MLL100_Experiment 9

Problem 1:

$$x^2 = Dt$$

x is distance, *D* is diffusivity, *t* is time.

The data set is: ‘diffData.csv’. 2nd column is: time, 3rd column is: distance, 4th column is: Diffusion, a classification among bulk diffusion (Db), and grain boundary diffusion (Dgb)

(1) Predict the bulk diffusivity value using linear regression
(2) Predict the grain boundary diffusivity value using linear regression

Problem 2:

(1) Use Logistic regression technique to classify bulk diffusion and grain boundary diffusion. Input data is: distance and time columns, target is diffusion column. Report the confusion matrix, classification report.
(2) Now scale or normalize the input data, i.e., $x_{new} = \frac{x - mean}{std}$

and repeat the same technique. Report the confusion matrix and classification report.
Hint: To scale the data, the code in python is:
newXinput=(xInput-xInput.mean(axis = 0))/(xInput.std(axis = 0))