
Machine learning for business Analytics

Report: Loan Prediction Machine Learning Project



Submitted to: Sir Hassan Mujtaba

Submitted by: Altaf ahmed (fa23-BBD-022)

Section: BBD-5B

Comsats University Islamabad

Contents

1. Introduction	3
2. Project Objective	3
3. Dataset Description	3
4. Methodology	3
4.1 Data Cleaning	3
4.2 Feature Selection	3
4.3 Model Training	4
4.4 Model Evaluation	4
5. Results and Discussion	4
6. Tools and Technologies	4
7. Conclusion	4
8. Future Improvements	4
9. Repository Link	4
10. Appendix	5

1. Introduction

Access to loans is a key part of the financial system. Banks and lending institutions must decide whether an applicant is likely to repay a loan. Machine learning can help support this decision by finding patterns in historical data.

This report presents a machine learning project focused on predicting loan approval status using applicant and loan related features. The project follows a complete workflow from data preparation to model evaluation.

2. Project Objective

The main objective of this project is to build a supervised machine learning model that can predict loan status based on applicant information. The project also aims to understand how feature selection and data quality affect prediction performance in financial datasets.

3. Dataset Description

The dataset contains real world loan related attributes. The key variables used in this project are:

- Person age
- Person income
- Employment experience in years
- Loan amount requested
- Credit score
- Loan status (target variable)

The target variable indicates whether a loan is approved or rejected.

4. Methodology

4.1 Data Cleaning

The data cleaning step involved checking for missing values, incorrect entries, and inconsistencies. Only relevant features were retained to reduce noise and improve model performance.

4.2 Feature Selection

Feature selection was performed to focus on variables that have a direct impact on loan approval decisions. This step helped simplify the model and improve interpretability.

4.3 Model Training

Supervised learning techniques were used to train the model. The dataset was split into training and testing sets. The model learned patterns between input features and loan status.

4.4 Model Evaluation

Model performance was evaluated using standard classification metrics such as Mae, MSE R2, . These metrics helped assess how well the model predicts loan approval outcomes.

5. Results and Discussion

The trained model showed reasonable performance in predicting loan status. The results highlighted that income, credit score, and employment experience play a major role in loan approval decisions. The project also demonstrated that clean data and proper feature selection significantly improve prediction reliability.

6. Tools and Technologies

The following tools and libraries were used in this project:

- Python
- Pandas and NumPy for data handling
- Scikit learn for model building
- Matplotlib for basic visual analysis

7. Conclusion

This project provided practical experience in applying machine learning to a real world financial problem. It reinforced the importance of data preparation, feature selection, and evaluation in building reliable prediction models.

8. Future Improvements

Future work may include:

- Using larger and more diverse datasets
- Testing additional machine learning algorithms
- Applying hyperparameter tuning
- Deploying the model as a simple web application

9. Repository Link

Project source code is available at:

<https://github.com/Altaf-code-ai/mL-models>

10.Appendix

<div>Age</div> <div>0</div>	<div>Loan Status</div> <div></div>
<div>Income</div> <div>0</div>	<div>Flag</div>
<div>Employment Experience (Years)</div> <div>0</div>	
<div>Loan Amount</div> <div>0</div>	
<div>Credit Score</div> <div>0</div>	
<div>Clear</div>	<div>Activate Windows</div> <div>Go to Settings to activate Windows.</div>
<div>Submit</div>	