BITS Pilani
Pilani | Dubai | Goa | Hyderabad
Work Integrated Learning Programmes

BSE Prediction

Group 134 – Vinayak Nayak, Shreysi Kalra, Niyati Gangwal

Introduction to Data Science

M.Tech Data Science and Engineering

Overview

- Regression Analysis for the prediction of Bombay Stock Exchange Sensex
- Regression analysis used

Dataset

- Number of features: 7 raw total
- Time period of data: 2015 Jan – 2021 June
- Numerical Timeseries Data
- Obtained as HTML from Yahoo Finance
- Munged, standardized and curated data in a time-dependent manner, made the frequency consistent, added holidays and forward filled df for missing value imputation

Feature Engineering Techniques

- Non-exogenous timeseries data like volume, open, close etc. were dropped
- One Hot encoding was done for the time-dependent features created from the date
- Time dependent features were created
  - Day of the Week
  - Quarter of the year
  - Month of the year etc.

Methodology

- Regressors Used
  - SARIMAX with 4 exogenous variables
  - Prophet
  - LSTM with 4 exogenous variables

Results

	Method	RMSE	RMSPE(%)
1	SARIMAX with 4 exogenous variables	16422	0.15
2	Prophet with a univariate time series	31917	7.5
3	LSTM with 4 exogenous variables	23216	4.722