Hw 1 Linear Regression & Logistic Regression

a) Plot made with a 0.02 learning rate for 100000 epochs.

b) My RMSE was around 1400 with the above hyper parameters

c) We have too high of a learning rate maybe.

Potentially we may be overfitting as it learned to specify itself to the training data. Lower the epochs.

d) That would mean our data is over fitting

e) I would say that the output has external factors that the dataset does not have and both models have gotten as close as they can to the output.

2) a) Yes we can interperet this.

T is linearly dependent on A

T can be represented by a logistic curve in case B

T has a linear represetation with c where c has an output range for each input

T is linearly dependent on D

We can say with the P-value of B that it does not have much of an influence on T

b) We can use linear regression to find the best fit and weights assigned from the regression. Then we can use this to predict future values

c) We can tell that A & D will have coefficients that are close to its multiples. B will have large changes and its coeffecient and will have less of an impact on target. This coefficient will be contributor to error.

3) With my implementation we get a great accuracy and f1 score at 99% and 98% respectively. This sounds amazing but no it actually is bad. We overfit to all yes. We also tested on a near all yes test set We trained with 1 no and we tested with 10 no's. My FP were 9 but there are only 10 data points of no's so we just made those contribute less to the overall score.

My code has a better set to fix this commented out.

iterations: 10000
learning rate: 0.0003

Standard normalized

TP: 357   Precision: 0.97
TN: 1   Recall: 1.0
FP: 9   F1:   0.98
FN: 0   Accuracy: 0.97