COMP3105 Assignment 1 Report

Name (Student ID):

Ho Nam Leung (101197127)

Ryan Lo (101117765)

Q1e) Looking at your tables from above, analyze the results and discuss any findings you may have and the possible reason behind them.

For the models, the losses shown are ranked by $\mathrm{L}_1 < \mathrm{L}_2 < \mathrm{L}_\infty$,

which in other words, the smaller the n for $L_n - model$,

the less loss is found on the data. This is because the smaller the n is for the model, the less affected the model would be by outliers,

which in turn lowers the average loss for the model as more data points will be closer to the model.

For the losses, they are ranked by $L_1 < L_2 < L_\infty$ in terms of size,

which in other words, the larger the n for $L_n - loss$,

the larger the loss becomes. This is because the n of a loss is determined by the power of the loss, therefore it enlarges the loss for every data point,

causing the higher n values resulting in higher loss values.

```
Table 1
[[0.0159739  0.14168156 0.41768878]
 [0.01795846 0.13337436 0.49642789]
 [0.0215959  0.17661856 0.311648  ]]
Table 2
[[0.84637828 0.1793525  0.7649751 ]
 [0.94894335 0.18938178 0.81441199]
 [1.07014054 0.19996144 0.85679535]]
```

Q2e) Looking at your tables from above, analyze the results and discuss any findings you may have and the possible reason behind them.

m (Number of Data Points) : With both the training and test accuracies, as the number of data points (m) increases, we can see that the accuracies also increases. This is pretty much expected because with more data points the model has a larger training set to learn from which leads to higher and better accuracies.

d (Dimensionality of Features): As the dimensions of the features, d, increases, the training and the test accuracies decreases. As the dimensional spaces increases, the data becomes more sparse which makes it harder to model meaningful patterns

eta (Learning Rate): The smaller the learning rate, eta, is the higher the training and test accuracies. The higher the learning rate is the more possibilities that the data can overshoot during the gradient descent.

```
table 5: training accuracies
[[0.943   0.84335 0.92465]
 [0.9243  0.9215  0.9209 ]
 [0.9209  0.98065 0.9251 ]
 [0.9186  0.99955 0.89825]]
table 6: test accuracies
[[0.904  0.8465 0.927 ]
 [0.925  0.918  0.9215]
 [0.92   0.9755 0.923 ]
 [0.9185 0.998  0.8945]]
```

Q2g) Looking at your results here, which learning rate do you recommend to use for this dataset and why?

Based on the results, the learning rate of ‘1’ has a good balance between both the training and the validation accuracies. The learning rate of ‘10’ achieves higher training accuracy but performs worse on the validation and the test sets. The learning rate of ‘0.1’ performs well with the training accuracy but not so well with the validation and test accuracies. Therefore, the learning rate of ‘1’ has the best balance of all the accuracies.

```
Learning Rate: 0.1
Train Accuracy: 0.8434
Validation Accuracy: 0.7108
Test Accuracy: 0.7619

Learning Rate: 1
Train Accuracy: 0.8795
Validation Accuracy: 0.6506
Test Accuracy: 0.8095

Learning Rate: 10
Train Accuracy: 0.8916
Validation Accuracy: 0.6988
Test Accuracy: 0.7619

Learning Rate: 100
Train Accuracy: 0.6867
Validation Accuracy: 0.7108
Test Accuracy: 0.7619
```