

Bitcoin vs. Twitter — Zusammenhänge des Bitcoin-Kurses und zeitgleicher Tweets

Alicia Hirt,^{*} Tristan Kreuziger,^{*} and Jan-Philipp Praetorius^{*}

*Faculty of Mathematics and Computer Science, Friedrich-Schiller-Universität Jena,
Germany*

E-mail: alicia.hirt@uni-jena.de; tristan.kreuziger@uni-jena.de; jan-philipp.mueller@uni-jena.de

Einführung

Neben etablierten Wertpapieren und Indizes, lassen sich seit einiger auch digitale Zahlungsmittel an offiziellen Marktplätzen und Börsen handeln. Während Währungen wie der Dollar einen relativ gut messbaren Wert besitzen und dieser Wert durch aktive Maßnahmen reguliert und beeinflusst werden kann, sind Kryptowährungen weitestgehend unabhängig. Dabei haben digitale Währungen wie der Bitcoin eine extreme Volatilität. Im Rahmen dieser Arbeit soll ermittelt werden, ob es einen Zusammenhang zwischen dem Kursverlauf des Bitcoin und der Meinung von Twitter-Nutzern über den Bitcoin gibt.

Daten

Um einen Zusammenhang zwischen dem Bitcoin und Twitter-Nutzern ermitteln können, werden Daten aus dem gleichen Zeitintervall benötigt. Der Kursverlauf des Bitcoin lässt sich unkompliziert über mehrere Quellen beschaffen. Insbesondere Internetseiten wie www.yahoo.com bieten gute Möglichkeiten sekundengenaue Informationen über den Kursverlauf zu erhalten. Für diese Arbeit wurde der Kurs im Minutentakt vom 10. Januar 2018 bis zum 2. Februar 2018 von der Seite www.coindesk.com extrahiert.

	A	B	C
1	Date	Open Price	Close Price
2	2018-01-10 00:00	14073,16	14073,16
3	2018-01-10 01:00	14073,16	14193,43
4	2018-01-10 02:00	14193,43	14290,4
5	2018-01-10 03:00	14290,4	14139,18
6	2018-01-10 04:00	14139,18	14374,32
7	2018-01-10 05:00	14374,32	14182,03
8	2018-01-10 06:00	14182,03	14240,25
9	2018-01-10 07:00	14240,25	14115,24
10	2018-01-10 08:00	14115,24	13500,39
11	2018-01-10 09:00	13500,39	13790,67
12	2018-01-10 10:00	13790,67	13862,86
13	2018-01-10 11:00	13862,86	14073,29
14	2018-01-10 12:00	14073,29	13898,03
15	2018-01-10 13:00	13898,03	13907,03
16	2018-01-10 14:00	13907,03	13778,42
17	2018-01-10 15:00	13778,42	14420,99
18	2018-01-10 16:00	14420,99	14498,19

Figure 1: Bitcoin-Kursverlauf

In welchem Format die Bitcoin-Kurse vorliegen ist in Abbildung 1 zu sehen. Den Bitcoin-Kursen werden im Rahmen dieser Arbeit Tweets gegenübergestellt. Ein Tweet hat maximal 280 Zeichen. In Abbildung 2 ist ein Beispiel für ein Tweet zu sehen.



Figure 2: Beispiel-Tweet

Um eine repräsentative Stichprobe zu haben, wurden knapp 120.000 Tweets untersucht. Neben dem Text eines Tweets, werden außerdem Informationen über denjenigen ausgelesen,

der den Tweet verfasst hat, wann diese Kurznachricht verfasst wurde, eine eindeutige Identifikationsnummer, retweets und jede Menge andere Informationen. Für diese Arbeit wurde jedoch nur das Erstellungsdatum und der Text eines Tweets gesammelt. Allgemein wurden nur Nachrichten aufgezeichnet, in denen die Zeichen: "Bitcoin", "#Bitcoin" und "cryptocurrency" vorkommen. Außerdem sind nur Tweets erfasst worden, die innerhalb der USA und einem Großteil von Kanada veröffentlicht wurden. Da die Sentiment-Analyse am effizientesten funktioniert, wenn alle zu untersuchenden Texte in einer Sprache sind, wurde mit Hilfe einer Bibliothek für jeden einzelnen Tweets die entsprechende Sprache ermittelt. Alle Nachrichten, die nicht in der Sprache Englisch waren, wurden daraufhin aus dem Datensatz entfernt.

Methode

Spark xyz ...

Auswertung

Ergebnisse xyz ...

Ausblick

andere Daten ... konsistente Daten ...