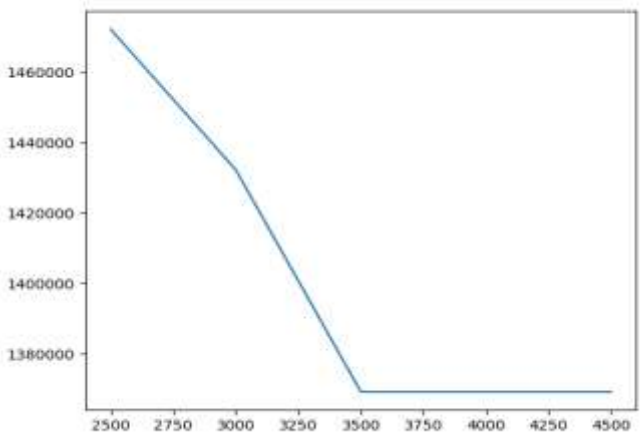
Assignment 1: Linear Regression

1. Gradient Descent implemented in code, for unregularized least squares model. Weight vector starts from all zeros vector.
    a. The absolute difference between the RMSE of closed solution and gradient descent solution is ***229058.869***
    b. The stopping criteria is either maximum iterations completed, or relative error between consecutive computed costs are low enough.
    c. The absolute difference between the RMSE of batch gradient descent and stochastic gradient descent is ***960155.773***

2. Implemented in code, set the value of lambda and learning rate to the value to be used for grading.

3. For basis functions, I have used the cube of mileage instead of the mileage, $(km\_driven)^2$ instead of km_driven and the square of engine instead of engine. The implementation is added to the get_features_basis() function and can be used for all 3 datasets to get basis feature matrix.

4. The plot of RMSE vs training examples used is given below. The Y Axis is the RMSE on the df_val.csv instances. The X axis is the size of the training set [2500, 3000, 3500, 4500]. I have used matplotlib to make this specific plot.



5. Using the weights of the vectors and some intuition, I would say that number of seats of the car and the seller are the 2 least useful features in my linear regression model. The task5() function was used to figure out the least useful features. Since many features are one hot encoded, It is not just so easy to see using the weights which are the least useful features, so looking through weights and using some intuition, the number of seats of the car and seller_type are the least useful features of this model.

6. The most useful enhancements according to me, for this model, are data enhancements using basis functions and normalization, these are already applied by me in the previous parts, so no additional enhancements have been applied.