Textual Analysis in Economics and Finance

Ioannis Pragidis
Democritus University of Thrace
Department of Economics
University Campus
P.C 69100 Greece
306.946.905.951 *voice*
gpragkid@econ.duth.gr

Paris Alexandros Karypidis
Democritus University of Thrace
Department of Economics
University Campus
P.C 69100 Greece
306.939.777.677 *voice*
karypidisparisalexandros@gmail.com

September 28, 2018

Abstract

It's almost a decade now that textual analysis or text mining is applied in economics and finance. There maybe a couple of reasons why this has been the case. The online availability of text from social media, news articles, announcements, earning filings provide ample folder for applying the technology. By quantifying texts, researchers get access to a new database that may be useful in analysing long standing questions; How does uncertainty affect markets? Are financial markets predictable? How monetary policy transmits in the economy? In this paper we show that tailored made word lists for specific domains of application can perform better than using vague algorithms and unsupervised (atheoretical) machine learning technology in terms of forecasting. Following Loughran et al., (2017), I test whether an oil prices related word list can better predict future oil prices than an unsupervised machine learning technique. Assuming linearity, the oil related word list used in an OLS framework, forecasts better oil prices than unsupervised machine learning. Under non-linearity we find that if the domain specific word list is used to train the machine learning algorithm, then forecasting results improve compared to the unsupervised case. Above results are also confirmed when we test if

European Central Bank announcements can predict stock market movements. Finally, domain specific word lists can also contribute to transparency since it's easy for other researchers to replicate the results.

Key words: Textual analysis; news stories; oil prices; machine learning; ECB announcements.

Loughran, T., McDonald, B. and Pragidis, I., 2017. Assimilation of Oil News into Prices. SSRN