COMPANY BANKRUPTCY PREDICTION
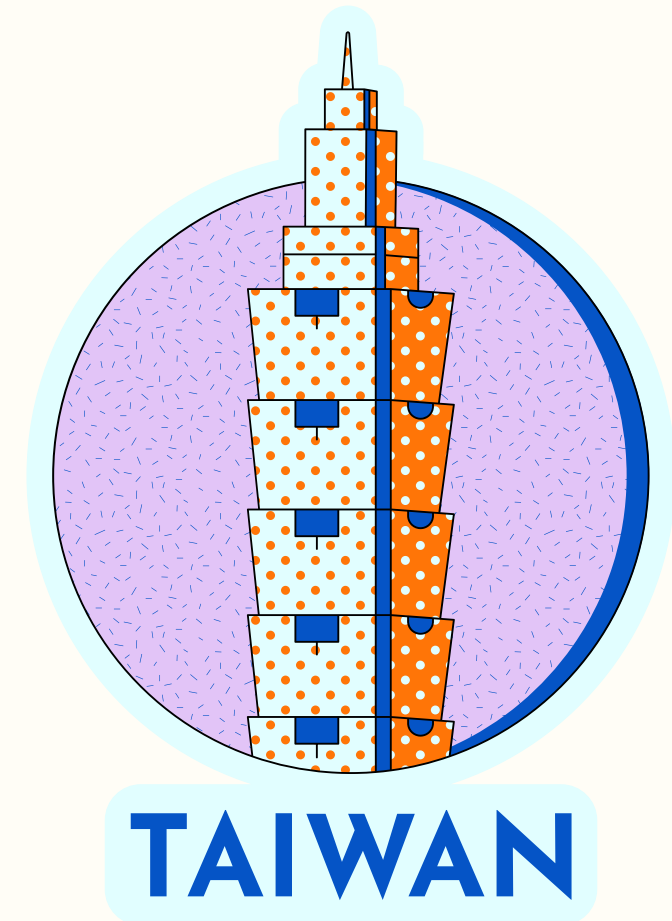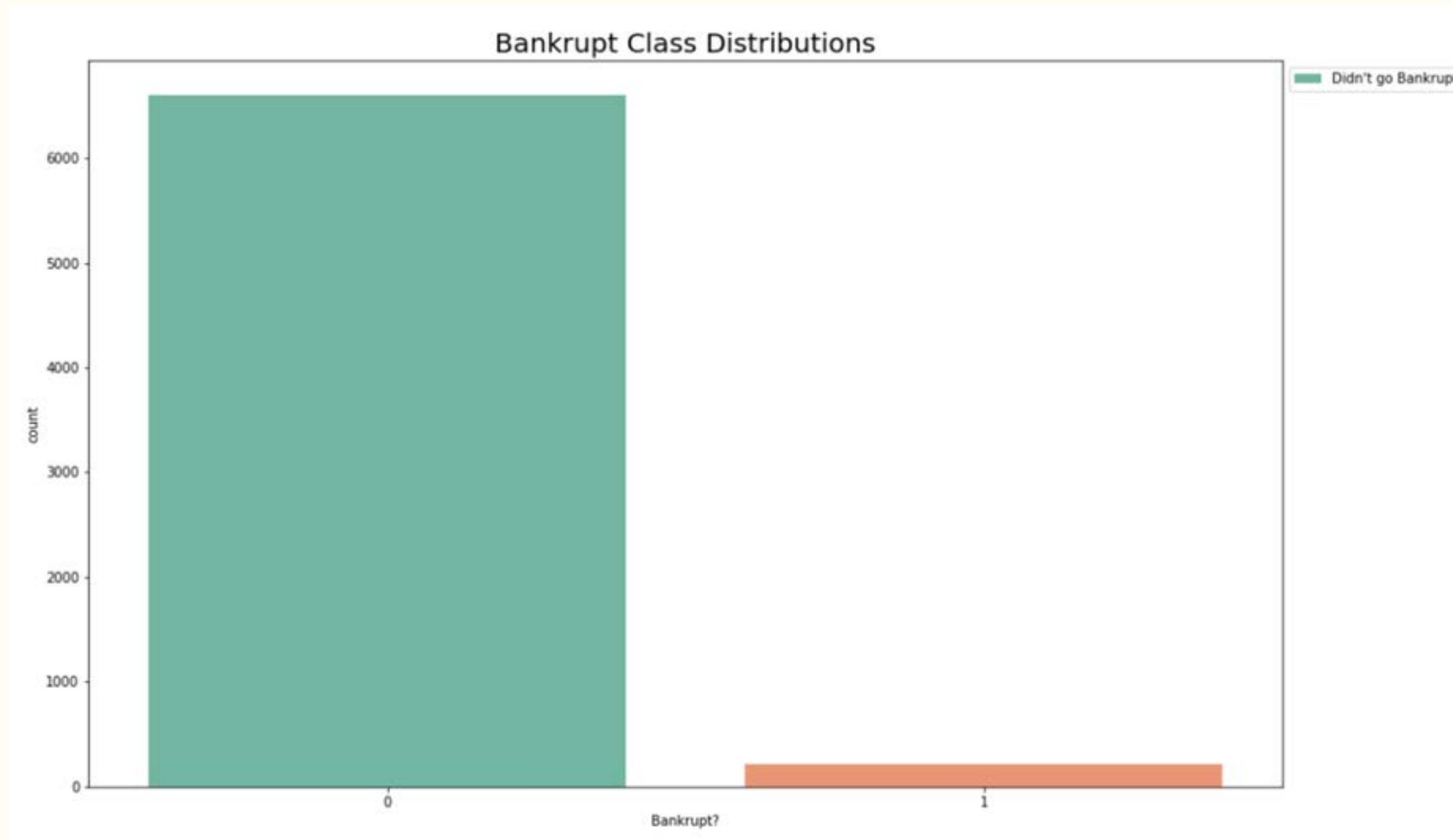
Bankruptcy prediction model

Data Cleaning

Exploratory Data Analysis

- **Financial information of 6800 Taiwanese companies**
- Between years 1999 and 2009
- Source: Taiwan Economic Journal
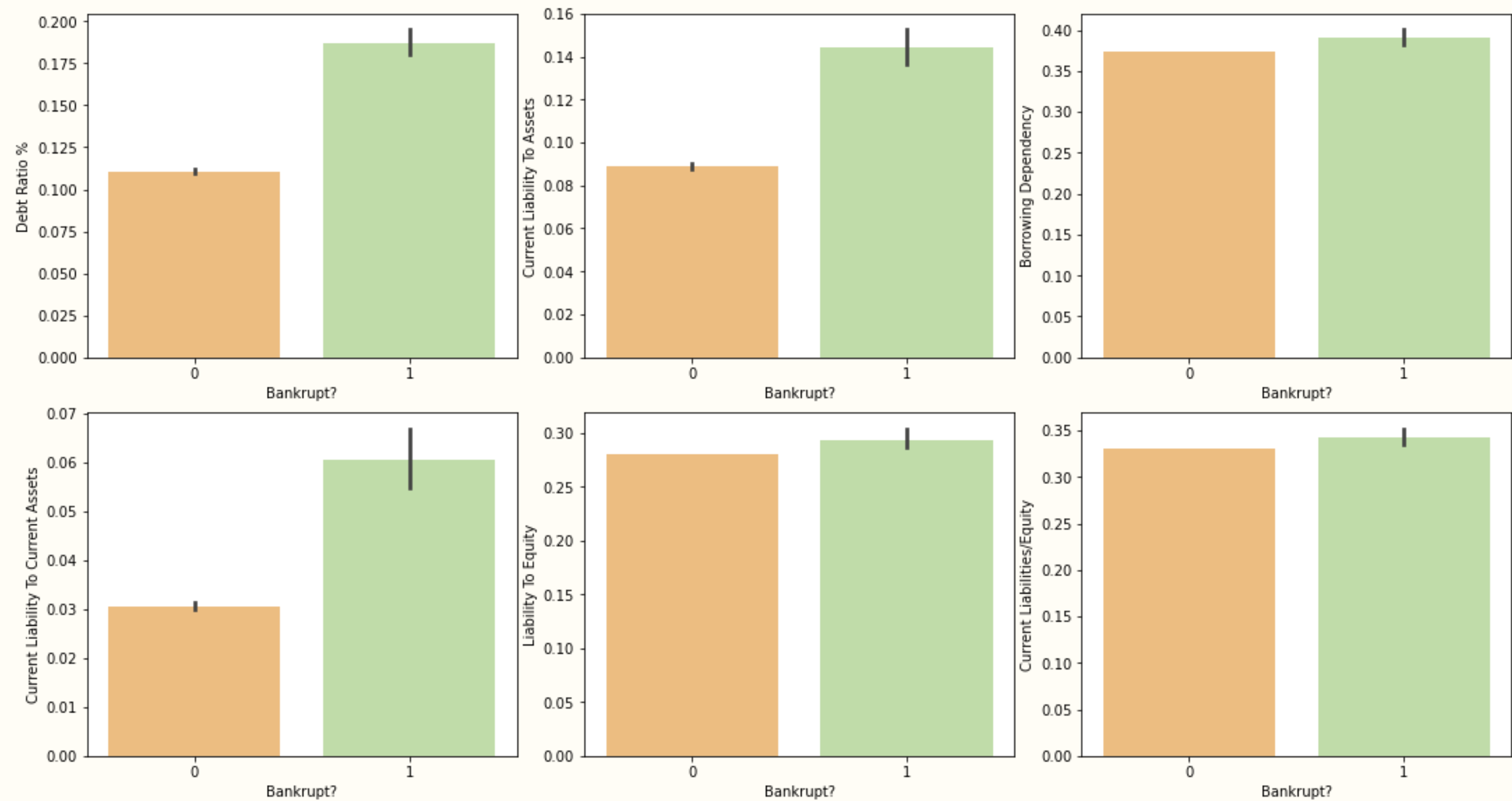
95 financial variables:
- Operating gross margin,
- Gross profit to sales,
- Equity to liability,
- Realized Sales Gross Profit Growth Rate,

TAIWAN

Bankruptcy: 220 companies
No bankruptcy: 6599 companies

Top six top positively correlated variables with bankrupcty:
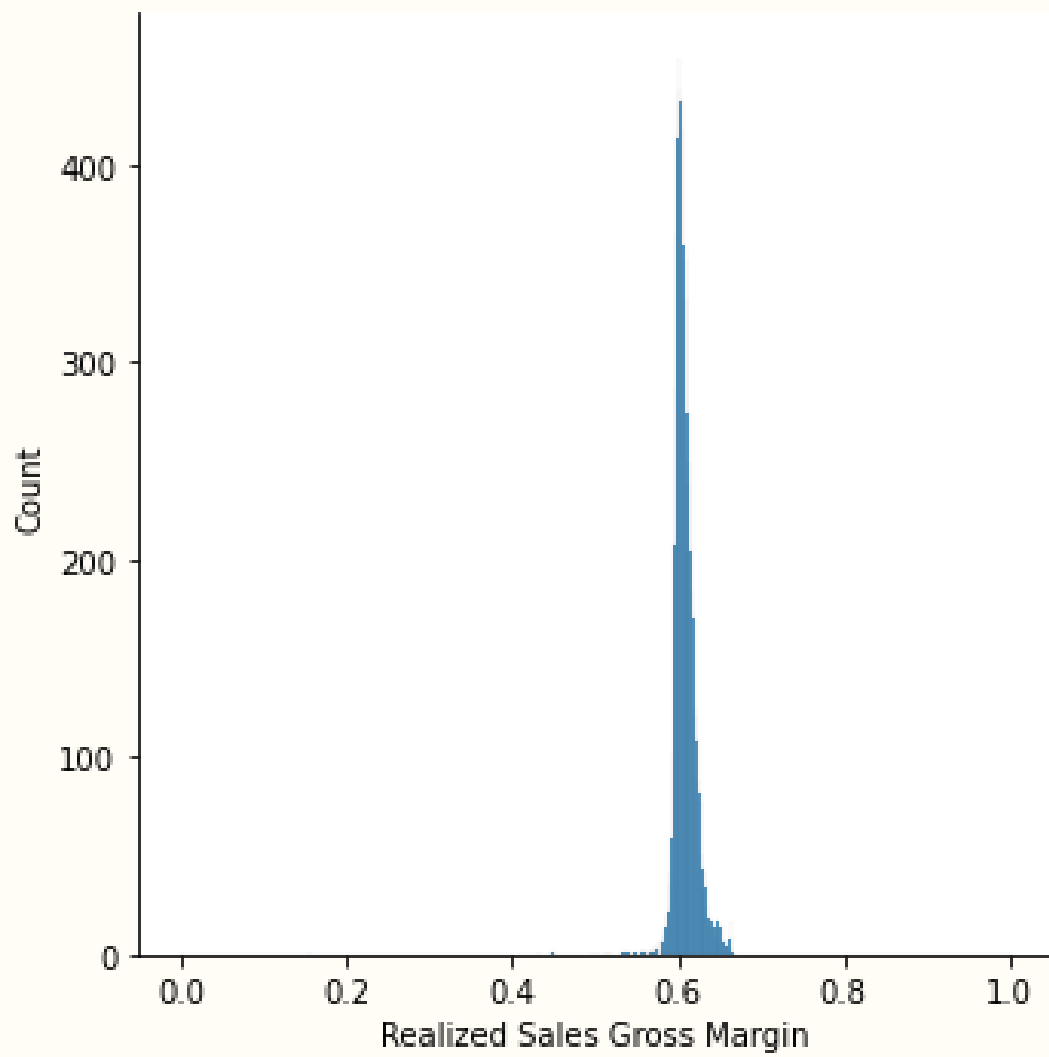
Debt Ratio, *Current Liability To Assets* and *Current Liability To Current Assets* are commonly high in bankrupt companies.
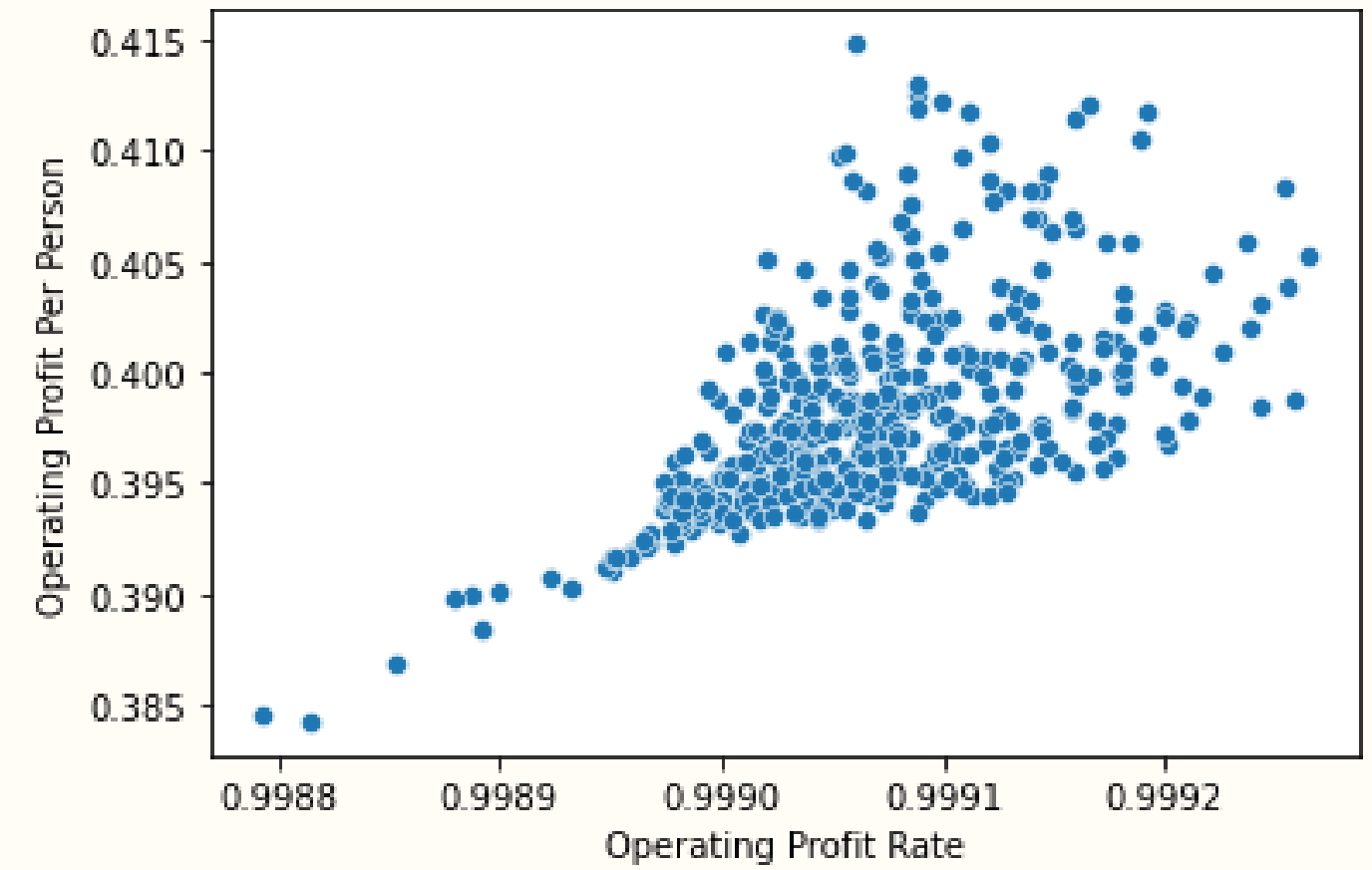
Top six negatively correlated variables with bankrupcty:
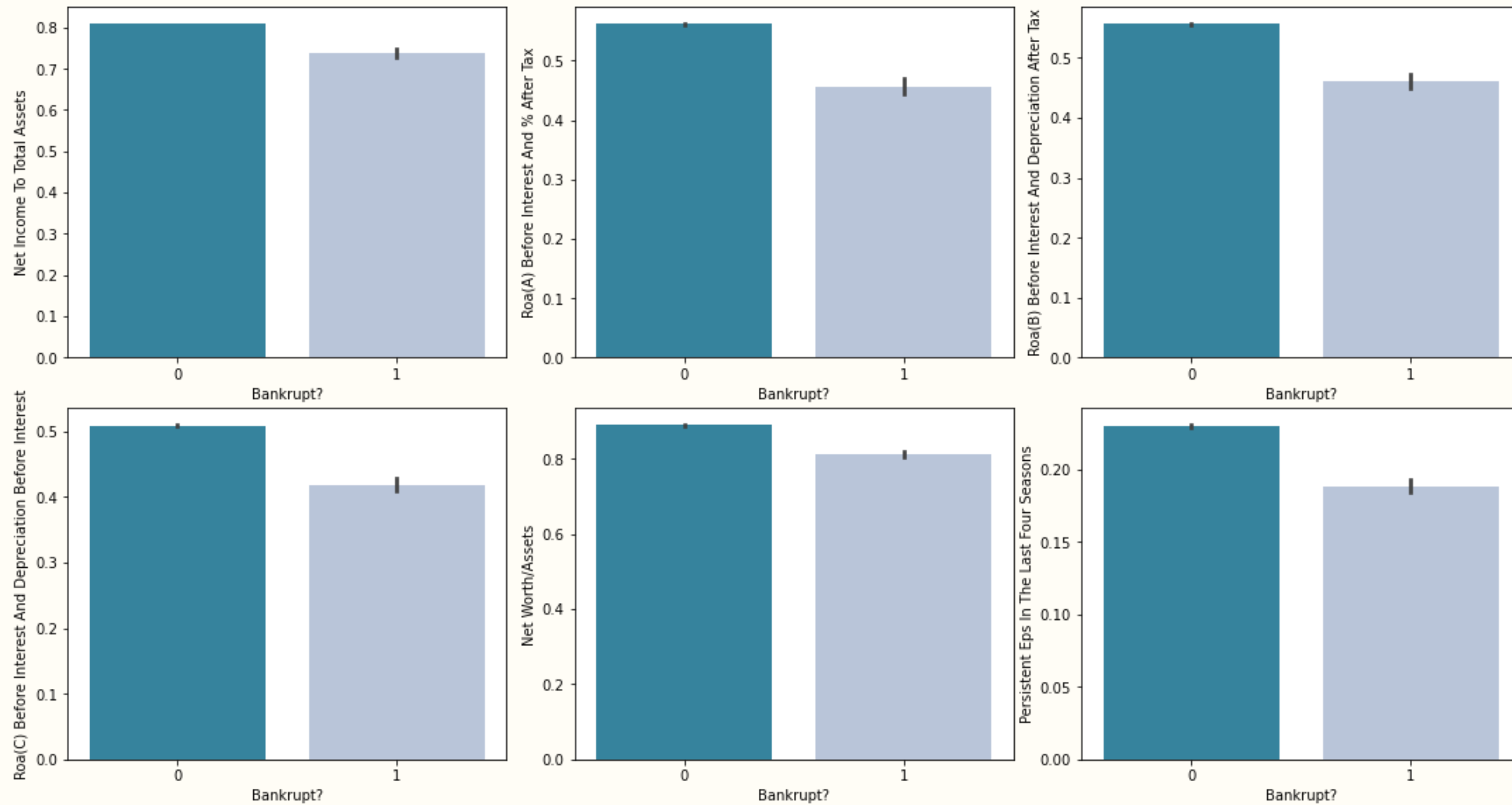
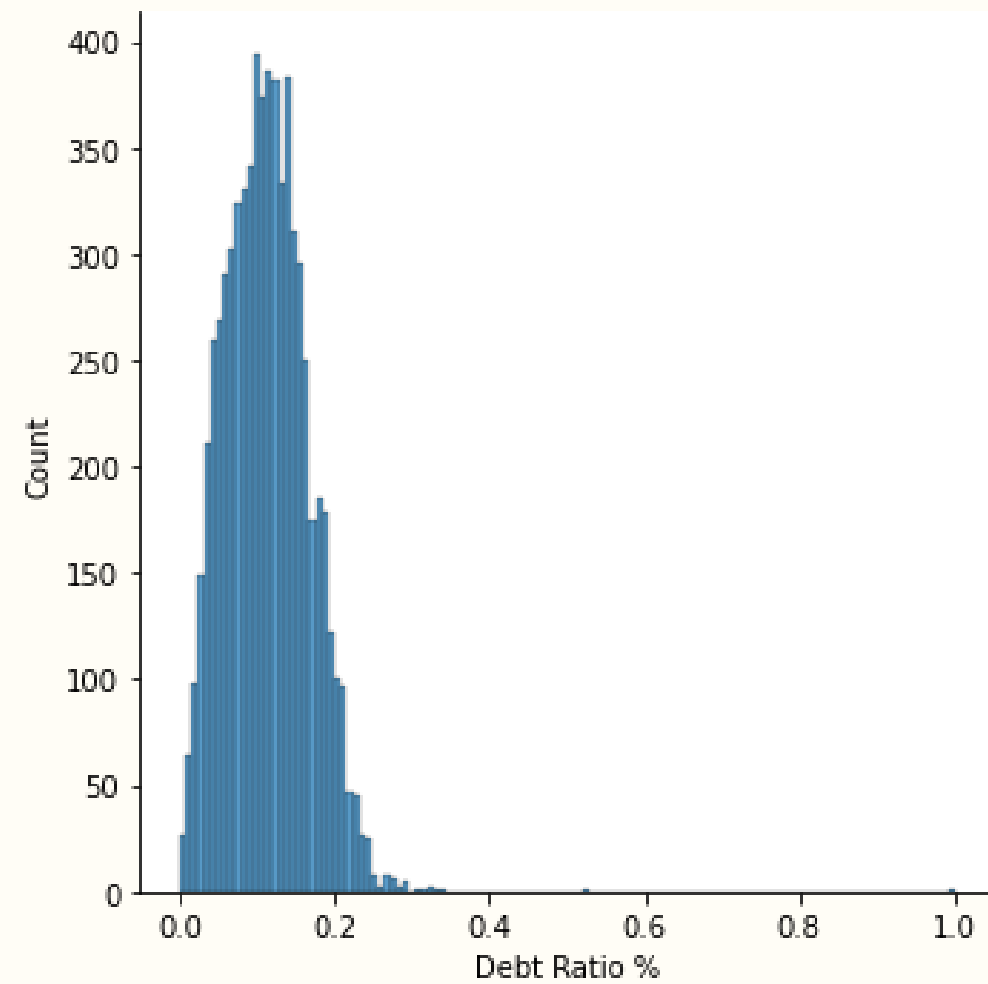The more **assets** and the highest the **earnings**, the less likely a company will be bankrupt.
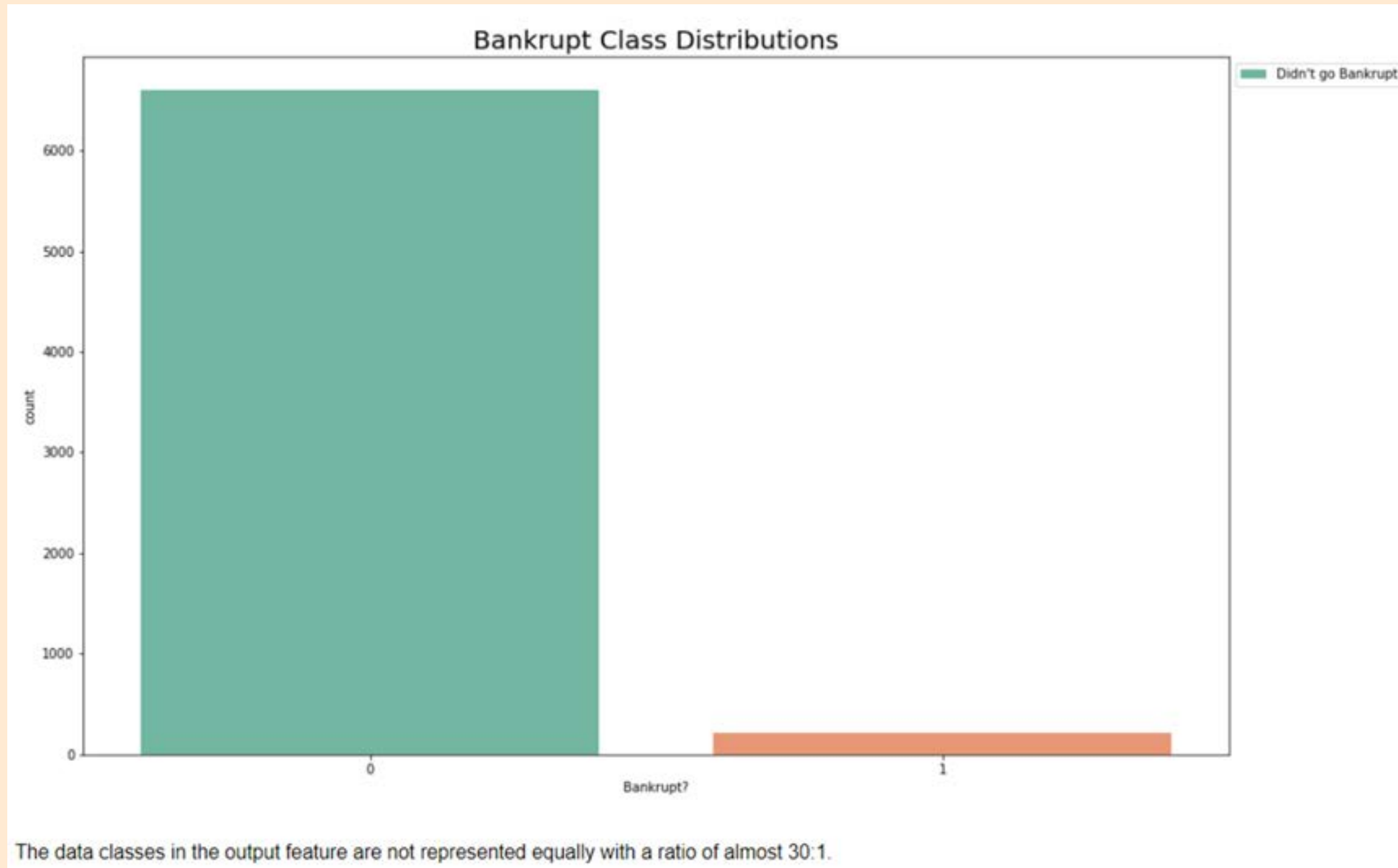
Most companies have a *sales gross margin* around 0.6

Most of the companies have a *debt ratio* between 0 and 20%

Operating profit per person and ***Operating profit rate*** are linearly correlated

DATA Imbalance Correlation MATRIX

Data Imbalance

Bankrupt Class Distributions

The data classes in the output feature are not represented equally with a ratio of almost 30:1.

Data Modeling
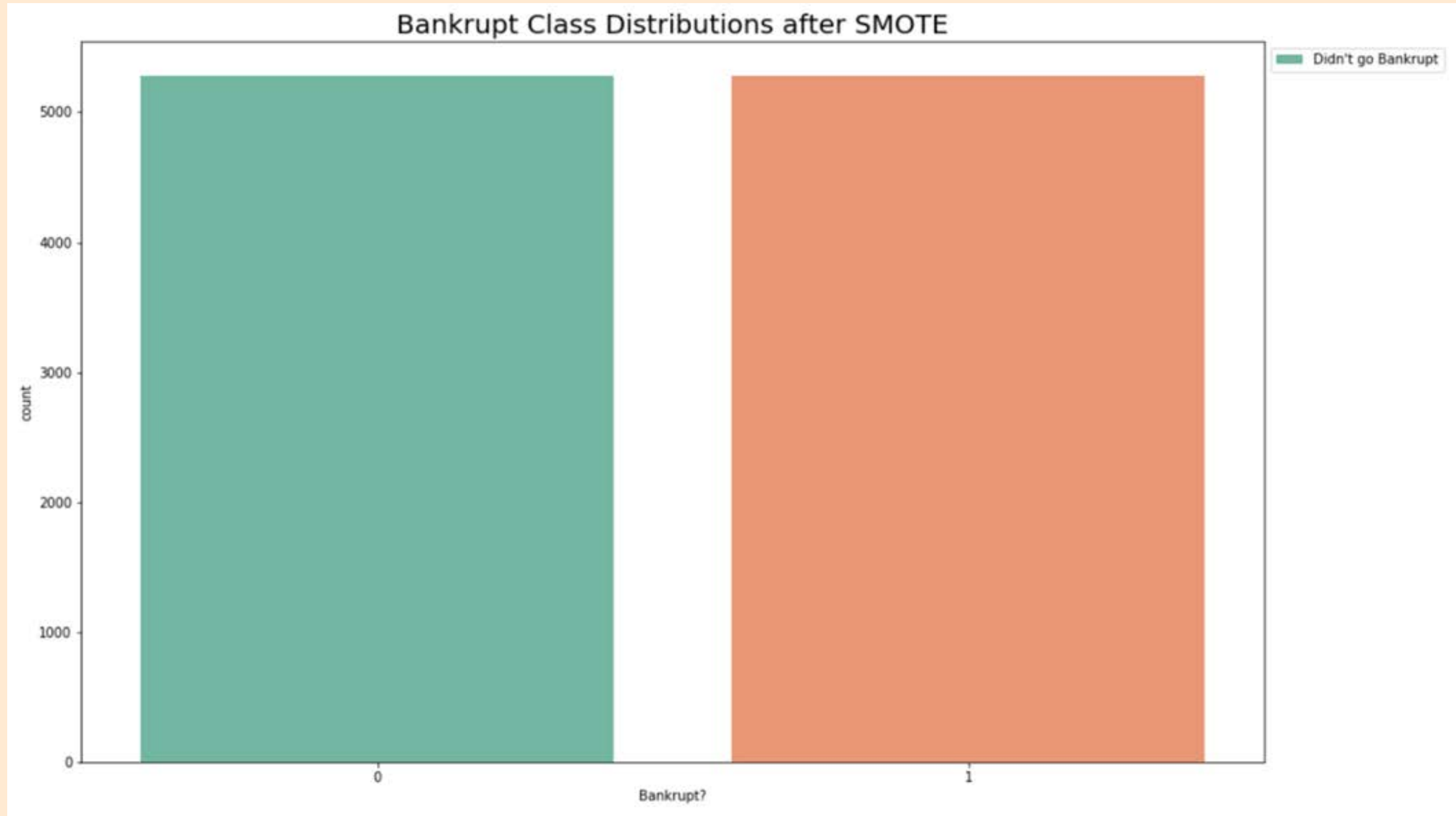
Correlation Matrix
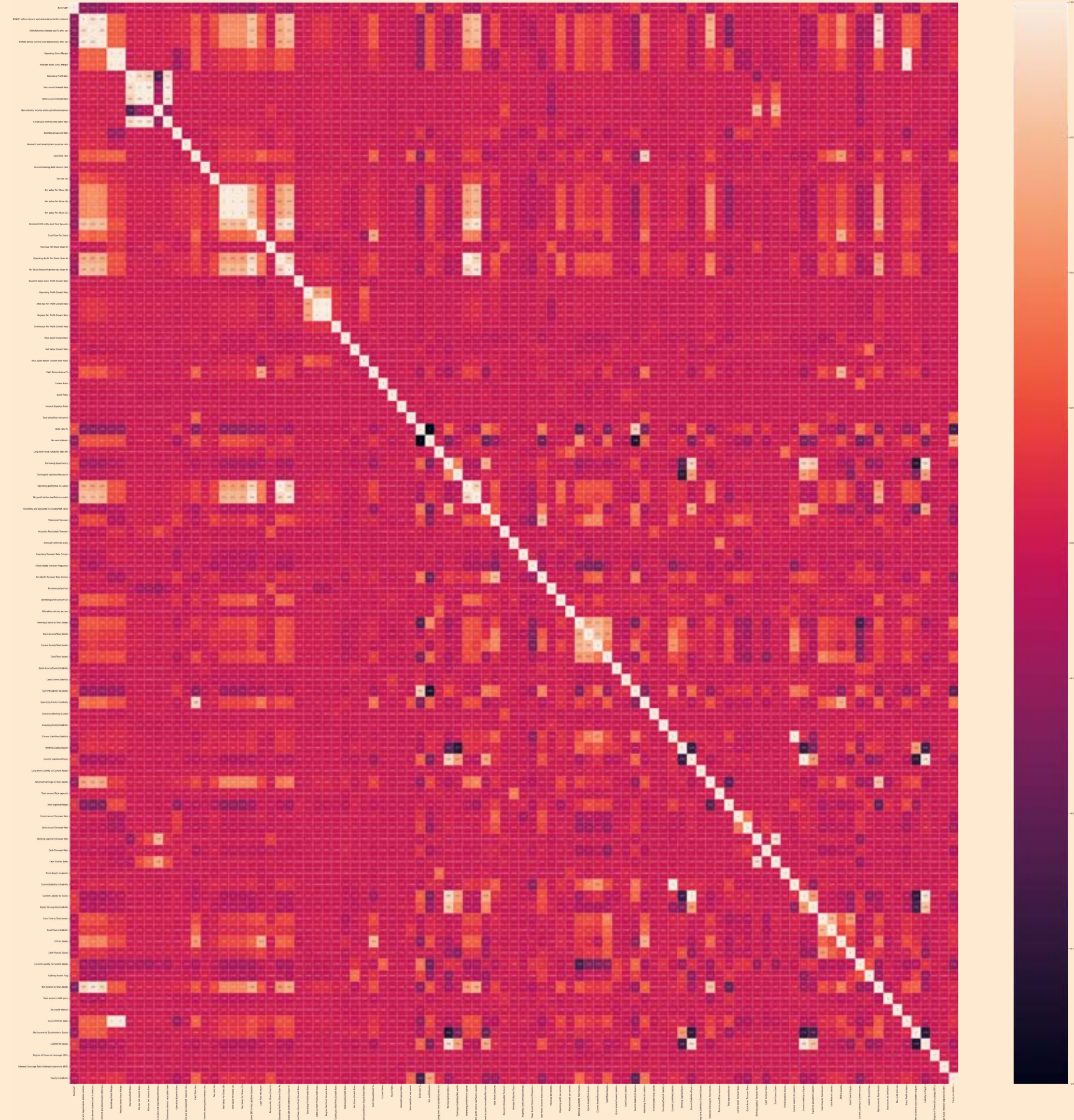
Machine Learning Models

Confusion Matrix

Logistic Regression Confusion matrix

	Predicted 0	Predicted 1
Actual 0	1168	152
Actual 1	10	34

Decision Tree Confusion matrix

	Predicted 0	Predicted 1
Actual 0	1241	79
Actual 1	23	21

Support Vector Machine Confusion matrix

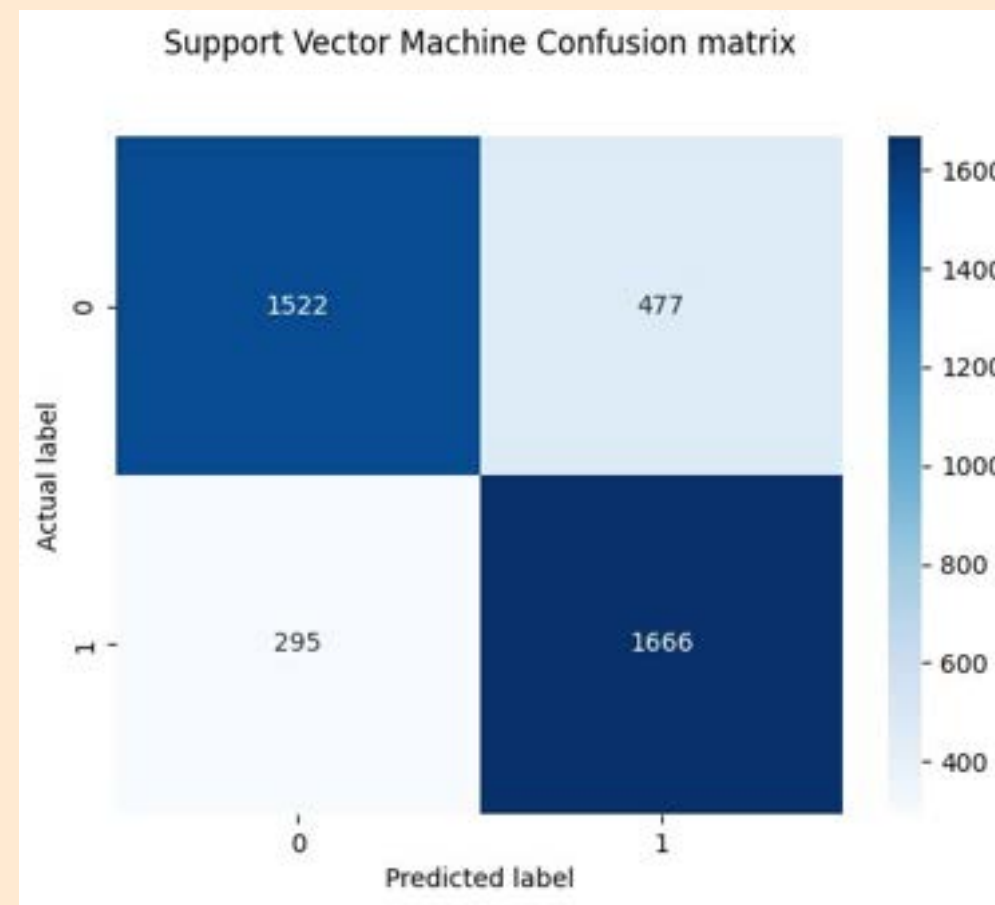|	Predicted 0	Predicted 1
Actual 0	1201	119
Actual 1	18	26

K-Nearest Neighbors Confusion matrix

	Predicted 0	Predicted 1
Actual 0	1224	96
Actual 1	21	23

Random Forest Confusion matrix

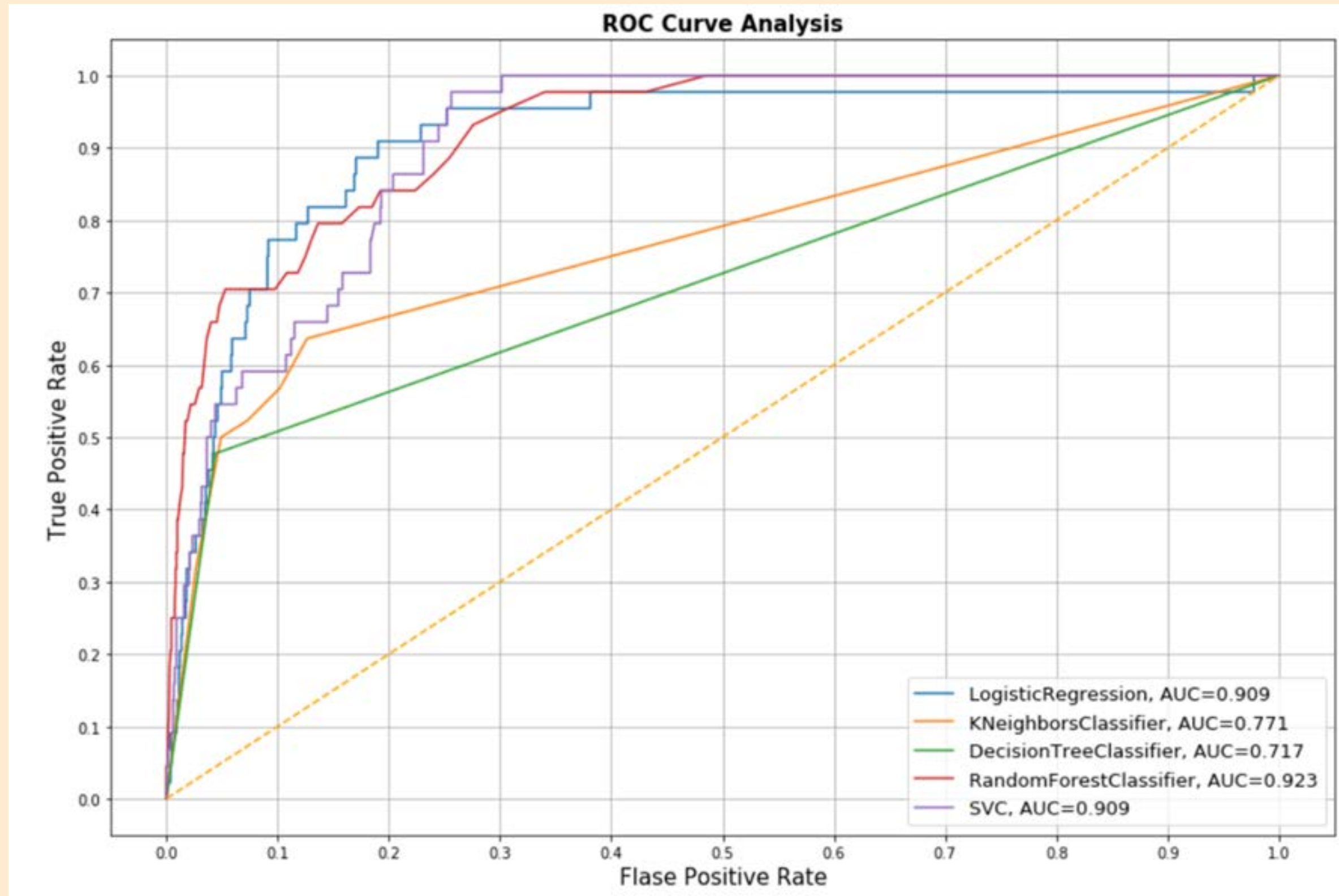|	Predicted 0	Predicted 1
Actual 0	1298	22
Actual 1	23	21

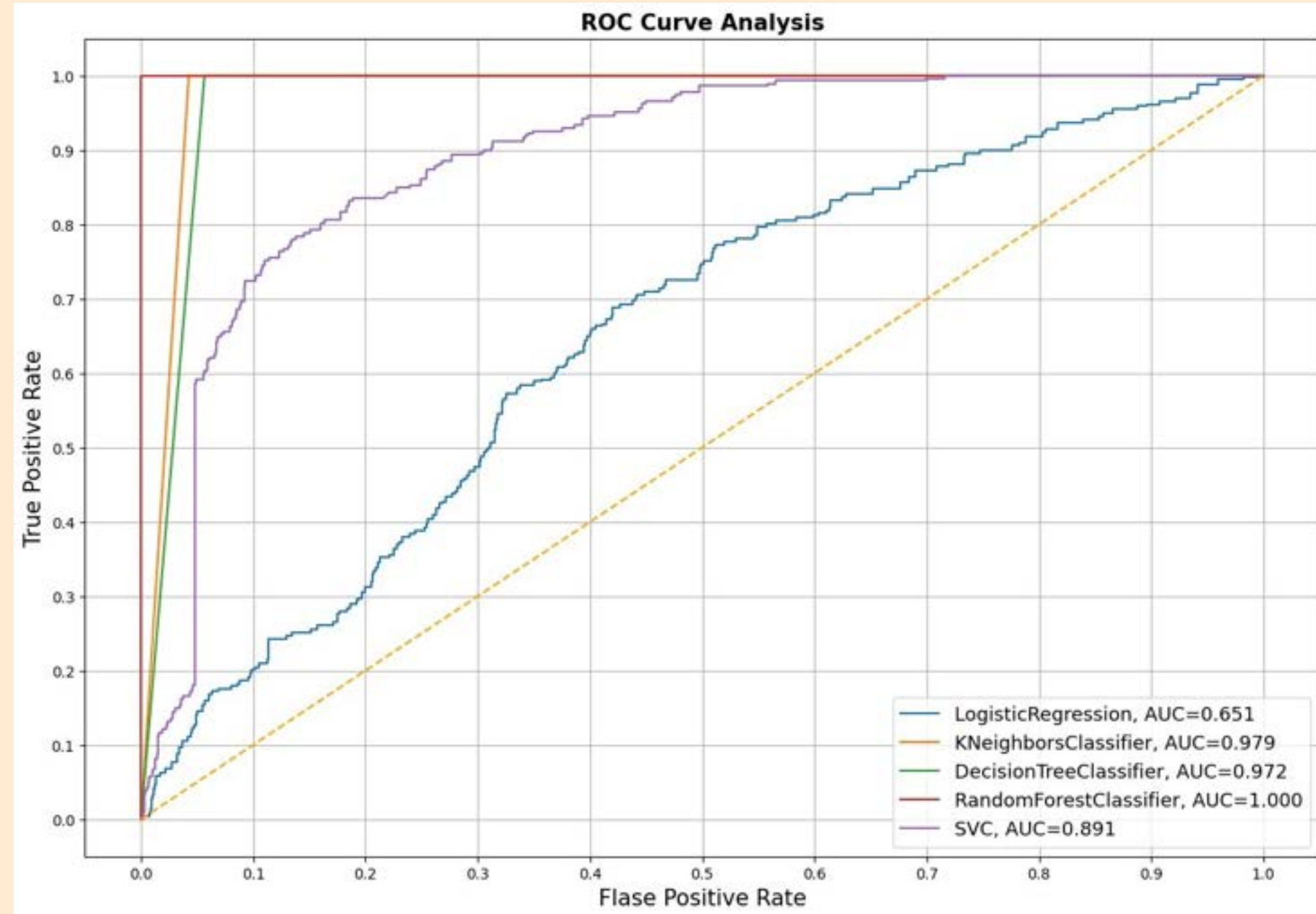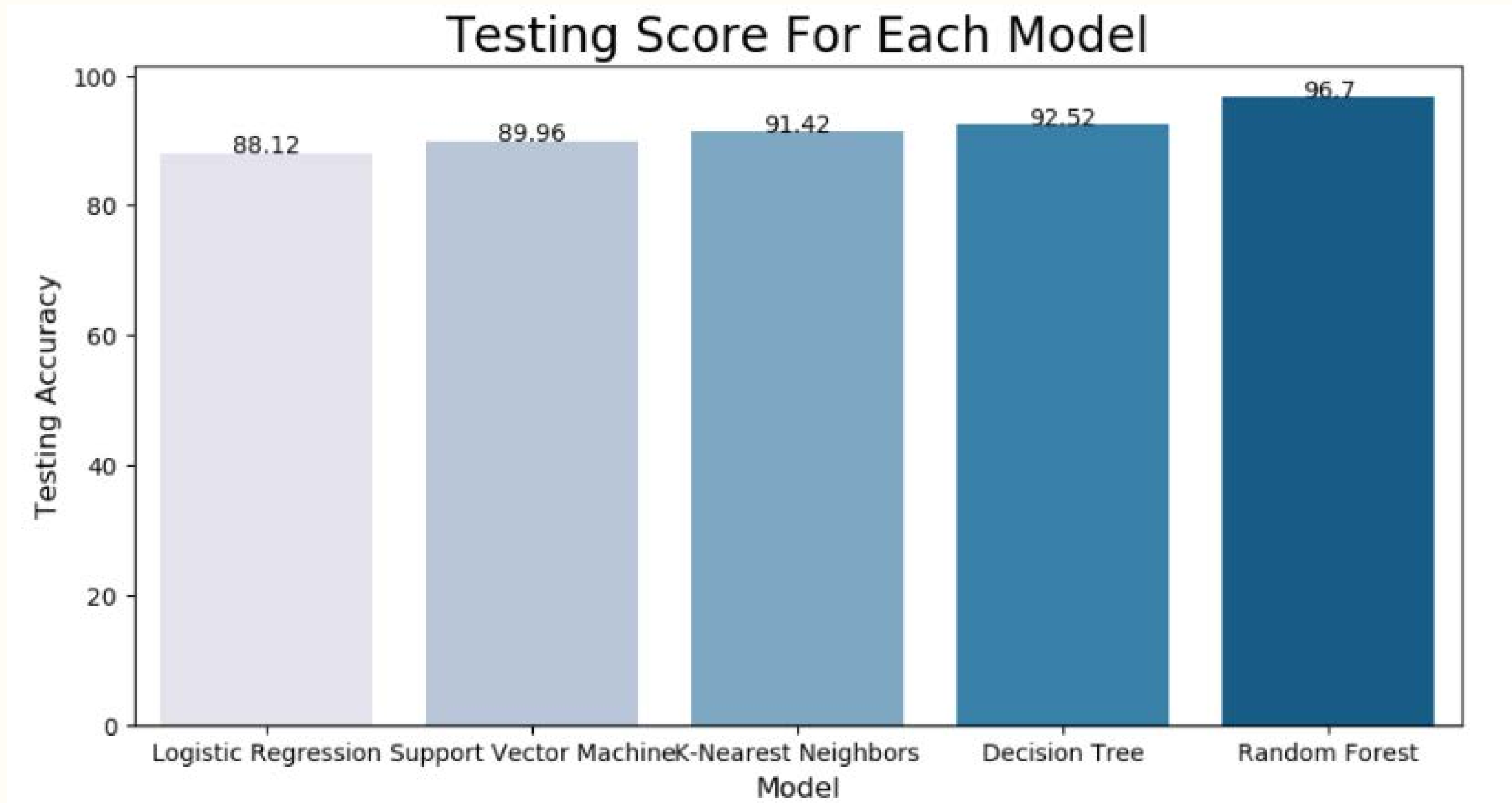# Confusion Matrix after Smote and PCA
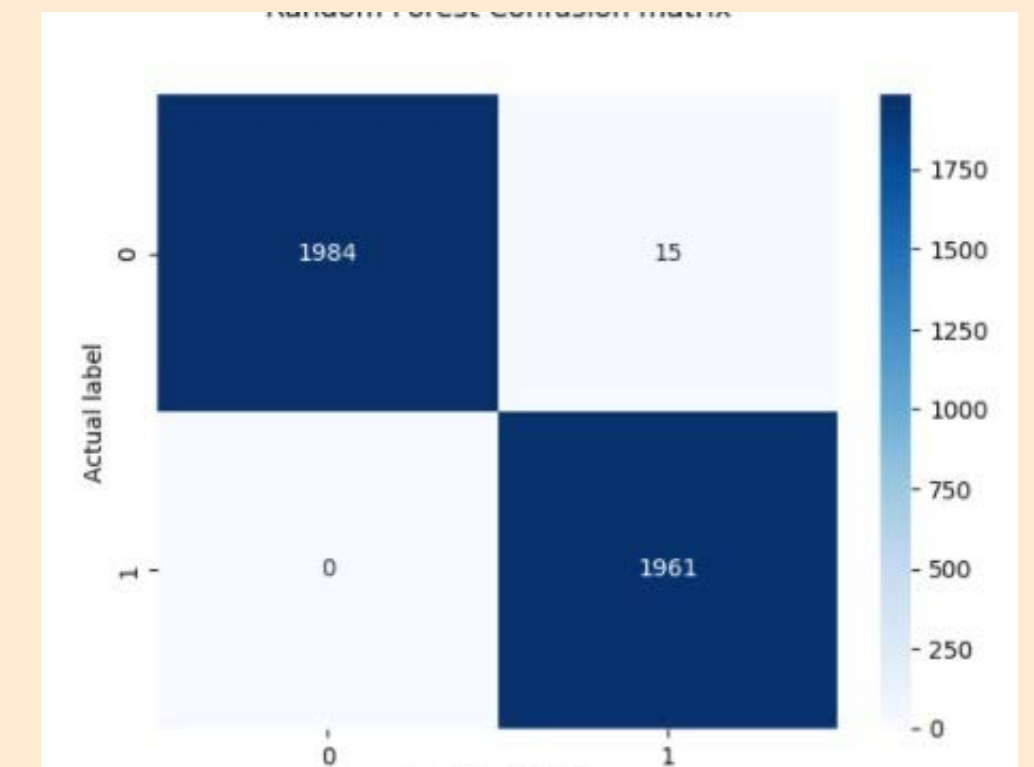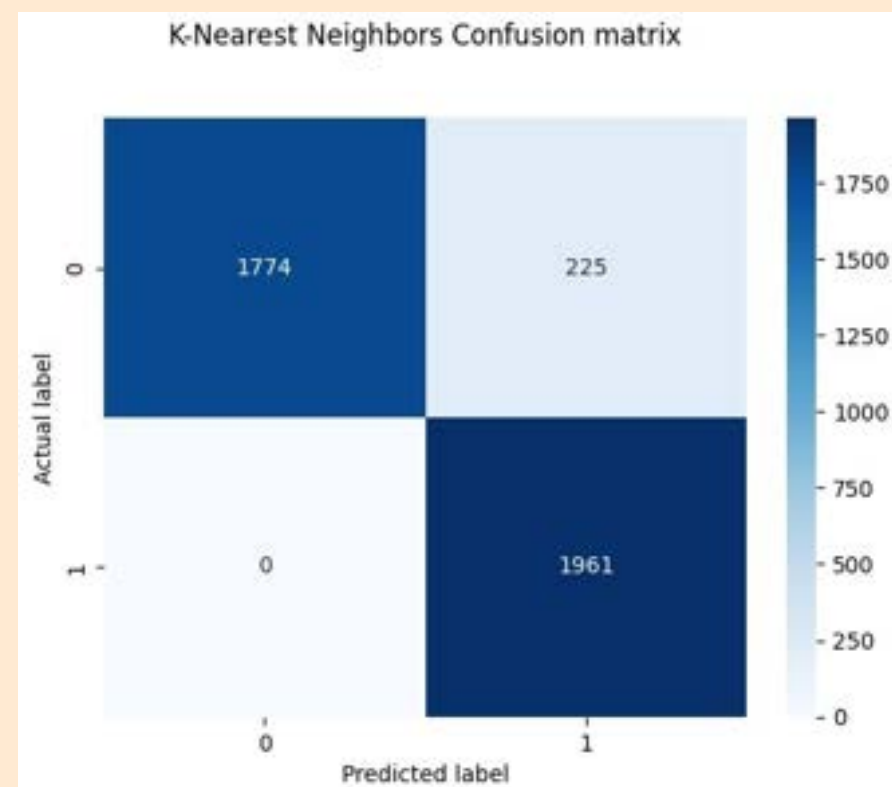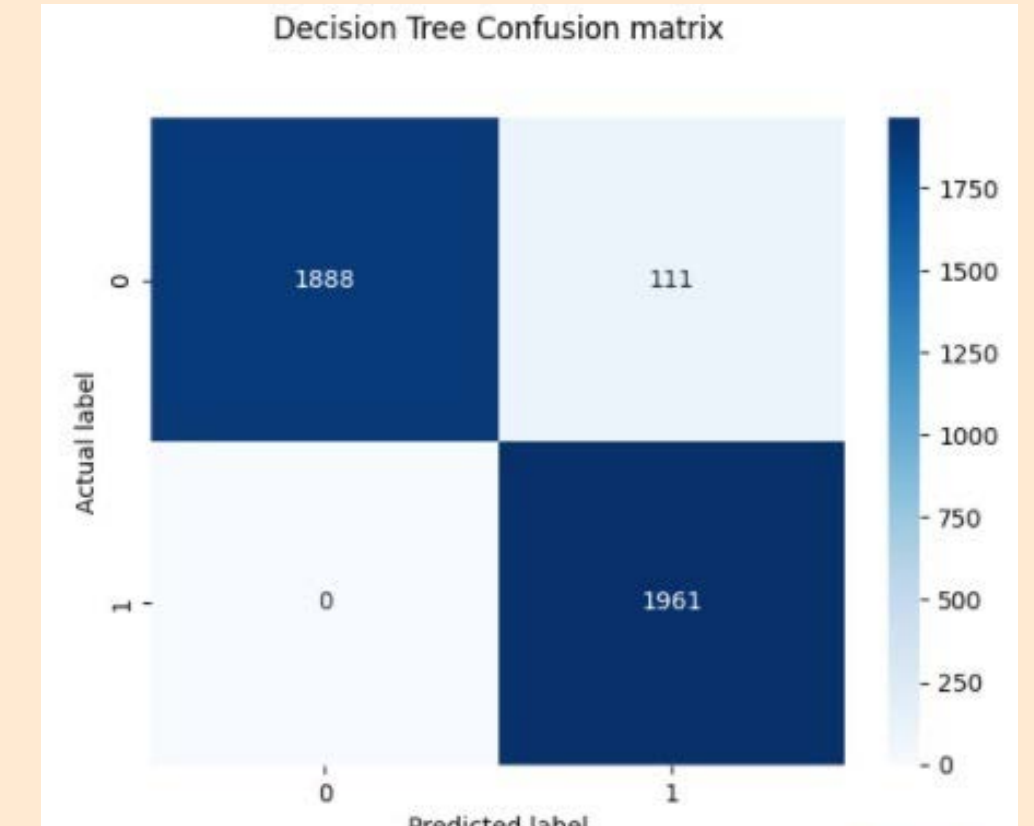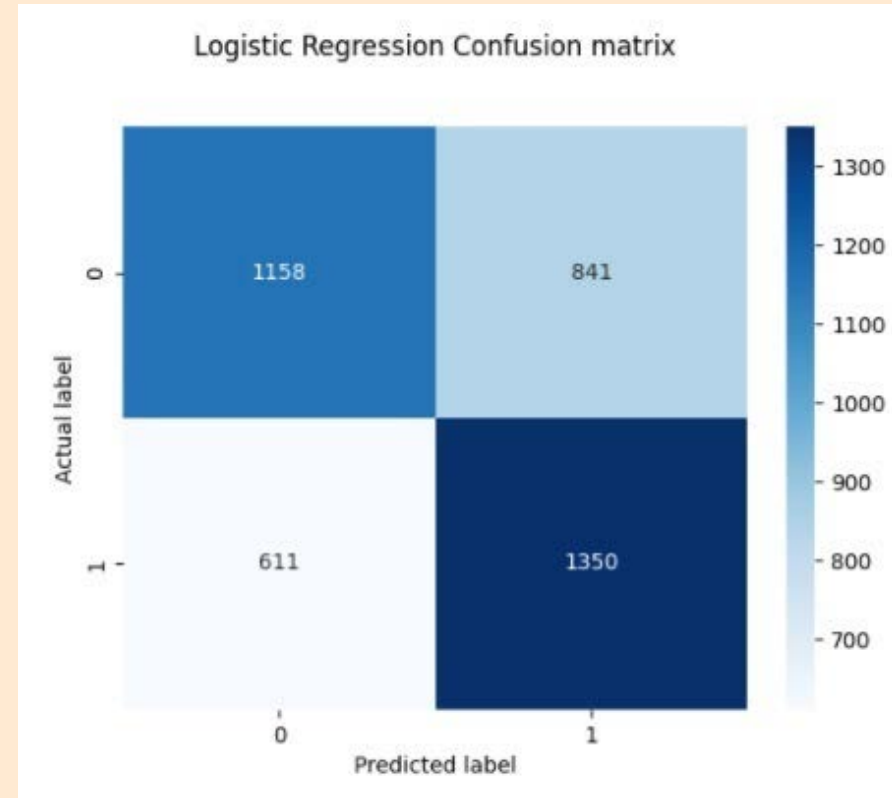
ROC Curve

ROC Curve after SMOTE and PCA

Comparing Models

Testing Score For Each Model

Conclusion