Syracuse University

Memo

|  |  |
| --- | --- |
| To: | Dr. Landowski |
| From: | George Smith |
| Date: | February 18, 2022 |
| Re: | Project Proposal |

**Topic: Analyze public workstation usage**

**Data Description:**

I plan on using two different data sets for this analysis. The first data set is provided by the Taiwan Economic Journal. This data set has distinct financial information and notes if a company has gone bankrupt or not with a value of 1 or 0. The second data set I will be using will be web scraped from yahoo finance. I will use similar financial information found in the first data set to make predictions whether a company is likely to go bankrupt or not.

|  |  |  |
| --- | --- | --- |
| **Fields** | **Description** | **Example** |
| Bankrupt | Boolean value that determines if a company went bankrupt | 0 , 1 |
| ROA before interest and depreciation before interest | Value taken off financial statements expected to be a decimal value | .37 |
| ROA before interest and % after tax | Value taken off financial statements expected to be a decimal value | .42 |
| ROA before interest and depreciation after tax | Value taken off financial statements expected to be a decimal value | .40 |
| Operating Gross Margin | Value taken off financial statements expected to be a decimal value | .60 |
| Realized Sales Gross Margin | Value taken off financial statements expected to be a decimal value | .60 |
| Operating Profit Rate | Value taken off financial statements expected to be a decimal value | .99 |
| Pre-tax net Interest Rate | Value taken off financial statements expected to be a decimal value | .79 |
| After-tax net Interest Rate | Value taken off financial statements expected to be a decimal value | .80 |
| Non-Industry income and expenditure/revenue | Value taken off financial statements expected to be a decimal value | .30 |
| Continuous interest rate (after tax) | Value taken off financial statements expected to be a decimal value | .78 |
| Operating Expense Rate | Value taken off financial statements expected to be a decimal value | .00012 |

**Research Questions:**

* What S & P 500 companies are likely to experience bankruptcy?
* What S & P 500 companies are not likely to experience bankruptcy?
* What are the key financial indicators that determine if a company is likely to go bankrupt?

**Data Preparation Plan**

1. Load initial data set into Python
2. Scrape data from yahoo finance to create a list of companies that will used for bankruptcy predications
3. Explore the data for anomalies, missing data, etc.
4. Resolve any issues related to the previous step (e.g., filling NA values).
5. Convert the data to appropriate data types.
6. Develop appropriate projection models