Impact of the COVID-19 outbreak on Italy's country reputation: a sentiment analysis approach

Gianpaolo Zammarchi[1], Francesco Mola[1], Claudio Conversano[1]
[1] Department of Economics and Business Science, University of Cagliari, Cagliari, Italy

Abbreviated abstract: Italy was one of the first European countries to be severely affected by the COVID-19 outbreak, potentially leading to country reputation damage. We investigated changes in opinions about Italy before and after the COVID-19 outbreak using sentiment analysis on Twitter data through lexicon-based and machine learning methods. Sentiment values showed a breakpoint after the first case of COVID-19 in Italy and were correlated with changes in the FTSE-MIB index, suggesting they might serve as early detection signals of changes in stock market.

Related publications:
– G. Zammarchi *et al*, arXiv https://arxiv.org/abs/2103.13871 (2021)

Università di Cagliari

3rd Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
October 28-30, 2021

Problem, Data, Previous Works

Challenge: assess the impact of a negative event such as the COVID-19 outbreak on country reputation in Italy, one of the first countries to be severely affected, and the potential economic repercussions

Previous works have shown how Twitter data might be used to gain insights into user opinion and reactions to events

Solution

1) Collect a large dataset of tweets to analyze the temporal evolution of the sentiment towards Italy before and during the COVID-19 outbreak through a sentiment analysis on tweets
2) Analyze changes in sentiment, as well as the correlation with stock market exchange index values, using different analytical approaches
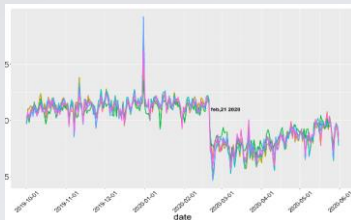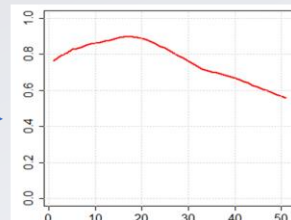
Università di Cagliari

3rd Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
October 28-30, 2021

Methods

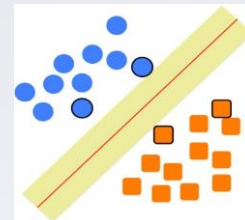| ID	Date	Tweet
1	2019-10-01	Tweet 1
2	2019-10-02	Tweet 2
…	…	…
n	2020-05-31	Tweet n

1,000 tweets/day
Oct 2019 - May 2020

243,846 tweets
after quality control

Sentiment analysis and
identification of breakpoints

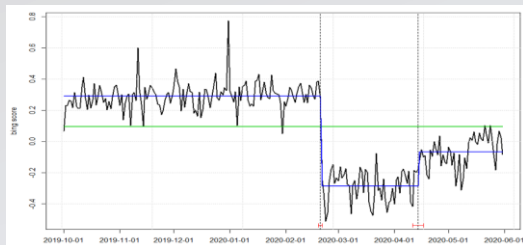Correlation between
sentiment and FTSE-MIB

Machine learning
analysis of tweets

- Sentiment analysis was conducted in R using different **dictionaries** (nrc, afinn, bing, syuzhet, sentimentR, vader) as well as the two **machine learning** methods naïve Bayes and Support Vector Machine (SVM)
- **Breakpoints** (i.e. sudden changes of values in a time series occurring at one or more specific dates) were identified using the strucchange R package
- The association between sentiment and FTSE-MIB index data was evaluated using **linear regression**, applying time lags from 0 to 50 days (sentiment scores from one day are compared with stock exchange values of 50 days later)

Università di Cagliari

3rd Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
October 28-30, 2021

Results and Conclusions

	Naïve Bayes		SVM	
	Pos	Neg	Pos	Neg
Precision	0.87	0.78	0.83	0.78
Recall	0.75	0.89	0.76	0.84
F1-score	0.81	0.83	0.79	0.81
Accuracy		0.82		0.80

Identification of a breakpoint in sentiment values on 21 February 2020, the date of the first case of COVID-19 in Italy
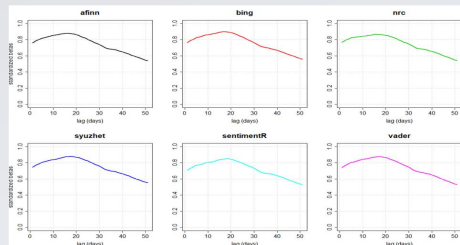
Strong association between sentiment and FTSE-MIB values at lag zero (β≥0.70, p<0.001, for all lexicons), also considering a lag up to 15 days

Similar performance for the two classifiers, with Naïve Bayes showing a slightly better performance for most metrics.

Take-home message: We observed a substantial rise in negative sentiment towards Italy in correspondence of the first Italian case of COVID-19 followed by a change towards slightly positive values starting two months later. Besides being useful to interpret the general sentiment as a proxy of the perceived country reputation, we suggest that sentiment scores can also be used to early detect changes in stock exchange values.