



Universität Paderborn
Fakultät für Wirtschaftswissenschaften
Department Wirtschaftsinformatik

Exposé zur Studienarbeit

Predictive Analytics

Daily Financial News for 6000+ Stocks

von

Benjamin Kasten

Matrikelnummer: 7154784

Warburger Str. 100, 33098 Paderborn

bkasten@mail.uni-paderborn.de

und

Dominik Höhr

Matrikelnummer: 7157684

Salentinstraße 16 A, 33102 Paderborn

dhoehr@mail.uni-paderborn.de

vorgelegt bei

Prof. Dr. Oliver Müller

Prof. Dr. Erika Mustermann

15. August 2020

Expose

Forschungsfrage und Thema

Zu Grunde gelegter Datensatz: Daily Financial News for 6000+ Stocks (see Developer, 2020)

Der Datensatz liefert rund 1.410.000 Überschriften von Internetartikeln, über die Entwicklung von zugeordneten Aktien, welche an der Amerikanischen Börse notiert sind. Zu beachten ist, dass die Internetartikel hauptsächlich von auf benzinga.com veröffentlicht wurden, dabei bestehen die Artikel häufig nur aus den Überschriften oder einer Sammlung von aktuellen Zahlen der Aktie(n).

Ziel der Studienarbeit ist es, die Headlines im ersten Schritt mittels eines Algorithmus (bspw. Topic Modeling oder Sentiment Analyse) zu klassifizieren oder zu kategorisieren, und anschließend zu untersuchen, ob das Ergebnis einem objektivem Bewertungskriterium folgt und beispielsweise mit dem Image, dem Aktienkurs oder den Umsatz assoziiert werden kann. Zur Findung eines Bewertungskriterium müssen weitere Datensätze ergänzt werden.

Wie sind Sie auf das Thema gekommen? Welche Hintergrundgeschichte führte zu der Forschungsfrage?

Der Markt wird immer schnelllebiger, insbesondere der Aktienmarkt erlebt einen immer schnelleren Wandel.

Die Einflussnahme von Image und Schlagzeilen nimmt dabei ebenfalls ständig zu und die Auswirkungen sind nicht immer vollständig zu erfassen. Für Unternehmen ist es von essentieller Bedeutung zu wissen, wie sich Headlines langfristig auf sie selber auswirken.

Warum ist dieses Thema bzw. diese Fragestellung untersuchenswert?

Dieses Thema ist untersuchenswert, Unternehmen in der Öffentlichkeit stehen und abhängig von Kund:innen sind, die Nachrichten konsumieren. Dementsprechend beeinflussen Nachrichten das Kaufverhalten und damit das Unternehmen. Durch das steigende Interesse der Menschen an Anlagestrategien, werden auch Aktiennews immer wichtiger. Deshalb ist es entscheidend für Unternehmen, erkennen zu können, was für einen Effekt Headlines auf KPIs haben.

Wer sind die möglichen Akteure? Wer könnte von unserer Forschung profitieren?

Die Akteure sind die börsennotierten Unternehmen. Diese können durch die Analyse der Headlines, die ihr Unternehmen betreffen, langfristige Trends erkennen, und so präventiv Maßnahmen treffen, um negative Auswirkungen auf ihre Unternehmen durch bestimmte Headlines zu verhindern.

Was wurde in diesem oder einem ähnlichen Gebiet bereits getan?

Es wurden bereits einige Analysen bezüglich der Auswirkungen von Headlines auf den Aktienkurs gemacht. Basierend auf Headlines aus der Vergangenheit und der darauf folgenden Kursentwicklung soll vorhergesagt werden, wie sich der Aktienkurs entwickelt, mit dem Ziel, Investitionsentscheidungen etc. zu erleichtern.

Wir möchten abschließend darauf hinweisen, dass eine konkrete Bewertung der Überschriften anhand der Aktienentwicklung nur mit entsprechenden Finanzdaten möglich ist. Die entsprechenden APIs sind dabei in der Größenordnung des Datensatzes nur kostenpflichtig zu erreichen.

(see Stegeman, 2020) (see Zhang, 2020)

Literaturverzeichnis

Developer, B. (2020). Daily Financial News for 6000+ Stocks. <https://www.kaggle.com/miguelaelnle/massive-stock-news-analysis-db-for-nlpbacktests>.

Stegeman, D. (2020). YahooFinance. <https://www.kaggle.com/danielstegemann/yahoofinance>.

Zhang, L. (2020). ECon 7930 Topic Modelling. <https://www.kaggle.com/luckyzx/econ-7930-topic-model>.