German Credit Data

Evaluating Lending Risk Using Multilayer Perceptron and Logistic Regression

Tanay Jayant Kayastha (173059008)

The Problem

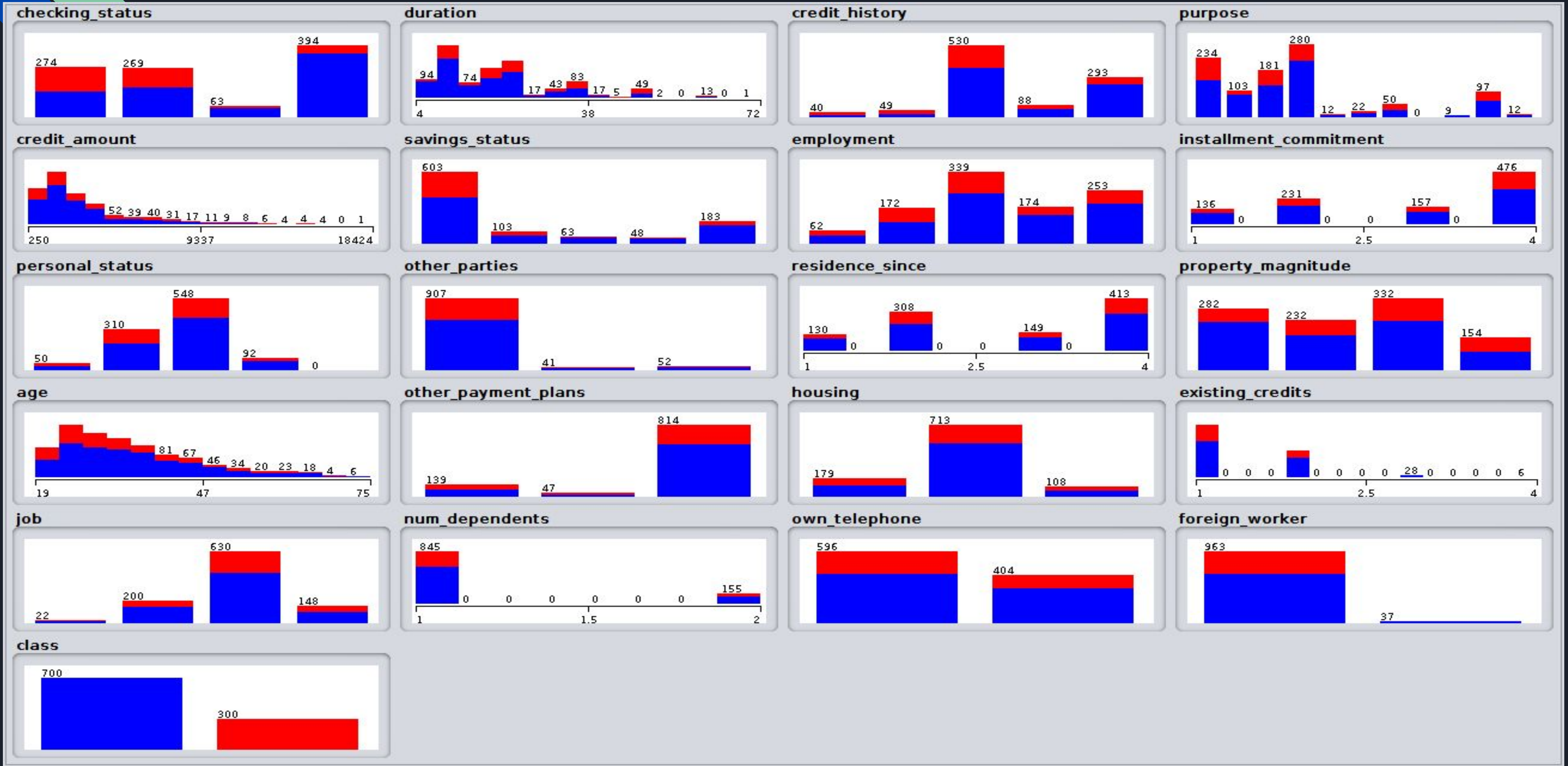
1. Is the applicant a GOOD lending risk?
    - Failure to detect this case implies a possibility of business loss and losing a worthy client.
2. Is the applicant a BAD lending risk?
    - Failure to detect this case is EVEN WORSE, possibility of loan defaulters!

Solution must not just be more ACCURATE in detecting the risk but also the number of FALSE NEGATIVES (ok to lose “some” clients) and FALSE POSITIVES (“highly risky” to encourage loan defaulters!).

The Dataset

- Number of instances: 1000
- Output Classes: yes, no
- Number of instance features: 20 specifying clients social, economic and demographic characteristics.
- Feature domains:
  - Nominal - e.g. purpose (car home, education, vacation, etc)
  - Numeric - e.g. age
- No missing values
- Challenges: nominal features, class-imbalance (70:30), un-normalized features values, feature importance.
- Let’s look at data distribution!

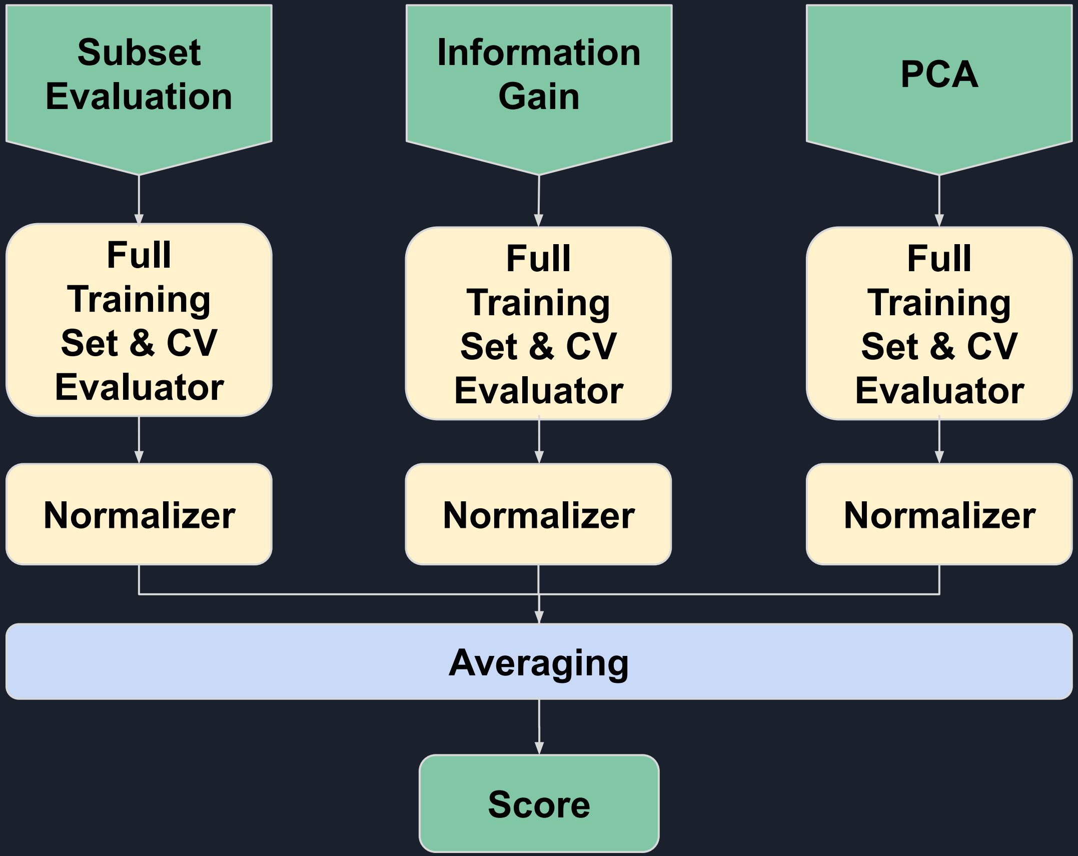
How does the data look like?

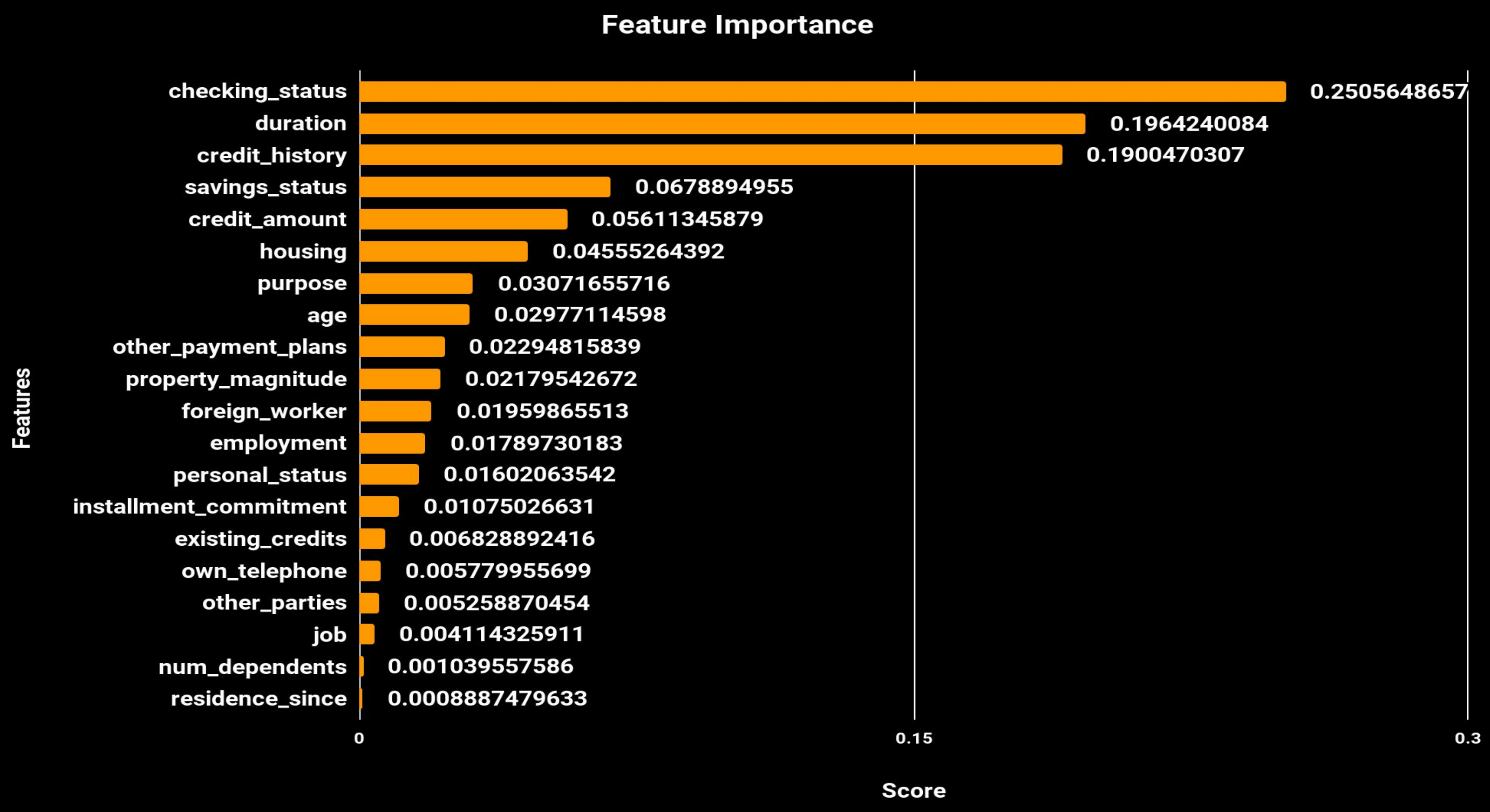


Preprocessing

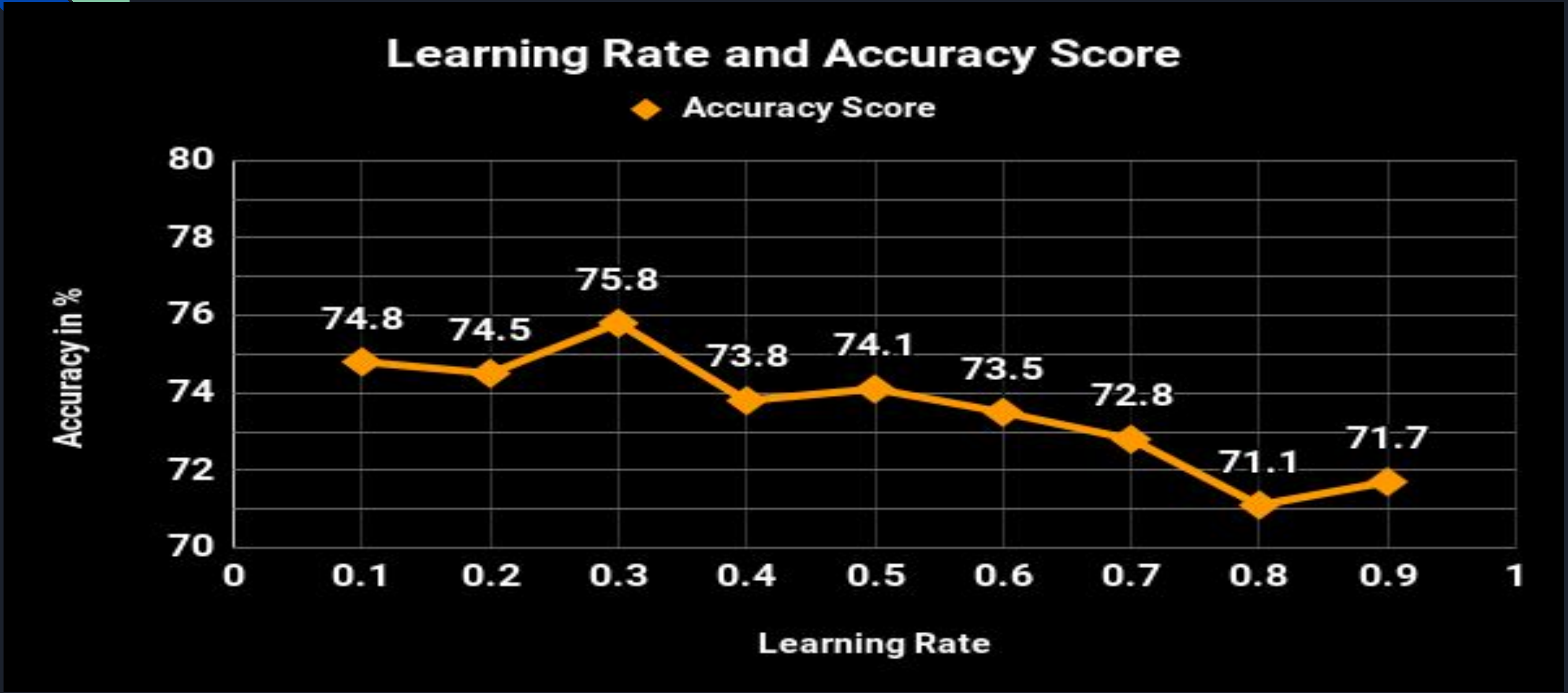
- Feature Importance: Among 20 features, 18 features with highest prediction importance were selected using normalized averaging of:
  - CFSSubsetEval Individual predictive ability of each feature along with the degree of redundancy between them.
  - InfoGainAttributeEval: Worth of an attribute by measuring the information gain with respect to the class.
  - CorrelationAttributeEval: Worth of an attribute by measuring the correlation between it and the class.
  - PCA: correlated feature set -> un-correlated feature set
- All these scores were considered with and without 10-fold cross validation.
- Nominal features were encoded using one-hot encoding.
- Instance classes were balanced using SMOTE.
- This resulted in dataset with 1300 instances with 55 attributes.

Feature Importance Score

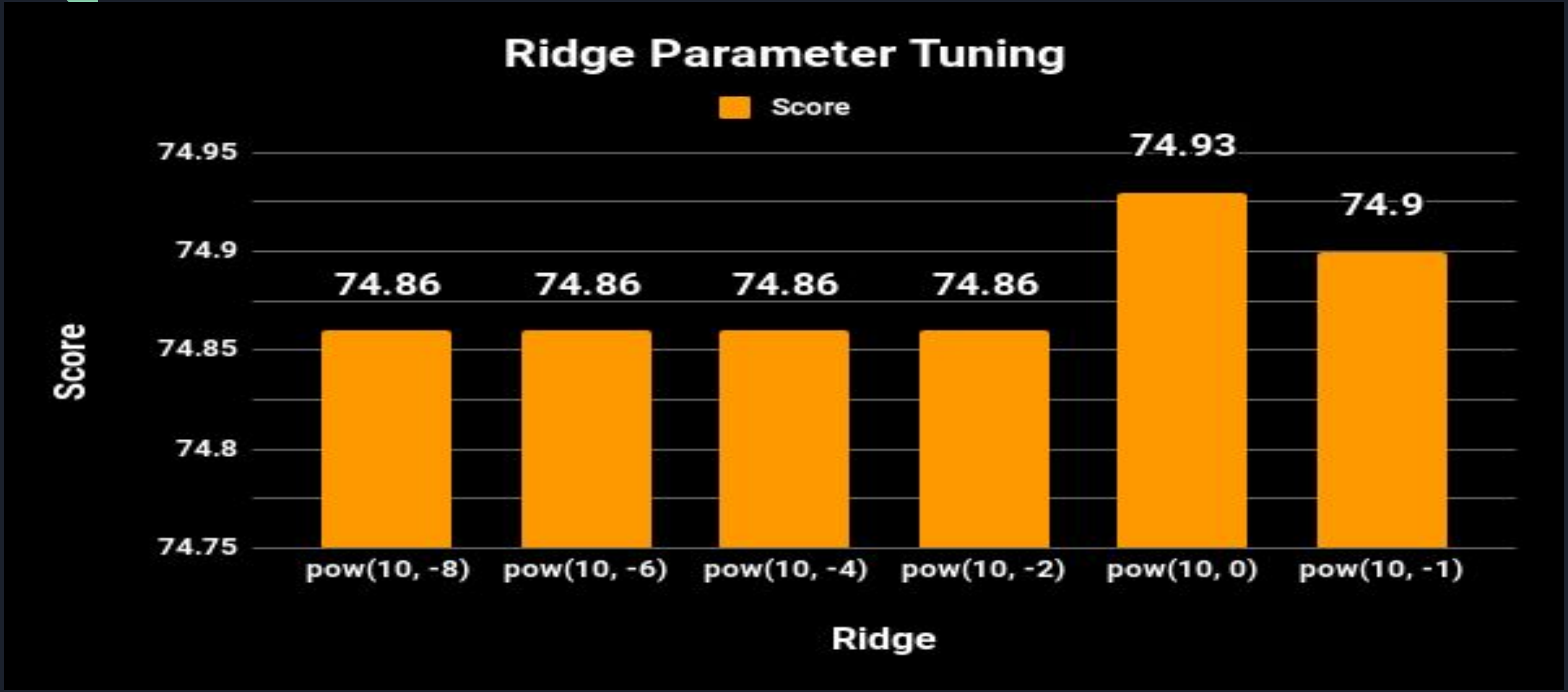



Feature Importance
checking_status
duration
credit_history
savings_status
credit_amount
housing
purpose
age
other_payment_plans
property_magnitude
foreign_worker
employment
personal_status
installment_commitment
existing_credits
own_telephone
other_parties
job
num_dependents
residence_since
0.2505648657
0.1964240084
0.1900470307
0.0678894955
0.05611345879
0.04555264392
0.03071655716
0.02977114598
0.02294815839
0.02179542672
0.01959865513
0.01789730183
0.01602063542
0.01075026631
0.006828892416
0.005779955699
0.005258870454
0.004114325911
0.001039557586
0.0008887479633
Features
0
0.15
0.3
Score

MLP Parameter Tuning



Logistic Regression Parameter Tuning



Results Comparison

Transformed Dataset

Logistic(77.2308 %)

P	R	
0.786	0.800	good
0.755	0.739	bad

MLP(86.6 %)

P	R	
0.885	0.864	good
0.846	0.868	bad

Original Dataset

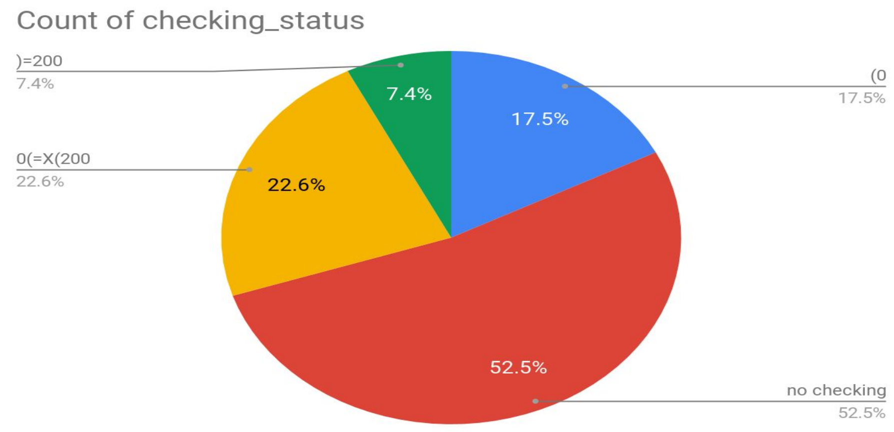
Logistic(75.1 %)

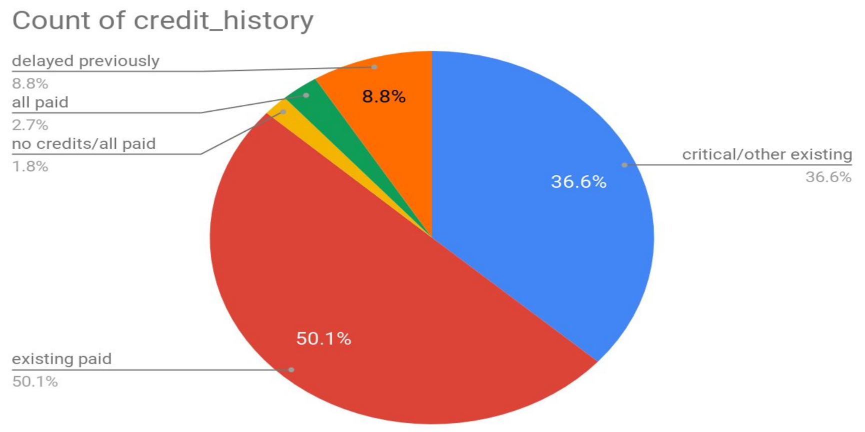
P	R	
0.796	0.866	good
0.607	0.483	bad

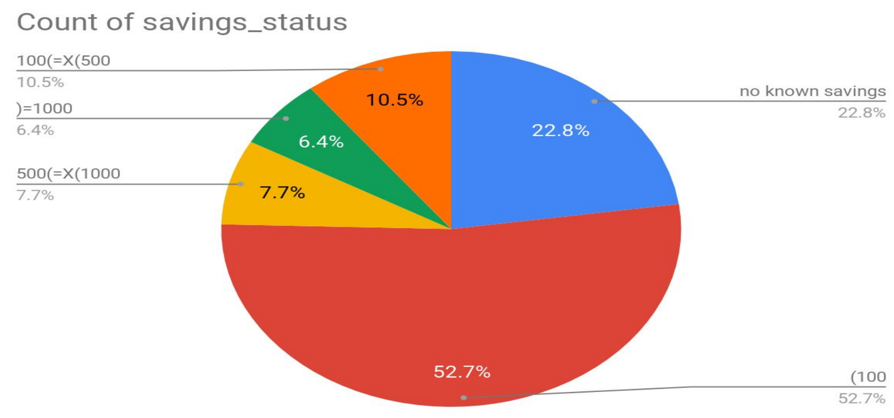
MLP(75.3 %)

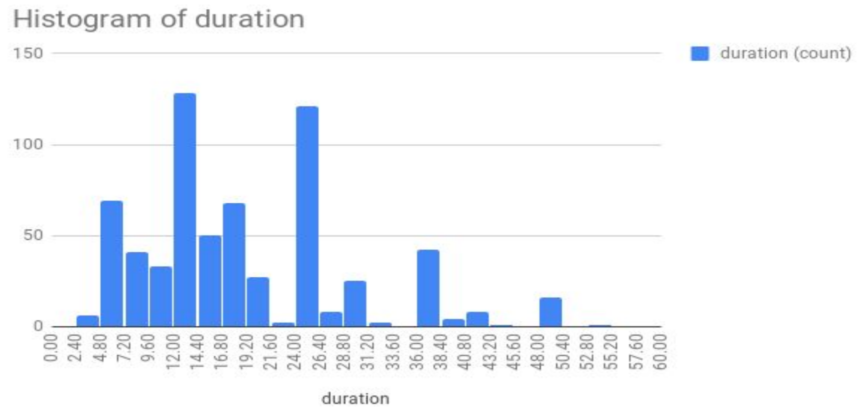
P	R	
0.812	0.843	good
0.597	0.543	bad

Failure Analysis : Good

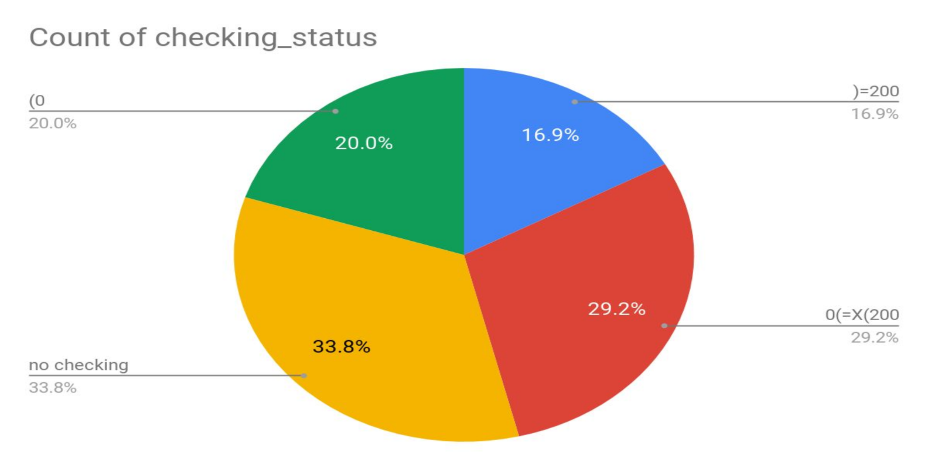


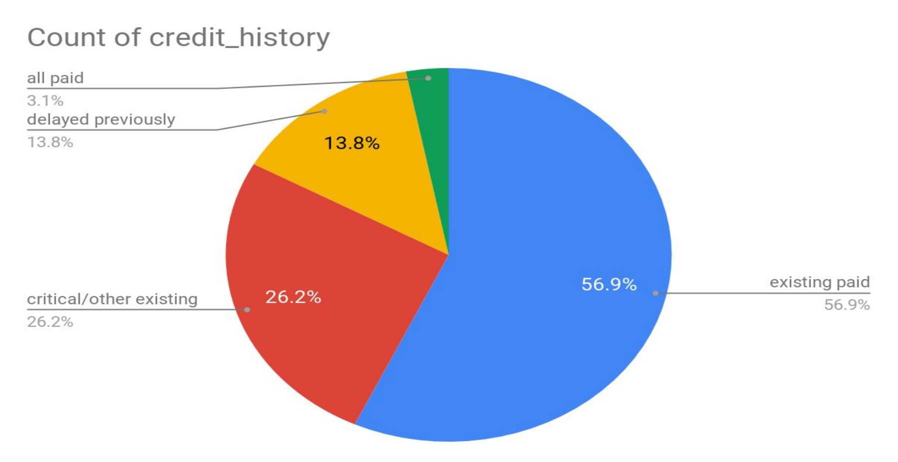


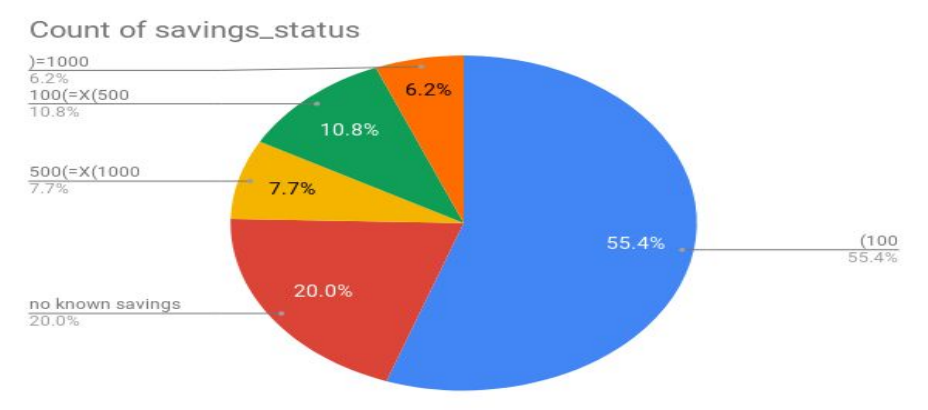


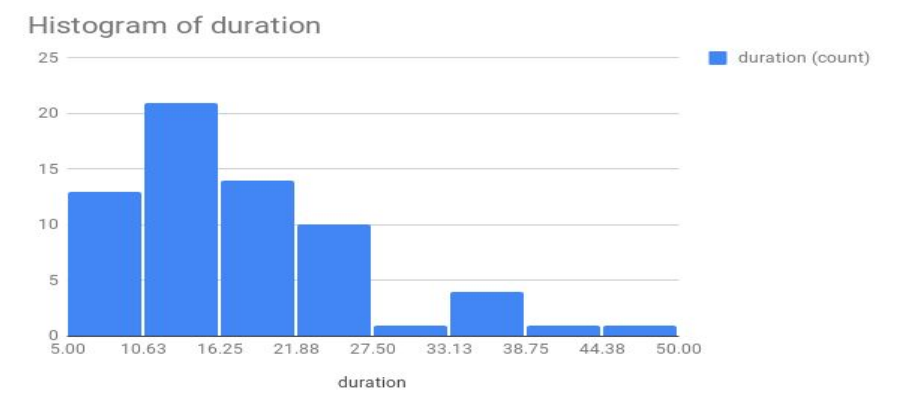


Failure Analysis : Bad->Good

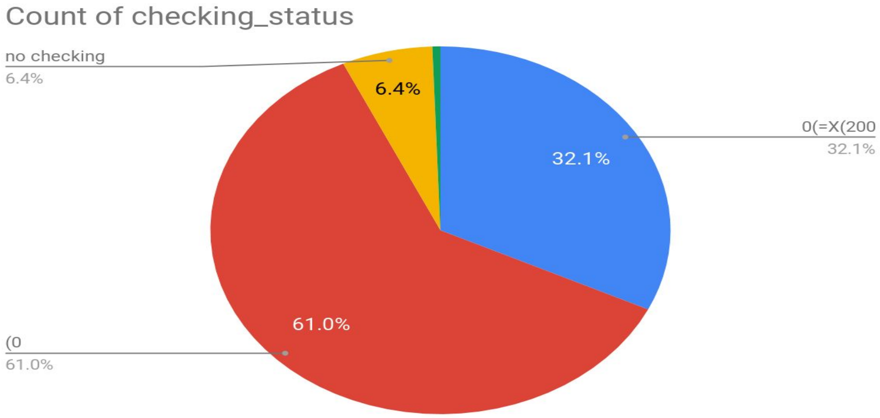


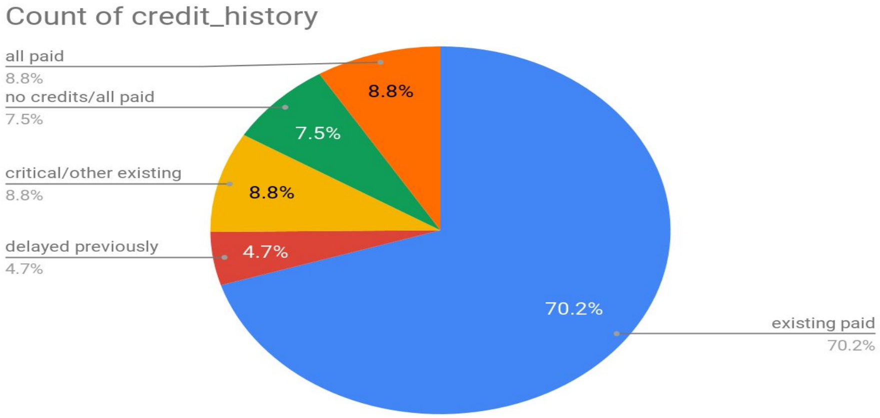


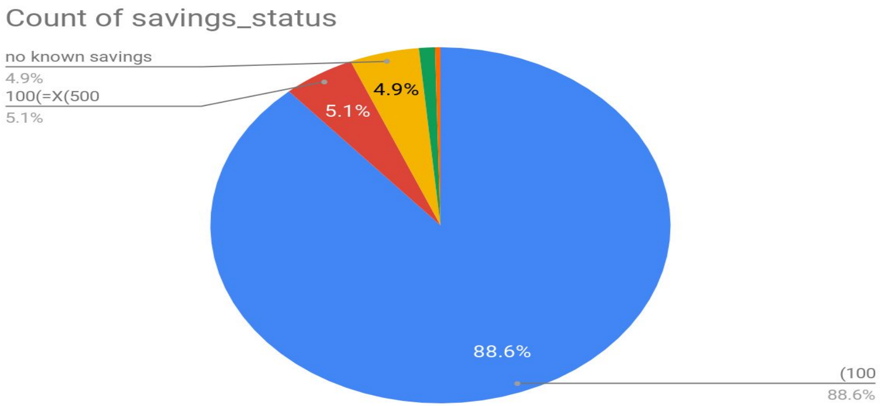


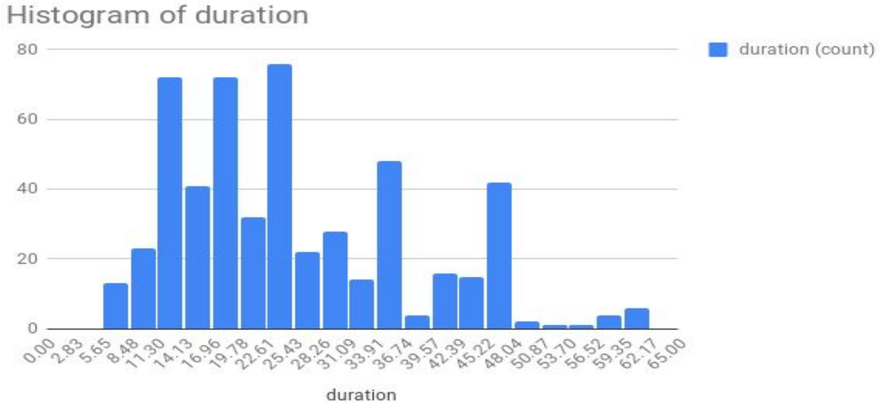


Failure Analysis : Bad

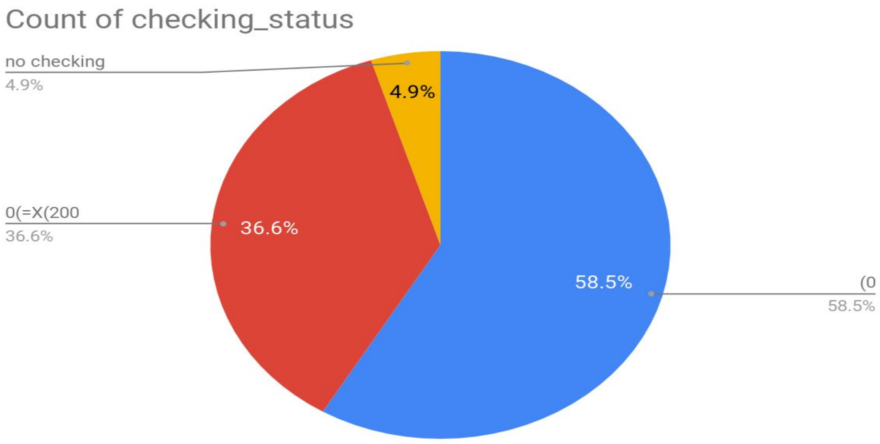


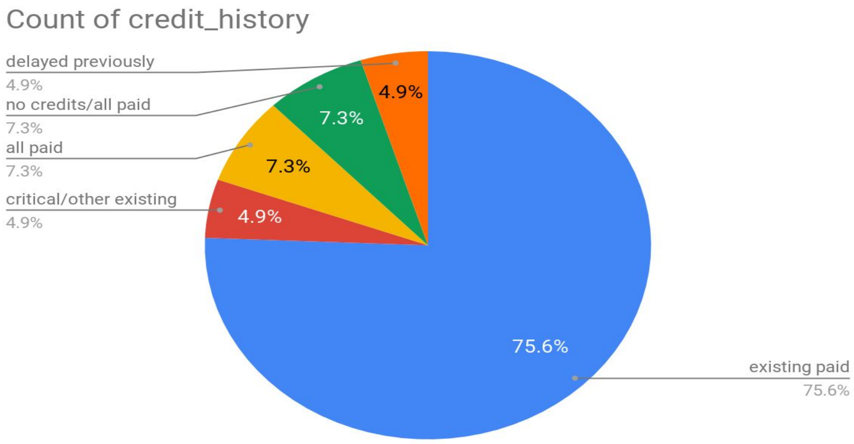


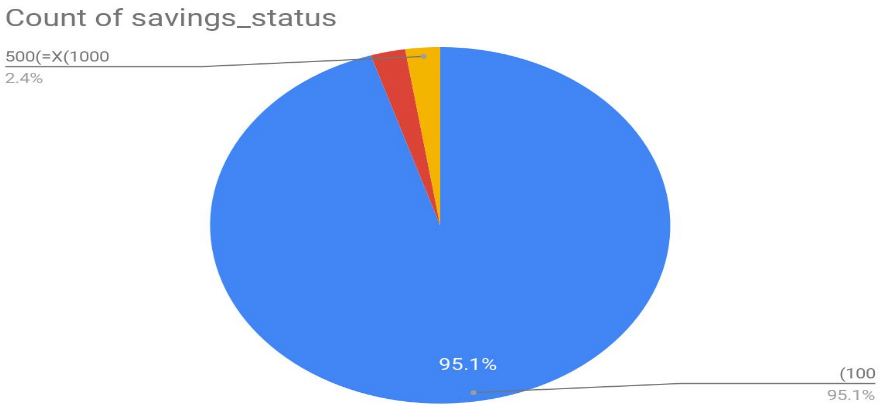


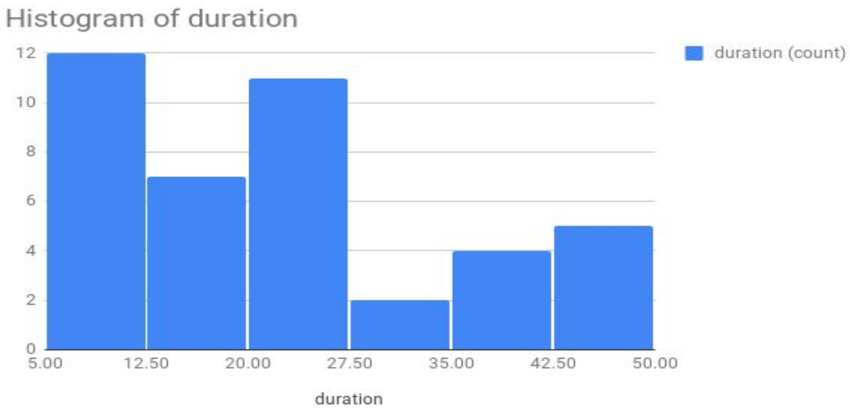


Failure Analysis : Good->Bad









Failure Observations

- Majority contribution by (70 %): checking_status, credit_history, saving_status and duration.
- Feature value distribution against checking_status for “Good” and misclassifications of the form “Bad->Good” matches well. Similarly for “Bad” and misclassifications of the form “Good->Bad”.
- This could have lead classifier to treat Bad as Good in case 1 and vice versa in case 2.
- Important observation is that feature value distributions against saving_status and credit_history resembles across “Good”, “Bad”, “Bad->Good” and “Good->Bad”.
- Possible solution is to raise these two features to higher powers to enable higher order dependency.
- These modifications are not resulting in any better performance. This implues there could be a possibility of some errors.