**Logistic Regression CPSC 6430 Report**

**Lake Summers**

**Problem Description**

Given a data set representing capacitor quality control testing data, develop a logistic regression algorithm that will determine if a capacitor will fail or pass quality control.

**Data Description**

The initial data was a set of 118 examples showing capacitors failing and passing the quality control tests (Figure 1). Each record had three tab-separated entries. The first is a float representing the results of one test, the second being the result of another test. The third is either a 1.0 if the capacitor passed QC and a 0.0 if it failed QC. The data was then split up into two data sets: a training set of 85 examples, and a test set of 33 examples.

Figure

**Model Description**

The input files were altered similar to how you showed in the project 4 description. I went with thePower=4 and ended up with the relationship between the new features and the original features as shown in below.

A picture containing text, scoreboard, blackboard

Description automatically generated

**Initial Values:**

Weights:

[[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]

[0.]]

Alpha: 0.01

J: 0

**Final Values:**

Alpha: 0.01

Weights:

[[ 3.25661719]

[ 2.14481154]

[-6.11343653]

[ 2.03284481]

[-4.44128612]

[ 4.19820937]

[-3.88643454]

[-2.34653885]

[ 0.93768656]

[-1.06326493]

[-4.72565897]

[-1.21701778]

[-2.61145927]

[ 0.77493135]

[-1.14189647]

[-0.97221444]

[-1.85155329]

[-1.6442211 ]

[ 0.34796851]

[-0.68663579]

[-4.29876347]

[-1.72368966]

[-1.43744731]

[ 0.21530624]

[-0.57067768]]

Iterations: 110000

Final J on Training Set: 0.35020586661571024

**J vs Number of Iterations:**

A picture containing graphical user interface

Description automatically generated

Value of J on Test Set: 0.35327338

Calendar

Description automatically generated

**Results**

A confusion matrix for the results of the logistic regression is shown above.

**Final Values:**

Given the initial training data of 85 capacitors and the results of their quality control tests, the logistic regression model was able to correctly predict the outcome of a test of the capacitors in the training data (33 capacitors) 81.8% of the time.

Accuracy: 0.8181818181818182

Precision: 0.8181818181818182

Recall: 0.8235294117647058

F1: 0.8235294117647058