Forecasting Business Cycle Direction using NLP

Executive Summary

Q3 2021

Problem identification - Context

Business cycle forecasting is valuable to businesses so that they can make informed business decisions.

- If anticipating expansion:
  - Decrease cash holdings
  - Increase capital investments and operations spending
  - Increase hiring
- If anticipating contraction:
  - Increase cash holdings
  - Decrease capital investments and spending
  - Decrease hiring or restructure (layoff) personnel

Problem identification - Context

- Existing forecasting approaches include qualitative models and quantitative models.
- Thousands of approaches exist for business cycle forecasting, but which ones are useful?

Problem identification - Context

- There may be value in using a model that examines public data in a novel way.
- Can we use a natural language processing model to analyse forward-looking statements by supply chain managers to forecast the business environment for the upcoming quarter?

Problem identification - Success Criteria

- The best model will maximize accuracy.
  - Stakeholders can best understand and make business decisions using accuracy.
  - No anomalies in data that warrant the extra complexity--relative to accuracy--of the other metrics.
- Training time and prediction time: not important.
  - New data is released infrequently.
  - Training and predictions will be run infrequently (once per month).

Problem identification - Scope

- Binary classification problem
- Target: direction of change in GDP in next quarter's GDP release vs current quarter (positive or negative)
- Features: variable, up to 300 features (depending on vector mode) derived from 5 text features
- Prediction: there exists a model that is significantly more accurate than the baseline (0.5>>accuracy>=1.00)

Problem identification - Scope

- Sample of the data: “Summary”.

APRIL 2021 MANUFACTURING INDEX SUMMARIES

PMI®

Manufacturing grew in April, as the Manufacturing PMI® registered 60.7 percent, 4 percentage points lower than the March reading of 64.7 percent. Although the Manufacturing PMI® has cooled compared to March, it remains at historically high levels. “The Manufacturing PMI® continued to indicate strong sector expansion and U.S. economic growth in April. Four of the five subindexes that directly factor into the Manufacturing PMI® were in growth territory. All of the six biggest manufacturing industries expanded, in the following order: Fabricated Metal Products; Chemical Products; Food, Beverage & Tobacco Products; Computer & Electronic Products; Transportation Equipment; and Petroleum & Coal Products. The New Orders and Production indexes continued to expand at strong levels. The Supplier Deliveries Index continued to reflect suppliers’ difficulties in maintaining delivery rates, due to factory labor-safety issues, transportation challenges and increased demand. Nine of 10 subindexes were positive for the period; a reading of ‘too low’ for Customers’ Inventories Index is considered a positive for future production,” says Fiore. A reading above 50 percent indicates that the manufacturing economy is generally expanding; below 50 percent indicates that it is generally contracting.

A Manufacturing PMI® above 43.1 percent, over a period of time, generally indicates an expansion of the overall economy. Therefore, the April Manufacturing PMI® indicates the overall economy grew in April for the 11th consecutive month following contraction in April 2020. “The past relationship between the Manufacturing PMI® and the overall economy indicates that the Manufacturing PMI® for April (60.7 percent) corresponds to a 5-percent increase in real gross domestic product (GDP) on an annualized basis,” says Fiore.

Problem identification - Scope

- Sample of the data: “WRAS”.

WHAT RESPONDENTS ARE SAYING

- “The current electronics/semiconductor shortage is having tremendous impacts on lead times and pricing. Additionally, there appears to be a general inflation of prices across most, if not all, supply lines.” (Computer & Electronic Products)
- “Upstream producers/suppliers are back online and working towards full rates. Demand is outpacing supply and will continue into the third quarter, when the supply chain is expected to be refilled. Supply/demand should be more balanced in Q3/Q4, but demand will continue as customers run hard to meet their demand and rebuild inventory.” (Chemical Products)
- “Continued strong sales; however, we have had to trim some production due to the global chip shortage. Hasn’t affected inventories greatly yet, but a continued decrease will begin to reduce available inventories if we don’t recover chip supply shortly.” (Transportation Equipment)
- “Business is picking up as restaurants open.” (Food, Beverage & Tobacco Products)
- “Oil production has been steady, along with market prices and capital expenditures.” (Petroleum & Coal Products)
- “Steel prices are crazy high. The normal checks on the domestic steel mills are not functioning — imported steel is distorted by the Section 232 tariffs.” (Fabricated Metal Products)
- “It’s getting much more difficult to supply production with materials that are made with copper or steel. Lots of work on the floor, but I am worried about getting the materials to support.” (Electrical Equipment, Appliances & Components)
- “Market capacity in most areas is oversold, with no realistic improvement on the horizon. In fact, it appears that demand will continue to strengthen, leading to more significant disruptions.” (Furniture & Related Products)
- “In 35 years of purchasing, I’ve never seen everything like these extended lead times and rising prices — from colors, film, corrugate to resins, they’re all up. The only thing plentiful at present, according to my spam filter, is personal protective equipment [PPE].” (Plastics & Rubber Products)
- “The metals markets remain very challenging at best. Shortages of raw materials have increased, especially in aluminum and carbon steel. Prices continue to rapidly increase. Transportation and trucking [are] also a big challenge.” (Primary Metals)
- “Demand continues to be very strong. Supply chain delays hamper our availability and ability to sell more.” (Machinery)

Problem identification - Scope

- Sample of the data:
  - “Commodities Up in Price”
  - “Commodities Down In Price”
  - “Commodities In Short Supply”.

COMMODITIES REPORTED UP/DOWN IN PRICE AND IN SHORT SUPPLY

Commodities Up in Price

Adhesives; Aluminum (14); Aluminum Products (4); Ammonia; Cable & Harnesses; Capacitors; Caustic Soda (2); Coatings (4); Copper (14); Corrugate (10); Corrugated Packaging (9); Crude Oil (2); Diesel Fuel (7); Electrical Components (8); Electrical Motors; Electronic Components (8); Ethylene; Freight (9); High-Density Polyethylene (HDPE) (7); Hydraulic Components; Labor – Temporary (3); Lumber* (13); Medium-Density Fiberboard (MDF); Natural Gas; Ocean Freight (8); Packaging Film; Packaging Supplies (8); Pallets; Paper Products; Phosphates; Plastic Resins (11); Polyethylene (6); Polyols; Polypropylene (13); Polyvinyl Chloride (PVC) Products (2); Printed Circuit Board Assemblies (PCBAs); Propylene; Resin-Based Products (6); Resins; Resistors; Semiconductors (6); Shipping Containers; Solvents; Steel (12); Steel – Carbon (8); Steel – Hot Rolled (11); Steel – Scrap (3); Steel – Stainless (9); Steel Products (11); Steel Wire; Sulfuric Acid; Totes; Wood (2); and Wood Products (2).

Commodities Down in Price

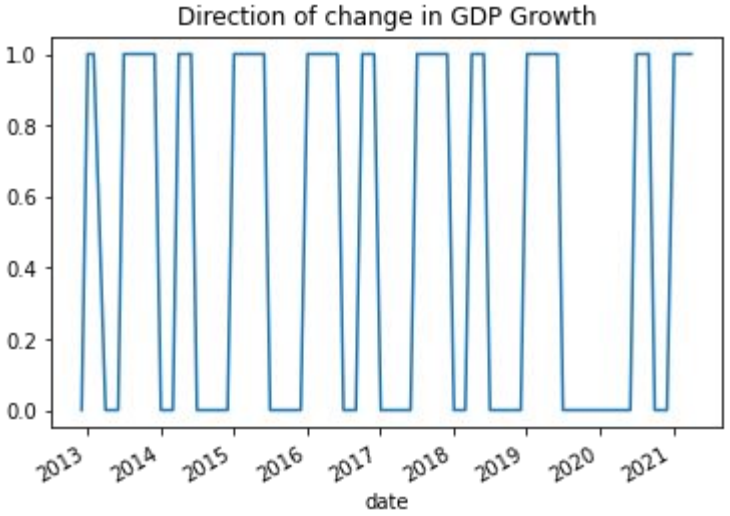
Copper; and Lumber*.

Commodities in Short Supply

Adhesives & Paint; Aluminum (4); Aluminum Products (3); Capacitors; Chlorine; Corrugate (2); Corrugated Packaging; Electrical Components (10); Electronic Components (8); Freight; Hardware; Hydraulic Components; Labor – Temporary (3); Lumber; Nylon Polymer; Ocean Freight (4); Packaging Supplies (2); Paper (2); Phosphates; Plastic Products (6); Plastic Resins – Other (5); Polyols; Polyvinyl Chloride (PVC) Resin (3); Personal Protective Equipment (PPE) – Gloves; Resistors;

Problem identification - Scope

- Target: “Direction of change in GDP growth”



Problem identification - Scope

- Target variable is derived from rate from GDP growth by taking the sign of it's slope (positive or negative).
- "Will the quarter's growth be greater than or less than last quarter's growth?"

Key findings

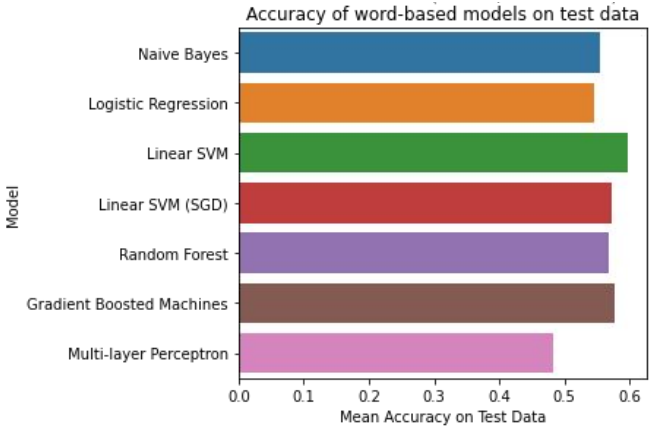
Among 6 models tested:

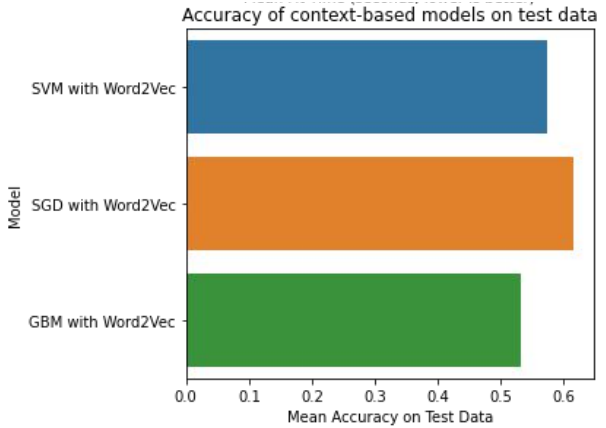
- SGD model with context-based preprocessing is top-ranked, and
- Linear SVM model with word-based preprocessing is the second-ranked

in accuracy.

Modeling results and analysis - Performance

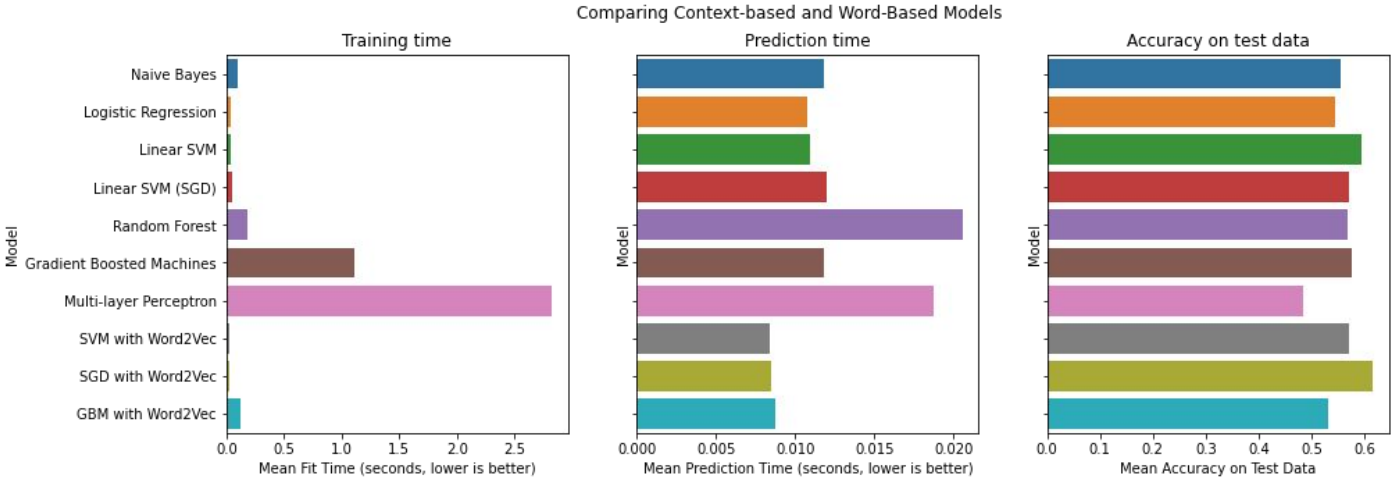
- Context-based (Word2Vec) preprocessing performed slightly better than word-based preprocessing (TF-IDF) on this data set





Modeling results and analysis - Performance

- Most models performed slightly better than the baseline (Accuracy = 0.5)
- Most models had similar training and prediction time.



Summary and conclusion

- Conditional recommendation: use the SVM with context-based preprocessing model to forecast the next quarter's direction of change in GDP, but don't rely solely on it's prediction.

- Conditional recommendation: Add the SVM with context-based preprocessing model to an existing uncorrelated forecasting approach to yield a better approach.

Further Work

- License the older data for research purposes (from pre-2012).
- Compare models with short-term and long-term memory.
- Compare models with newer context-based preprocessors like GloVe (2014) and fastText (2016)
- Compare models that use topic embeddings