Credit Score Prediction

Supervised Machine Learning

Hello & welcome to presentation

Credit Score Prediction

Welcome to the Credit Score
Prediction ML classification . This is designed
For Bank and credit card companies calculate your
Credit score to determine your credit worthiness.
It helps bank and credit card companies immediately
To issue loans to customers with good credit
worthiness

LETS GET STARTED

Project Timeline

Step 1

Introduction, Understanding Data

Step 2

Data Wrangling & Data Visualization

Step 3

Data separating & Data Preprocessing

Step 4

Feature Manipulation & Model Implementation

Step 5

Future Work & Conclusion

1

Introduction

Information about Credit Score Classification Prediction

Project Architecture

The Credit Score Classification Prediction ML model is based on supervised machine learning techniques. It is trained using a large dataset of past credit data. The model is then tested on a separate dataset to evaluate its performance. The performance of the model is measured using metrics such as precision, recall, and accuracy.

One of the key advantages of the Credit Score Classification Prediction ML model is that it is highly accurate to predict credit score which is Good, Standard, poor.

The model can accurately detect up to **81%** of credit score of customers Classifying customers based on their credit scores helps banks and credit card companies immediately to issue loans to customers with good creditworthiness.

Another advantage of the model is that it is scalable and can be easily integrated into existing mobile security solutions. However, there are also some challenges that need to be addressed when using the Credit Score Prediction ML model. One of the challenges is the need for a large dataset of Credit Score data for training and testing the model. This requires a significant amount of resources and expertise to collect and analyze the data.

Credit Score Prediction Classification

Data

Machine Learning Model

Prediction values

2

Problem Statement

Problem Statement

Banks and Financial institutions play a crucial role in assessing the creditworthiness of individuals before extending loans or credit lines. To streamline and enhance this process, there is a growing need for a robust credit score classification model. The company has collected basic bank details and gathered lot of credit-related information. The management wants to build an intelligent system to segregate the people into credit score brackets to reduce the manual efforts.
The model should address the following key challenges:

- Data variety and Quality
- Feature selection and Engineering
- Scalability and Efficiency
- Model Interpretability
- Model Evaluation and Performance Metrics

3

Understanding Data

What are features and label in our data

- Customer ID: Represents a unique identification of a person
- Occupation: Represents the occupation of the person
- Annual Income: Represents the annual income of the person
- Monthly Inhand salary: Represents the monthly base salary of a person
- Number of Bank Accounts: Represents the number of bank accounts a person holds
- Number of Credit card: Represents the number of other credit cards held by a person
- Interest Rate: Represents the interest rate on credit card
- Number of Loan: Represents the number of loans taken from the bank
- Type of Loan: Represents the types of loan taken by a person
- Delay from due date: Represents the average number of days delayed from the payment date
- Number of Delayed payments: Represents the average number of payments delayed by a person
- Change Credit Limit: Represents the percentage change in credit card limit
- Number Credit inquiries: Represents the number of credit card inquiries
- Credit mix: Represents the classification of the mix of credits
- Outstanding debt: Represents the remaining debt to be paid (in USD)
- Credit utilization ratio: Represents the utilization ratio of credit card
- Payment of minimum amount: Represents whether only the minimum amount was paid by the person
- Total EMI per month: Represents the monthly EMI payments (in USD)
- Amount invested monthly: Represents the monthly amount invested by the customer (in USD)
- Payment behaviour: Represents the payment behavior of the customer (in USD)
- Monthly balance: Represents the monthly balance amount of the customer (in USD)

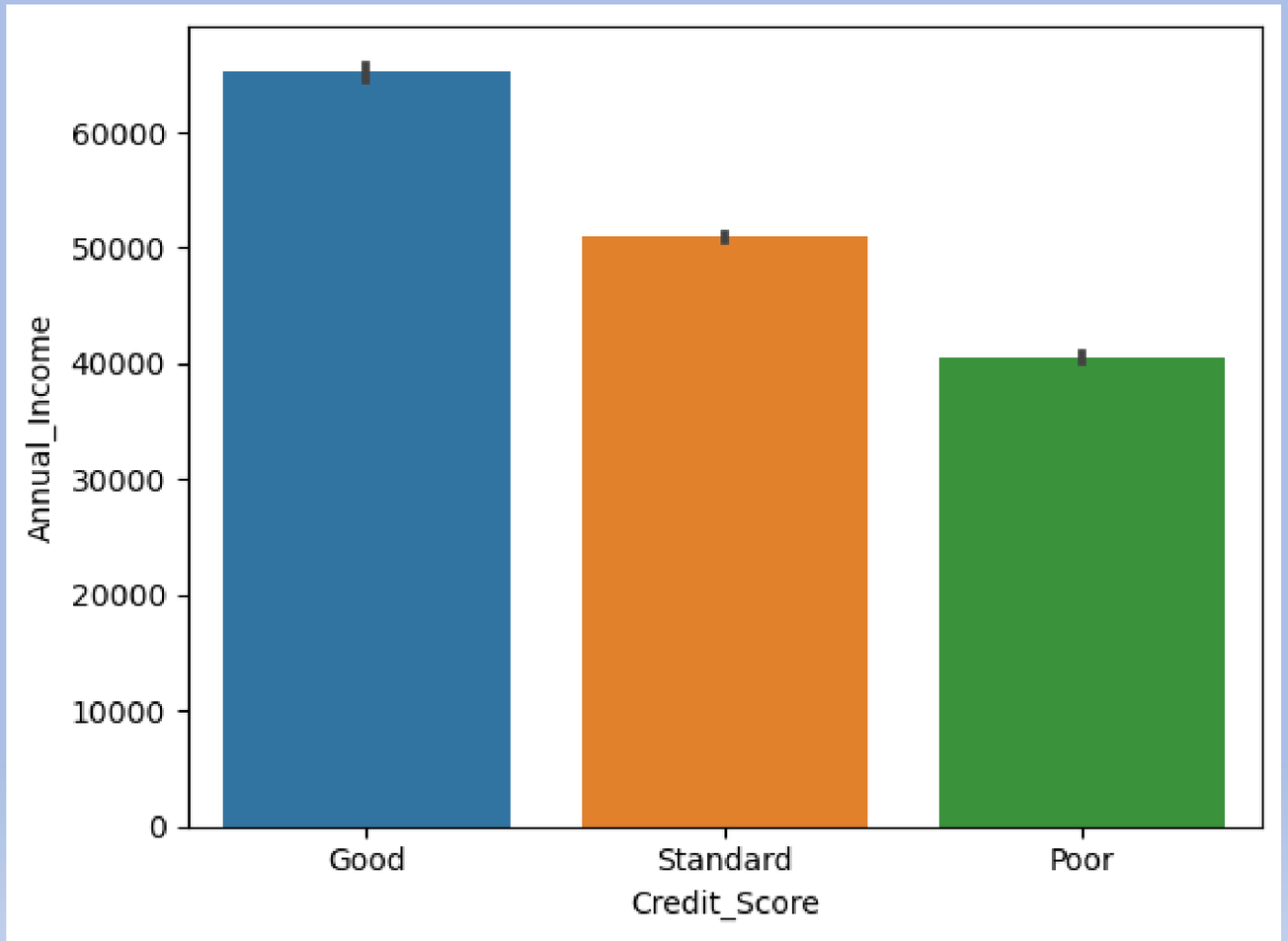
4

Data Wrangling

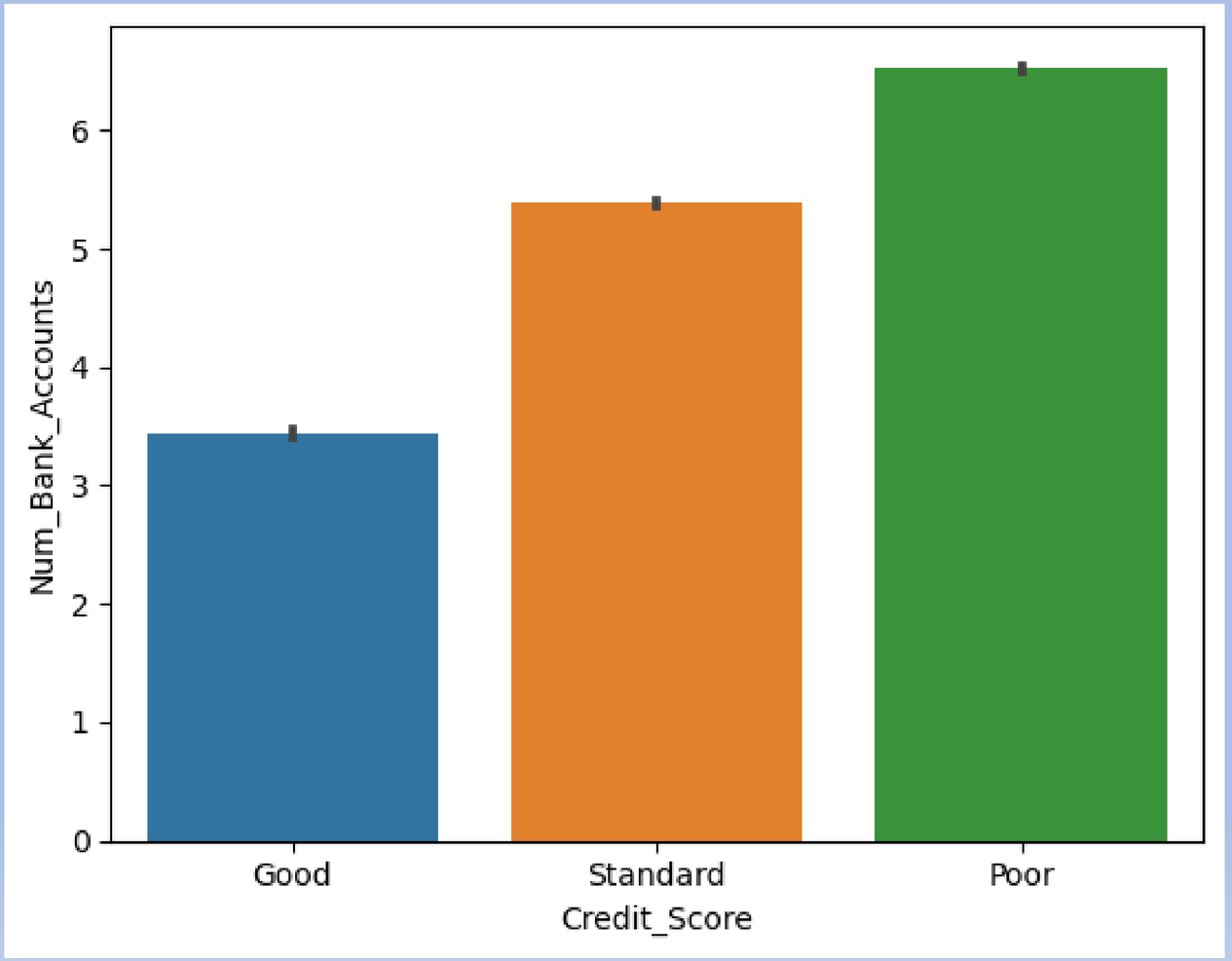
Finding meaningful insights

Count Credit Score Classification Values

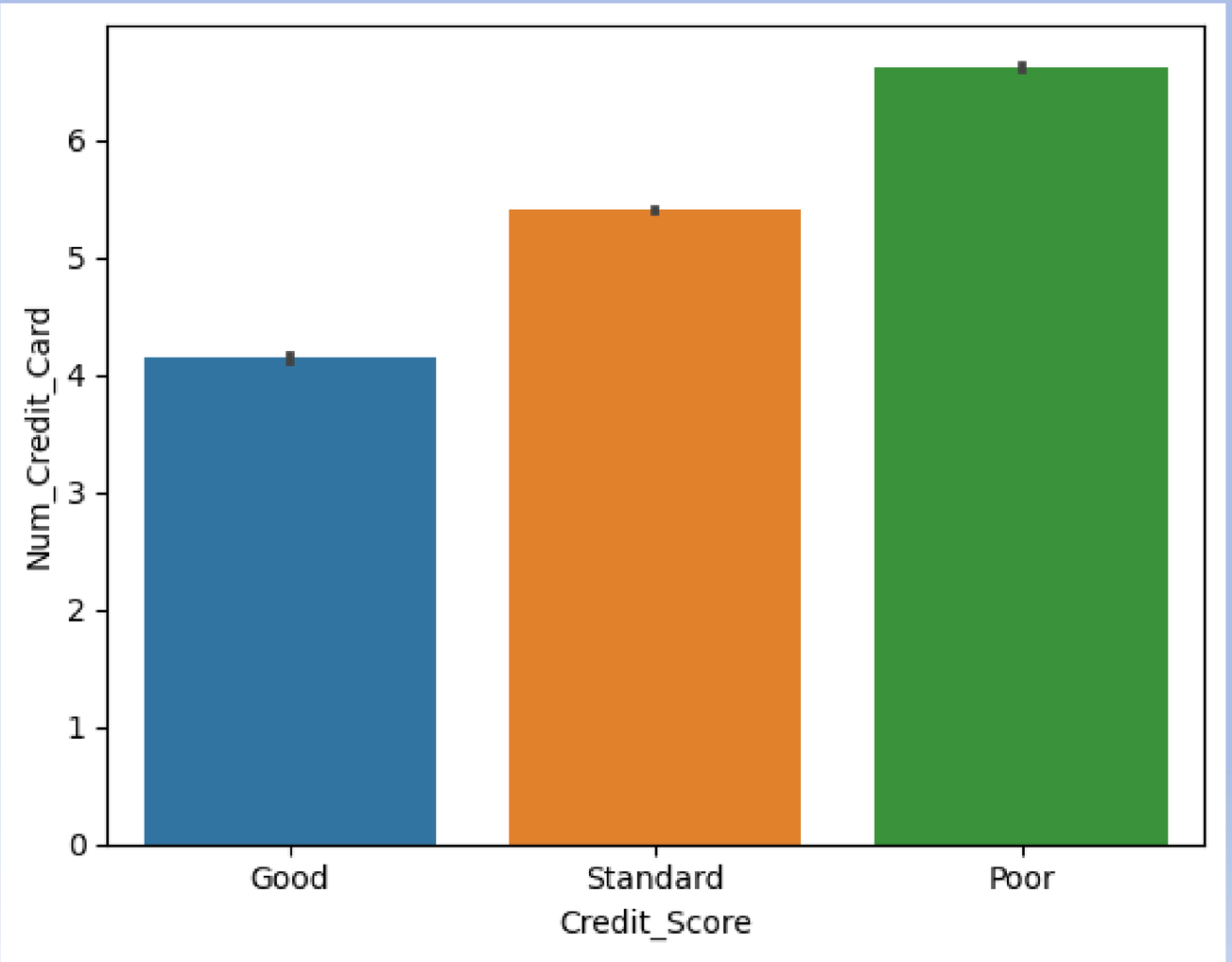
Plotting check Annual Income of the person impacts your credit scores or not



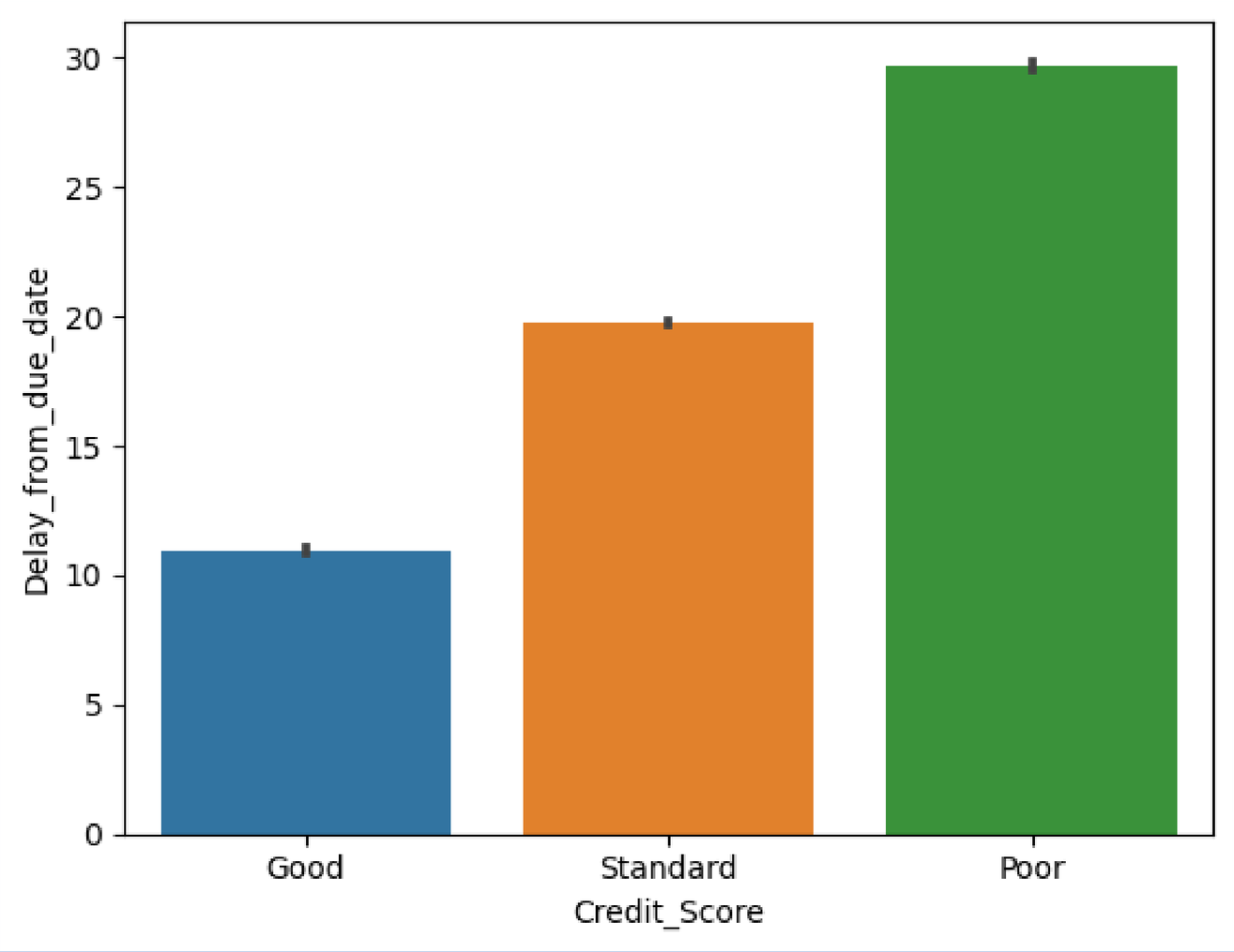
Check more bank accounts impacts credit scores or not



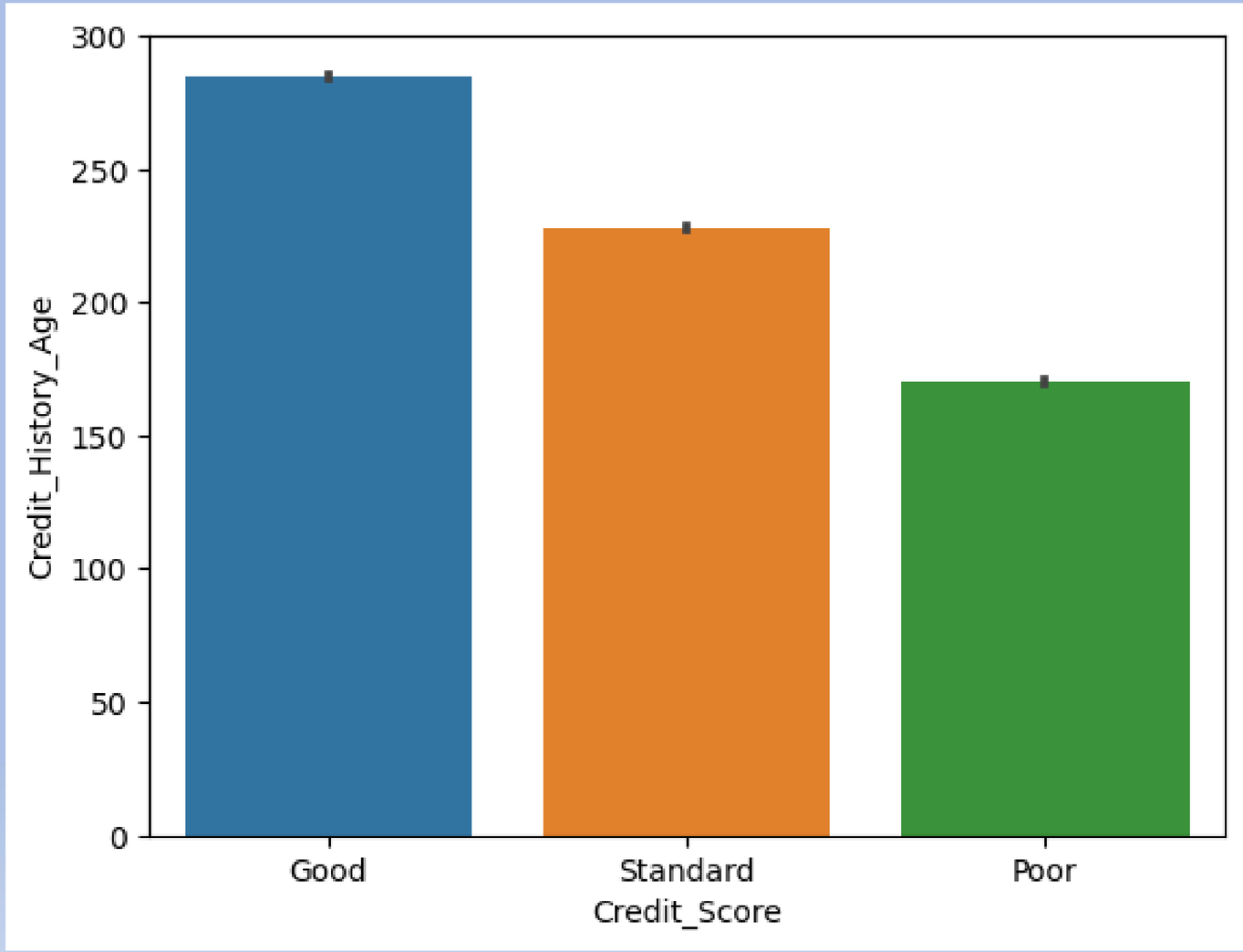
Check credit scores based on the number of credit cards of customer



Check delaying payments on the due date impacts your credit scores or not



Check how the credit history age of a person affects credit scores



Conclusion of above charts

- 1st chart shows Having high annual income may impacts on credit score that means high annual income means good credit score and low annual income indicates poor credit score.
- 2nd chart shows the increasing no. of bank accounts affects the credit score which is poor.
- 3rd chart shows the increasing no. of credit cards affects the credit score which is poor.
- 4th chart shows Delaying payments from due date may more impact on credit score.
- 5th chart shows credit history age of a person affects credit scores

Feature Engineering

Feature Manipulation , & Feature selection

Categorical Column Encoding

In our data set 28 columns in which 2 columns are object type column

Here we are using Label encoding is a technique used to represent categorical variables as numerical values in a machine learning model. As the category column in our dataset is an important feature for evaluation, we have to convert it into numeric for model implementation so we have used this technique. The advantages of using Label encoding includes:

It allows the use of categorical variables in models that require numerical input.

It can improve model performance by providing more information to the model about the categorical variable.

It can help to avoid the problem of ordinality, which can occur when a categorical variable has a natural ordering (e.g. "small","medium", "large").

Model Implementation

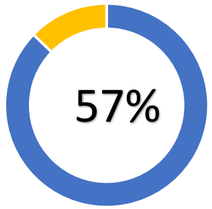
Implement various ML models

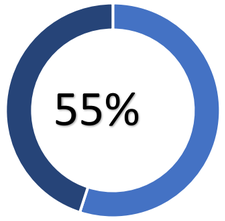
Model 1: Logistic Regression

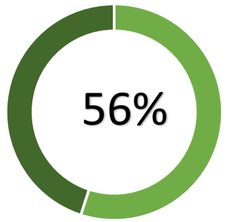
Good

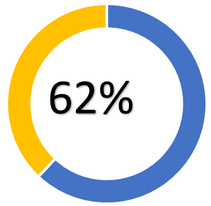
poor

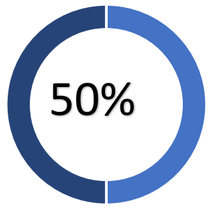
STANDARD

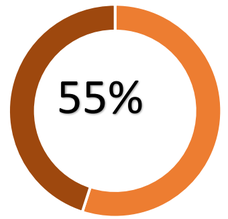


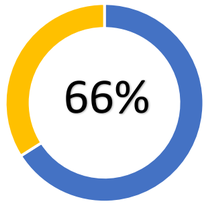


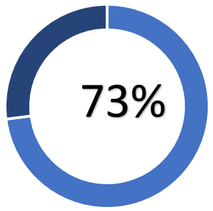


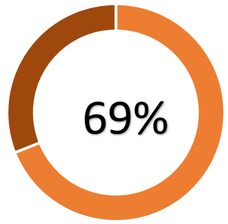


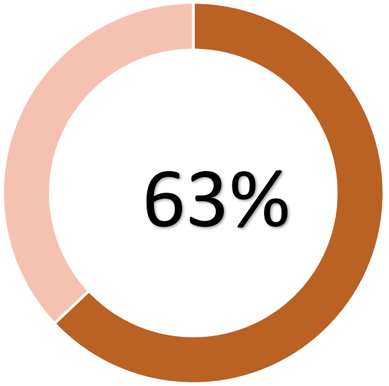






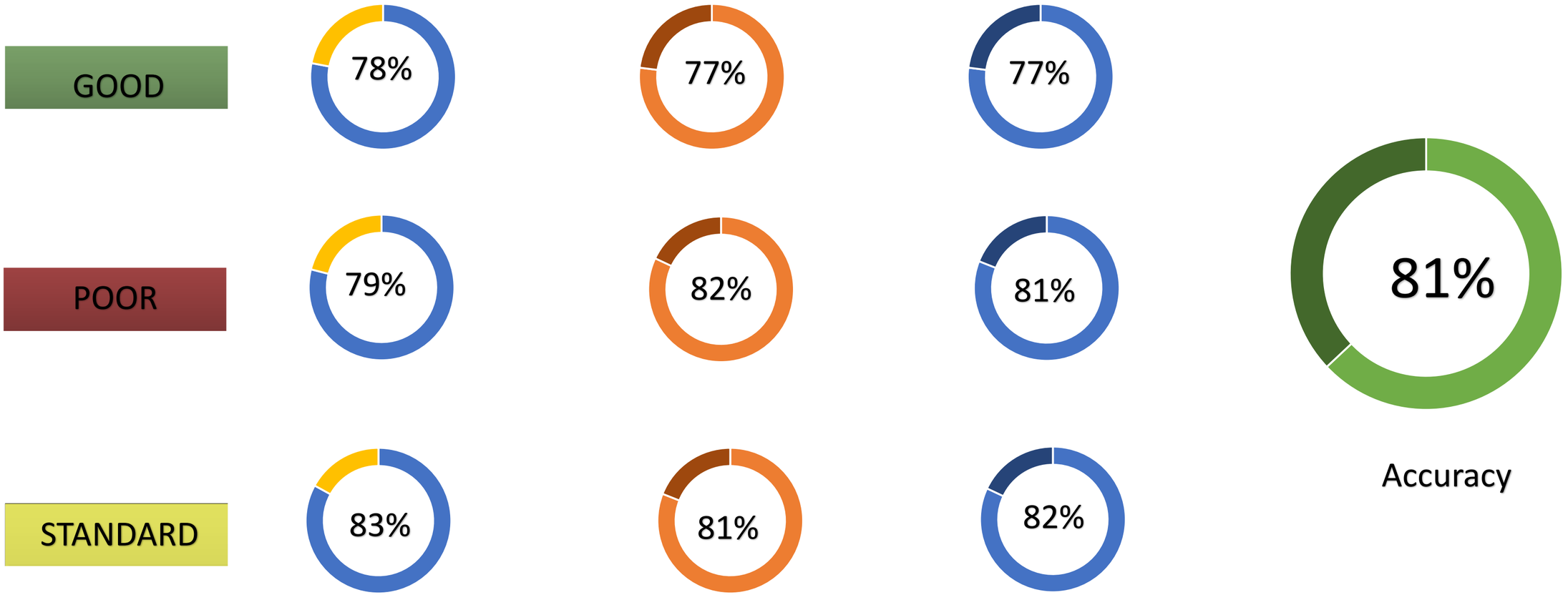




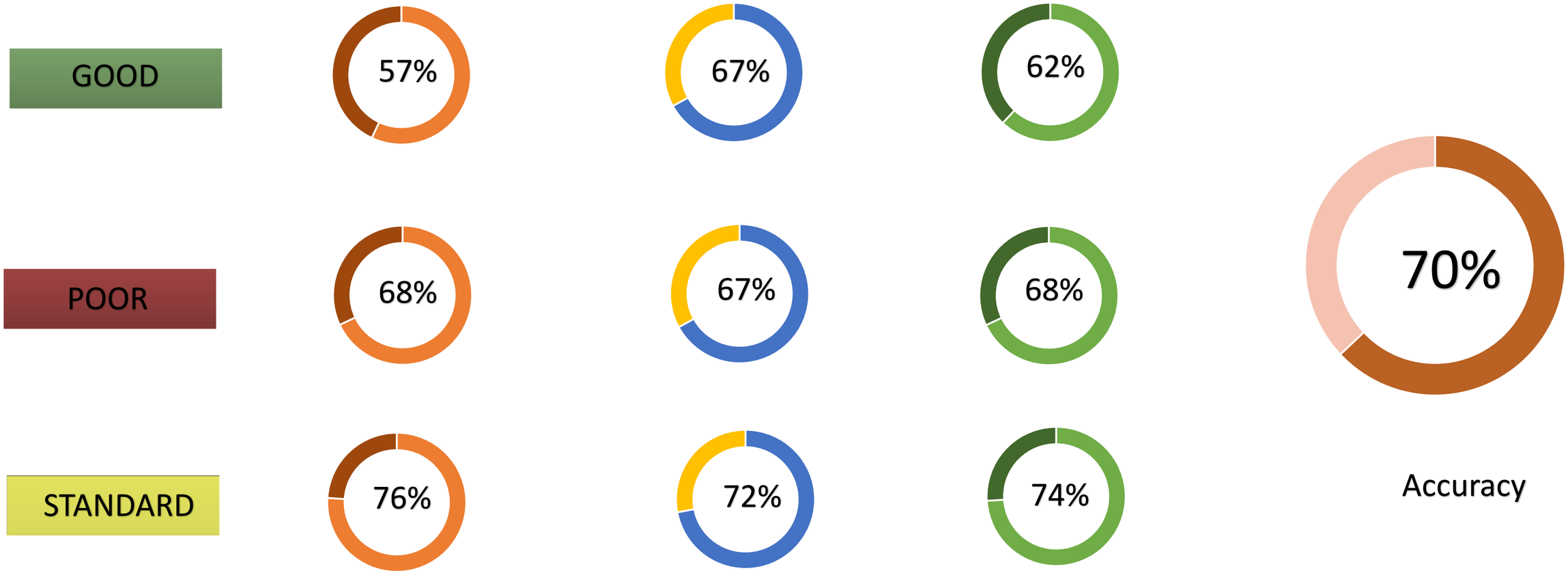


Accuracy

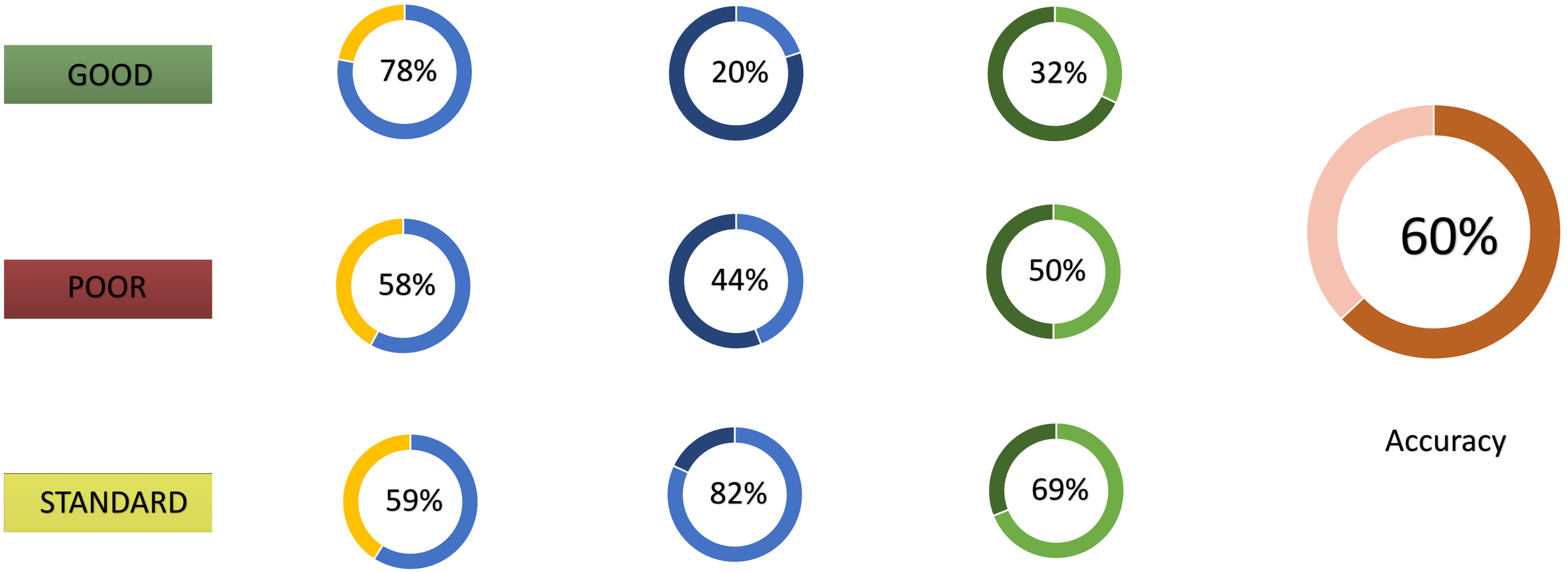
Model 2: Random Forest Classifier



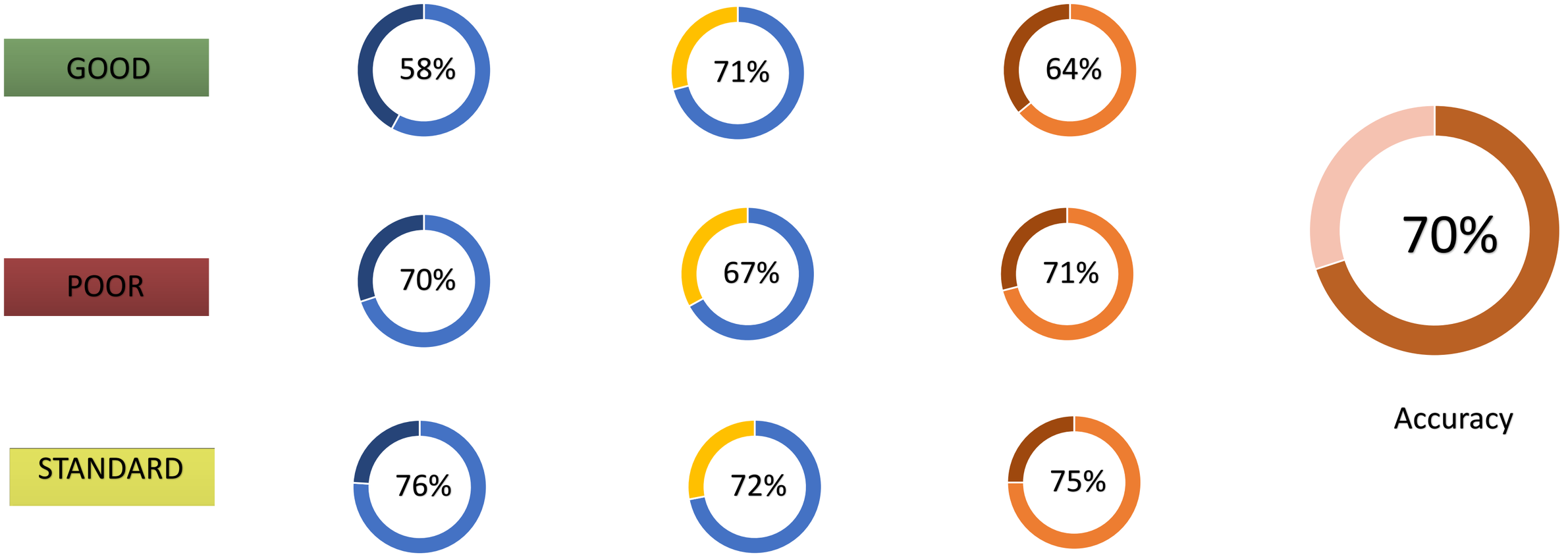
Model 3: Decision Tree Classifier



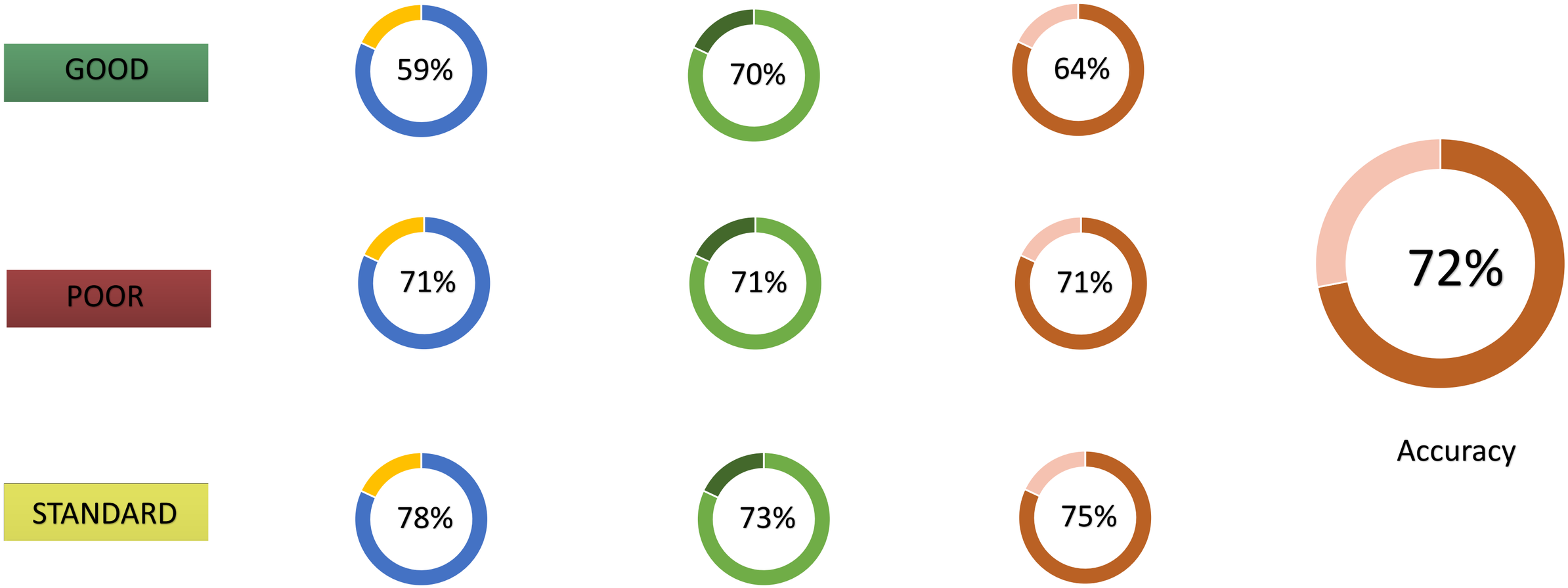
Model 4: Ada Boost Classifier



Model 5: Gradient Boost Classifier

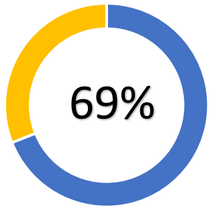


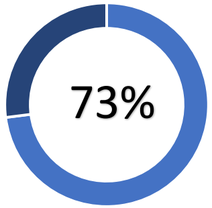
Model 6: XG Boost Classifier

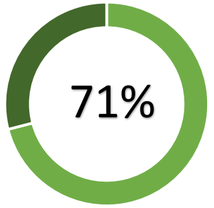


Model 7: K-Nearest Neighbours

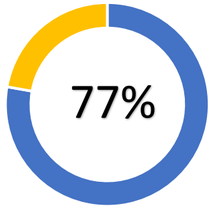
GOOD

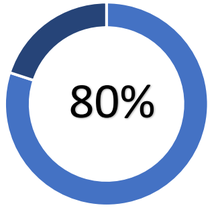


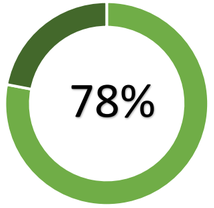




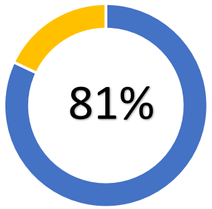
POOR

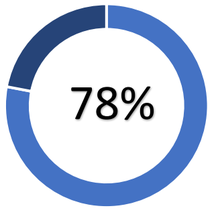


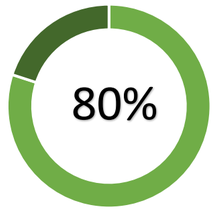


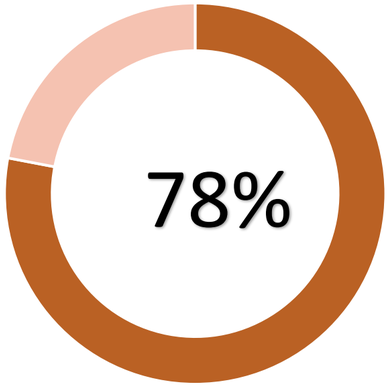


STANDARD





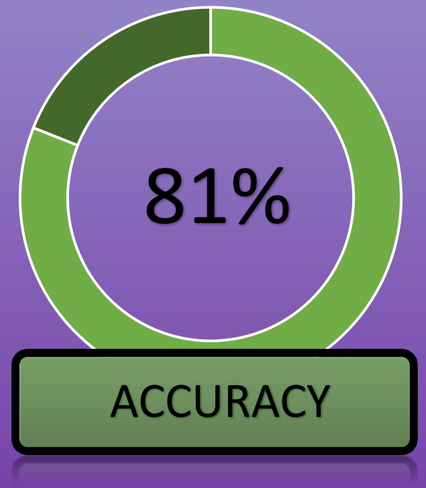




Accuracy

BEST MODEL

After apply various ML model, Random Forest Classifier comes out the best model



8

Future Work

Future Work

Future directions for improving Credit Score Classification model

1. **Using advanced machine learning techniques:**
Advanced machine learning techniques, perceptron, Data Mining can be applied to Credit Score prediction. These techniques can help capture complex patterns in the data that may be difficult to detect using traditional modelling approaches.

2. **Developing AI bases Real-time Approach**:
based approaches of real time Deep learning models, can be used to identify and examine credit score . These methods can be used to recognize credit score by analysing the customer history, payment transaction, and income structure.

3. **Cloud Computing for Huge Data**-
Cloud computing can be used to store a large volume of Customer Credit data. So that we can get accurate prediction of customer Credit Score Classification .

9

Conclusion

Final conclusion

Conclusion

In conclusion, the development and implementation of the credit score classification model represent a significant stride towards enhancing financial decision-making processes. This robust model, built on comprehensive data analysis and machine learning algorithms, demonstrates its potential to accurately assess an individual's creditworthiness. By considering various factors such as payment history, credit utilization, length of credit history, and other relevant variables, the model provides a nuanced and reliable evaluation.

The successful deployment of this credit score classification model holds the promise of optimizing lending practices, promoting financial inclusion, and mitigating risks for financial institutions. It not only streamlines the evaluation of credit applicants but also aids in identifying potential areas of improvement for individuals seeking to enhance their credit profiles. Furthermore, the model's adaptability and scalability make it well-suited for evolving financial landscapes, ensuring its continued effectiveness in a dynamic economic environment.

As we move forward, continuous monitoring, refinement, and ethical considerations will be essential to maintain the model's accuracy and fairness. Striking a balance between innovation and responsible use of data is paramount to uphold the trust of both consumers and financial institutions. In essence, the credit score classification model stands as a testament to the transformative power of technology in shaping a more inclusive and efficient financial ecosystem.

THANKS!

Any questions?

You can find me at:

- Sumit-Patil-1997 (Sumit Patil) (github.com)

- **Sumitkpatil97@gmail.com**