Textual Analysis and Financial Statements

Isaac Liu

March 24, 2024

Introduction

high-level subject area info

(Das et al., 2023)

problem statement and question

Can incorporating the text of earnings calls improve predictions of corporate credit ratings?

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

high level data description

roadmap we then

Data and Exploratory Data Analysis

sources

all data formats come as CSV though we use parquet files for efficient intermediate data storage throughout the project

Credit Ratings

Long-term credit rating issuances from S and P Rating Services, 2010-2016

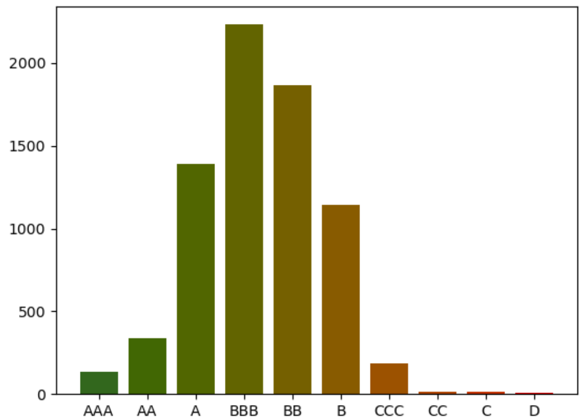
Combination of Kaggle datasets

Can be a change in rating (upgrade, downgrade) or reaffirmation

Finer grades (+, -) sometimes assigned but removed for this project

BBB and above is investment grade (one-year default 0 to 1%), below is junk (1 to 30, 40, 50%)

Figure 1: Credit Ratings



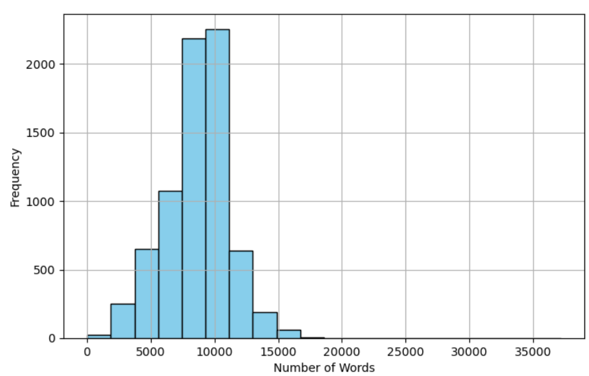
Earnings Calls

API source

Quarterly conference call transcripts that contains speaker remarks and Q and A session from 2010 - 2016.

Remove all calls that happened more than 250 days prior and after the fixed quarter date

Figure 2: Number of Words in Earnings Calls



Financial Statements

API source

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

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

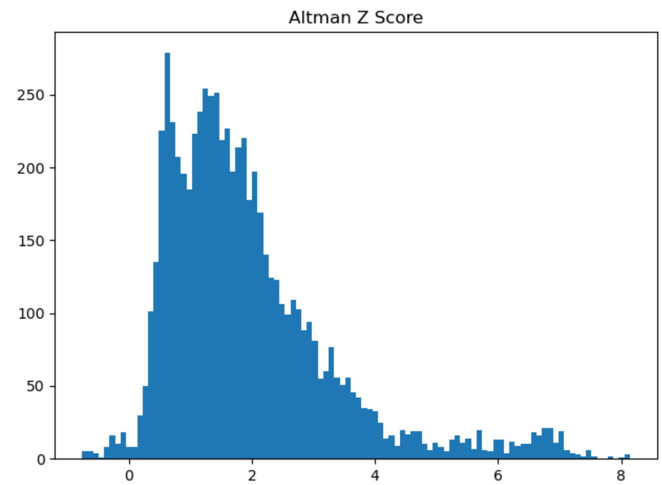
Limit to items reported in USD

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

Figure 3: Altman Z-Score

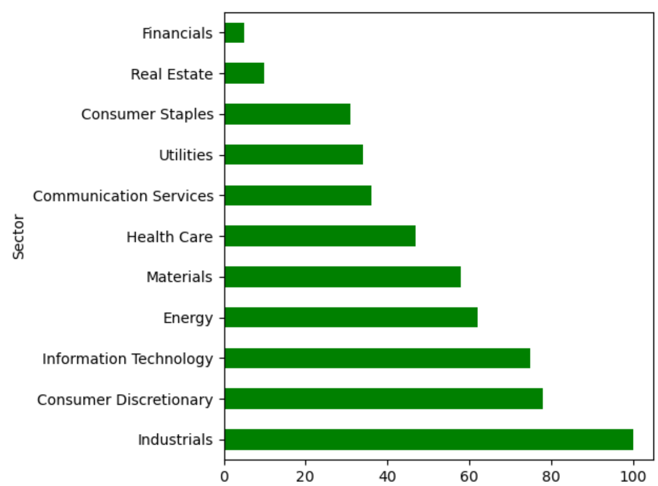


Sector

GCIS developed by S and P

Obtained from Kaggle with supplementary manual lookup

Figure 4: Firms by Sector



sectoral imbalance

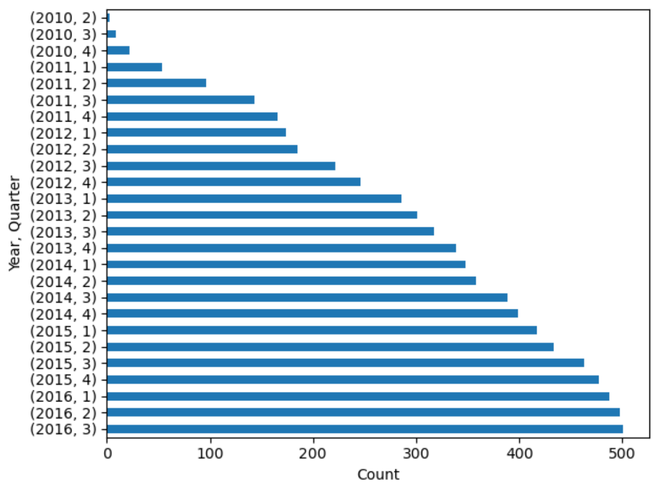
Merged Data

Data is at the level of

XXX quarters, XXX companies

temporal imbalance

Figure 5: Observations by Quarter and Year



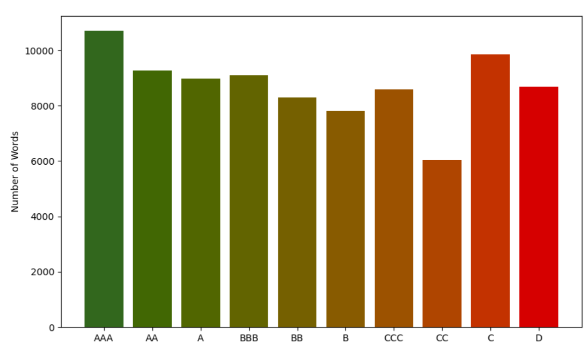
Quality Control

quality control code review of all data cleaning code numerous investigations

NLP Features

Average call length

Figure 6: Average Call Length by Credit Rating



outliers and errors

correlations and patterns

identification of good machine learning methods

Modelling

Our overall model architecture is of the form

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

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

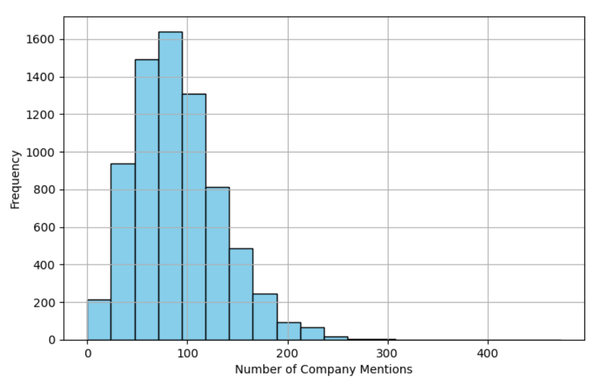
a good starting point for diving deep on more algorithms

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

Figure 7: Company Mentions



Fine tune the pre-trained LLMs for NLP feature construction

References

Das, S., X. Huang, S. Adeshina, P. Yang, and L. Bachega (2023, October). Credit Risk Modeling with Graph Machine Learning. *INFORMS Journal on Data Science 2*(2), 197–217. Publisher: INFORMS.

Appendix