

FINAL PROJECT

Description of the Data

The data was imported from Alpha Vantage API connection it is the Stock price data for IBM for past 20+ years with data showing Open, Close, Low and high for every day.

Data Preprocessing

The API is giving by default Json file in order to demonstrate the learned pandas data frame I have kept it in Json format and transformed the data in such a manner as shown in the fig 1

	open	high	low	close	volume
1999-11-05	92.7500	92.9400	90.1900	90.2500	13737600
1999-11-04	94.4400	94.4400	90.0000	91.5600	16697600
1999-11-03	95.8700	95.9400	93.5000	94.3700	10369100
1999-11-02	96.7500	96.8100	93.6900	94.8100	11105400
1999-11-01	98.5000	98.8100	96.3700	96.7500	9551800

Fig 1

Learning Outcomes

I have demonstrated how to clean the data as in to check for missing data use describe function to understand the statistics of the data which will be beneficial in developing a machine learning model.

Challenges

The main challenge was to keep it in Json format and make the preprocessing. Also when plotting the distribution chart I had to change the then string data into numeric format.

The project gave me idea to how to look for API and how to add it to your code the next aim is to use a already created model api for stock price prediction and also built a LSTM model and compare the results.

As for the ethical take the data used does not contain any personal information nor does it harms anyone.