CAPSTONE PROJECT SUBMISSION

Yes Bank is a well-known bank in the Indian financial domain. Since 2018, it has been in the news because of the fraud case involving Rana Kapoor. Owing to this fact, it was interesting to see how that impacted the stock prices of the company and whether any predictive models can do justice to such situations. This dataset has monthly stock prices of the bank since its inception and includes closing, starting, highest, and lowest stock prices of every month. The main objective is to predict the stock's closing price for the month. As the first step, performing data-wrangling over the raw data further, we divided the complete project into four main parts i.e. data observation and cleaning, exploratory data analysis, checking multicollinearity, and data modeling. In the first step, we collected the raw data from Almabetter. Then we did a basic data cleaning check and removed zero NA values because there were no null values or missing values in the dataset.

The second step was exploratory data analysis we briefly study on variables after that we found our dependent and independent variables. After knowing the dependent variable we study close price for time over a longer period and try to identify the basic movement of stock price. Next, we plot some distribution plots to check the skewness of the variables and we found that the data is highly skewed so we use log transformation to make normal distributions.

Now after transformation, the correlation has been checked with help of heatmap, distribution plots, and scatterplots and there is a very high correlation among all variables. So we check multicollinearity with VIF (variation inflation factor). Then we drop some features to prevent wrong predictions.

Finally, for data modeling, we prepared dependent and independent variables for the train test split method. We apply Linear, Lasso, Ridge, and Elastic net regression. As per model performance, linear regression and ridge perform well. After cross-validation and Hyperparameter tuning performance increase significantly.

Using this dataset in our ML-Regression project we found some relevant analysis which predicts the monthly stock closing price of Yes Bank. But due to academic regulation, we use the regression algorithms for the analysis. As we know stock price data is time-related so the accuracy will be better with time series algorithms.

In the stock market there's a catch phrase "chart won't lie" which means you can watch and understand everything through the chart but in the end a stock price move not only technical perspective but also fundamentally we don't have any fundamental data except the 2018 fraud case.

Contribution Roles:

i). Ajay Tiwari
> Data Cleaning
> Data Preprocessing
> General Analysis Perform EDA(Exploratory Data
    Analysis)
> Build Regression Model Such As Linear regression, Lasso
    regression, Ridge regression and Elastic net regression

GitHub Link:
https://github.com/AjayUnique/Yes-Bank-Stock-Closing-Prediction

Drive Link:
https://drive.google.com/drive/folders/1MdBmjh-3jKJJmV6upV3wRISby9bBpZ6F