**Projekt „Börse Analytics“**

**Business Understanding**

* Basierend auf ein Dataset von englischsprachigen Finanznachrichten wollen wir ein Börse-Analytics Model entwickeln, das die Entwicklung des Aktienkurses NASDAQ vorhersagt.
* Wir benutzen ein Dataset für Finanznachrichten: <https://github.com/EagleAdelaide/FinSen_Dataset>

**Data Understanding**

* Ergänzend zum Dataset wird ein Web Scrapper entwickelt, der die Daten über die NASDAQ-Kurswerte abruft. Diese Daten sind für die Evaluation des Modells notwendig. Es werden die tägliche Werte gesammelt und anschießend daraus werden die Daten zu Kursänderungen abgeleitet. -> Andrei, das Label für die Kursentwicklung bis zum 24.06
* Initial wird eine explorative Analyse des Nachrichten-Datasets durchgeführt, d.h. die Anzahl von Nachrichten, die Textlänge, sowie Schlagwörter und einfach Korrelationen (z.B. Häufigkeit des Wortes über eine Zeitperiode). -> Berke

**Data Preparation -> Muhammed**

* Nachrichtentexte müssen bereinigt werden.
* Einige Nachrichten werden mit den Labels vorgesehen, die die positive, negative oder schwankende Kursentwicklung kennzeichnen.

**Data Modelling -> Martin / Andrei**

* Wir nutzen ein Modell (Modell finden), das die finale Entscheidung über das Label trifft.
* Ergänzendes Modell für die Optimierung des Modells.

**Data Evaluation**

* Genauigkeit des Modells basierend auf die rohen und gelabelten Datensätzen.