Textual Analysis and Financial Statements

Isaac Liu

March 25, 2024

Introduction

high-level subject area info

(Das et al., 2023)

problem statement and question

Company ratings and creditworthiness are important information for investors - not just institutional investors and financially sophisticated bondholders, but also stockholders, who may be wiped out completely in the event of bankruptcy.

Are ratings based on hard numbers, or do company outlooks and sentiment also matter? Are they predictable?

note credit rating data access is limited and our model can be used to interpolate

In this project, we seek to explore whether incorporating the text of earnings calls improves predictions of corporate credit ratings.

high level data description

Though much literature has focused on financial statements and reports and credit ratings (CITATIONS HERE), our paper takes a relatively underexplored approach, instead incorporating earnings call transcripts. We believe calls offer a richer picture of a firm's financial prospects because they include two-way conversation between company management and financial analysts in form of a Q and A section. This section incorporates the broader beliefs and concerns of the financial community into our predictions. Prior work has recognized the 'wisdom of the crowd' in making predictions, and our approach fully accounts for it, going beyond the proclamations of a company's management.

Calls also require less parsing - no looking for specific section of financial statement reports.

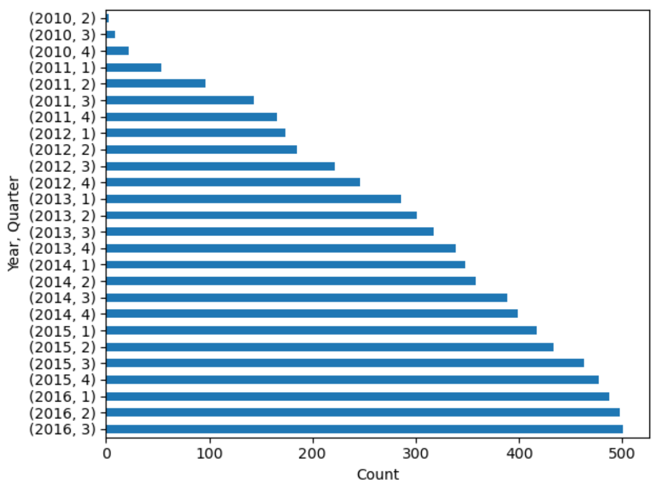
roadmap we then

Data and Exploratory Data Analysis

We combine a wide variety of data sources to support our predictions of credit ratings - combining rating data with company earnings calls, financial statement variables, and industry sector. In our final dataset, each observation represents a fixed quarter date (1/1, 4/1, 7/1, 10/1) for a company, with the company's most recent credit rating, earnings call and associated financial statement variables, and sector attached.

Our scope of interest is publicly traded companies from 2010-2016 (a limitation due to the availability of credit rating data). The data is temporally unbalanced, with many companies entering the dataset in later years after they first receive an observable credit rating (Figure 1).

Figure 1: Observations by Quarter and Year



In all, we have 7333 quarters for 536 unique companies.

Credit Ratings

We make use of long-term credit rating issuances from S and P Rating Services, provided from a combination of two credit rating datasets downloaded in CSV and Excel format from Kaggle (Gewerc, 2020; Makwana, Bhatt and Delwadia, 2022). Each issuance be a change in rating (upgrade, downgrade) or reaffirmation - they occur at ad-hoc intervals. We reshape these rating issuances to a dataset of ratings for each company on each fixed quarter date by creating a rating end date variable that is the date of the next issuance, and joining a list of the fixed quarter dates on the condition that the fixed quarter date is between the issuance date and the end date.

Figure 2 shows the distribution of rating grades used in our final dataset. Finer grades (+, -) are sometimes assigned by agencies, but these grades were removed for this project. Ratings of BBB and above are considered investment grade - these bonds carry empirical one-year default rates of 0 to 1%. Ratings below that are classified as junk, with default rates from 1 to 30, 40, or even 50% for some years (S and P Global Ratings, 2024). Most company-quarters have ratings around the BBB threshold, with very few cases on the extreme ends of the spectrum.

Earnings Calls

Our earnings call data comes from the Financial Modelling Prep API (**?**), a trusted source widely used in industry. We remove all calls that happened more than 250 days prior and after the year and quarter they are supposed to discuss the results from. Including both prepared remarks and analyst Q and A sessions, the overall average call length in our final data stands at 8688.25 words.

Financial Statements

API source

Figure 2: Credit Ratings

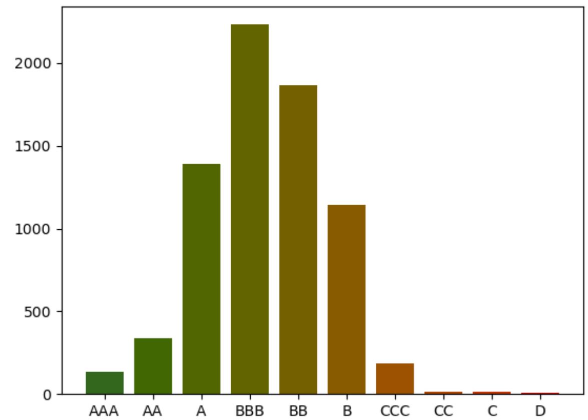
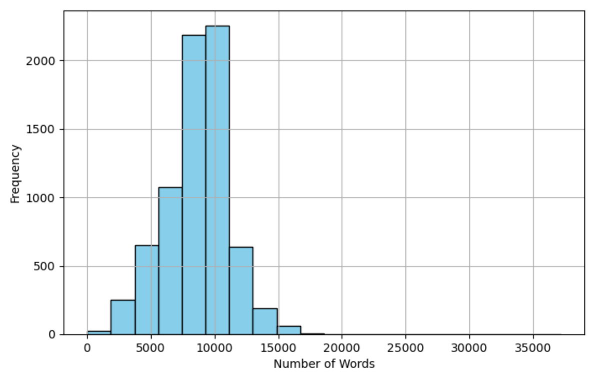


Figure 3: Number of Words in Earnings Calls



Items from balance sheet, cash flow statement, income statement, and company market capitalization

124 variables in total. Examples: revenue, total liabilities, net income, EBITDA

Limit to items reported in USD

Winsorizing: Check for items mis-multiplied by 1,000 in parsing - if last digits are “000.00” and item is above or below 2.5% and 97.5% quantile, divide by 1,000

Tests to ensure the value in income statement and balance sheet are consistent with each other.

Construct Altman Z-score

Sector

GCIS developed by S and P

Obtained from Kaggle with supplementary manual lookup

CSV and Excel format

sectoral imbalance

Figure 4: Altman Z-Score

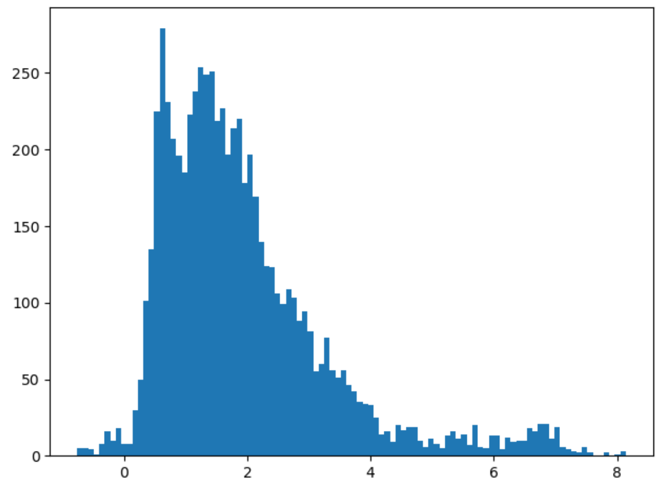
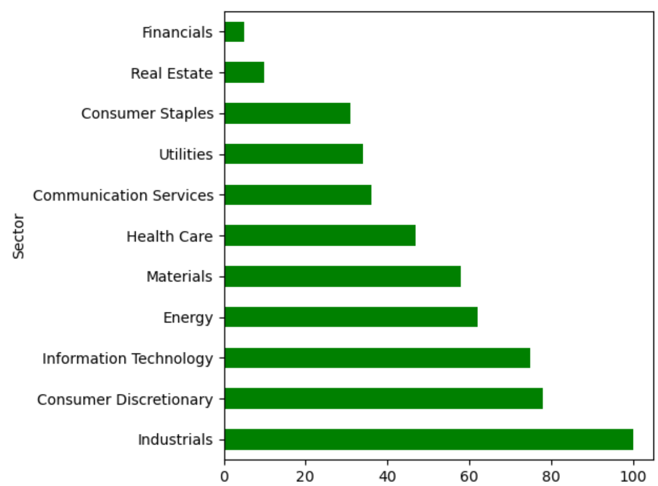


Figure 5: Firms by Sector



Quality Control

quality control code review of all data cleaning code numerous investigations

date gaps investigations

company dropout and the company that leaves

NLP Features

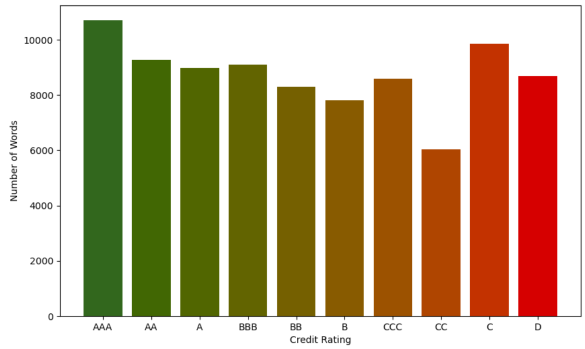
Average call length

outliers and errors

correlations and patterns

identification of good machine learning methods

Figure 6: Average Call Length by Credit Rating



Modelling

Our overall model architecture is of the form

$$\text{Predicted Credit Rating} = f(\text{Financial Statement Variables}, \text{Sector}, \text{NLP Features})$$

functions began with logistic regression

XXX logistic regression predictors

multinomial, balanced class weights, l1 penalty

table of predictions

fitting and output

assumptions

interpretation

Next Steps

Ensembling and Auto-ML

more classifiers

first steps using AutoML

a good starting point for diving deep on more algorithms

algorithms and accuracy from them

outputted feature importance

Graph Neural Network incorporating the relationships between companies, trained end-to-end with both tabular financial data and NLP features

Fine tune the pre-trained LLMs for NLP feature construction

References

Das, Sanjiv, Xin Huang, Soji Adeshina, Patrick Yang, and Leonardo Bachega. 2023. "Credit Risk Modeling with Graph Machine Learning." *INFORMS Journal on Data Science*, 2(2): 197–217. Publisher: INFORMS.

Gewerc, Alan. 2020. "Corporate Credit Rating with Financial Ratios."

Makwana, Ravi, Dhruvil Bhatt, and Kirtan Delwadia. 2022. "Corporate Credit Rating."

S and P Global Ratings. 2024. "S and P Global Ratings."