

# Report on Bachelor Thesis

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

<b>Student:</b>	<b>Jakub Doškář</b>
<b>Advisor:</b>	<b>Mgr. Nicolas Fanta</b>
<b>Title of the thesis:</b>	<b>Examining the Interaction between the Cryptocurrency Market Development and Activity on Leading Social Networks</b>

## OVERALL ASSESSMENT

### Short summary

The thesis of Mr. Doškář aims to clarify the role of sentiment on the pricing of crypto assets. For that, a decent dataset has been prepared by the author. However, the thesis was apparently elaborated under time pressure which unfortunately resulted into various shortcomings on multiple levels. The study fails to bring a model to predict Bitcoin prices based on aggregate tweets about the cryptocurrency.

### Contribution

For the analysis, a dataset covering the period from February 2021 until November 2022 has been created. In terms of social media coverage, more than 4.5 million Bitcoin related tweets have been collected. The dataset therefore forms a solid basis for the analysis and the effort to create it should be praised.

In terms of approach, the choice of NLP methods is reasonable. Common NLP methods cannot consider various aspects specific for social media communication (e.g. slang, emoticons, initialisms, acronyms or even the contractions as negations), therefore the presented study leverages a commonly used enhanced natural language processing approach to study the effects.

The contribution can be seen in the analysis of sentiment impact on the post-covid cryptocurrency market. Unfortunately, the study does not confirm any significant effects of sentiment on the Bitcoin asset pricing. This can at least partially be explained by flaws in the methods described below.

### Methods

A significant part of the analysis is run very likely on non-stationary data as the close price is used as the explanatory variable. In addition, any assessments on stationarity could not be found in the thesis. An estimation of returns based on sentiment change is included but is very limited (and without major significant results) as in this case it only examines the direction of the change in sentiment and does not deal with its magnitude. Again, even in this approach, a stationarity check is missing. Overall, this presents a serious flaw in the methods.

In terms of the tweets' processing, the Valence Aware Dictionary for Sentiment Reasoning (VADER) is used to overcome shortcomings of other standard NLP approaches with regards to the specifics of social media communication. This step seems reasonable and nothing to object here.

### Literature

The comparison of the different streams of research and the overall positioning of the thesis within those could have been done in a more structured way and more extensively. Further comparisons and analyses of the used methods in the relevant different strands of literature would also help.

### Manuscript form

A lack of structure can be observed throughout the thesis. In the methodology, a clear separation of the different approaches is missing, the literature review is not organised by different streams of research. Furthermore, the clarity of the explanations is low. For instance, in the part dedicated to methods, models' summaries containing the formula and description on the variables used would be useful.

From the graphical perspective, there are several objections to be made: graphs are not unified (not the same library was used to generate them), tables are inserted as pictures and not as text and overall, anything inserted as a picture is in low resolution. Last but not least, a more diligent spell check should have been done.

# Report on Bachelor Thesis

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

<b>Student:</b>	<b>Jakub Doškář</b>
<b>Advisor:</b>	<b>Mgr. Nicolas Fanta</b>
<b>Title of the thesis:</b>	<b>Examining the Interaction between the Cryptocurrency Market Development and Activity on Leading Social Networks</b>

## Overall evaluation and suggested questions for the discussion during the defense

A solid dataset has been elaborated for the purpose of the analysis. The literature summary is very brief, but still manages to introduce the field of research. There are gaps in the estimation methods used. Yet, I think the thesis is defensible and fulfills the requirements for a bachelor thesis at IES, Faculty of Social Sciences, Charles University. I therefore recommend it for the defense and suggest a grade E. The results of the Urkund analysis do not indicate significant text similarity with other available sources.

For the defense the following questions are suggested:

- Summarise your main findings of the models estimating price change and explain the case where you reach significant results.
- Why do you think positive sentiment is more frequent? On Figure 1 we can observe this disbalance.

## SUMMARY OF POINTS AWARDED

<b>CATEGORY</b>	<b>POINTS</b>
<i>Contribution (max. 30 points)</i>	17
<i>Methods (max. 30 points)</i>	13
<i>Literature (max. 20 points)</i>	15
<i>Manuscript Form (max. 20 points)</i>	7
<b>TOTAL POINTS (max. 100 points)</b>	<b>52</b>
<b>GRADE</b>	<b>E</b>

**NAME OF THE REFEREE:**

**Nicolas Fanta**

**DATE OF EVALUATION:**

*Digitally signed 16.1.2023  
Nicolas Fanta*

---

**Referee Signature**

