

Company Bankruptcy Prediction

ECE 143 Group 14

Members: Kendra Chen, Zhuosheng Song, Matthew Alegrado, Jinglin Cao

Motivation and Objective

- **Motivation:** Corporate bankruptcy can cause significant financial and economic disruption. Early detection helps investors, businesses, and policymakers take proactive measures to reduce risk.
- **Objective:** Conduct a thorough data analysis of bankruptcy data in the US and Taiwan, and provide interpretable thresholds for actionable insights.
- **Goal:** Enable informed decision-making for businesses, investors, and regulators by highlighting key risk factors and early warning signals.

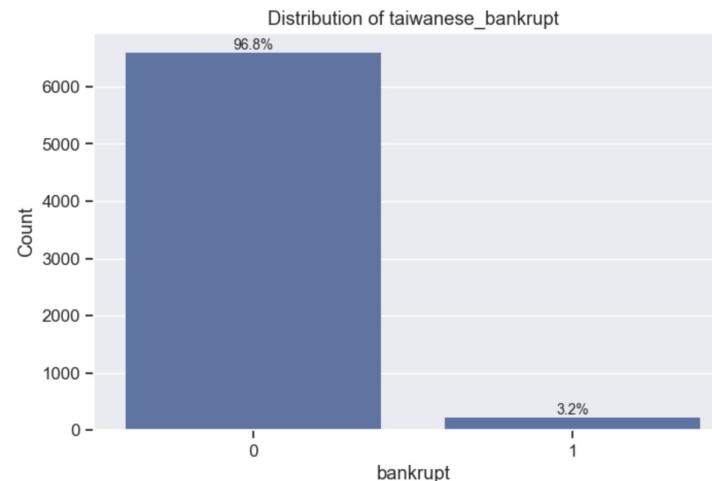
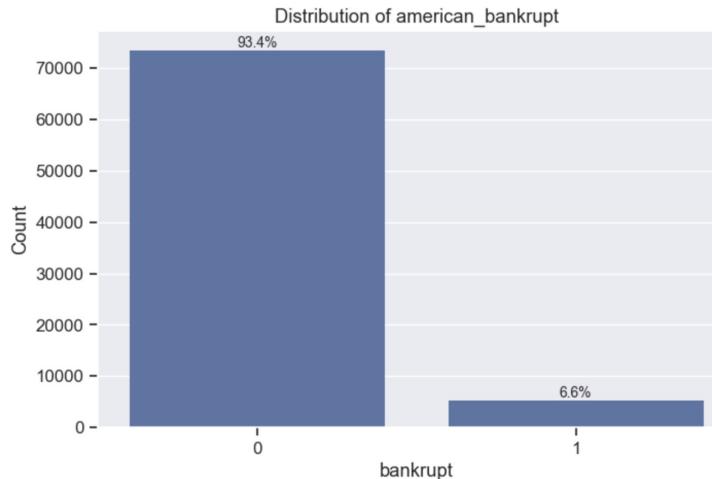
Methodology

Financial indicators are often highly correlated, adding noise and reducing prediction reliability; we address this with a robust pipeline that removes redundancy and provides interpretable early-warning insights.

- **Multicollinearity Removal:** Use VIF pruning to remove highly correlated features.
- **Robust Modeling:** Handle class imbalance and use train-test splits.
- **Interpretation & Monitoring:** Visualize key features with SHAP and extract simple, actionable rules via a decision tree.

Data Overview

- 1999-2009 Bankruptcy
Data from the Taiwan
Economic Journal
- 1999-2018 US
Bankruptcy Prediction
Dataset
- 3-6% bankrupt: **highly
imbalanced classes**



Initial Data Processing

Processing Steps

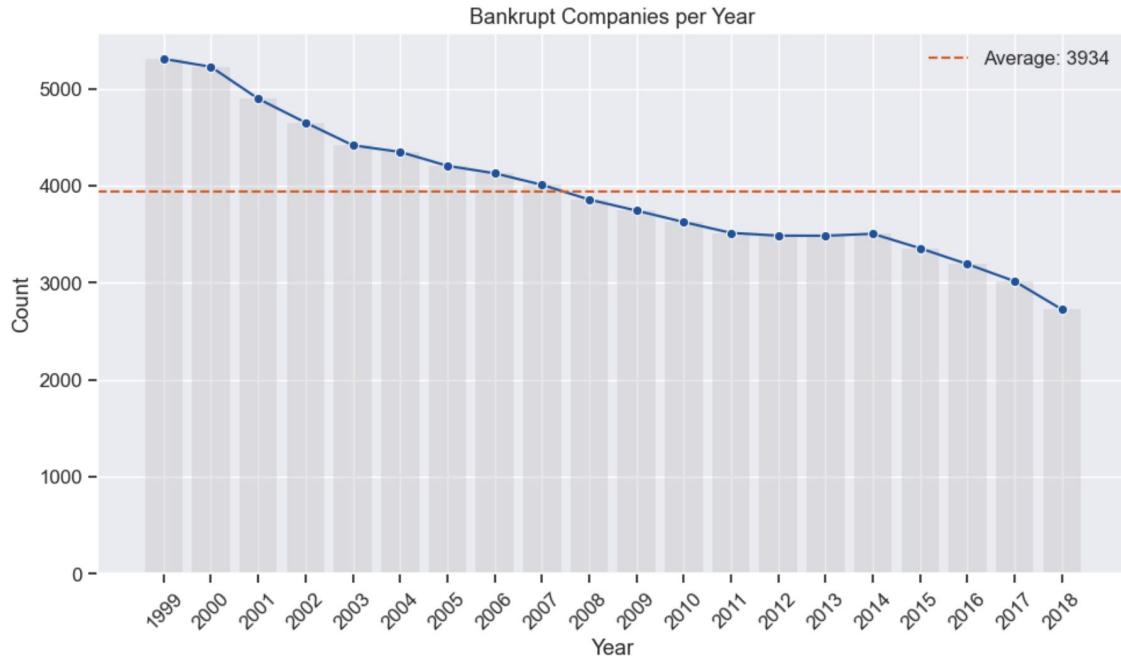
- Fix column names & convert to numbers
- Remove duplicates & zero-variance columns
- Drop messy or extreme-value columns

	Bankrupt?	ROA(C) before interest and depreciation before interest	ROA(A) before interest and % after tax	ROA(B) before interest and depreciation after tax	Operating Gross Margin	Realized Sales Gross Margin	Operating Profit Rate	Pre-tax net Interest Rate	After-tax net Interest Rate
0	1	0.370594	0.424389	0.405750	0.601457	0.601457	0.998969	0.796887	0.808809
1	1	0.464291	0.538214	0.516730	0.610235	0.610235	0.998946	0.797380	0.809301
2	1	0.426071	0.499019	0.472295	0.601450	0.601364	0.998857	0.796403	0.808388
3	1	0.399844	0.451265	0.457733	0.583541	0.583541	0.998700	0.796967	0.808966
4	1	0.465022	0.538432	0.522298	0.598783	0.598783	0.998973	0.797366	0.809304
5	1	0.388680	0.415177	0.419134	0.590171	0.590251	0.998758	0.796903	0.808771
6	0	0.390923	0.445704	0.436158	0.619950	0.619950	0.998993	0.797012	0.808960
7	0	0.508361	0.570922	0.559077	0.601738	0.601717	0.999009	0.797449	0.809362

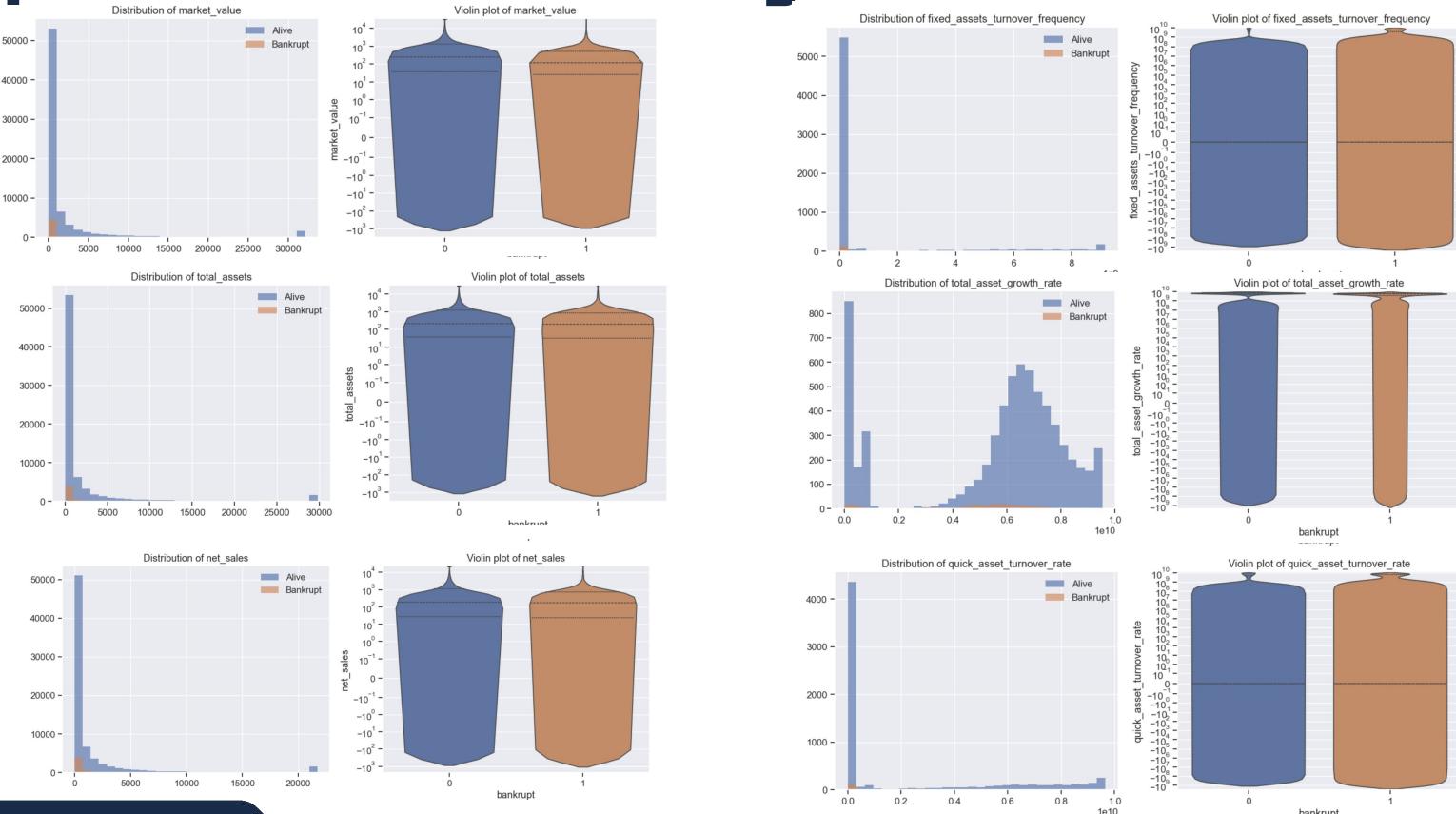
8 rows × 96 columns

Exploratory Data Analysis

- Is there variation in the number of **bankrupt companies per year**?
- American company bankruptcies over the time period **1999-2018**

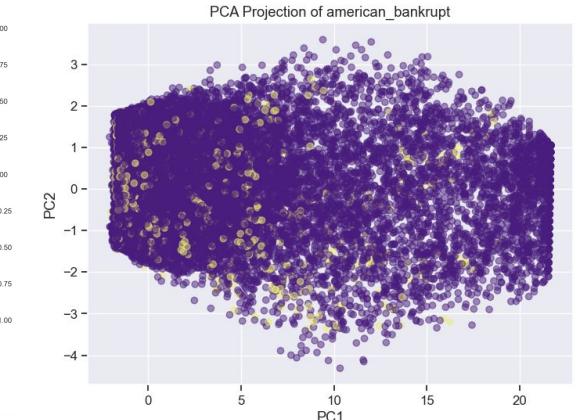
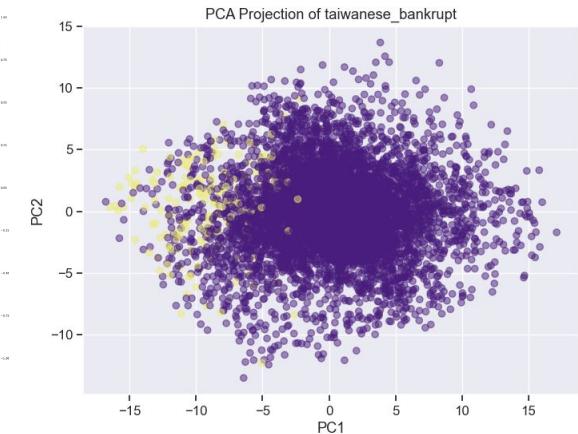
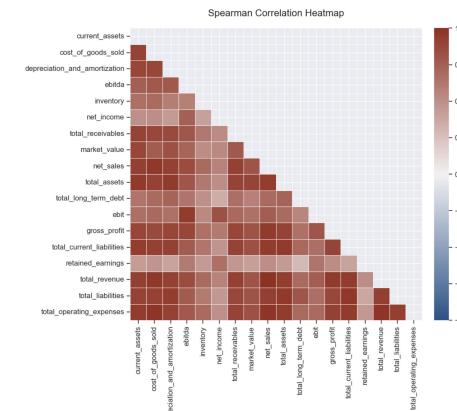


Top 3 Features With the Largest Differences

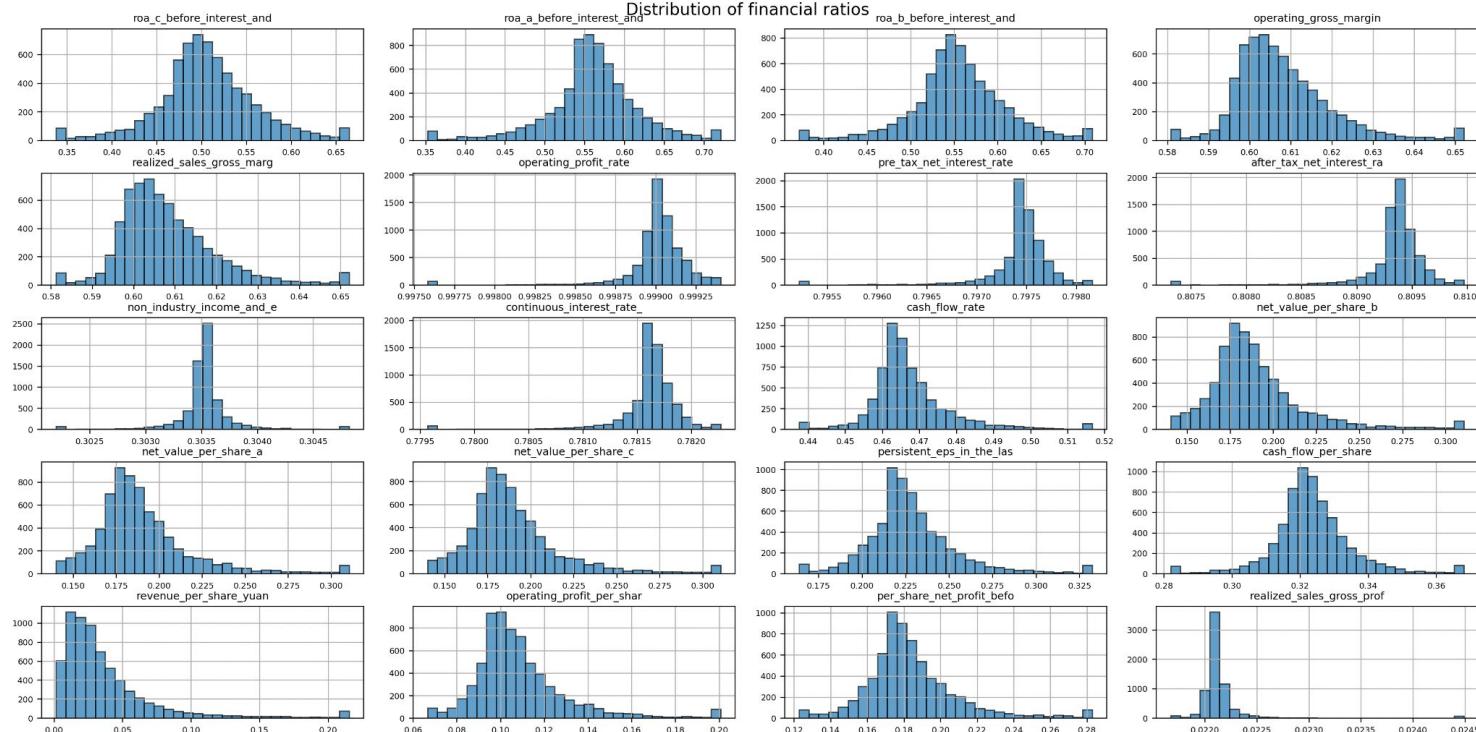


Exploratory Data Analysis

- Spearman heatmaps shows **strong multicollinearity**.
- Large red/blue blocks highlight **redundant features**.
- Dense PCA clusters show **high multicollinearity and poor separation** of bankrupt vs. healthy firms

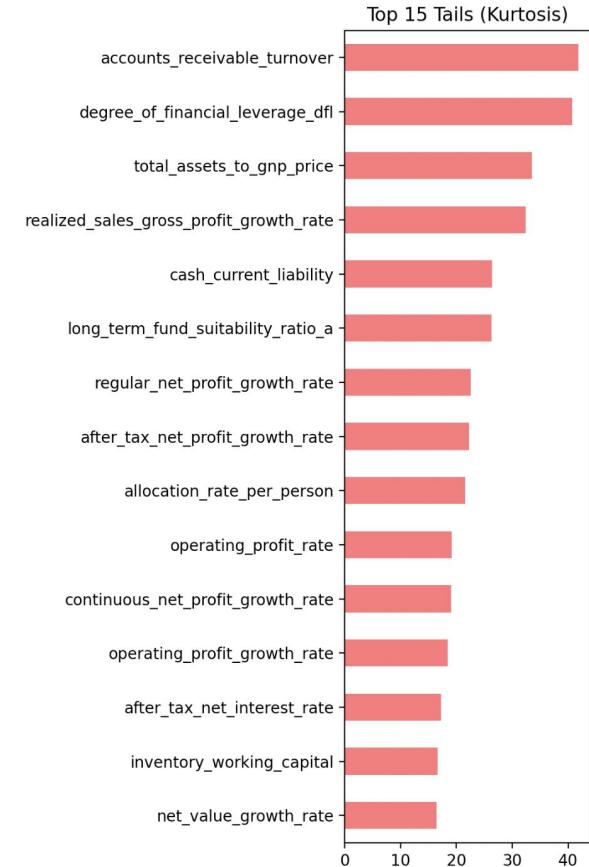
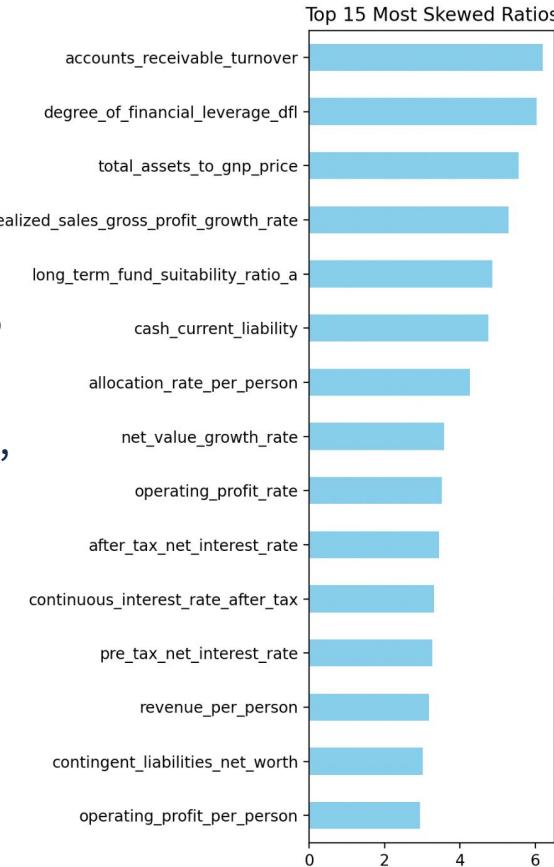


Strong right skew and heavy tails... extreme performers and potential outliers



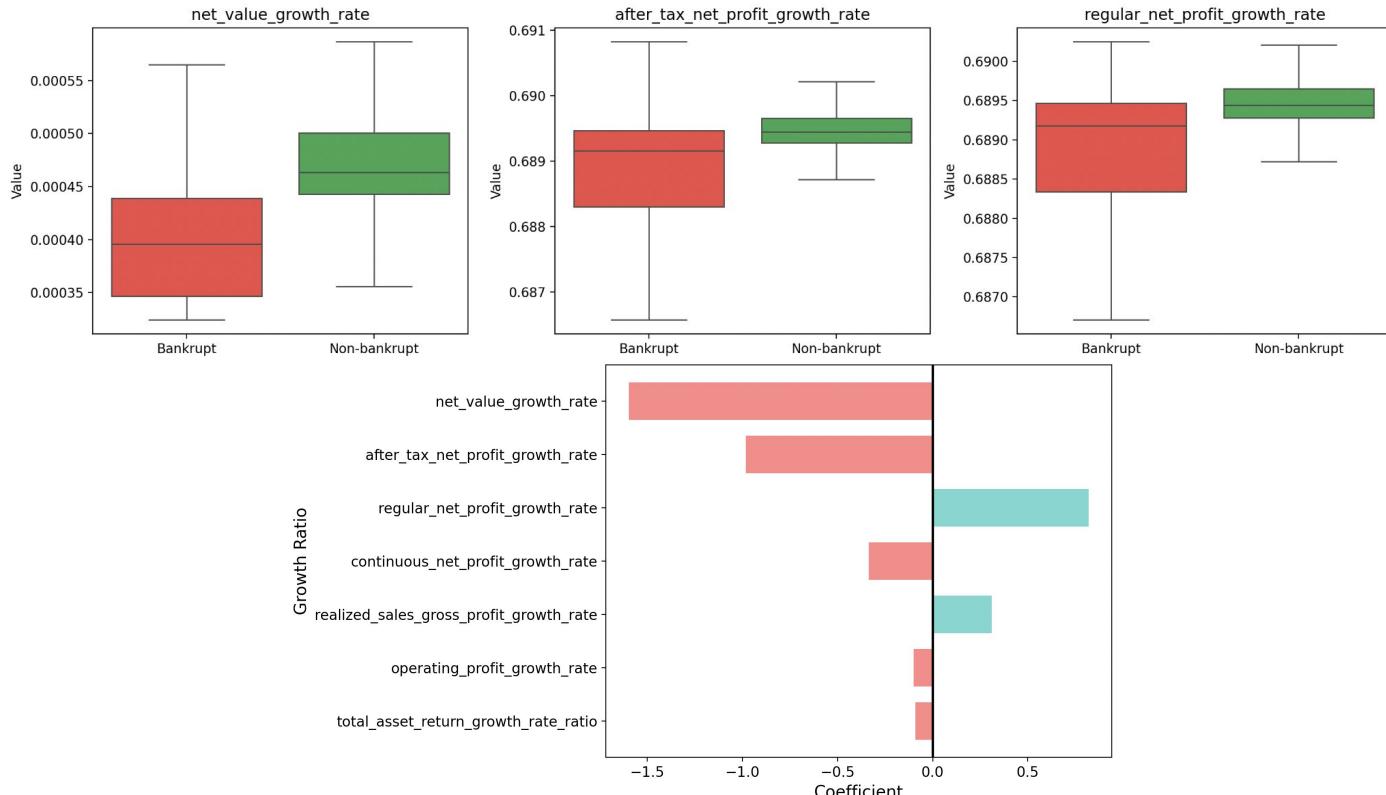
Skewness and tail distribution

- Skewed features may benefit from **log-transforming, winsorizing, or binning**
- Kurtosis identifies **volatile, risky, or noisy features**.
- Transform skewed features, normalize **tail distributions**.

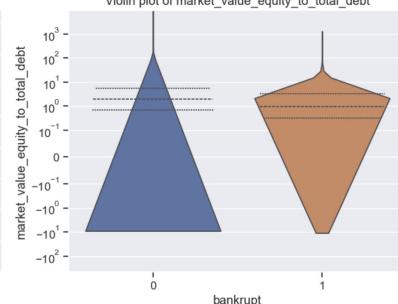
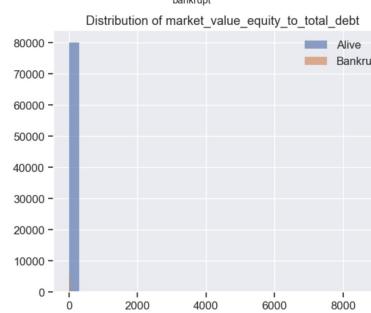
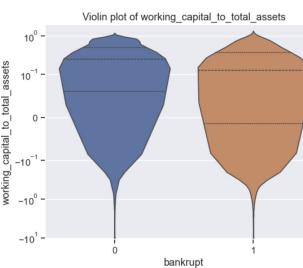
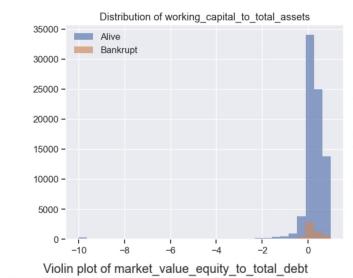
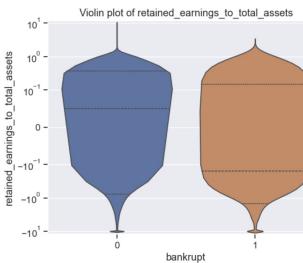
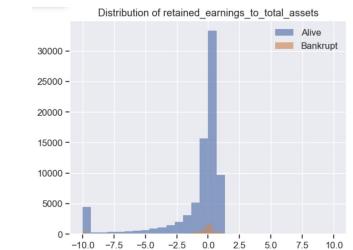
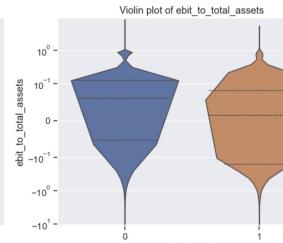
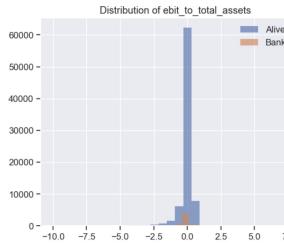
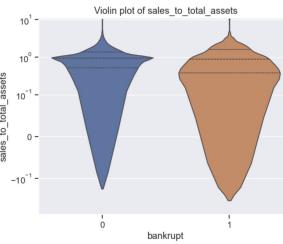
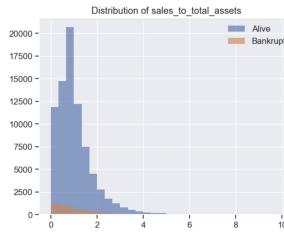


Do early-warning patterns appear in growth ratios for firms that go bankrupt?

- Net value growth rate:** Strongly protects against bankruptcy.
- After tax net profit:** Higher profit lowers bankruptcy odds.
- Regular net profit growth:** May reflect fragile or inconsistent performance.



Altman's 5 Key Financial Ratios...



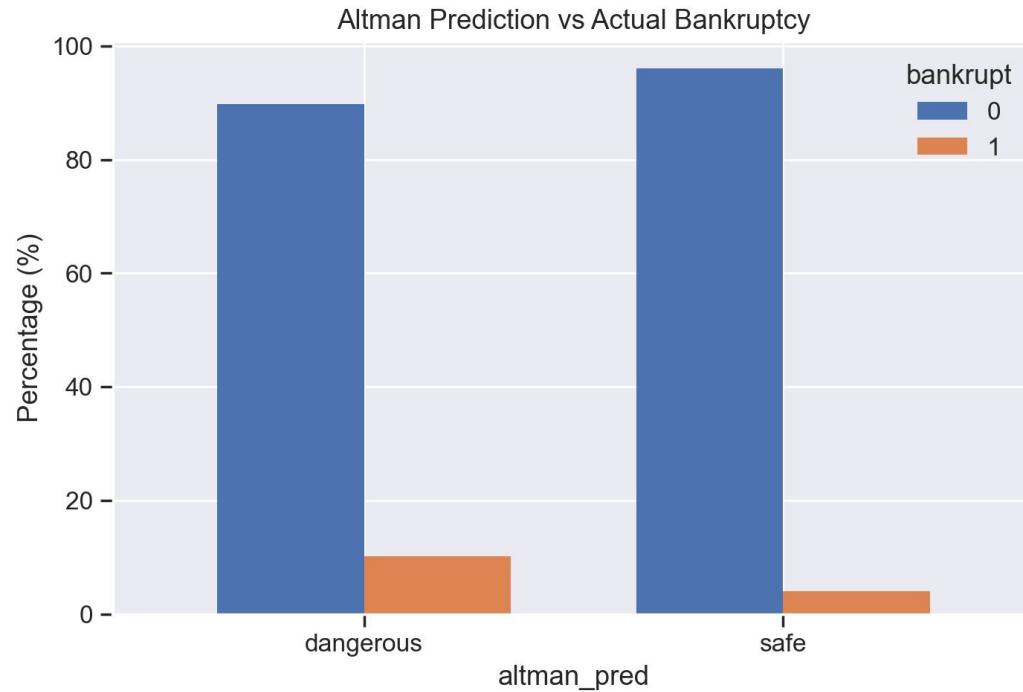
Applying Altman's Z-score model

From Altman's formula

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + X_5$$

$Z > 2.99$ company safe

$Z < 1.81$ company dangerous

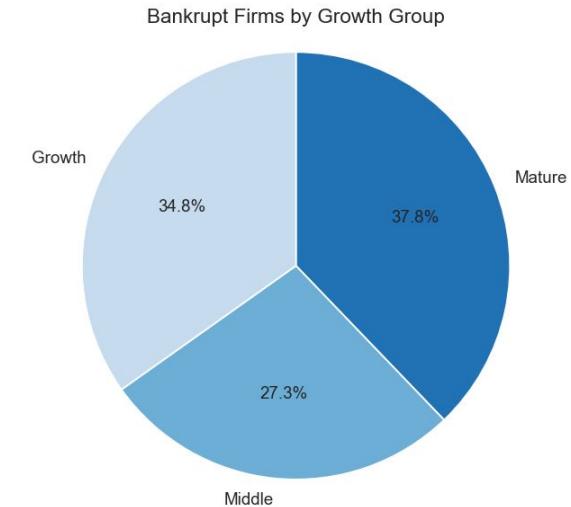
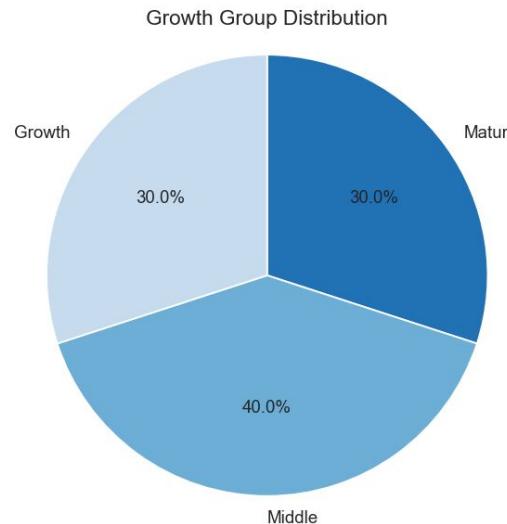


Is bankruptcy more common among high growth firms?

Growth firms expand aggressively and take on higher financial risk; mature firms tend to be more stable.

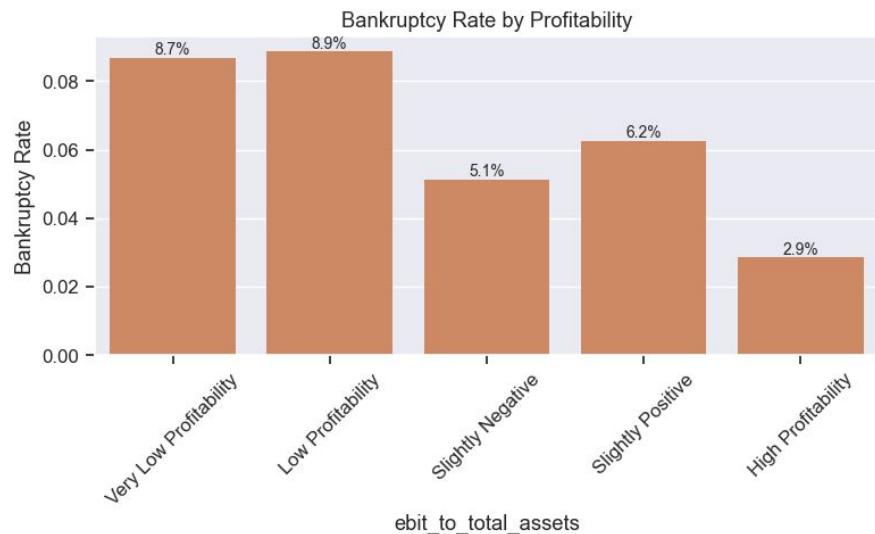
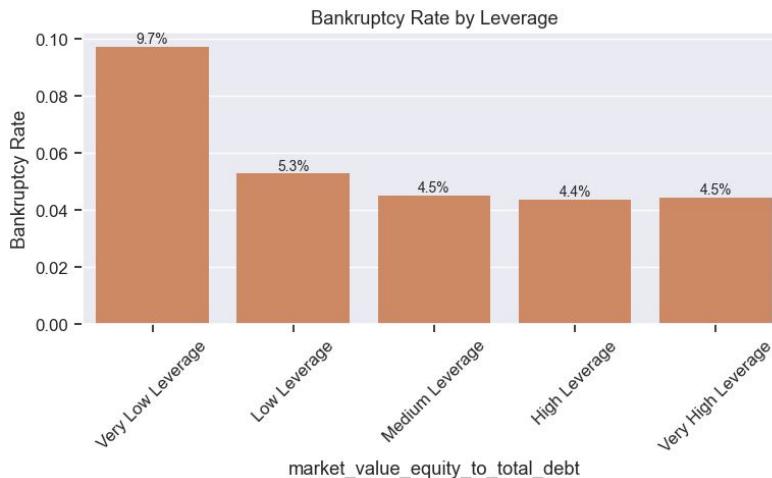
Growth is measured using the **Sales-to-Total-Assets** ratio.

Result: Bankruptcy rates are similar across growth and mature firms — an unexpected finding.

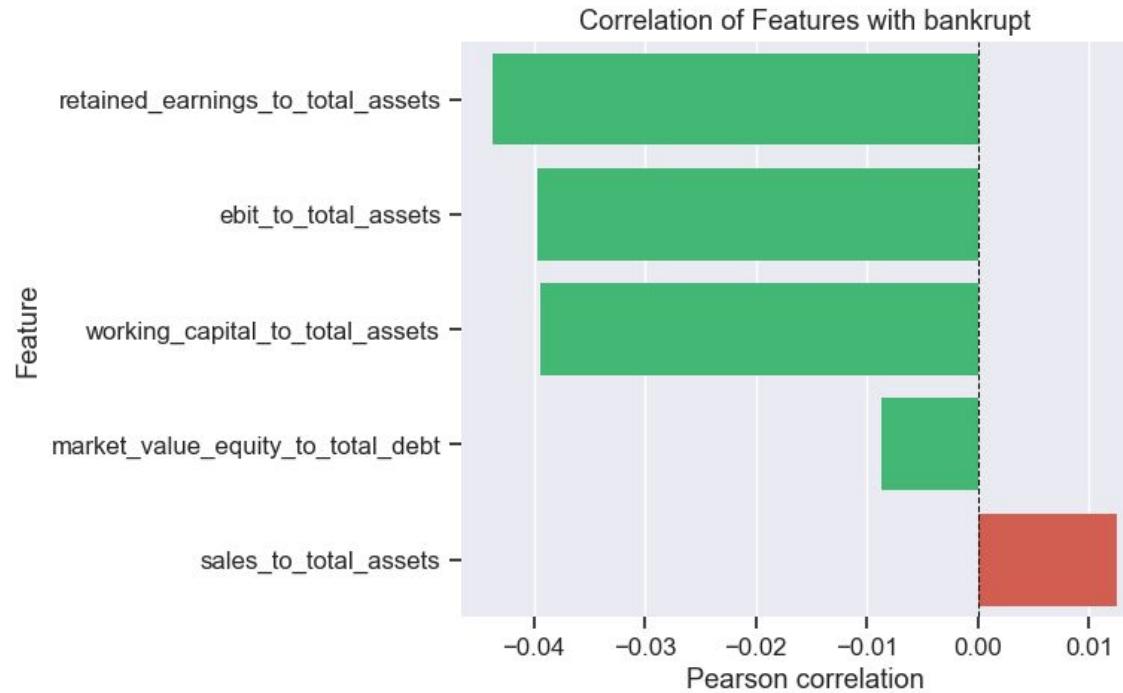


Effect of leverage and profitability?

Impact of **leverage level** & **profitability level**



Which features best predict bankruptcy?



How can we model bankruptcy risk?

Classification Model

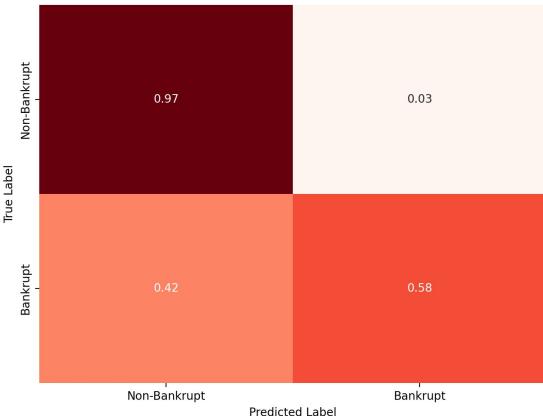
We select commonly used financial ratios from the literature for the baseline model

- Current & Quick Ratios
- Working Capital / Total Assets
- Retained Earnings / Total Assets
- EBIT / Total Assets
- Market Value of Equity / Total Liabilities
- Sales / Total Assets
- Cash Flow / Liability & Assets
- Net Income / Total Assets
- Liability / Equity

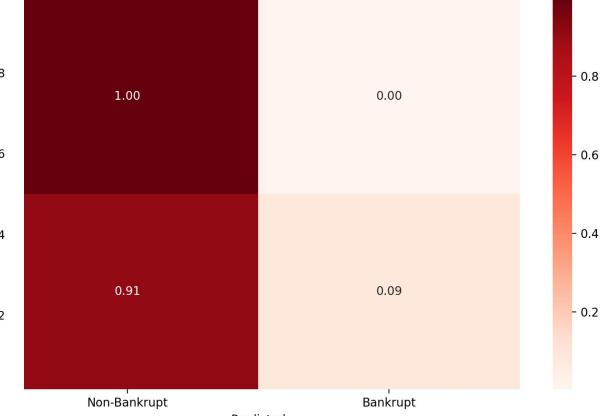
Improved model:

- Safe/Risky + Domain Features
- VIF pruning
- SMOTE oversampling
- Threshold tuning

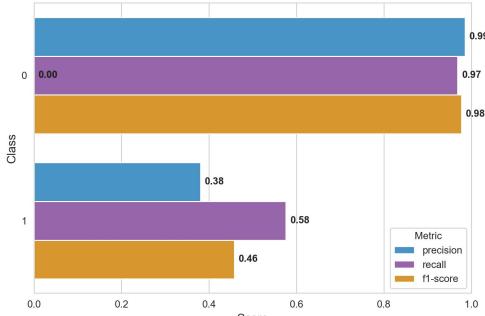
Improved Model



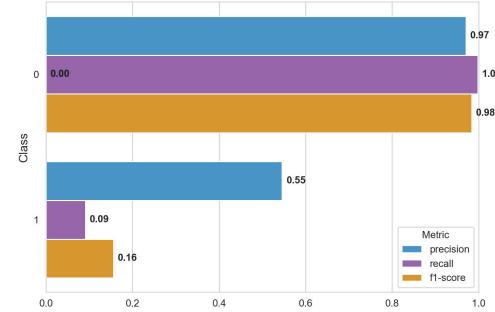
Baseline Model



Classification Metrics for Classes 0 and 1

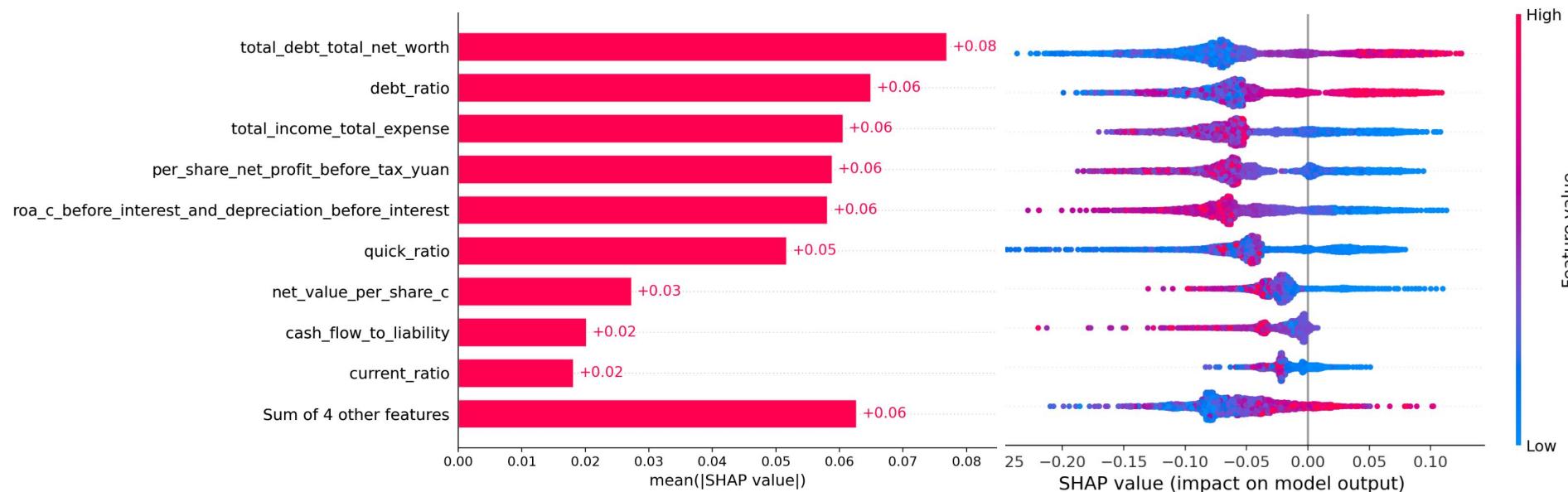


Classification Metrics for Classes 0 and 1



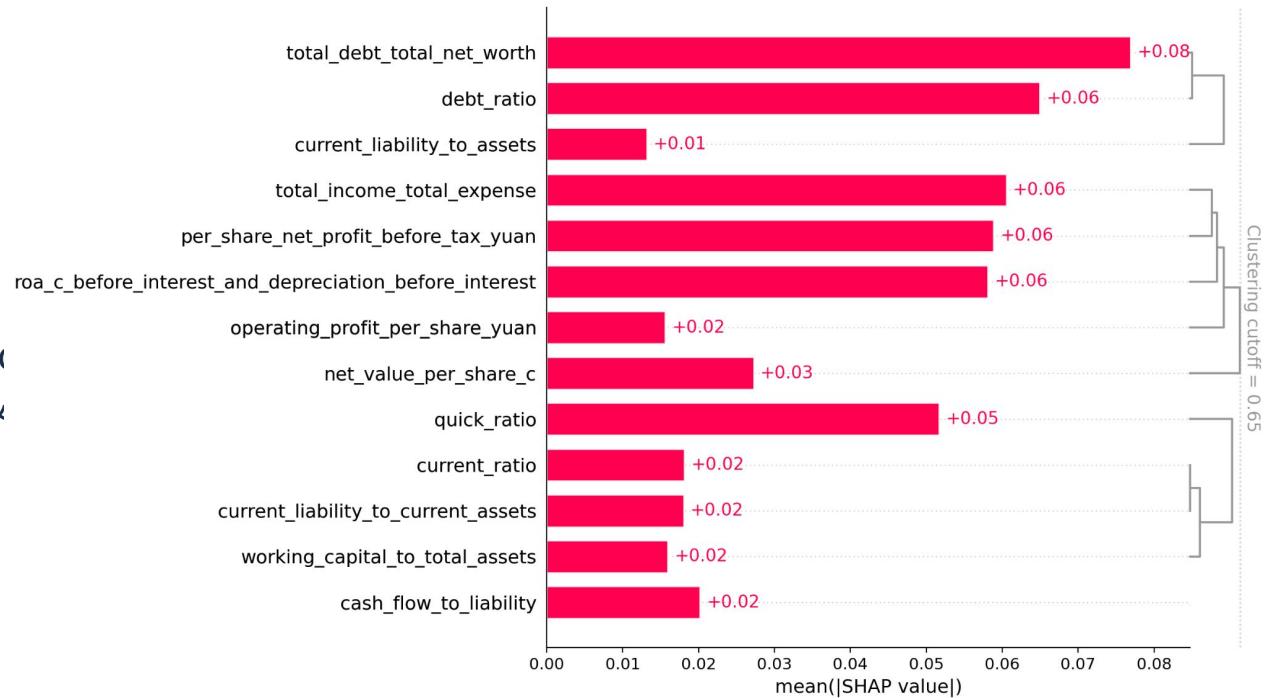
SHAP Analysis

- Compared feature Shapley values, which notes the **predictive impact** each feature has on current model.
- Left: Mean absolute value of SHAP values
- Right: SHAP value given feature value



SHAP Analysis

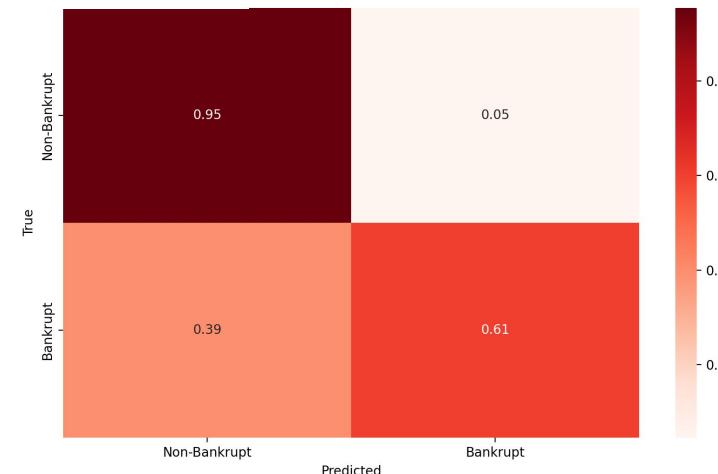
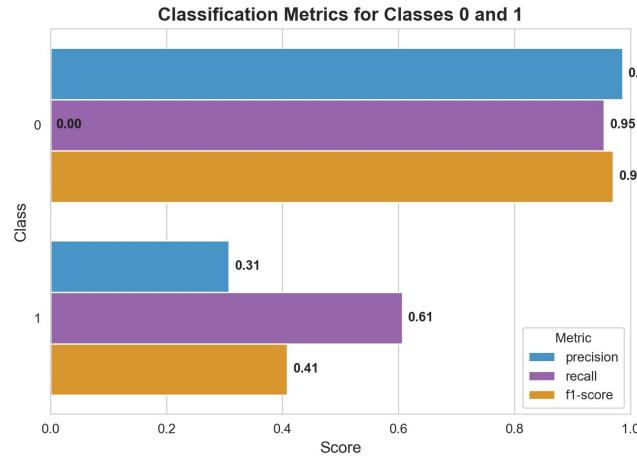
- We also cluster uncorrelated features and calculate mean SHAP scores by cluster.
- We can use a further refined feature set based on the most impactful & uncorrelated features.



SHAP Model

Using SHAP Analysis, we further reduce the features to following set:

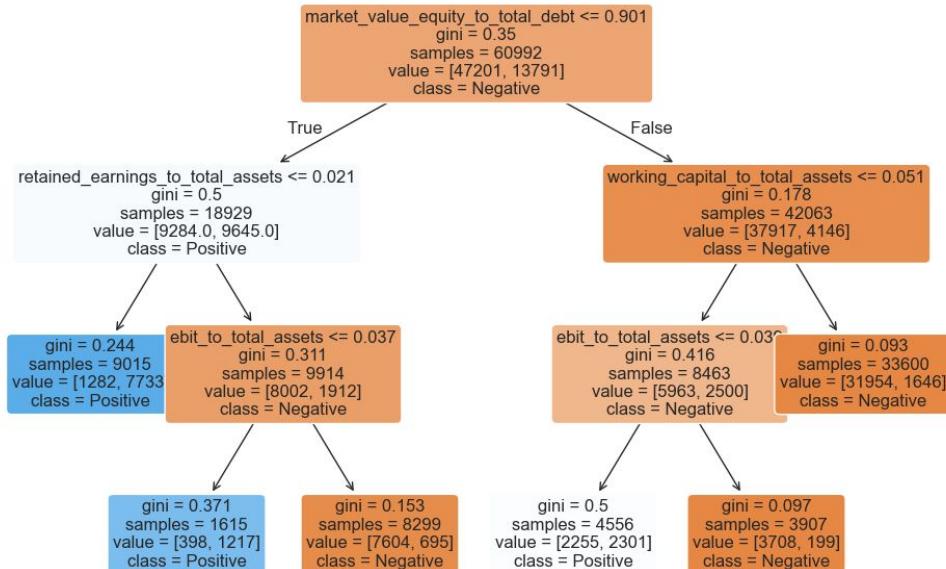
- Debt Ratio
- EPS (Earnings Per Share)
- Quick Ratio
- ROA (Return of Asset)
- Total Debt - Total Net Worth
- Total Income - Total Expense



Monitoring Threshold

Decision Tree Thresholds for Bankruptcy

- Decision Tree Rules
 - `market_value_equity_to_total_debt <= 0.901`
 - `ebit_to_total_assets < 0.037`



Conclusion

Our comprehensive early-warning system

- **Key Drivers:** Low profitability, low liquidity, high leverage
- **Flagging Rule:** Company flagged if **model predicts bankruptcy** or
 - $\text{market_value_equity_to_total_debt} \leq 0.901$
 - $\text{ebit_to_total_assets} < 0.037$
- Insight: Combining model predictions with simple thresholds provides **interpretable, actionable early-warning signals**

Thank you!