

Projektarbeit

Lutz Seidel

Python KI Kurs

Plauen, den 11.7.2022

Zielsetzung

In dem Projekt möchte ich mit diversen Vorhersagemodellen den Kurs von Aktien und Kryptowährungen vorhersagen und mit den realen Kursen Vergleichen.

Versuchsaufbau

Es werden zur ersten Betrachtung die Daten einer Aktie verwendet die in dem Tutoriell

<https://www.analyticsvidhya.com/blog/2018/10/predicting-stock-price-machine-learningnd-deep-learning-techniques-python/>

verwendet wurden. In einem zweiten Schritt verwende ich dann einen Datensatz des BTC Kurses im Stunden Chart über 24 Stunden und betrachte die Vorhersagegenauigkeit gegenüber den ersten Datensatz.

Verwendete Vorhersage Modelle

1. Linear Regression

Der grundlegendste maschinelle Lernalgorithmus, der auf diesen Daten implementiert werden kann, ist die lineare Regression. Das lineare Regressionsmodell gibt eine **Gleichung zurück, die die Beziehung zwischen den unabhängigen Variablen und der abhängigen Variablen bestimmt.** Die Gleichung für die lineare Regression kann geschrieben werden als:

$$Y = \theta_1 X_1 + \theta_2 X_2 + \dots \theta_n X_n$$

Dabei stellen x_1, x_2, \dots, x_n die unabhängigen Variablen dar, während die Koeffizienten $\theta_1, \theta_2, \dots, \theta_n$ repräsentiert die Gewichte

2. Auto ARIMA

ARIMA ist eine sehr beliebte statistische Methode für Zeitreihenprognosen. ARIMA-Modelle berücksichtigen die vergangenen Werte, um die zukünftigen Werte vorherzusagen. Es gibt drei wichtige Parameter in ARIMA:

p (vergangene Werte, die zur Vorhersage des nächsten Werts verwendet werden)

q (vorherige Prognosefehler, die zur Vorhersage der zukünftigen Werte verwendet werden)

d (Ordnung der Differenzierung)

Die Parameterabstimmung für ARIMA nimmt viel Zeit in Anspruch. Ich verwende Auto-ARIMA, das automatisch die beste Kombination von (p,q,d) auswählt und den geringsten Fehler liefert.

3. Prophet

Es gibt eine Reihe von Zeitreihentechniken, die auf dem Bestandsvorhersagedatensatz implementiert werden können, aber die meisten dieser Techniken erfordern eine umfangreiche Datenvorverarbeitung, bevor sie an das Modell angepasst werden. Prophet, entworfen und entwickelt von Facebook, ist eine Zeitreihen-Prognosebibliothek, die keine Datenvorverarbeitung erfordert und extrem einfach zu implementieren ist. Die Eingabe für Prophet ist ein Datenrahmen mit zwei Spalten: Datum und Ziel (ds und y).

Prophet versucht, die Saisonalität in den vergangenen Daten zu erfassen und funktioniert gut, wenn der Datensatz groß ist

4. Long Short Term Memory (LSTM)

LSTMs werden häufig für Sequenzvorhersageprobleme verwendet und haben sich als äußerst effektiv erwiesen. Der Grund, warum sie so gut funktionieren, liegt darin, dass LSTM wichtige Informationen aus der Vergangenheit speichern und die unwichtigen Informationen vergessen kann. LSTM hat drei Tore:

Das Eingangsgatter: Das Eingangsgatter fügt Informationen zum Zellenzustand hinzu

Das Vergessenstor: Es entfernt die Informationen, die vom Modell nicht mehr benötigt werden

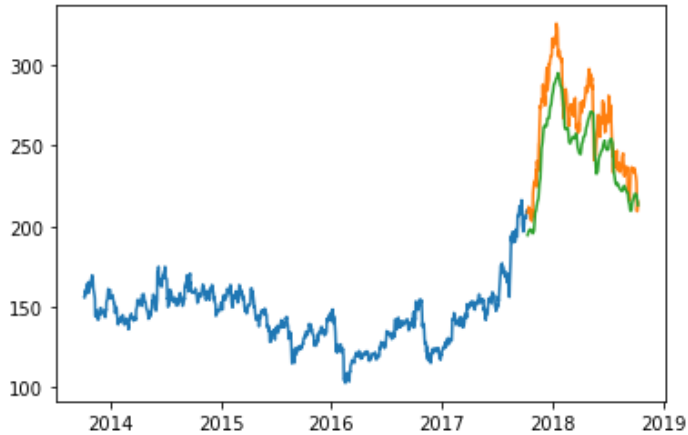
Das Ausgangsgatter: Das Ausgangsgatter bei LSTM wählt die Informationen aus, die als Ausgang angezeigt werden sollen

Weitere Informationen dazu unter

<https://www.analyticsvidhya.com/blog/2017/12/fundamentals-of-deep-learning-introduction-to-lstm/>

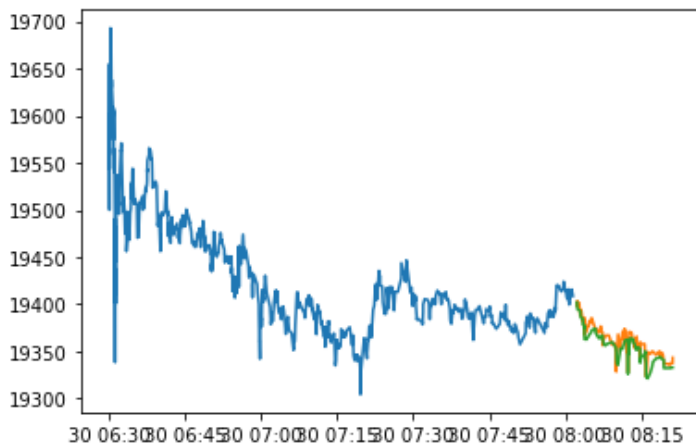
Ergebnisse

In der Betrachtung der Vorhersage zeigte sich, dass das LSTM Modell das genaueste der Betrachteten Modelle ist und eine sehr gute Vorhersage bei den Aktien Kurs Liefert



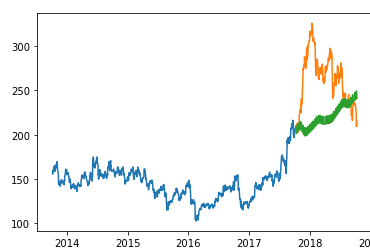
Blau ist der Lernbereich Orange der original Kurs wie er nach den Lernbereich weiter läuft und Grün die Vorhersage. Angewandt auf die Aktien Daten

Hier nun dasselbe mit den Daten für den BTC Kurs

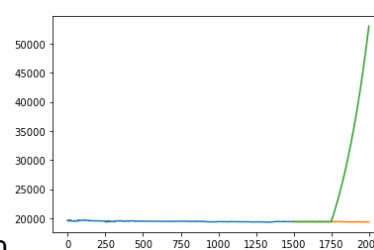


Die Ausgabe der anderen Modelle ist nicht verwertbar da sie viel zu weit von den Originaldaten abweichen. Außerdem ist bemerkenswert die Genauigkeit der Vorhersagen bei den Zwei Datensätzen unterscheidet.

Lineare Regression

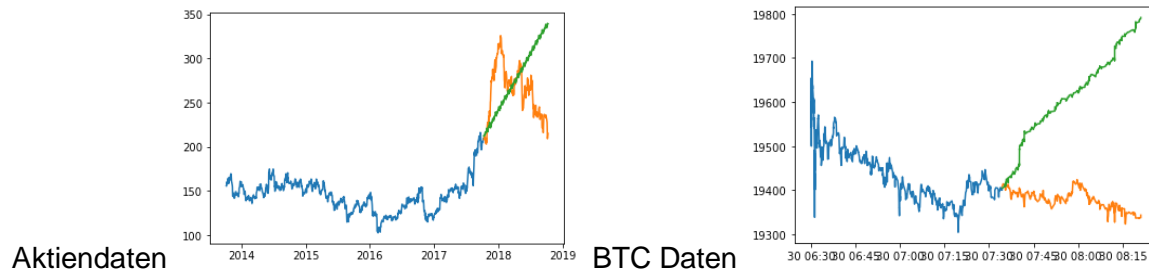


Aktien Daten



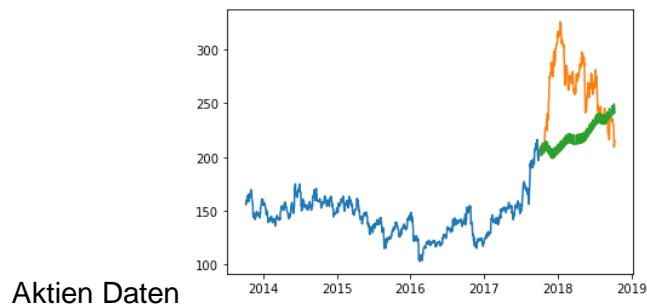
BTC Daten

Auto ARIMA



Prophet

Lies sich nur mit den Aktiendaten verwenden, da die Verarbeitung der BTC Daten für meine Hardware zu Zeit intensiv bzw. zu rechen intensiv darstellte



Fazit

Im Fazit kann man sagen, dass eine Vorhersage möglich ist mit dem LSTM, die auch über einen kurzen Zeitraum eine sehr gute Korrelation mit den realen Werten hat. Um das zu Nutzen als Signalgeber für Kauf und Verkaufs Entscheidungen bedarf es einer längeren Betrachtung und gegebenen Falls eine weitere Überprüfung der Plausibilität mittels TA und anderen Finanztechnischen Möglichkeiten