



PERSONAL LOAN ACCEPTANCE MODEL REPORT

Author
Priyanka Regmi



Introduction

Objective

The objective of this project is to develop a machine learning model to predict the personal loan acceptance based on the customer data. This report summarizes the methodology, key finding, and insights from the analysis.

Scope

The report covers data description, preprocessing steps, model selection and evaluation, key findings from the analysis, and insights.

Approach

Data Collection and Description

The dataset used for this analysis was provided by the Company 'Treeleaf' for this assessment. The data consist of total 5000 entries with total 15 columns. The columns included on the data with their description are:

- ID: ID of the customer
- Age: Age of the customer
- Gender: M for Male, F for Female and O for Others
- Experience: Amount of work experience in years
- Income: Amount of annual income (in thousands)
- Home Ownership: Home Owner, Rent and Home Mortgage.
- Zipcode: Postal code in which the client lives
- Family: Number of family members
- CCAvg: Average monthly spending with the credit card (in thousands)
- Education: Education level (1: bachelor's degree, 2: master's degree, 3: advanced/professional degree)
- Mortgage: Value of home mortgage, if any (in thousands)
- Securities Account: Does the customer have a securities account with the bank?
- CD Account: Does the customer have a certificate of deposit account (CD) with the bank?
- Online: Does the customer use the internet banking facilities?
- CreditCard: Does the customer use a credit card issued by the bank?
- Personal Loan: Did this customer accept the personal loan offered in the last campaign?
- (Target Variable)

Data Preprocessing

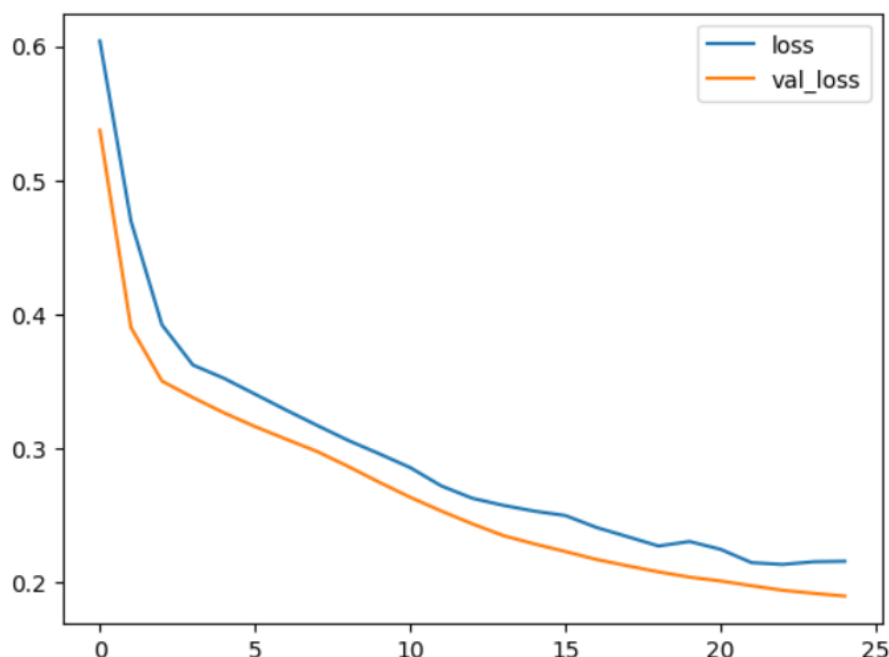
- Handled the missing values with relevant values like 0, according to data requirements
- Made a dummy variable for the columns which have specific categories like Home Ownership and Gender
- Dropped irrelevant columns like ID, Zip Code
- Split the dataset into training (60%) and testing (40%) sets

Model Selection

- Model used for ML model building is TensorFlow's Sequential Model which is well-suited for developing and training deep learning models and best way to build neural networks layer by layer
- Architecture
The Sequential model consist of multiple layers:
 - An input layer with the number of nodes equal to the number of features
 - Three hidden layers with a specified number of neurons and activation functions like ReLU.
 - An output layer with a single neuron for binary classification using a sigmoid activation function

- Training

The model was compiled using the Adam optimizer and binary cross-entropy loss function. It was trained over several epochs with a validation split to monitor performance. The training loss and validation loss were both decreasing which shows model is learning on data and doing quite well.



- Evaluation

The model's performance was evaluated using metrics such as accuracy, precision, recall and confusion matrix.

	precision	recall	f1-score	support
0	0.90	1.00	0.95	1795
1	0.00	0.00	0.00	203
accuracy			0.90	1998
macro avg	0.45	0.50	0.47	1998
weighted avg	0.81	0.90	0.85	1998

```
confusion_matrix(y_test,predictions)
```

```
array([[1795,    0],
       [ 203,    0]], dtype=int64)
```

Key Findings

Exploratory Data Analysis (EDA):

- personal loan is highly correlated with income, CCAvg, and CD amount and negative correlation with Other Genders and Other Home Ownership
- Experience and Age are highly correlated with each other i.e. 0.98
- Income and CCAvg are positively correlated with each other
- Income has high negative correlation with family and education
- Here zip code is not much related to any data so we can remove as it is no contributing on model prediction.
- Higer educated people seems to take personal loan more

Feature Importance:

The most influential features for predicting personal loan acceptance were:

1. Income
2. CCAvg
3. CD account
4. Education

Insights and Observations

Income:

Higher income levels are strong indicator of loan acceptance, suggesting that financial stability is a critical factor.

Average spending through Credit Card (CCAvg):

The positive credit history greatly influences acceptance, highlighting the importance of trustworthiness and repayment ability. And same for the credit account.

Education:

Highly educated people are likely to accept loan, suggesting higher the education may be linked to financial stability and responsible credit behavior.

Summary

The developed personal loan acceptance prediction model is quite good and give mostly accurate result with the accuracy of 90%. The key features to analyzes this prediction can be income, credit history, credit account and level of education.

Recommendations:

Data Enhancement: Incorporate additional customer behavior data to improve model accuracy.

Target High-Income Customers: Focus marketing efforts on individuals with higher income levels, as they are more likely to accepted for loans.

Leverage Credit History: Use the model to emphasize history in the loan approval process. Customers with good credit histories likely to accept loans.

Optimize Loan Amount Offers: Analyze the distribution of loan amounts that are frequently accepted and adjust the offering to align with these amounts.

Focus on Education Outreach: Implement financial literacy program and education outreach for customers, particularly targeting those with lower education levels.

Regular Model Updates: Regularly update and retrain the model with new data to maintain its accuracy and relevance.