

## **Approach :: Analytics Vidhya - Loan Default Hackathon**

By Praveen : [praveensai266@gmail.com](mailto:praveensai266@gmail.com)

### **Data exploration :**

1. Education column has null values

### **Approach :**

Tried different models like lightgbm, catboost, randomforest, extra trees, ada boost, svc, etc. Finally ensemble of lightgbm model gave me better classification based on CV validation

### **Feature Transformation :**

- One hot encoded

proof\_submitted variable

### **Feature Selection :**

Tried lightgbm (plot\_importance) attribute to select important features.

### **Important Features:**

Top 10 features that turned out to be important are

'loan\_amount', 'no\_of\_loans', 'no\_of\_curr\_loan', 'age', 'asset\_cost', 'education', 'proof\_submitted\_Driving', 'proof\_submitted\_PAN', 'proof\_submitted', 'proof\_submitted\_VoterID'

