**Task-1**

Download the following dataset (Remove the column **‘extracurricular activities’**) and apply multiple linear regression using the inbuilt method and find the coefficient (beta) matrix.

[Student Performance (Multiple Linear Regression) (kaggle.com)](https://www.kaggle.com/datasets/nikhil7280/student-performance-multiple-linear-regression?resource=download)

Now divide the dataset into train and test sets and implement the multiple linear regression from scratch using Batch Gradient Descent algorithm, to find the optimized values of coefficients.

Compare the outcomes of both methods.

Further, implement the Stochastic gradient descent and mini-batch (batch size = 32) gradient descent on the above dataset from scratch and observe the differences in the outputs.

Plot the models found by each variant of gradient descent on the train and test data. Also calculate the errors, R2 score, etc. for each method.

According to your experiment, which version of gradient descent is performing best on this dataset and why?