**HOME LOAN PREDICTION**

**Problem Statement:**

Dream Housing Finance company deals in all home loans. They have presence across all urban, semi urban and rural areas. Customer first apply for home loan after that company validates the customer eligibility for loan. Company wants to automate the loan eligibility process (real time) based on customer detail provided while filling online application form. These details are Gender, Marital Status, Education, Number of Dependents, Income, Loan Amount, Credit History and others. To automate this process, they have given a problem to identify the customers segments, those are eligible for loan amount so that they can specifically target these customers.

This is a standard supervised classification task. A classification problem where we have to predict whether a loan would be approved or not. In a classification problem, we have to predict discrete values based on a given set of independent variable(s).Classification can be of two types:

**Supervised:** The labels are included in the training data and the goal is to train a model to learn to predict the labels from the features.

**Binary Classification:** In this classification we have to predict either of the two given classes.

**Given below is the description for each variable.**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Description** |  |
| Loan\_ID | Unique Loan ID | Independent Variables |
| 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 meets guidelines |
| Property\_Area | Urban/ Semi Urban/ Rural |
|  |  |  |
| Loan\_Status | Loan approved (Y/N) | Dependent Variable |

**Results:**

|  |  |
| --- | --- |
| **Algorithm** | **Results** |
| KNeighborsClassifier | 0.84 |
| RandomForestClassifier | 0.78 |
| GradientBoostingClassifier | 0.80 |
| LogisticRegression | 0.66 |