**DATE: 24-8-2021**

Course 7 Final Activity

Prediction of marks of a student based on the number of hours he/she studies

<https://raw.githubusercontent.com/Kushal997-das/THE-SPARKS-FOUNDATION/master/Prediction%20using%20Supervised%20ML/student_scores%20-%20student_scores.csv>

| Task |
| --- |
| Read the data from the provided link |
| Understand the data by reading first 5 and last 5 records from the dataset |
| Check the shape of data |
| Check if there is any null values |
| Display the summary of the dataset |
| Group the data using score and hours, then Using bar chart ,plot score vs hours. The next phase is to enter distribution scores and plot them according to the requirement, here we are going to enter the title, x\_label, and y\_label, and show it according to the desired result. |
| Visualize distribution of data using scatter plot |
| Prepare the data by splitting it into Training and Test set |
| Plot the regression line |
| Plot the training and test data . Predict the score and compare predicted and actual score |
| Predict accuracy using r2square |
| Predict the percentage grade of a student if study hours are 9.25 |
| Evaluate the model: Calculate Mean absolute error, mean squared error, root mean squared error |
| Import seaborn and plot the distribution as well |
| Print actual vs predicted percentage score |
| Apply ridge and grid search to improvise your model |