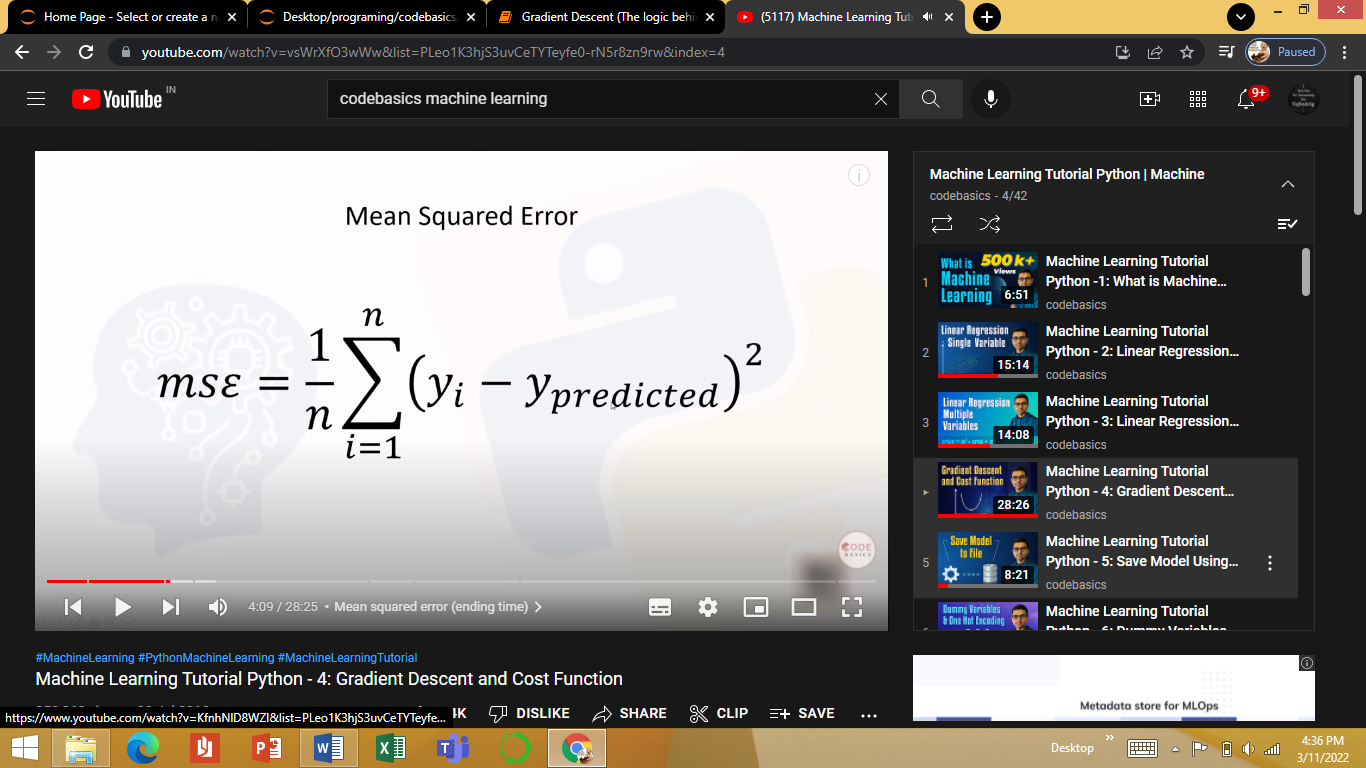
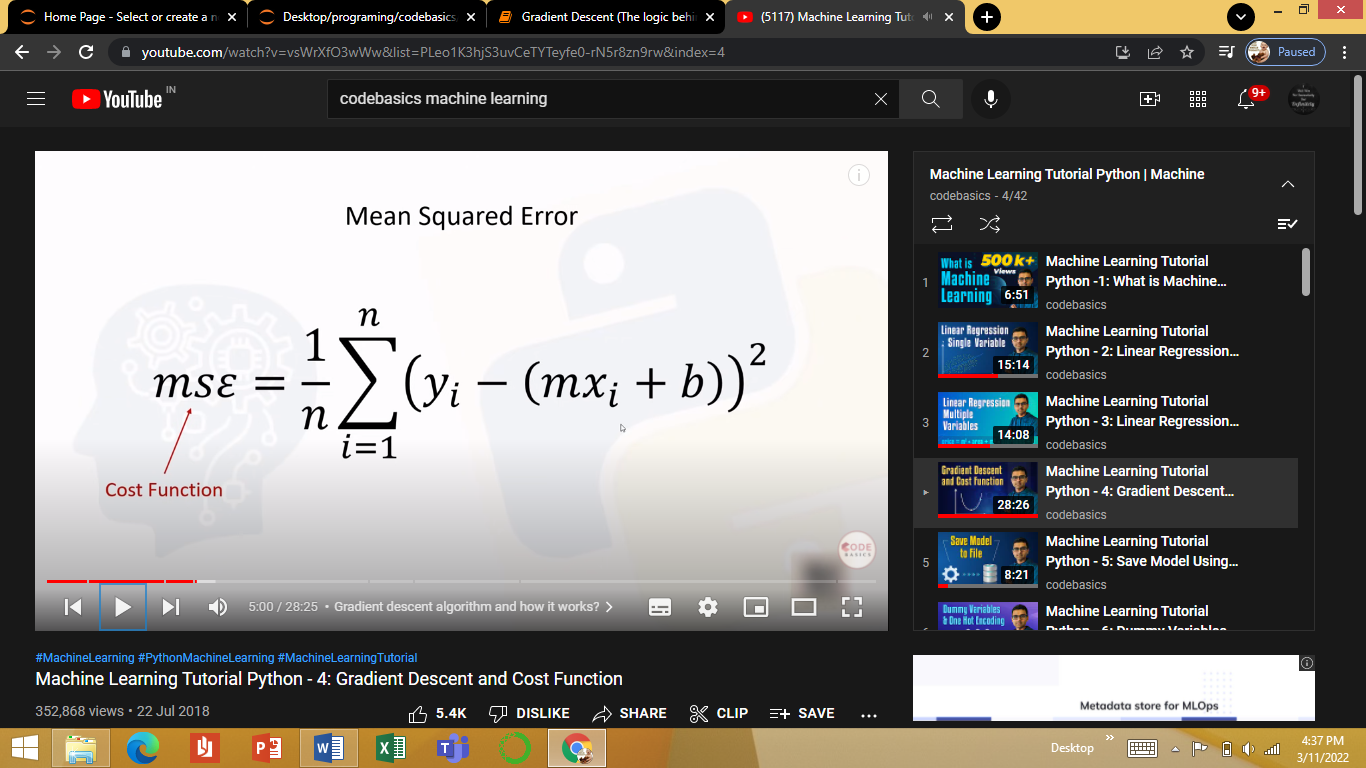
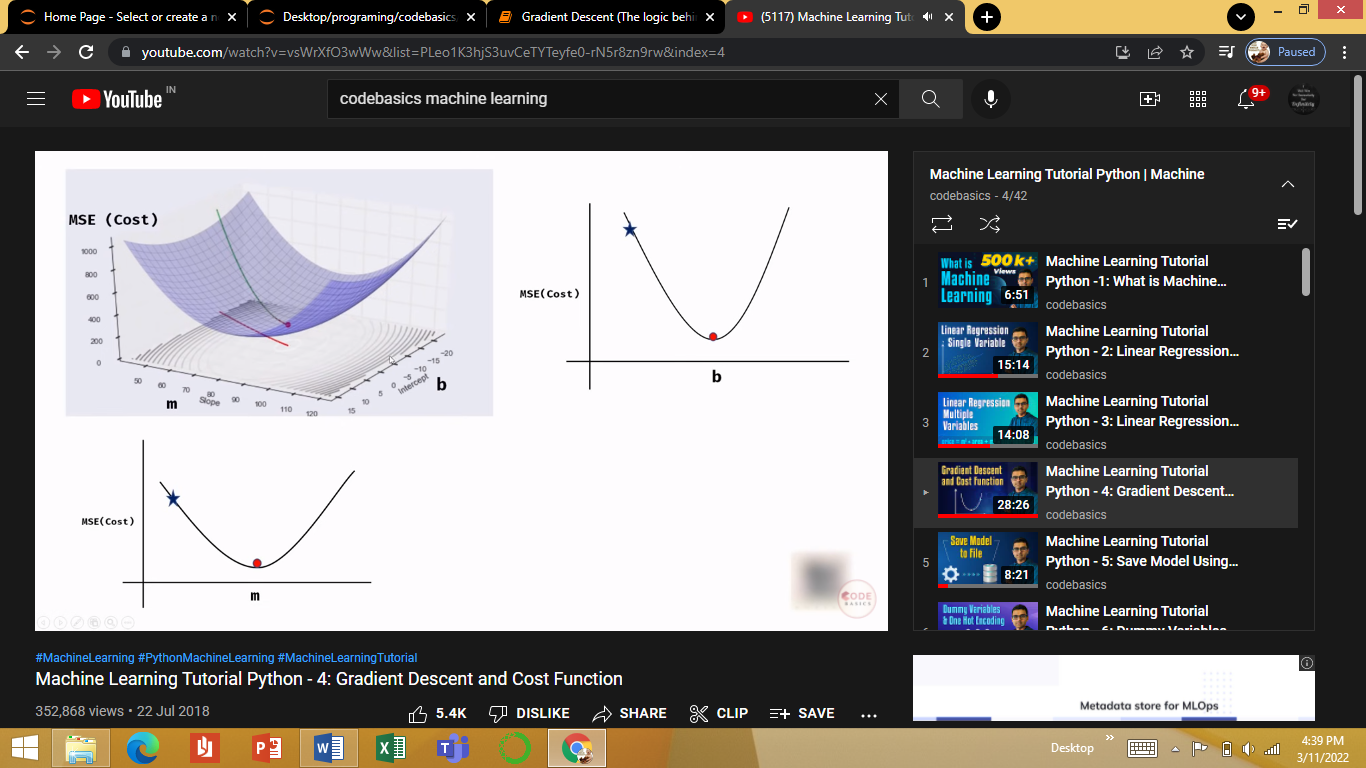
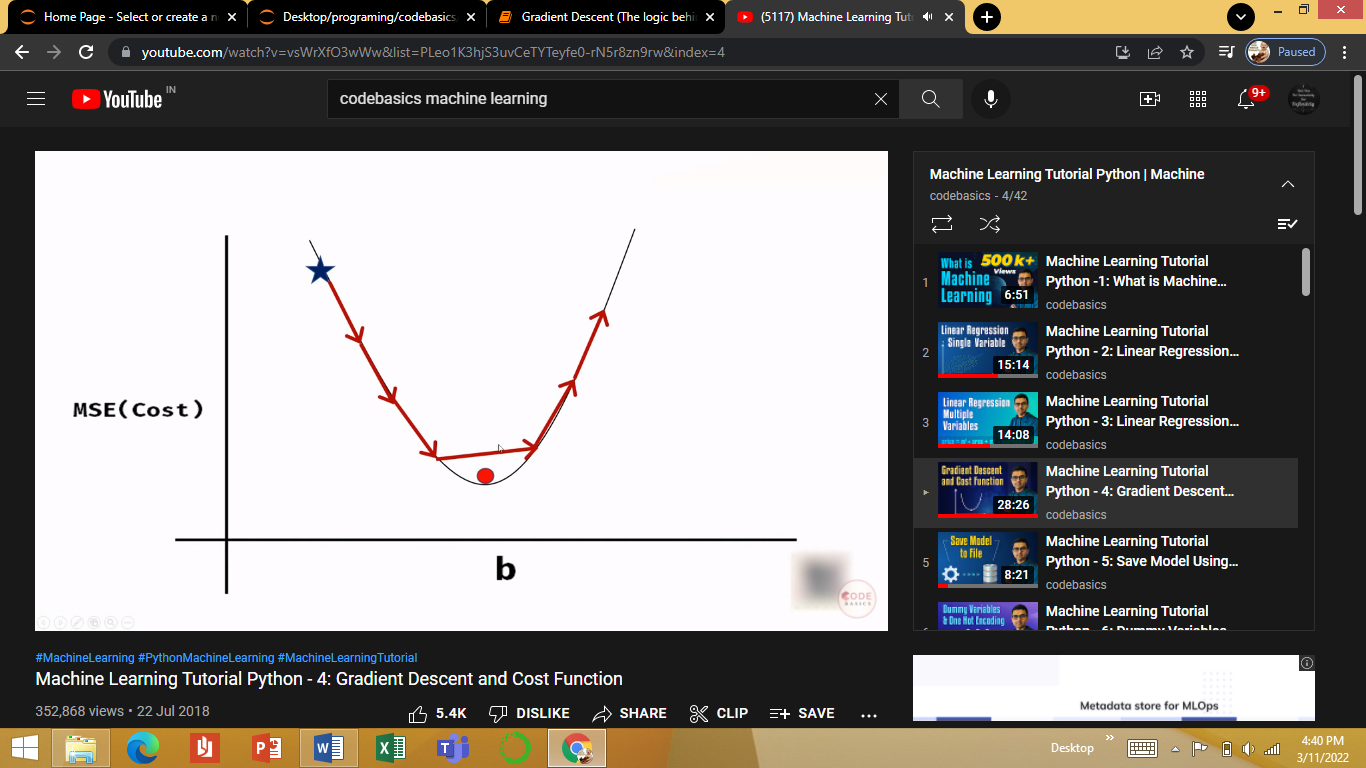
For Linear Regression, the best fit line is the one where the mean squared error between the line (i.e. predicted) and the actual points is the minimum.

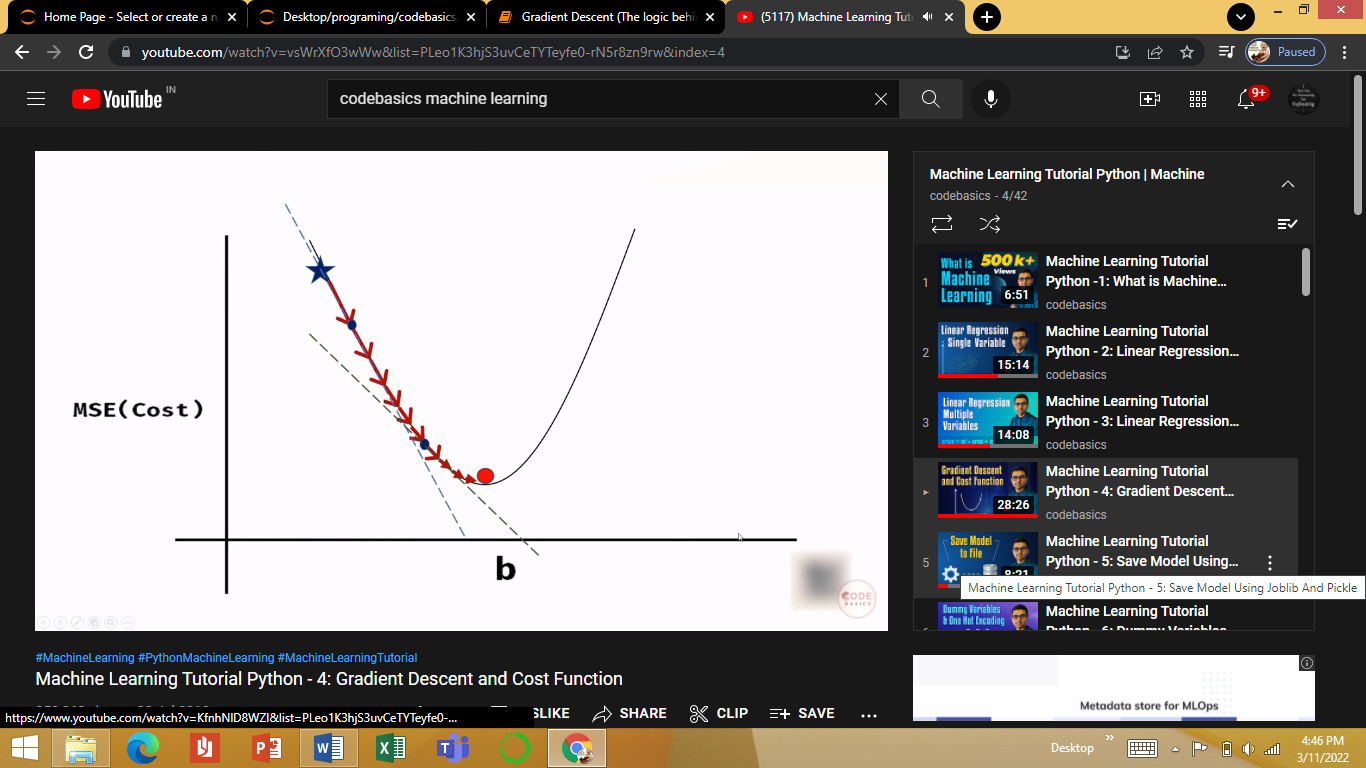








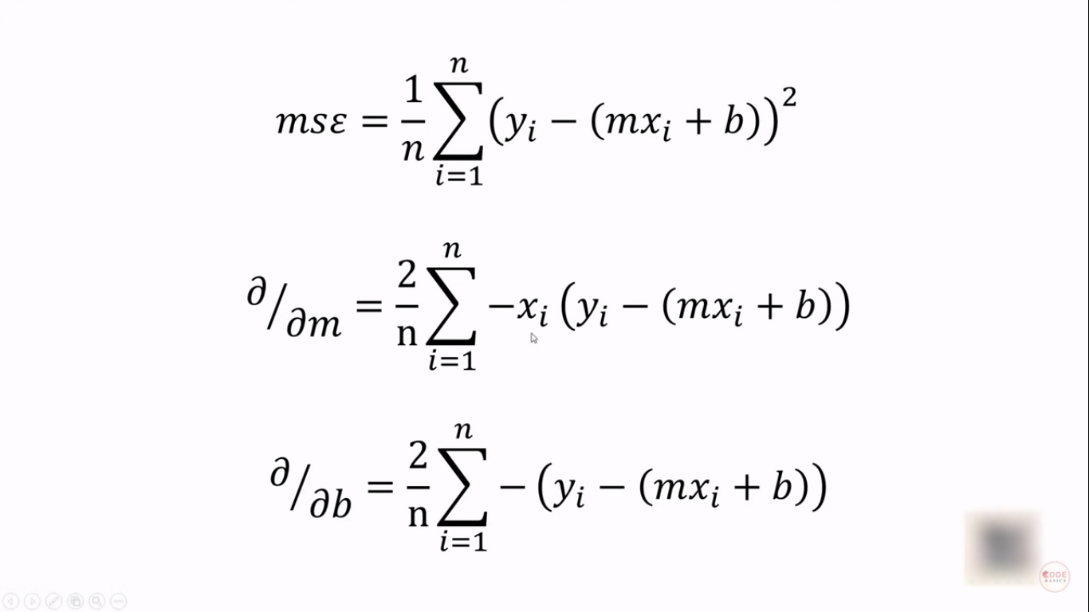
There are 2 approaches to minimize the cost

(A) by taking fixed sized steps along slope line…but in this case, there is a possibility of missing the global minima

(B) by taking baby steps… i.e. reducing the step size as we approach towards the global minima. After a threshold point after which the values of ‘m’ and ‘b’ along with ‘cost’ do not change significantly and be assumed more or less to be the minimum point

The size of each baby step is decided by takin the partial derivative of slope (m) and intercept (c) in the equation

Y = mX+c

 Starting values for ‘m’ and ‘c’ need to be assumed (say 0)

Learning rate for each model is also assumed / found by trial and error unless and until the cost does not increase after a particular point (which means that we have missed the global minima)

Gradient Descent is an iterative process so as to find the best fit line for a given set of points

