Abstract:

Predicting the financial performance of companies using various data sources has been the focus of a large number of studies. However, current models can still be improved to achieve even better predictions. This research makes three contributions to knowledge.

First, this research introduces a new method to more accurately predict the probability that a company's credit rating stays the same or will change. More specifically, this research exploits the vast increase in social media data put out by companies, that enables to develop more predictive models than those developed in the past. By analysing the tweets put out by different companies, we show, that there is a correlation between specific words in the tweets and the credit rating. Furthermore, we will show that using these words in a predictive model leads to an increase in predictive power when comparing with a model not containing these words as predictors.

Second, we are using differential language analysis for identifying linguistic features in tweets that are most predictive of a change in credit ratings, a method which is usually used in psychology and health and has not been used in a financial context before.

Third, we use a very large and unique dataset of variables from different frequencies for the analysis. For instance, tweets from 2016 to April 2018 are considered as well as data on the companies, such as their size or their sector of a large sample of NYSE listed companies.