

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	Dejan Lazeski
Advisor:	Evžen Kočenda
Title of the thesis:	Stock Market Prediction: A Multiclass Classification on Emotions and Sentiment Analysis for Tweets and News Headlines

OVERALL ASSESSMENT (provided in English, Czech, or Slovak):

Please provide your assessment of each of the following four categories, summary and suggested questions for the discussion. The minimum length of the report is 300 words.

Contribution

This thesis analyzes whether and how stock market prediction works when emotions and sentiment from various sources are put into work. The stocks of the top 5 Big Tech companies take the center stage.

There are two key methodological contributions. One, emotion information content is divided into a fine granular data set. Results show that such a finer-grained emotion classification, as an important contribution to the existing literature, delivers also finer-grained findings. Second contribution rests in comparison of both headlines and tweets from same source in parallel, a feature not investigated so far.

The empirical contributions show that there is no strong correlation between daily stock price movements and the distribution of sentiment and emotions. Further, it was shown that tweets are less neutral than news headlines. Finally, results indicate that sentiment polarity can effectively predict future stock price movements compared to finer-grained emotion classification. The detailed results of the thesis are simply too numerous to be reviewed within the scope of this report.

The results of the thesis are based on a meticulously executed and very detailed analysis.

Methods

The very rich sample consists of the tweet and headline news data, in thousands of items; the key details are summarized in tables 3.2 and 3.4. The tweets and headlines are drawn from well-known financial newspapers, explicitly addressing the top 5 Big Tech companies. Collection of the data is an achievement on its own. In order to analyze the impact of sentiment and emotions on stock prices, a multiclass emotion and sentiment classifiers are developed by utilizing a supervised learning approach. The information from the data is transformed into information items representing positive, negative, and neutral sentiment, and emotions defining anger, joy, surprise, and sadness. An empirical investment strategy shows that when sentiment polarity is extracted, then sentiment polarity can effectively predict future stock price movements

Literature

The literature review section summarizes the current state of research in the field. It is quite exhaustive, and it is also conveniently divided into several parts. The section is not only a standard literature review, but it offers quite useful overview that provides further motivation and justifies the analysis. Literature is reviewed from the following angles relevant to stock performance and investor sentiment, news, and Twitter releases. It also discusses emotion classification. Hence, the literature is reviewed in a detail and covers all relevant papers and angles.

Manuscript form

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	Dejan Lazeski
Advisor:	Evžen Kočenda
Title of the thesis:	Stock Market Prediction: A Multiclass Classification on Emotions and Sentiment Analysis for Tweets and News Headlines

The manuscript conforms to formal requirements for the master thesis. It reads well and the flow and grammar are fine. The results are presented with expansive detail. Introduction is short but clear, to the point and it motivates well the researched topic. Data are fully described. Hypotheses are stated in the thesis proposal, but they deviate from the completed thesis for a good reason. Tables and figures are presented in an organized and legible manner. References are complete.

Summary and suggested questions for the discussion during the defense

The thesis represents a solid piece of empirical work on the subject of stock prices and market sentiment, tackled from very uncharted angle. The results of the Urkund analysis do not indicate significant text similarity with other available sources. In my view, the thesis fulfills the requirements for a master thesis at IES, Faculty of Social Sciences, Charles University, and I recommend it for the defense and suggest a grade A.

SUMMARY OF POINTS AWARDED (for details, see below):

CATEGORY	POINTS
<i>Contribution</i> (max. 30 points)	29
<i>Methods</i> (max. 30 points)	27
<i>Literature</i> (max. 20 points)	19
<i>Manuscript Form</i> (max. 20 points)	20
TOTAL POINTS (max. 100 points)	95
GRADE (A – B – C – D – E – F)	A

NAME OF THE REFEREE: *Evžen Kočenda*

DATE OF EVALUATION:

Digitally signed (29.08.2020)
Evžen Kočenda

Referee Signature

EXPLANATION OF CATEGORIES AND SCALE:

CONTRIBUTION: *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

METHODS: *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

LITERATURE REVIEW: *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

MANUSCRIPT FORM: *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

Overall grading:

TOTAL	GRADE
91 – 100	A
81 - 90	B
71 - 80	C
61 – 70	D
51 – 60	E
0 – 50	F