

Grundmotivation

Für Broker und Analysten von grossen Institutionen, vor allem aber auch für kleinere Vermögensverwalter sowie Privatpersonen ist trotz riesigen Datenmengen noch schwierig, gute Vorhersagen zum zukünftigen Verlauf von Aktienkursen zu machen, oder deren aktuellen Verlauf richtig nachvollziehen zu können.

Ziel

Das Ziel ist es daher, ein Tool zu erstellen, welches die oben beschriebene Zielgruppe dabei unterstützt, fundiertere Entscheidungen zum Finanzmarkt, im Speziellen zum Verlauf von Aktienkursen zu treffen. Dies soll unter Einbezug verschiedener Datenquellen sowie Datenarten geschehen.

Problemstellung

Ein Hauptproblem bei Vorhersagen besteht momentan oft darin, dass nicht alle Datenquellen, welche für die Analyse relevant wären, miteinbezogen werden. Zudem ist es weiterhin schwierig, die relevanten Datenpunkte aus den Datensätzen herauszufiltern, diese systematisch zu analysieren und sodann auch noch zu kombinieren.

Eine weitere Herausforderung solcher Analysen ist der relativ grosse Zeitaufwand um die erwähnten Daten zu finden & zusammenzuführen. *Oft geschieht dies z.T. noch manuell und ist daher ineffizient und fehleranfällig.*

Vorgehen

Zuerst werden die historischen Datenpunkte ausgewählter Aktienkurse heruntergeladen. Dann werden die Aktienkurse für die darauffolgende Periode vorhergesagt. Dies geschieht basierend auf dem historischen Verlauf sowie unter Einbezug verschiedener anderer Datenquellen. Um die Genauigkeit der Vorhersage zu prüfen, werden die Vorhersagen mit den tatsächlichen Aktienkursentwicklungen verglichen.

Für das Erstellen des Tools werden folgende Schritte unternommen:

1. Datenbeschaffung

1. Herunterladen der Aktienkurse für eine Periode $t-2$ ($T-2$ = relativ viele historische daten)
 - a. Firmennamen werden mittels Webcrawler ausgelesen
 - b. Aktienkurse werden mittels einem Yahoo sowie Google-API extrahiert
2. Beschaffung qualitativer Daten für die Periode $t-2$
 - a. Veränderung der LinkedIn Job Postings / Openings mittels API
 - b. Twitter Sentiment Analysis mittels Twitter API (Google News oder Expertenmeinungen)

2. Vorhersage

3. Vorhersage (Forecasting) der Aktienkurse für die Periode $t-1$
 - a. Mittels Machine Learning basierend auf den historischen Daten $t-2$
4. Miteinbezug der qualitativen Daten in die Vorhersage.
 - a. Twitter: Allgemeine Stimmung im Markt (Sentiment in Forecast einfließen lassen)
5. Unterstützen die qualitativen Daten die Vorhersage für ausgewählte Firmen
 - a. LinkedIn: Mehr Job Postings = positiv (und umgekehrt)
 - b. Twitter: Gute Stimmung = Positiv (und umgekehrt) (optional)
 - c. (Optional Headlines from News articles from Google News)

3. Validierung

6. Erstellte Prognose für $t-1$ mit den tatsächlichen Werten von $t-1$ vergleichen
 - a. Datenbeschaffung von $t-1$ ebenfalls wieder mittels Yahoo/Google API

4. Präsentation der Daten

7. Präsentation von
 - a. Der Genauigkeit der Vorhersage
 - b. Der verschiedenen Korrelationen (oder auch nicht)
 - c. Einer tatsächlichen Vorhersage für $t+1$
 - i. Alles mittels Tableau

Offene Fragen:

- Länge der zu analysierenden Perioden?
- Machbarkeit allg.?