Capstone Project Submission

Instructions:
i) Please fill in all the required information.
ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
• Contributor Role: PARUL SRIVASTAVA(srivastava98parul@gmail.com) 1. Data Analyzing ➢ Performing Data Wrangling, Data Gathering, processing and transforming. 2. Data Cleaning ➢ Finding missing values and removing null/replicated values. 3. Drawing Inferences from the data. 4. Data Transformation 5. Performing Regression Analysis ➢ Linear Regression ➢ Support Vector Machine ➢ Lasso Regression ➢ Ridge Regression ➢ K- Nearest Neighbor Regression ➢ Elastic Net Regression 6. Conclusion ➢ All model's visualization ➢ Checking Accuracy and Results
Please paste the GitHub Repo link.
Github Link:- https://github.com/ParulSrivastava98/Yes-Bank-Stock-Closing-Price-Prediction Google drive Link:- https://drive.google.com/drive/folders/1UWmYt7SMD06YK9tW9FwQ45PLPILhqbhX?usp=sharing
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Yes Bank is a commercial bank which provides variety of services like corporate banking, internet banking and personal banking for its customers. Here we are provided with data set which comprises of following:

OPEN: Open is the price at which the financial security opens in the market when trading begins. It may or may not be different from the previous day.

HIGH: It is the highest price at which the stock traded that day in the market.

LOW: It is the lowest price at which the stock traded that day in the market.

CLOSE: Close is the price at which the financial security closes in the market.

DATE: Monthly observation of stock is provided in the data.

Firstly, in my capstone project I started with data analyzing by using data wrangling operations to know more about our data. Further I performed various operations on raw data to gather process and transform it and draw insights from it.

Secondly, I did data cleaning and checked for any null, nan or replicated values
Thirdly, after that I performed data transformation operations by splitting it into dependent and independent variables using train test splitting method. Followed by this I performed Linear Regression, Support Vector Machine, Lasso Regression, Ridge Regression, K- Nearest Neighbor Regression and Elastic Net Regression.

After that I was able to draw various conclusions from regression analysis as following: -

1. The mean square error values for linear regression, support vector machine, ridge regression and elastic test prediction were somewhat same.
2. The root mean square values of linear regression and ridge regression were somewhat same.
3. Lasso regression gave highest mean square and highest root mean square value.
4. While on the other hand lasso regression gave lowest mean absolute error an R2 score.
5. The highest R2 value was of K-nearest neighbour as it gave 98% accuracy.
6. Also, for mean absolute error K-nearest neighbour gave the highest accuracy of 99%.
7. Accuracy was mostly same in the case of elastic net regression and support vector machine.
8. The maximum accuracy attained was 99%.