# 1 Written Exercises: Linear Regression and Precision/Recall

**Problem 1**
Q1 Linear Model with Laplace Error

**Solution**
Steps:

1. $y_i \sim Laplace(\mu = w^T x_i, b)$
2. $P(y_i \mid x_i, w) = \sum_{i=1}^{N} \frac{1}{2b} e^{\frac{-\left|y_i - w^T x_i\right|}{b}}$
3. Likelihood $= \prod_{i=1}^{N} P(y_i \mid x_i, w)$
4. Log-Likelihood $= \sum_{i=1}^{N} ln(\frac{1}{2b} e^{\frac{-\left|y_i - w^T x_i\right|}{b}})$
5. $= \sum_{i=1}^{N} [\frac{-\left|y_i - w^T x_i\right|}{b} - ln(2b)]$
6. $= \sum_{i=1}^{N} [-\frac{1}{b} \left|y_i - w^T x_i\right|] - Nln(2b)$
7. $= -\frac{1}{b} \sum_{i=1}^{N} [\left|y_i - w^T x_i\right|] - Nln(2b) \sim SAE = \sum_{i=1}^{N} \left|y_i - w^T x_i\right|$
8. ***Answer***: The probabilistic model with Laplace error also minimizes the SAE because their likelihoods are proportional.

**Problem 2**
Q2 Computing Recall and Precision

**Solution**

1. Case t=0: FP=8, FN=0, TP=8, TN=0. Recall=1.0, Precision=0.5.
2. Case t=0.2: FP=6, FN=2, TP=8, TN=0. Recall=0.8, Precision=0.57.
3. Case t=0.4: FP=3, FN=5, TP=6, TN=2. Recall=0.55, Precision=0.67.
4. Case t=0.6: FP=1, FN=7, TP=6, TN=2. Recall=0.46, Precision=0.86.
5. Case t=0.8: FP=1, FN=7, TP=4, TN=4. Recall=0.36, Precision=0.8.
6. Case t=1: FP=0, FN=8, TP=0, TN=8. Recall=0.0, Precision=0.0 (No TP or FP).

# 2 Implementing Logistic Regression for Tumor Diagnosis

**Problem 3**
Q3 Negative Log-Likelihood

**Solution**
Steps:

1. Completed

**Problem 4**
Q4 Gradient Descent for Logistic Regression

**Solution**
Steps:

1. Train Accuracy: 86.27%
2. Learned Weight Vector: [-0.2464, 0.8689, 0.2001, 0.2785, -0.6766, -0.3326, 0.434, -0.3499]

**Problem 5**
Q5 Adding a Dummy Variable

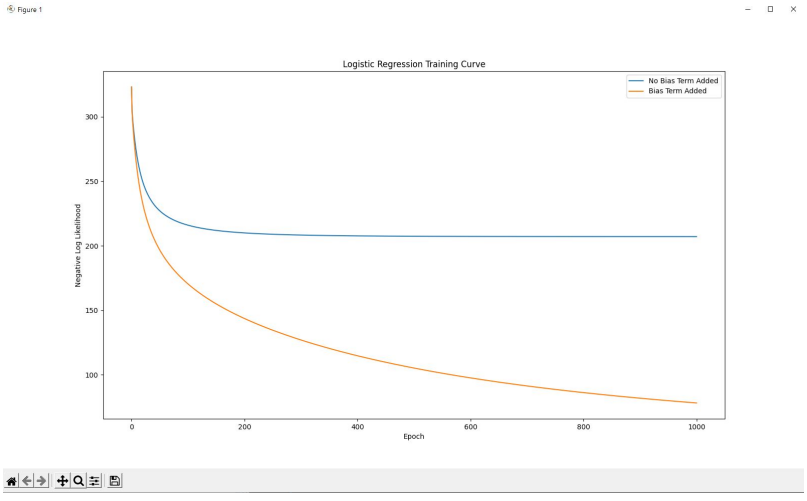
**Solution**
Steps:

1. Train Accuracy: 96.35%
2. Learned Weight Vector: [-3.4144, 0.08, 0.4192, 0.2177, 0.2745, -0.2522, 0.0561, 0.2724, -0.0528]
3. There is significant improvement from the No-Bias accuracy in the new Bias accuracy.

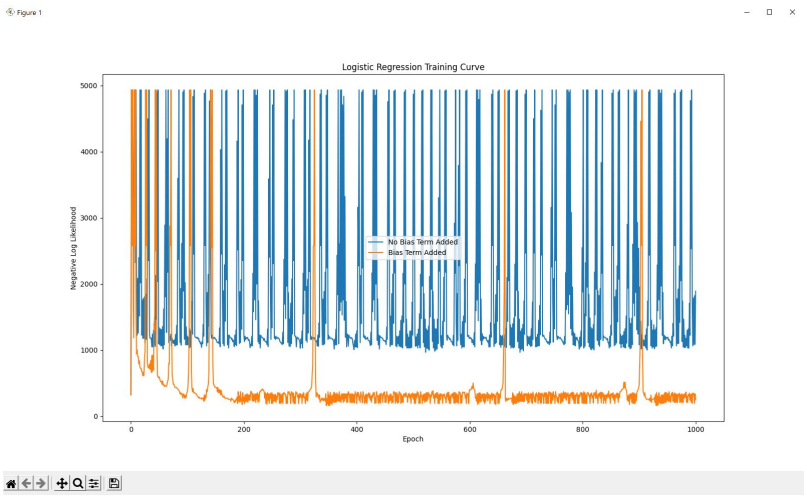
**Problem 6**
Q6 Learning Rates/ Step sizes

**Solution**
Steps:

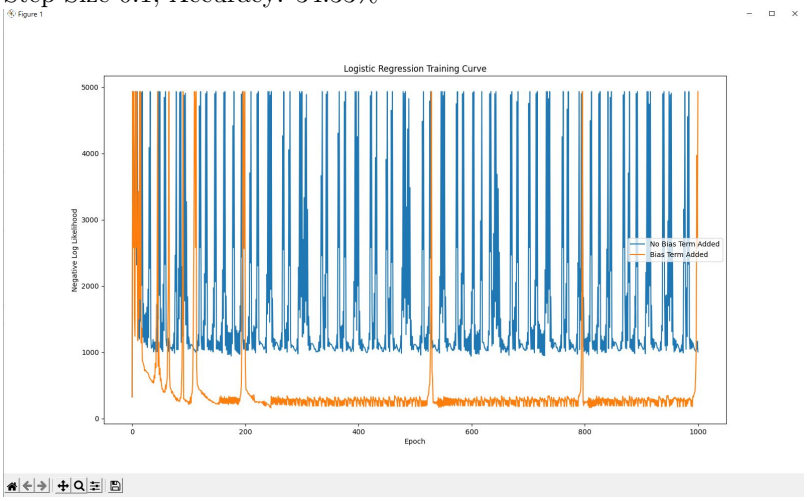
1. It looks like as the number of iterations increases, the gradient descent algorithm converges. It also appears that the Bias case converges more quickly than the non-Bias case.
   - Step Size 0.0001, Accuracy: 96.35%



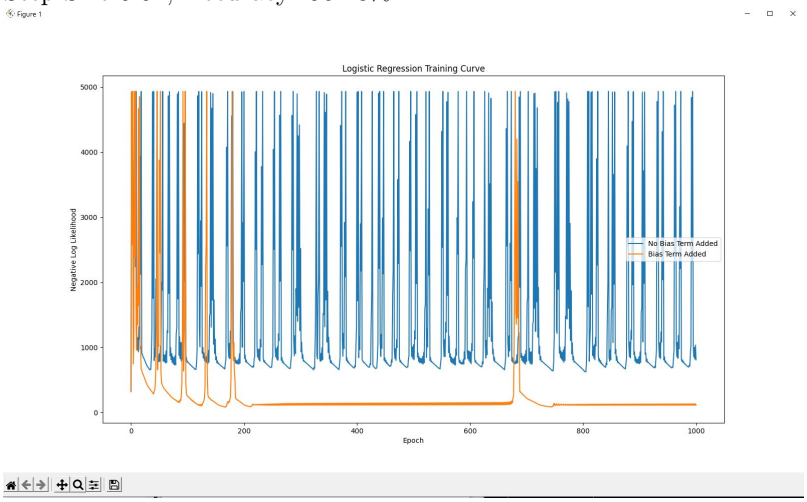

   - Step Size 1, Accuracy: 96.57%



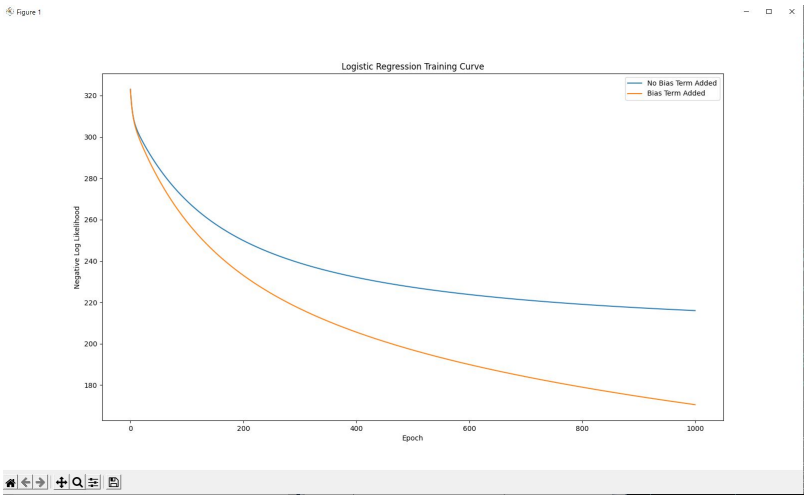

- Step Size 0.1, Accuracy: 34.33%




- Step Size 0.01, Accuracy: 95.28%




- Step Size 0.00001, Accuracy: 90.77%




2. It appears that as the step size gets smaller, the volatility of the Negative Log Likelihood also shrinks across all epochs but the accuracy still varies.

**Problem 7**
Q7 Evaluating Cross Validation

**Solution**
Steps:

1. The means and standard deviations from each K in K-fold cross validation roughly capture the performance of my submission on the Kaggle leader-board. In general, the higher the mean cross val accuracy, the better my score will be.

**Problem 8**
Q8 Kaggle Submission

**Solution**
Steps:

1. I modified the K-fold to run through each step size with a k of 50, compare their mean accuracy and standard deviations, and return the step with the best mean accuracy and standard deviation.

# 3 Debriefing

**Problem 1**
Approximately how many hours did you spend on this assignment?

**Solution**

1. About 10 hours.

**Problem 2**
Would you rate it as easy, moderate, or difficult?

**Solution**

1. Moderate.

**Problem 3**
Did you work on it mostly alone or did you discuss the problems with others?

**Solution**

1. Worked on it alone.

**Problem 4**
How deeply do you feel you understand the material it covers (0%–100%)?

**Solution**

1. I felt I understood about 85 percent of the material covered.

**Problem 4**
Any other comments?

**Solution**

1. None.