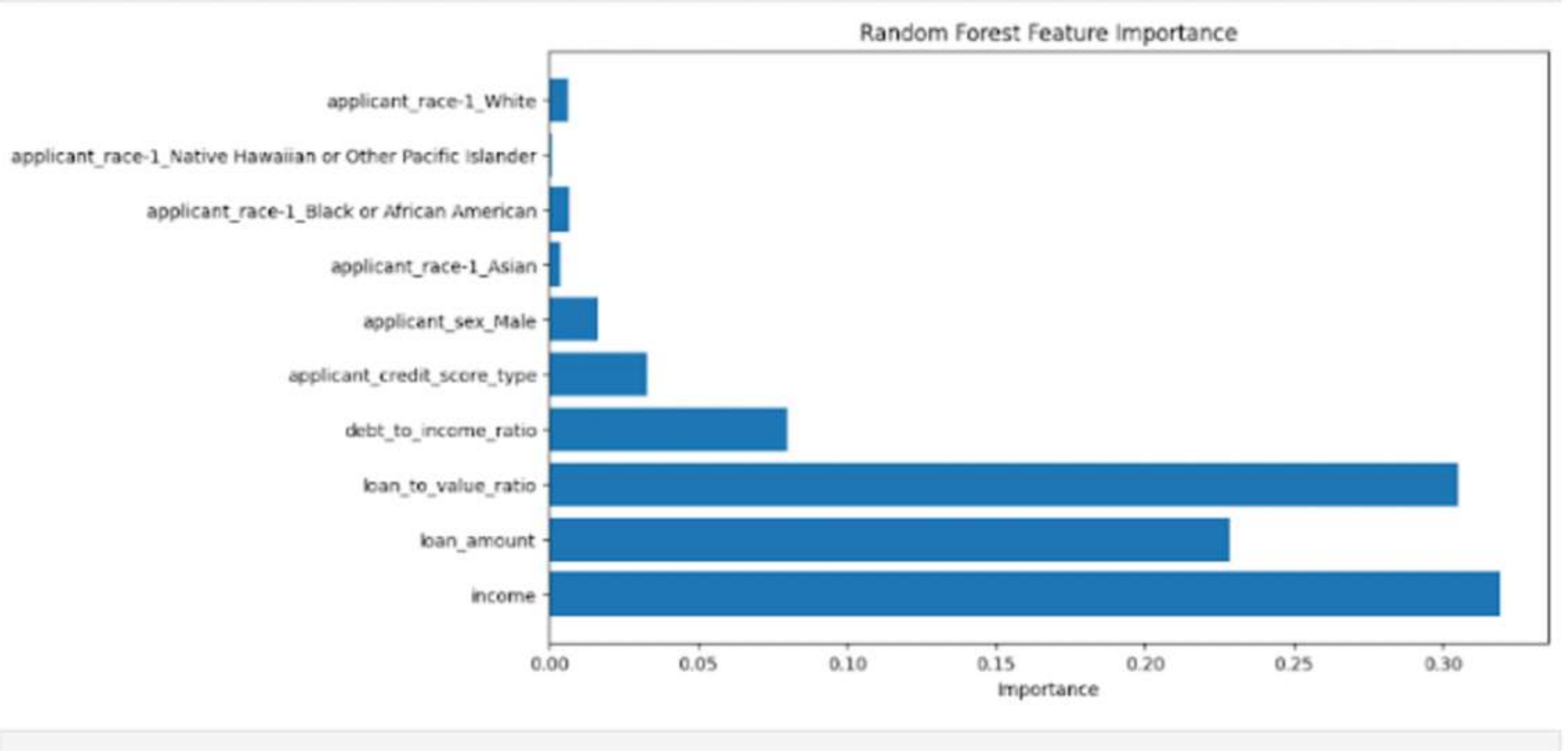


Figure 6: Random Forest Feature Importance Analysis

Findings

Loan Amount and Debt-to-Income Ratio: Key reasons for loan denials

Income: A primary driver for approvals.

Demographic Influence: Race and gender play a smaller role but may still cause bias.

Conclusion

Bias Exists: Significant biases related to race and income were identified in loan approval processes.

XAI Helps: Using XAI techniques such as SHAP improves model transparency, but it’s not a complete solution to bias.

Next Steps: Additional work is needed to improve recall for loan rejections and to further refine fairness algorithms in loan approval systems.

Thank You!