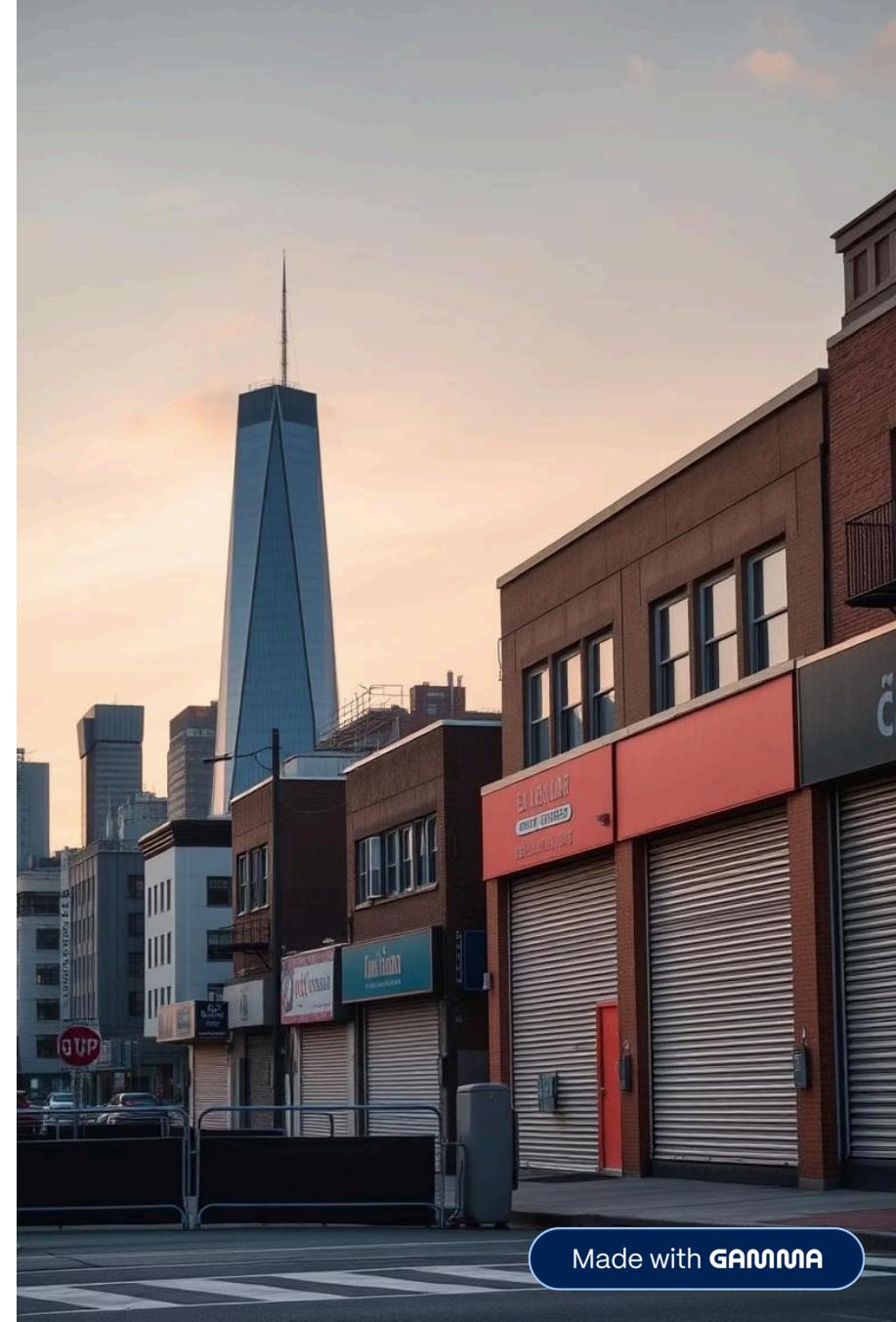


Predicting Company Bankruptcies

- The model will help identify companies at high risk of bankruptcy, allowing for proactive measures. Accurate predictions can inform lending and investment strategies. Understanding bankruptcy risk can assist in maintaining a balanced and low-risk financial portfolio.

by Rishita Shah



Introduction :

Objectives :

1. To explain the business problem & Project Motivation .
2. To showcase the Machine Learning Approach for bankruptcy prediction .
3. To present Insights Through Visualization .
4. To Deliver Actionable Outcomes & Project Deliverables .

Dataset :

- Total Rows : 6819 Companies
- Total Columns : 96 Columns
- Target Variable : Bankrupt? (Binary : 1 = Bankrupt , 0 = Not Bankrupt)

Importance :

 Early Risk Detection

 Investor & Lender Performance

 Portfolio Management

 Economic Impact Mitigation

 Strategic Decision-Making

Here is a Data Dictionary Summarizing the Key Columns in the Dataset :

Dataset contains **96 columns**, including 1 target variable and 95 financial indicators. Below is a sample of the **Data Dictionary** showing the first 10 features.

Column Name	Data type
• Bankrupt?	• int64
• ROA(C) before Interest and depreciation before Interest	• float 64
• ROA(A) before interest & % after tax	• float 64
• ROA(B) before interest & depreciation after tax	• float 64
• Operating gross margin	• float 64
• Realized sales gross margin	• float 64
• Operating profit rate	• float 64
• Pre-tax net interest rate	• float 64
• After-tax net interest rate	• float 64
• Non-industry income & Expenditure/revenue	• float 64

Steps :

1

Data Collection

Gathering & Organizing data to train machine learning models .

2

Preprocessing

Preparing and cleaning data to enhance its quality and suitability .

3

Exploratory Data Analysis

Analyzing and visualizing data patterns to understand its characteristics .

4

Split Train & Test Data

Dividing the dataset into training and testing sets to evaluate .

5

Model selection & Model Training

Choosing a suitable machine learning algorithm and optimizing its parameters .

6

Model Evaluation

Assessing the performance of machine learning model using metrics .

Exploratory Data Analysis :

- **Descriptive Statistics :**

- Mean, median, mode
- Min, max, quartiles
- Standard deviation, variance

(Using methods like `.describe()` in Python)

- **Correlation Analysis :**

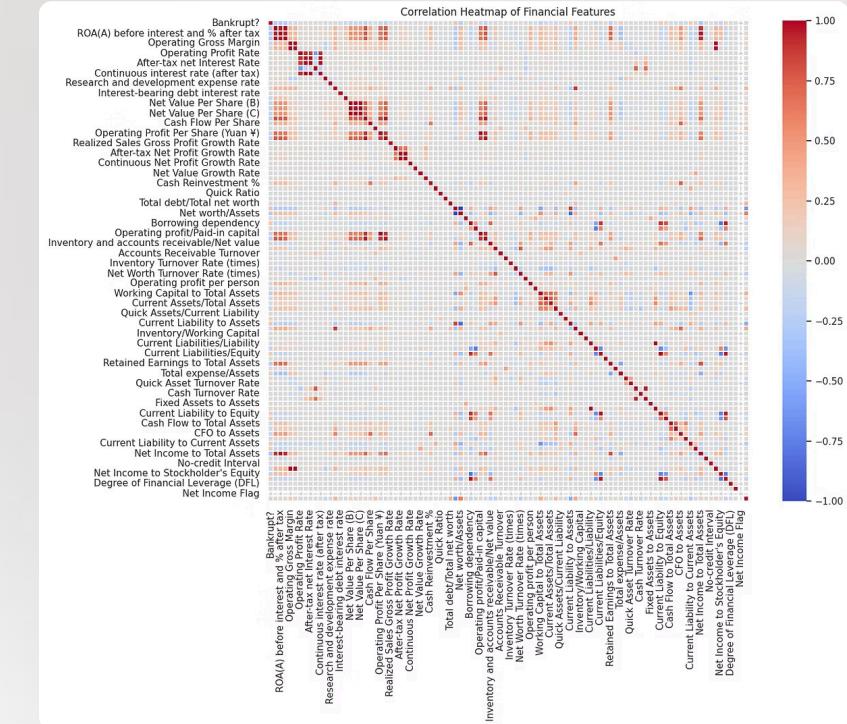
- Examining correlations between features reveals potential relationships and helps identify which features are most relevant for segmentation.
- Examining correlations between features reveals potential relationships and helps identify which features are most relevant for segmentation.

- **Visualizations :**

- Makes complex data simple
- Highlights key insights quickly
- Supports decision-making
- Engages the audience better than raw numbers

List of visualizations which i choose :

- 1 Class Distribution : Bankrupt VS Non-Bankrupt
- 2 Top Features Correlated with Bankruptcy
- 3 Correlation Heatmap of Financial Features





Bankruptcy Prediction Models



Logistics Regression

Accuracy : 0.75



SVM (Support vector machine)

Accuracy : 0.79



Voting Classifier

Accuracy : 0.93



Random Forest

Accuracy : 0.96



KNN (K-Neighbors Classifier)

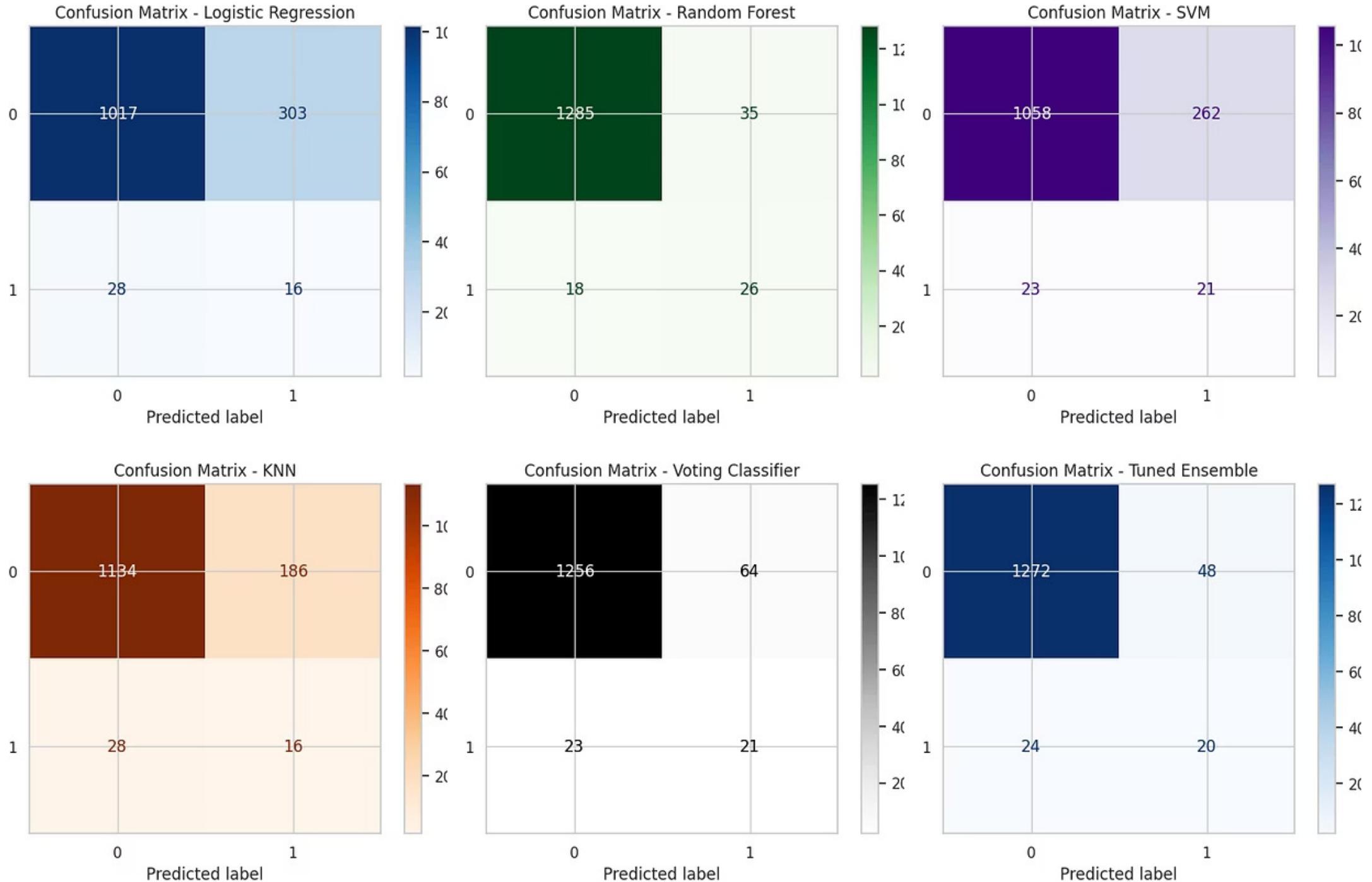
Accuracy : 0.84



Tunned voting classifier

Accuracy : 0.94

Confusion Matrix of Models :





Conclusion :

Developed machine learning models to predict company bankruptcies using financial, operational, and risk-related indicators .

Key Insights:

- Financial ratios and past delinquency are strong predictors of bankruptcy risk .
- Early identification of at-risk companies can significantly reduce financial losses .
- Data quality and recency are critical for accurate predictions .

Future Work:

- Incorporate real-time data and macroeconomic indicators .
- Explore ensemble models for further accuracy improvements .

Thank You !!