Credit Card Approval System

Abhinav Saurabh MT20127
Rahul Meena MT20117

Problem Statement

In the banking system, there is a major task to determine if the person is eligible for the credit card. Profit to the credit card providers largely depend on if person is paying back.

By predicting correct credit card customers ,the banks can maximize the profits.

We here automated this process by applying machine learning techniques. Here we need to make sure to target the right customers.

Importance of Project

- Credit card approval can very helpful for organisations that lend credit cards and due to increase in huge number of applicant there is need to automate the task and classify the applicants into if they are eligible for credit card or not.
- This helps to avoid organisation losses by avoiding potential defaulters.
- Here we are not just looking into bank balance but into their personal attributes like gender,married,age,Occupation etc.We account for these personal attributes to evaluate if given is applicant is good customer.
- This gives benefit by cutting down costs on credit analysis and faster credit decisions.

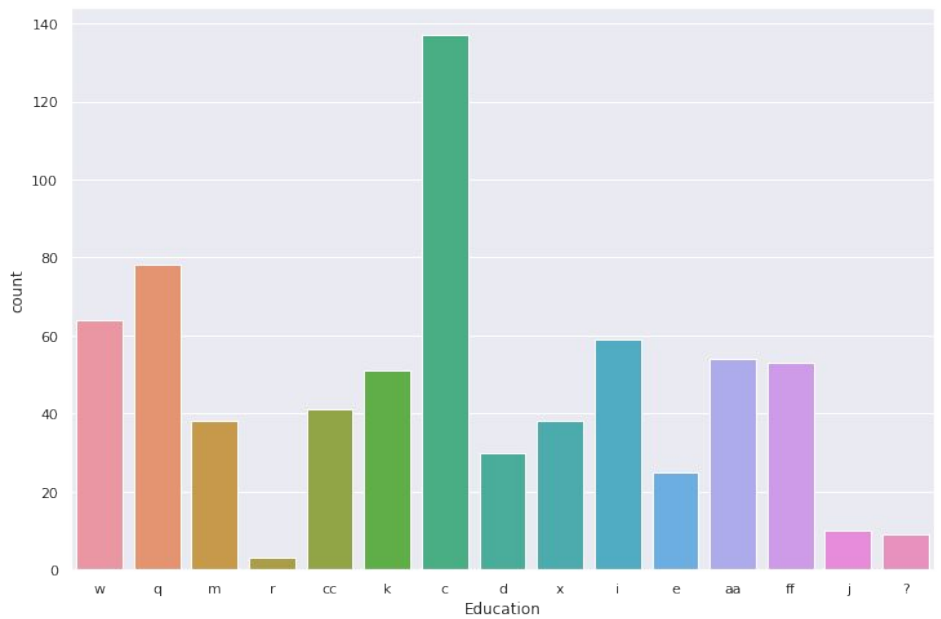
Data Analysis and Preprocessing

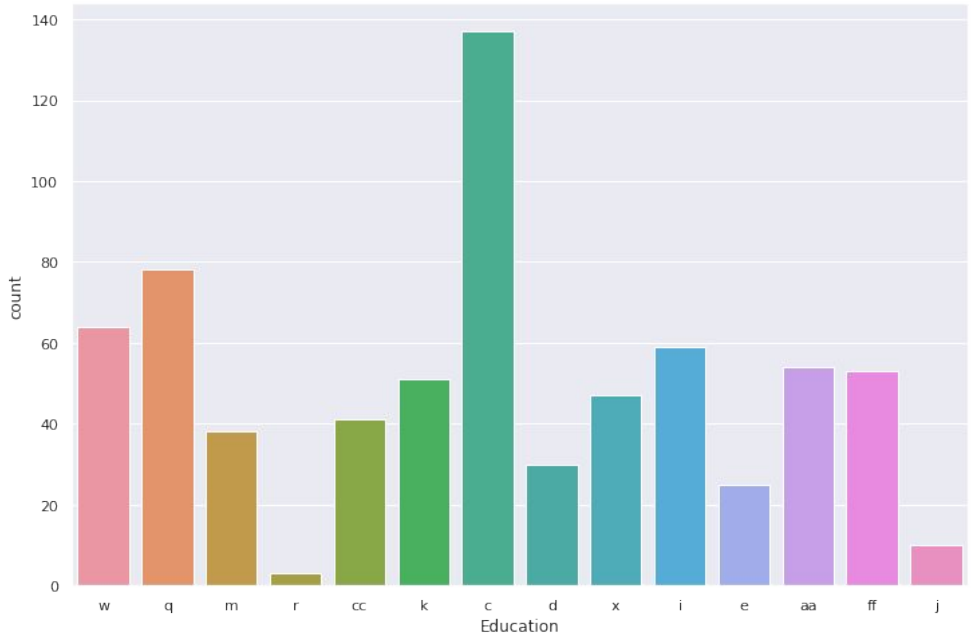
Data Analysis and Preprocessing

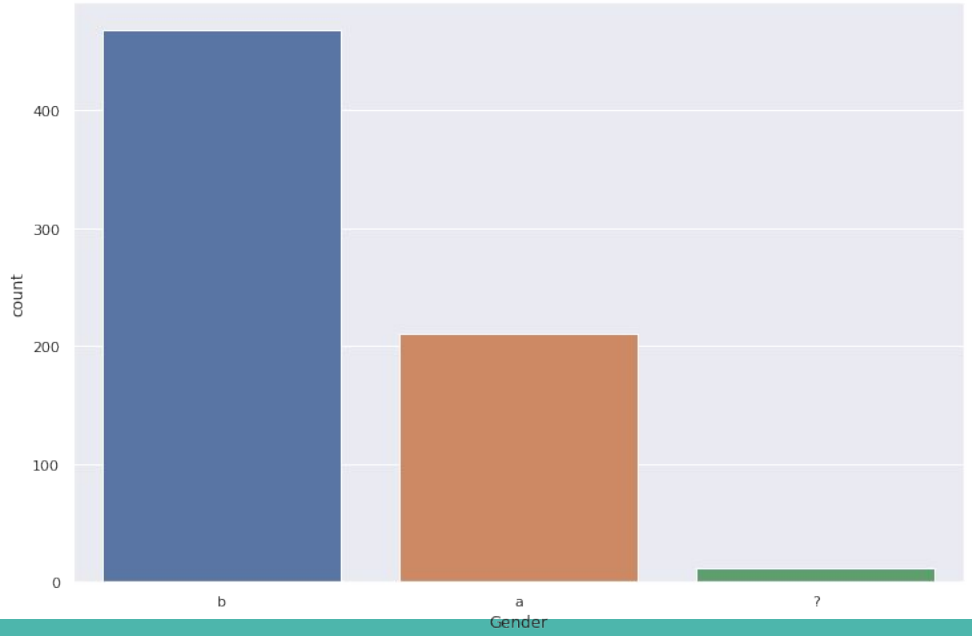
Anamolies and Outliers

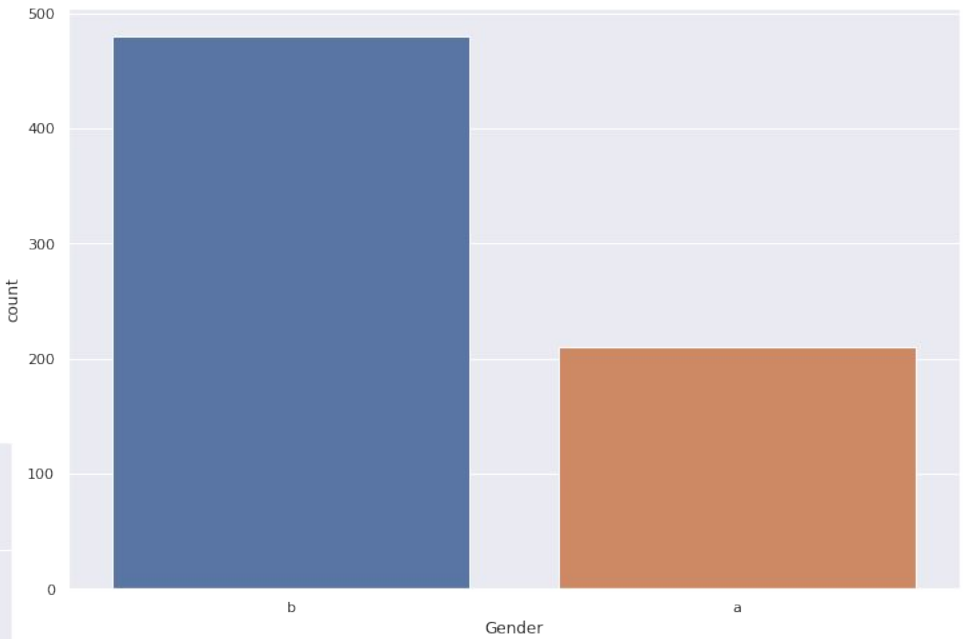
Missing Data

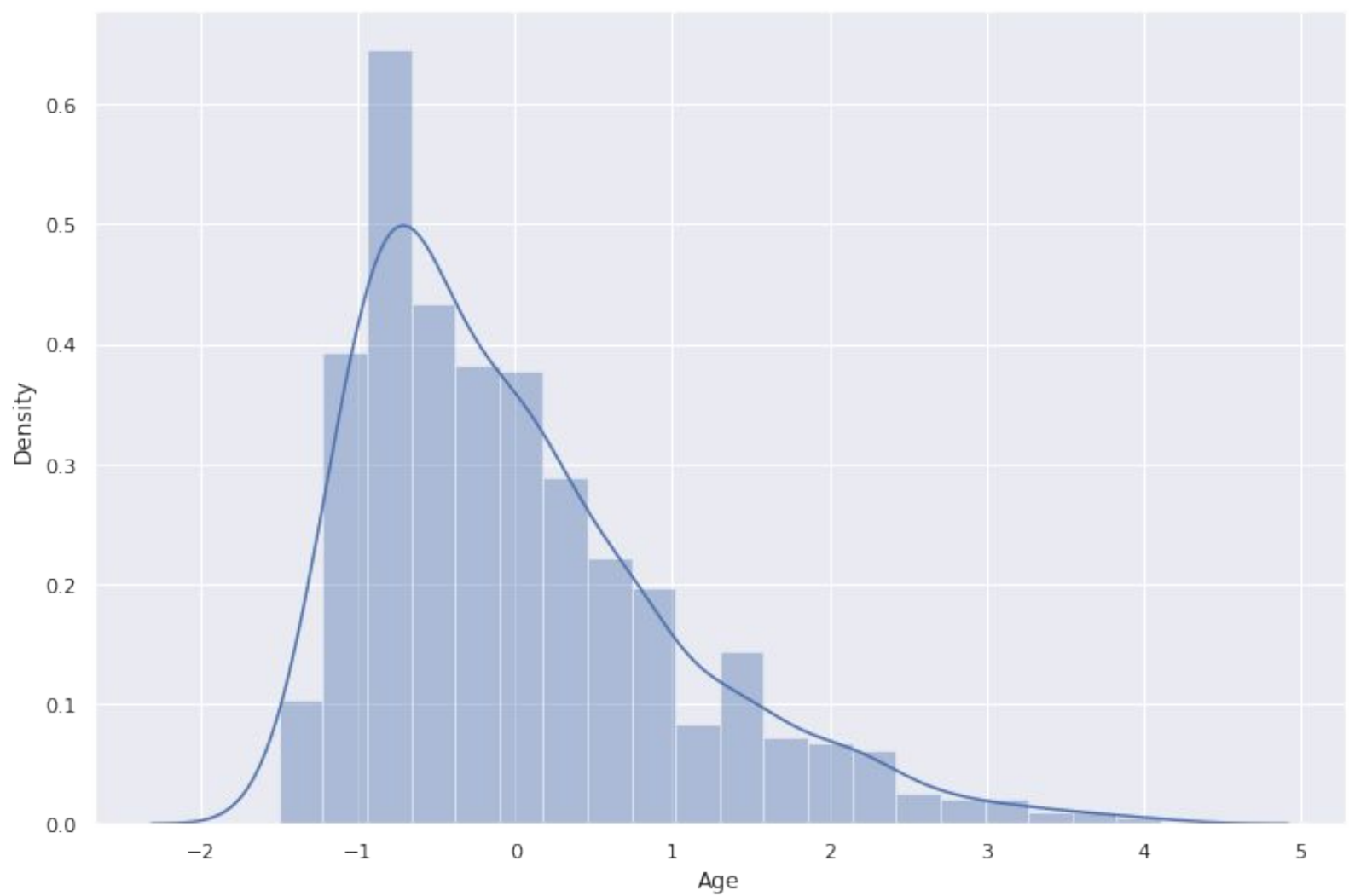
Normalization


count
0
20
40
60
80
100
120
140
w
q
m
r
cc
k
c
d
x
i
e
aa
ff
j
?
Education


count
0
20
40
60
80
100
120
140
w
q
m
r
cc
k
c
d
x
i
e
aa
ff
j
Education


count
400
300
200
100
0
b
a
?
Gender

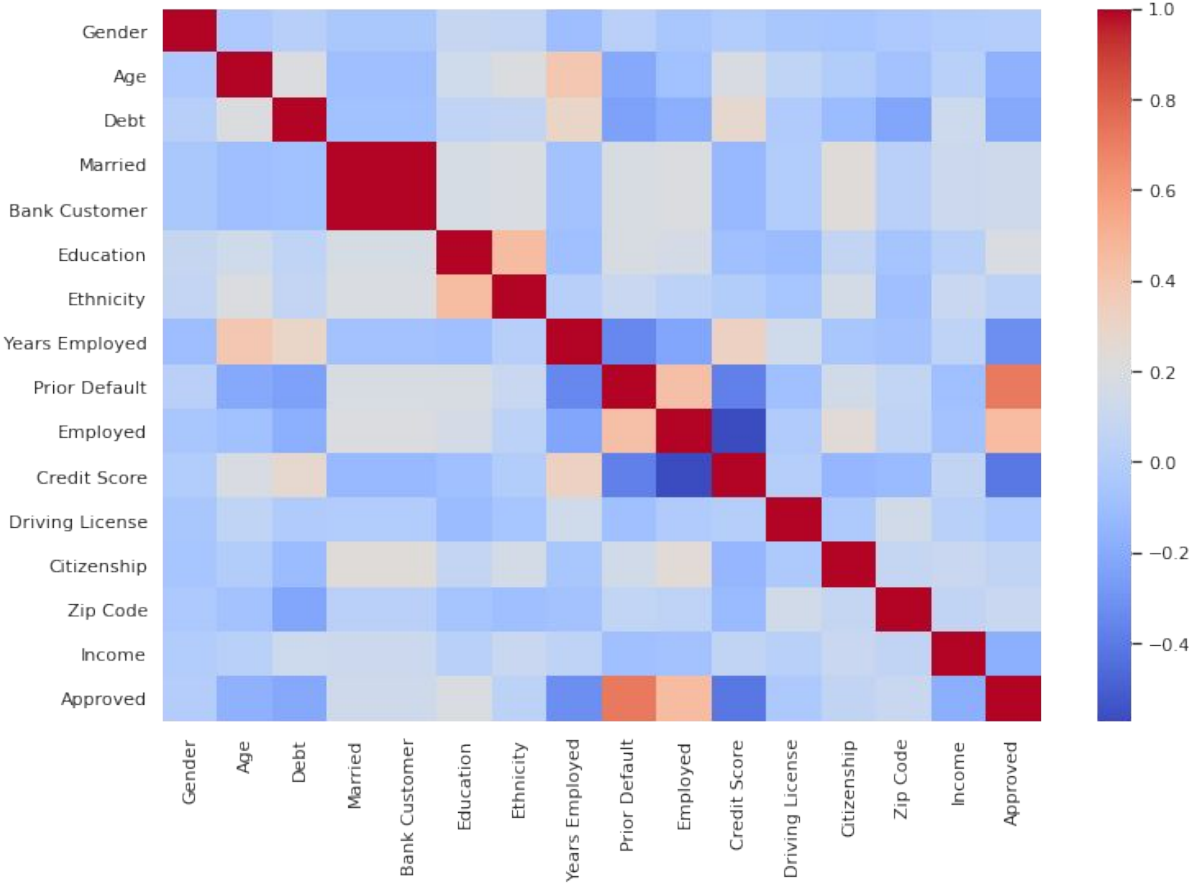

count
500
400
300
200
100
0
b
a
Gender


Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
−2
−1
0
1
2
3
4
5
Age

Correlation Matrix

The correlation matrix tells us that some features like Married, Gender, Citizenship does not play any important role in the credit card approval.

We can simply drop these attributes

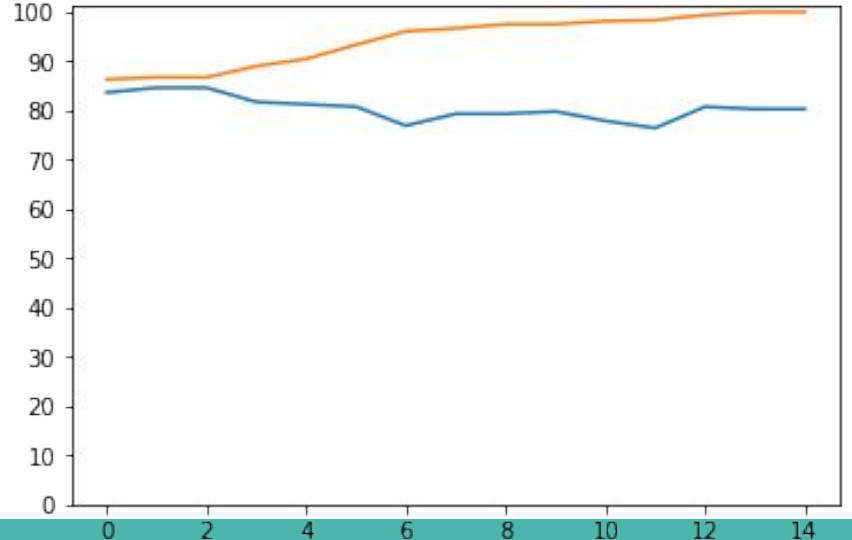


Model Selection

Because the output is classification type so we can't use Linear Regression here, but can use Decision Tree Classifier, Logistic Regression

Using **Decision Tree Classifier** we got the following accuracies

Depth	Training	Testing
1	0.8631	0.8029
2	0.8672	0.8365
3	0.8672	0.8462
4	0.89	0.8462
5	0.9046	0.8173
6	0.9336	0.8125
7	0.9606	0.8077
8	0.9668	0.7692
9	0.9751	0.7933
10	0.9751	0.7933
11	0.9813	0.7981
12	0.9834	0.7788
13	0.9938	0.7644
14	1	0.8077
15	1	0.8029



Using Logistic Regression

We used Logistic Regression and tuned its parameters to get optimal, and obtained following Tolerance : 0.0001 & Max Iterations : 150

With the above optimal parameters we got

Train Accuracy : 84.1393%

Test Accuracy : 86.7052

Gradient Boosting Classifier

We have further implemented Gradient Boosting Classifier which is a type of ensemble algorithm.

The training accuracy obtained in this algorithm was 95.5%.While the testing accuracy obtained was 87.28%.

This is the best result we have obtained from a model so far. Therefore we conclude this model to be our best model.

Summary

1) Data Source
2) Data Analysis
3) Data Cleaning
4) Feature Selection
5) Model Selection

Conclusion

Humans who approve request of credit card sometimes they issue for those persons which might not pay the bill on time which is called bankruptcy and is perfectly legal, then bank clear their debts.

This cause huge loss to bank, as banks are backbones of our country so it is loss for our economy so for us also.

With our system it is possible to pre-determine the eligible person whom to issue credit card whom to not which saves our country’s economy and human time.