Modeling financial analysts’ decision making via the pragmatics and semantics of earnings calls

Katherine A. Keith
University of Massachusetts Amherst

Amanda Stent
Bloomberg L.P.

Summary

- Overall motivation: Examine financial analysts’ decision making as it pertains to the language content of earnings calls.
- We correlate analysts’ pre-call recommendations (buy/sell) with the questions they ask during calls and find bullish analysts tend to be called on earlier, ask questions that are more positive, more concrete, and less about the past.
- Our prediction task indicates the textual content of earnings calls is moderately predictive of changes in analysts’ price target forecasts and more predictive than market features alone.

Data

Earnings call transcripts:
- S&P 500 companies from 2010-2017
- ~12,000 earnings call documents, temporal train/dev/test split
- ~600,000 total Q&A sets
- 10.9 average unique analysts speaking per call

Descriptive study of analysts’ questions

Pearson correlations of features with the type of analyst:
- bearish (sell) -1
- neutral (hold) 0
- bullish (buy) 1

Predicting changes in analysts’ price targets

Null hypothesis: Earnings calls are NOT predictive of analysts’ price target changes.
Rationale: analysts have access to private communication with executives and current events.
Research hypothesis: The pragmatic and semantic content of earnings calls ARE predictive of analysts’ price target changes.

Classification task set-up

Y = percent change in price target averaged over all analysts
Class 1: 0.0% < Y ≤ 1.67%
Class 0: -1.67% ≤ Y < 0.0%
Class -1: Y < -1.67%

Results

Feature type | Feature | Model | Acc. | F1 | % err.
--- | --- | --- | --- | --- | ---
Baselines | Random (ave. 10 seeds) – | – | 0.340 | 0.338 | –
 | Predict majority class – | – | 0.387 | 0.186 | 0.0
Market | Market LogReg | – | 0.435 | 0.408 | 12.4
 | Market LogReg | – | 0.482 | 0.475 | 24.8
Semantic | Bag-of-words LogReg | – | 0.479 | 0.460 | 23.8
 | doc2vec LogReg | – | 0.442 | 0.400 | 14.2
Pragmatic | Pragmatic lexicons LSTMs | – | 0.415 | 0.398 | 7.2
 | Fusion doc2vec + prag LSTMs | – | 0.461 | 0.460 | 19.1
 | Ensemble doc2vec + prag + market Ensemble | – | 0.460 | 0.461 | 18.9

Error analysis

Per-industry breakdown of errors on the validation set for 11 GICS industries

References