Applications to finance

Today we will discuss 2 applications of NLP for finance:

1) Credit risk reporting
2) Interpreting central bank statements

Credit Risk Scoring & Sentiment Analysis

Based on the work "Identifying Corporate Credit Risk Sentiments from Financial News" (2022) by researchers at Moody's

Credit risk management traditionally follows either:

a) a structural model approach such as Black-Scholes to compute the probability of default based on modeling of assets & liabilities

b) a reduced form approach (also called default intensity models) which measures the default event as a statistical process without considering assets + liabilities

These historical approaches primarily focus on assessing the probability of default

They are <u>NOT</u> meant for gaining insights about a company's overall credit situation or identifying events the company has experienced (or might likely experience) which impact credit risk

This task is usually accomplished by subject matter experts
Often rely on news to make determinations
Extremely time consuming to accomplish

<u>Goal</u>: Automate this process to rapidly + reliably update credit risk using news data

"The paper focuses only on negative / credit adverse signals"

Approach

1) Topic classification (NOT topic modelling) to determine if a given document is relevant for credit risk

2) Target entity identification to determine the relevant companies

3) Sentiment analysis to determine if it is a positive or negative event (Recall: this paper only studies negative events)

4) Risk categorization to classify the document into different risk levels

5) Aggregation into a company-relevant credit risk score

1) Topic Classification

Goal: Train a binary classifier to determine if a document is relevant to credit risk

Idea: Reuters news comes with topic codes

Select codes that are relevant as class 1 (ex. Merger/Acquisition)

all other codes are given class 0 (ex: Sports)

↪ Train a linear support vector machine (or other method) on TF-IDF features

This method can be used on new unlabeled text to determine relevance

[Topic modeling is inappropriate because the ex ante interpretation is important

↪ Topic modeling may determine clusters which don't map so neatly to our goal]

2) Entity Identification

If the document is determined to be relevant by the text classifier, we want to determine which company is being discussed

↳ Hard problem but multiple approaches with large language models (named entity recognition)

3) Sentiment Analysis

Goal: Determine if the document is describing a positive or negative event

⌈This research used the Electra Base Model, but you can use your favorite NLP tool such as BERT or GPT⌋

Difficulty: Training data

↳ This paper had 5 subject matter experts annotate text

↳ If there was consensus then it was added to the training data

If no consensus then it was not used for training

Total of 9,859 sentences were annotated with consensus

↳ This is a labor intensive process

↳ As discussed previously in lecture, we could look at the bond market to automatically label a document based on interest rate changes

↳ Need to think about different maturities, etc...

4) Risk Categorization

In addition to the sentiment analysis, we want to classify the text as being related to:

- bankruptcy
- default
- credit downgrade
- profit warning
- compliance issue
- other

Note: all are only relevant if negative sentiment detected

Each type is given a fixed score

↳ Instead of a continuum as we can get in sentiment analysis

5) Aggregation

Use a weighted average (decreasing weight as news gets older) to determine a credit risk score

Interpreting Central Bank Communications

Based on the work "How You Say It Matters: Text Analysis of FOMC Statements Using Natural Language Processing" (2021) by researchers at the Federal Reserve

The Federal Reserve (and other central banks) frequently issue public reports + statements

These statements can be very informative, e.g., forward guidance about the future of policy rates or quantitative easing

Though sometimes quantitative information is included (ex: inflation data)
These statements include qualitative descriptions of economic conditions

Goal: Use NLP to measure how changes in qualitative descriptions of the economy affects bond prices
(Can be documents with or without quantitative information)

Assessing & Quantifying the Tone of FOMC Statements

Background

- FOMC statements can include both qualitative & quantitative information
- Because of "Fed Speak", manual tagging usually performs very poorly
- Since March 2004, the Federal Reserve prepares 3 or 4 alternate versions of FOMC statements with a rationale for each

↪ These drafts are released 5 years later, but can serve for training or

Counterfactual analysis

↳ This paper focuses on counterfactual studies

Approach

At a high-level, the authors take a similar approach to FinBERT

↳ implement a pre-trained NLP method to quantify text

↳ Universal Sentence Encoder (USE)

[As before, other more complex tools can be used instead]

USE is similar to BERT in that it is "context-aware" to understand not just words but entire sentences

This is important because terms like "inflation" will show up in both positive & negative texts

As opposed to the transformer model of BERT, USE is only an encoder

↳ each document is encoded into a numeric value

Because it is difficult to say if a statement is hawkish or dovish ex ante
they use the alternate statements to quantify this sentiments

↳ Alt. A - Typically more lenient / dovish
Alt. B - Actual statement
Alt. C - Typically tighter / hawkish

↳ Measure similarity between B & A/C
↳ if more similar to A then this is a "dovish" statement
C "hawkish"

↳ Measure change over time to determine how much "new information" / "new direction"
there is

Idea: (In paper) such metrics improve performance of bond market prediction

(In reality) the time lag to get the alternative statements causes issues for real-time use

Possible if a large enough data set of alternates we can learn the sentiment

Possible to use large language models to construct the alternatives in near real-time