**Minor Project Report**



**Loan Approval Prediction Dashboard**

**Submitted By:**

Arpit Gupta (500122077)

Vatsal Garg (500124447)

Chetanya Jain (500124521)

Course: BCA AIML

Semester: IV

Institution: UPES, Dehradun

 Date: 6th March, 2025

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Heading** | **Remarks** |
| **1** | **Abstract** |  |
| **2** | **Introduction** |  |
| **3** | **Project Overview** |  |
| **4** | **Progress Made** |  |
| **5** | **Methodology Used** |  |
| **6** | **Results & Findings** |  |
| **7** | **Output** |  |
| **8** | **Future Plan** |  |
| **9** | **Conclusion** |  |
| **10** | **Remarks** |  |
|  |  |  |
|  |  |  |
|  |  |  |

**Abstract**

This project presents a **Loan Approval Prediction Dashboard** designed using **Streamlit, Pandas, NumPy, Scikit-learn, and Plotly**. The system aims to automate **loan approval decisions** based on key financial parameters. The methodology involves **data preprocessing, feature engineering, model training, and performance evaluation** using **machine learning models** such as **Random Forest, Gradient Boosting, and Logistic Regression**. The dashboard provides **interactive insights, predictive analytics, and user-friendly visualizations**. Future enhancements include **deploying the model online, integrating real-time data, and improving model performance with deep learning techniques**.

**Introduction**

In the modern financial landscape, ensuring **efficient, accurate, and data-driven loan approval** is critical for banks and financial institutions. Traditionally, loan approvals involved manual reviews based on fixed eligibility criteria, leading to **human biases, inconsistencies, and inefficiencies**. To overcome these challenges, we propose a **Loan Approval Prediction Dashboard**, an **AI-driven solution** that automates loan approval decisions using machine learning techniques.

This project is built using **Streamlit, Pandas, NumPy, Scikit-learn, and Plotly** and aims to provide:  
✔ **Automated loan approval predictions** based on applicant data.  
✔ **Interactive dashboards** with financial insights and trends.  
✔ **Accurate and efficient decision-making** powered by machine learning algorithms.  
✔ **User-friendly visualization** for better interpretation of approval patterns.

The project integrates multiple **machine learning models**, including **Random Forest, Gradient Boosting, and Logistic Regression**, ensuring that financial institutions can **minimize risks and maximize approval efficiency** while maintaining transparency.

### ****Project Overview****

### 

The **Loan Approval Prediction Dashboard** is designed to:

1. **Upload and process loan applicant data** from various sources.
2. **Analyze data** using statistical summaries and graphical visualizations.
3. **Train and compare machine learning models** to predict loan approvals.
4. **Provide real-time loan approval predictions** based on trained models.
5. **Visualize feature importance and model performance** for better decision-making.

**Progress Made**

### ****Completed Tasks:****

✔ Developed core modules: **DataProcessor, ModelTrainer, and Visualizer**.  
✔ Implemented **data preprocessing**: Cleaning, feature engineering, and categorical encoding.  
✔ Trained and evaluated **machine learning models**: Random Forest, Gradient Boosting, and Logistic Regression.  
✔ Built an **interactive UI** using **Streamlit**.  
✔ Integrated **Plotly** for visualizing feature importance, correlation matrix, and confusion matrix.

### ****Pending Tasks:****

❌ Deploying the dashboard on a cloud platform.  
❌ Integrating with a **real-time financial database**.  
❌ Optimizing models for **higher accuracy** and scalability.

**Methodology Used**

**Data Processing**

* **Cleaning:** Removing duplicates, handling missing values.
* **Feature Engineering:** Debt-to-income ratio, loan-to-asset ratio.
* **Encoding:** Categorical variables converted into numerical format using LabelEncoder.

**Model Training**

* **Algorithms Used:**
* Random Forest Classifier
* Gradient Boosting Classifier
* Logistic Regression
* **Cross-validation:** Evaluated using accuracy, precision, recall, F1-score, and ROC AUC.

**Visualization**

* Feature importance plots.
* Correlation matrix.
* Confusion matrix for model evaluation.

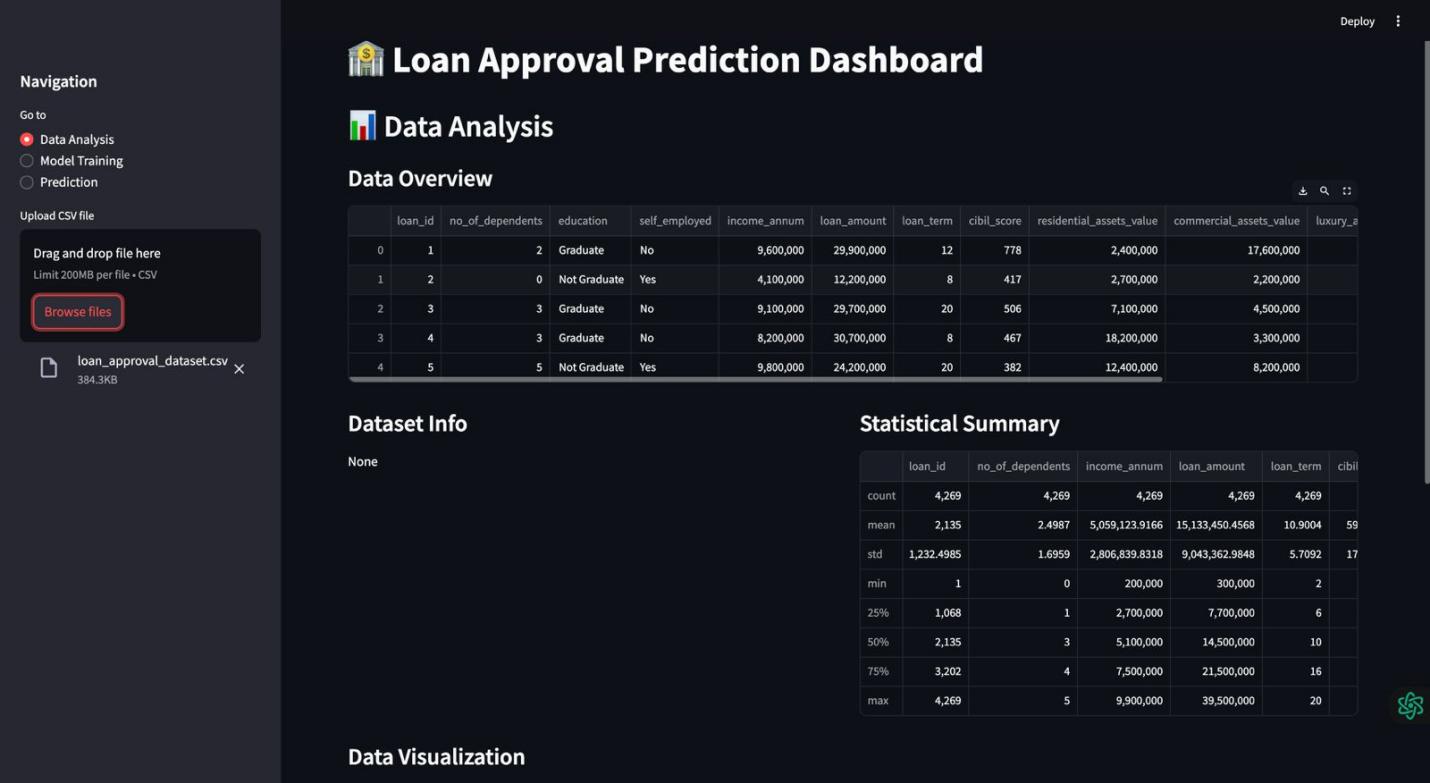
**Results & Findings**

* **Best Performing Model:** [Gradient Descent]
* **Accuracy Achieved:** [100%]
* **Key Influencing Features:** [CIBIL Score]
* **Graphical Analysis:** Visualizations provide insights into key financial factors.

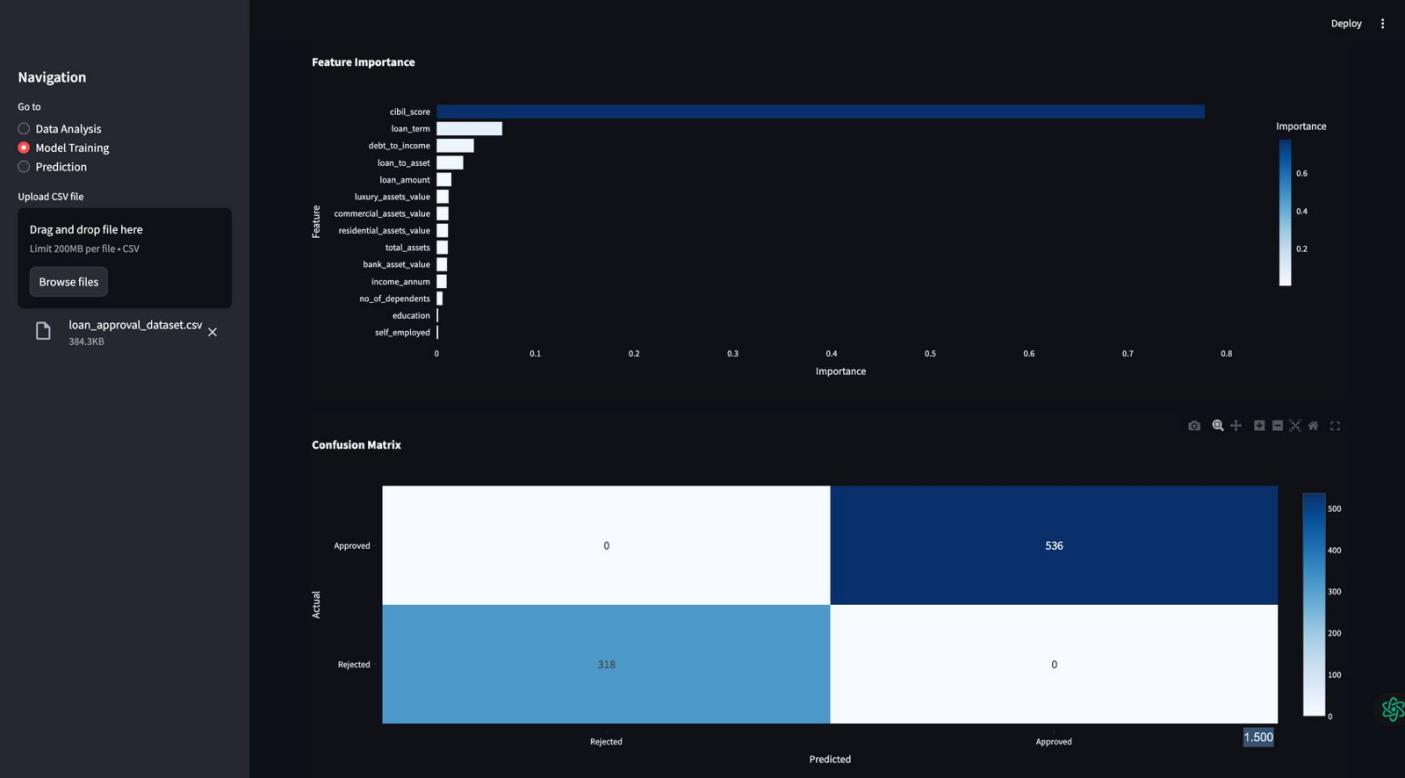
**Output**

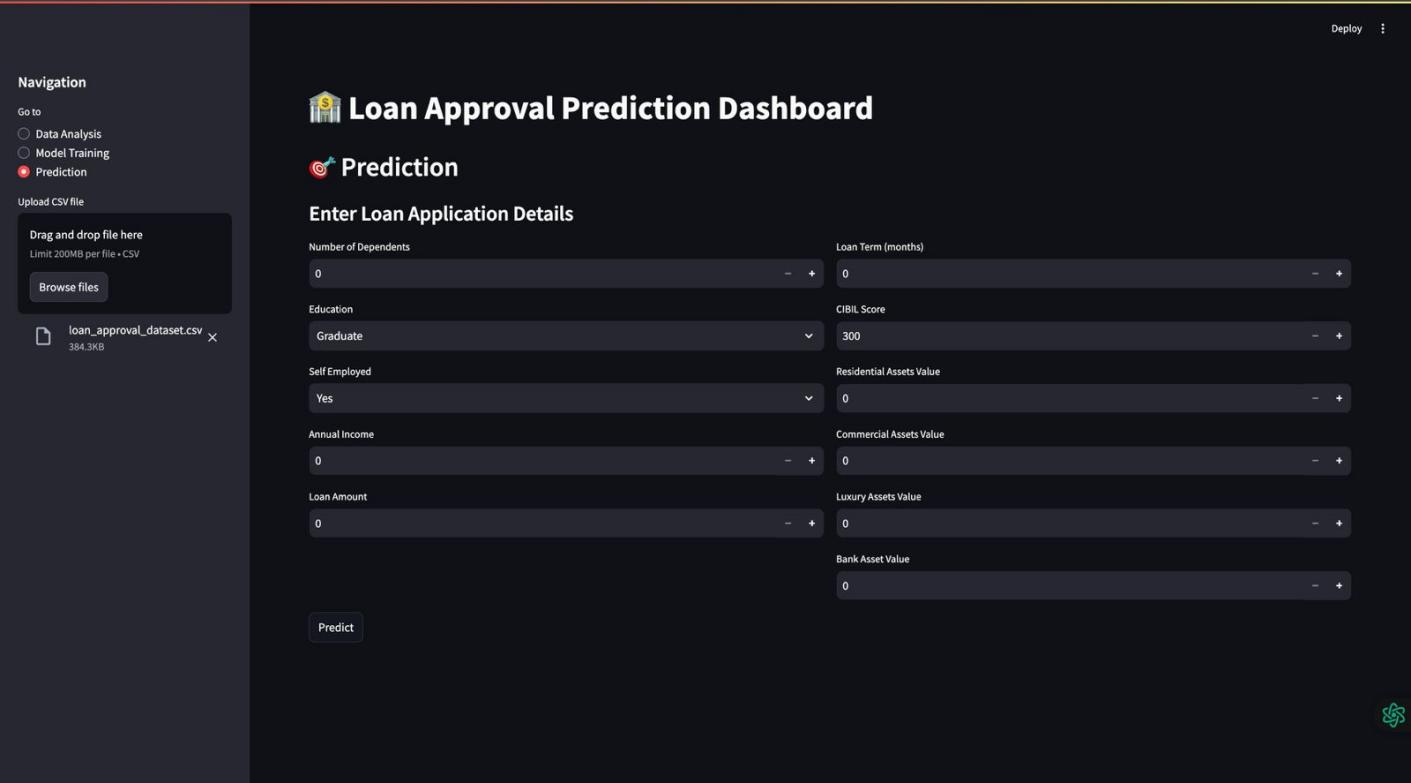
**Interface Screenshots:**

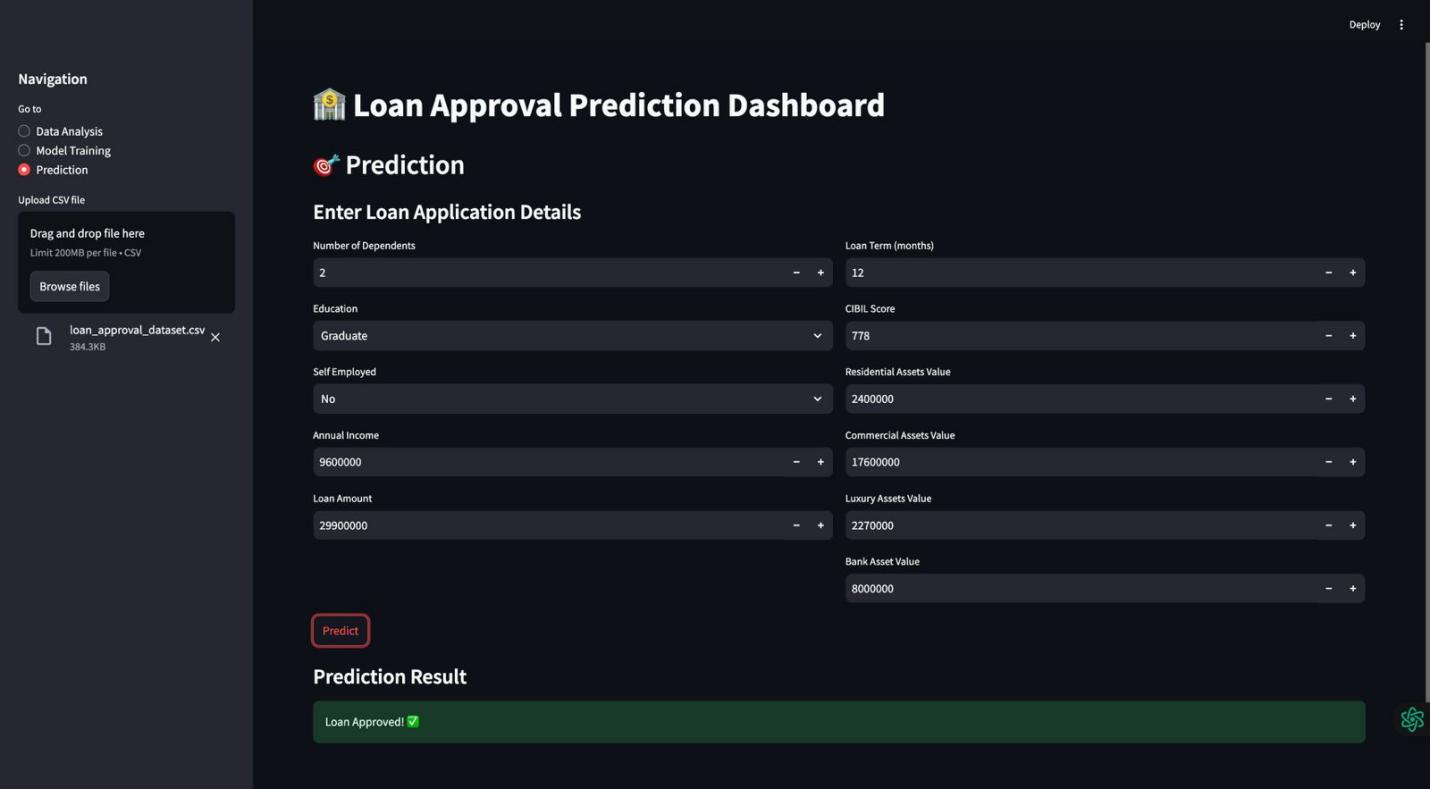
****

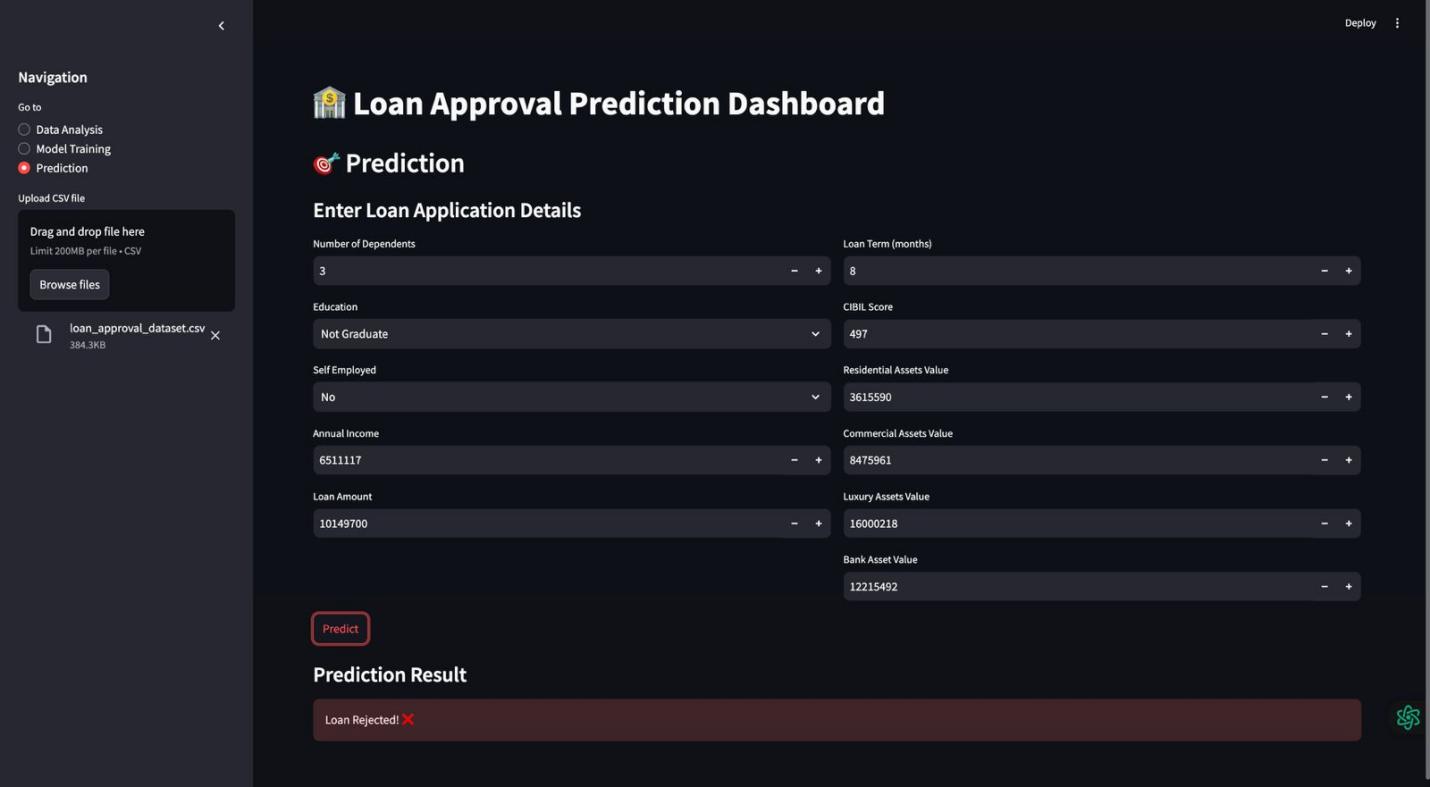
****

****

****

****

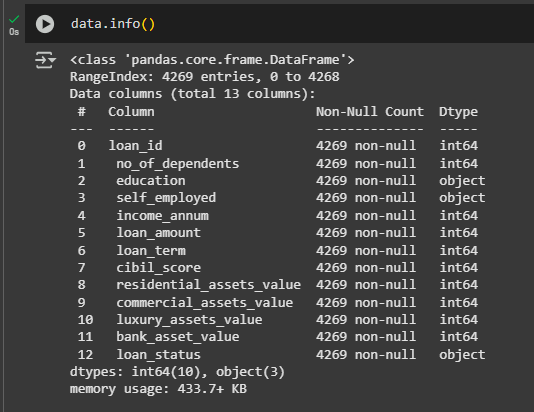
****

****

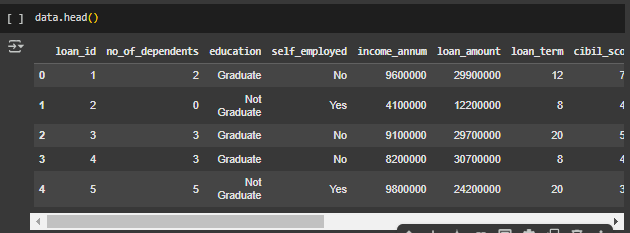
**Output on unseen data**

**Google Colab Screenshots:**

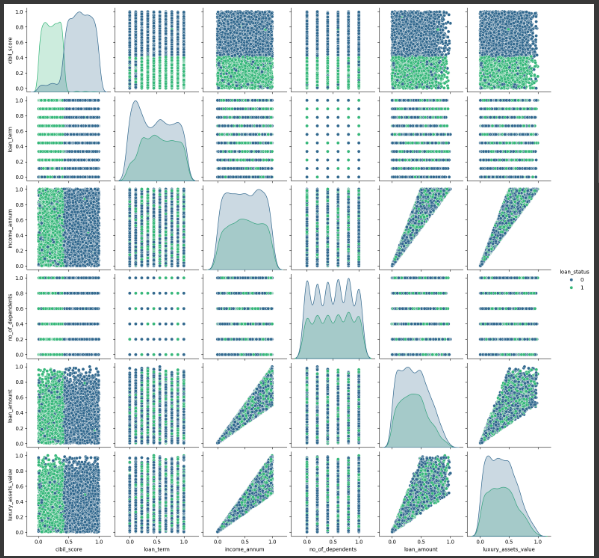
* Data Info



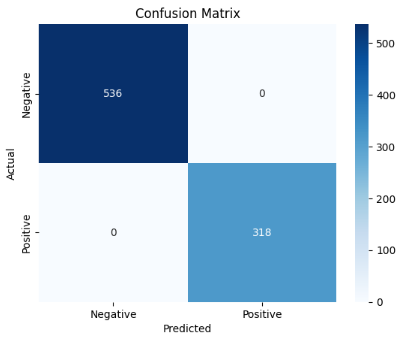
* Data Head



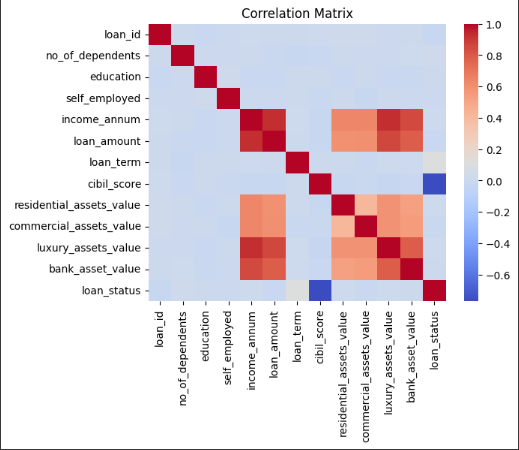
* SNS Pairplot



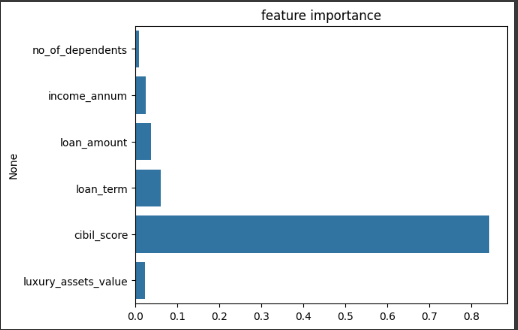
* Confusion matrix where FP is 0.



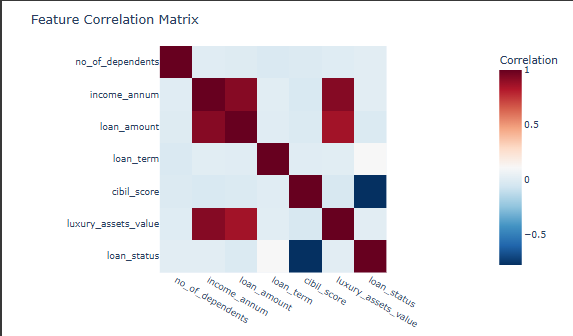
* Heatmap



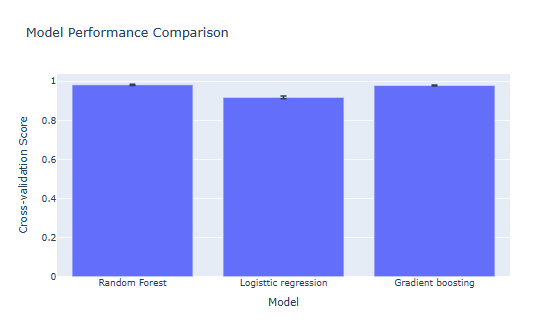
* Barplot which shows that prediction depend on cibil score



* Correlation Matrix



* Model Performance Matrix



**Future Plan**

🔹 **Deploy the dashboard** on **AWS, Google Cloud, or Heroku**.  
🔹 **Integrate an API** for **real-time data processing** and banking system integration.  
🔹 **Enhance model performance** using **deep learning techniques** (Neural Networks).  
🔹 **Improve UI/UX** for better user interaction and accessibility.  
🔹 **Expand dataset** to enhance **model generalization** across diverse applicants.

**Conclusion**

This project successfully demonstrates an **AI-powered loan approval prediction system** that enhances **accuracy, efficiency, and decision-making transparency** in financial institutions. The dashboard ensures **real-time insights and data-driven approvals**, reducing **human bias and operational delays**. Future improvements will focus on **scalability, cloud deployment, real-time data integration, and deep learning enhancements**, making the system **more robust, intelligent, and widely applicable**.

**References**

* **Scikit-learn Documentation:** [https://scikit-learn.org/](https://scikit-learn.org/" \t "_blank)
* **Streamlit Official Guide:** [https://docs.streamlit.io/](https://docs.streamlit.io/" \t "_blank)
* **Pandas & NumPy Documentation:** [https://pandas.pydata.org/](https://pandas.pydata.org/" \t "_blank)
* **Plotly Visualization:** [https://plotly.com/python/](https://plotly.com/python/" \t "_blank)
* **Loan Prediction Dataset:** <https://www.kaggle.com/code/blatalia/loan-approval-prediction/input>
* **Coding:** <https://replit.com/ai>
* **Idea & Roadmap: Perplexity**