

# Major Project

You are provided with a loan sanction dataset where you will have to identify whether the loan of a particular person is approved or not depending on the information the individual has provided.

The dataset consists of the following attributes:

Attribute	Significance
Loan_id	Unique loan Id
Gender	Male/female
Married	Applicant married (Y/N)
Dependents	Number of dependents
Education	Applicant education (Graduate/ Under graduate)
Self-employed	Self employed (Y/N)
ApplicantIncome	Applicant income
CoapplicantIncome	Coapplicant income
LoanAmount	Loan amount in thousands
Loan_Amount_Term	Term of loan in months
Credit_history	Credit history meet guidelines
Property_area	Urban / Semi urban / Rural
Loan_status	Loan approved (Y/N)

Steps for you to follow:

1. Import all the necessary libraries
2. Import the dataset provided
3. Understand the data
4. Deal with the missing values if any
5. Do some visualization if necessary
6. Divide the dataset into training and test datasets
7. Build the machine learning model which ever is suitable for the dataset
8. Fit the model on the training dataset

9. Test the model and find the accuracy of the model on the test and the training datasets
10. Create a confusion matrix

At last, draw conclusions based on the dataset provided and document the same on the jupyter notebook

Most importantly, please comment the usefulness of each cell.