George Smith

Final Project Report

# Introduction

The purpose of this project is to analyze the financial metrics of previously bankrupt and current S&P 500 companies. As part of this analysis, key financial metrics that influence if a company is likely to go bankrupt will be identified and analyzed. Additionally, financial metrics that do not influence bankruptcy will also be identified.

## Data

Two datasets were utilized to conduct the bankruptcy analysis:

1. Taiwan Economic Journal Bankruptcy data
   1. This dataset included 95 financial metrics for 6,820 companies.
   2. A Boolean indicator was included in this dataset that represents if a company has gone bankrupt.
2. Web scraped financial metrics for every ticker represented in the S&P 500 index.
   1. This dataset was scraped from <https://finviz.com>.
   2. This dataset included 18 financial metrics for the 505 tickers in the S&P 500 index.

## Code

The code utilizes the Pandas, Numpy, Beautiful Soup, and Urlib.Request packages to scrape financial metrics for all the S&P 500 tickers.

Image 1 depicts the necessary imports



As the financial metrics for each company were found on different pages of the Finviz website, it was necessary to concatenate 25 different data frames into 1 data frame, so I could have all 505 tickers represented in a single dataset for analysis. Creating this data frame was very lengthy process but was necessary for comparison purposes. This data also required a bit of manipulation to be useable. All % signs needed to be removed, and values were transformed to be decimals. I then converted every column to be a number rather than an object so that I could conduct descriptive statistics on each of the financial metrics.

Image 2 depicts a subset of concatenation process to make 1 complete data frame

Graphical user interface, text, application, email

Description automatically generated

The data provided by the Taiwan Economic Journal although large, was relatively easy to work with. It required less data cleaning than the S&P 500 scraped data. As the dataset included a Boolean indicator to represent if a company has gone bankrupt, this dataset was perfect for a machine learning model. It was also very easy to subset this data to include only bankrupt companies.

Image 3 depicts subsetting the Taiwan data to include only bankrupt companies

Text

Description automatically generated

A correlation heatmap was created to view each financial metrics relationship to one another.

Image 4 represents a correlation heatmap for the financial metrics found in the bankruptcy data

A picture containing diagram

Description automatically generated

Sklearn was utilized to conduct both a logistic regression as well random forest machine learning models.

Image 5 depicts the key packages utilized to conduct the machine learning models

Graphical user interface, text, application

Description automatically generated

The machine learning models proved to be successful, as I was able to obtain an accuracy of 96% with both models.

Image 6 depicts a confusion matrix for the logistic regression model

Chart

Description automatically generated

Image 7 depicts the classification report for the Random Forest model

Table

Description automatically generated

## Data Questions

The discussed code was utilized to answer the following questions:

1. Are there any companies in the S&P 500 with financial similarities to bankrupt companies?
2. What Financial indicators heavily influence if a company is likely to face bankruptcy?
3. What Financial indicators have no influence on the likelihood of a company facing bankruptcy?

**Question 1:**

In order to answer if there are any companies in the S&P 500 with financial similarities to bankrupt companies the following financial metrics were considered:

* ROA
* Current Ratio
* Quick Ratio
* Debt / Equity
* Gross Margin

Additionally, the Taiwan Economic Journal Bankruptcy data was subset to include only instances of companies that have gone bankrupt. The mean value of each financial metric was calculated and compared to current S&P 500 companies.

**ROA**

The average ROA of Bankrupt companies was found to be 50%. EBAY was on the only company in the S&P 500 with a similar ROA of 51.5%.

Image 8 depicts the output of the mean ROA of bankrupt companies



**Current Ratio**

The average Current Ratio of bankrupt companies was 416,729. No companies in the S&P 500 were found to have a current ratio near this value, as the maximum was a value of 14 by Aadi Bioscience. An outlier value of 2,750,000,000 was found in the bankruptcy data which far exceeds any current ratios of S&P 500 companies. Due to the discrepancy in values, it was difficult to draw any conclusions about current ratios effect on bankrupt companies.

Image 9 depicts the output of mean current ratio on bankrupt companies

Text

Description automatically generated with medium confidence

**Quick Ratio**

The average quick ratio of bankrupt companies was 7,257,160. Like Current ratio this value far exceeded the maximum quick ratio value of S&P 500 companies, as the largest value again was a value of 14 by Aadi Bioscience.

Image 10 depicts the output of the mean quick ratio of bankrupt companies

Table

Description automatically generated with low confidence

**Debt/ Equity**

The average debt to equity ratio of bankrupt companies was 28%. 10 companies in the S&P 500 had similar debt/equity ratios within .01 of bankrupt companies. These companies include Chubb Ltd, Coterra Energy Inc, Harford Financial Services Group Inc., Metlife Inc, Mohawk Industries, Inc, New Corp Class A, Paychex Inc, Principal Financial Group, Progressive Corp, Exxon Mobil Corp. Due to the number of companies with similar debt/equity ratios it doesn’t appear that debt/equity ratio plays a crucial role in determining if a company is likely to go bankrupt.

Image 11 depicts the output of the mean debt/equity ratio of bankrupt companies

Text

Description automatically generated with low confidence

**Gross Margin**

The average gross margin of bankrupt companies was 60%. 13 companies in the S&P 500 had similar operating gross margin values within 1% of the average gross margin of bankrupt companies. These companies include Brown-Forman Corporation, Colgate-Palmolive Company, Equifax, Hasbro, Coca-Cola Co, Mid-American Apartment Communities, Healthpeak Properties, Teradyne Inc., Take-Two Interactive Software, WEC Energy Group Inc. Due to the number of companies with similar gross margin values it doesn’t appear that gross margin plays a crucial role in determining if a company is likely to go bankrupt.

Image 12 depicts the output of the mean debt/equity ratio of bankrupt companies

**Text

Description automatically generated with low confidence**

After analyzing each of these financial metrics for bankrupt and S&P 500 companies it appears difficult to leverage only financial metrics in order to predict if a company will go bankrupt. I believe news related to fraud or regulatory fines would be a much better indicator to determine if a company is likely face bankruptcy.

**Question 2:**

In order to determine the financial indicators that influence if a company is likely to face bankruptcy two machine learning models were leveraged. These models include a logistic regression model as well as a random forest model. All 94 financial metrics were considered in these models and the features were ranked in order of importance. The top 10 features include net income to stockholder’s equity, borrowing dependency, net profit before tax/paid-in capital, net income to total assets, degree of financial leverage persistent EPS in the last four seasons, net worth/assets, interest expense ratio, net value per share, non-industry income and expenditure/ revenue. Each of these financial metrics has an importance value between 2% and 3%. Based on these results it does not appear that there is one single financial metric that can be leveraged to determine if a company is likely to go bankrupt.

Image 13 depicts the output of feature importance’s related to predicting if a company is likely to go bankrupt

A screenshot of a computer

Description automatically generated with medium confidence

**Question 3:**

In order to determine the financial indicators that have no influence on if company is likely to face bankruptcy the same machine learning models were leveraged. Feature importance values of 0 indicate financial metrics that have no influence on if a company is likely to face bankruptcy. As a 1result, Revenue per person, revenue per share, quick assets/ current liability, net value growth rate, net income flag, average collection days, and allocation rate person were found to have no usefulness in predicting bankruptcy.

Image 14 depicts the output of feature importance’s with values of 0

Text

Description automatically generated