Problem 1

Below are some results of the trained model overlaid on the testing set with different choices
of orders:

Note: The "order" here refers to the total number of terms in

$$f(x; \boldsymbol{\theta}) = \sum_{i=0}^{N} \theta_i x^i$$

Therefore, the term with the highest power of x of a model with order d is
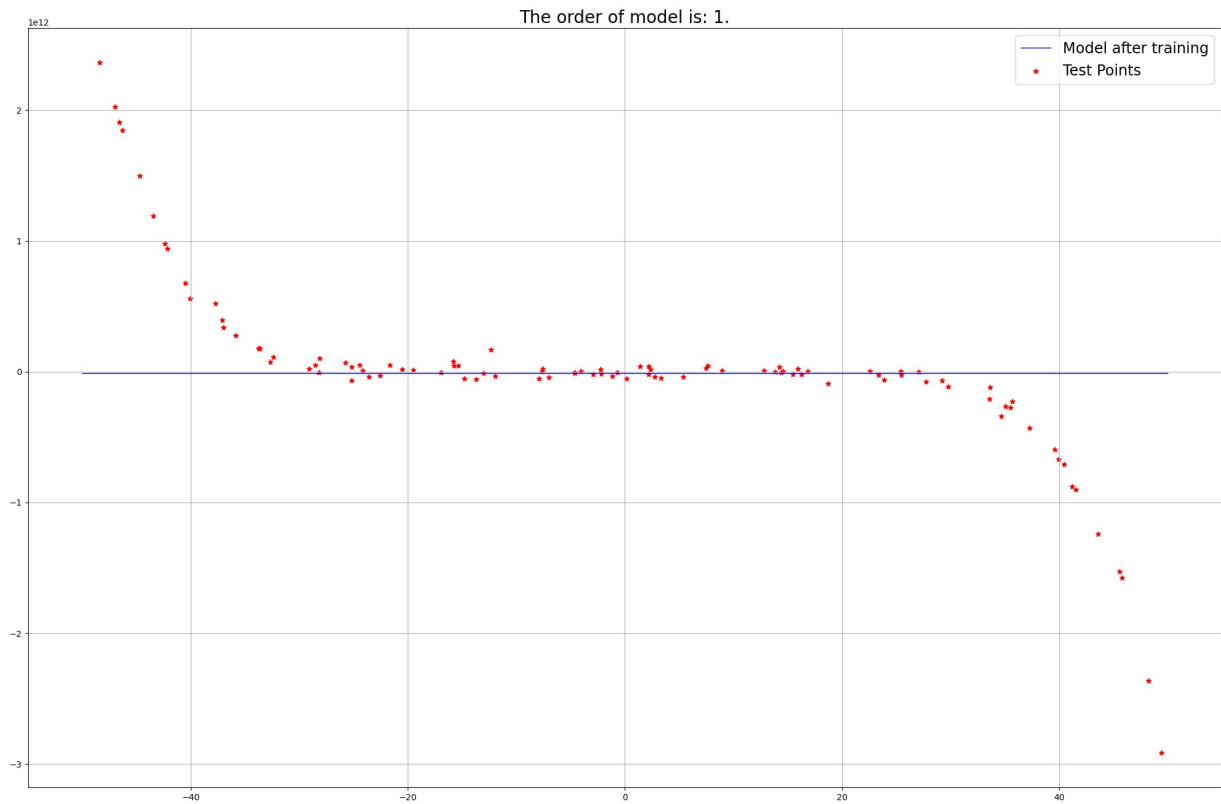
$$\theta_{d-1} x^{d-1}$$

Figure 1.1 Trained model with order $= 1$ overlaid on testing set
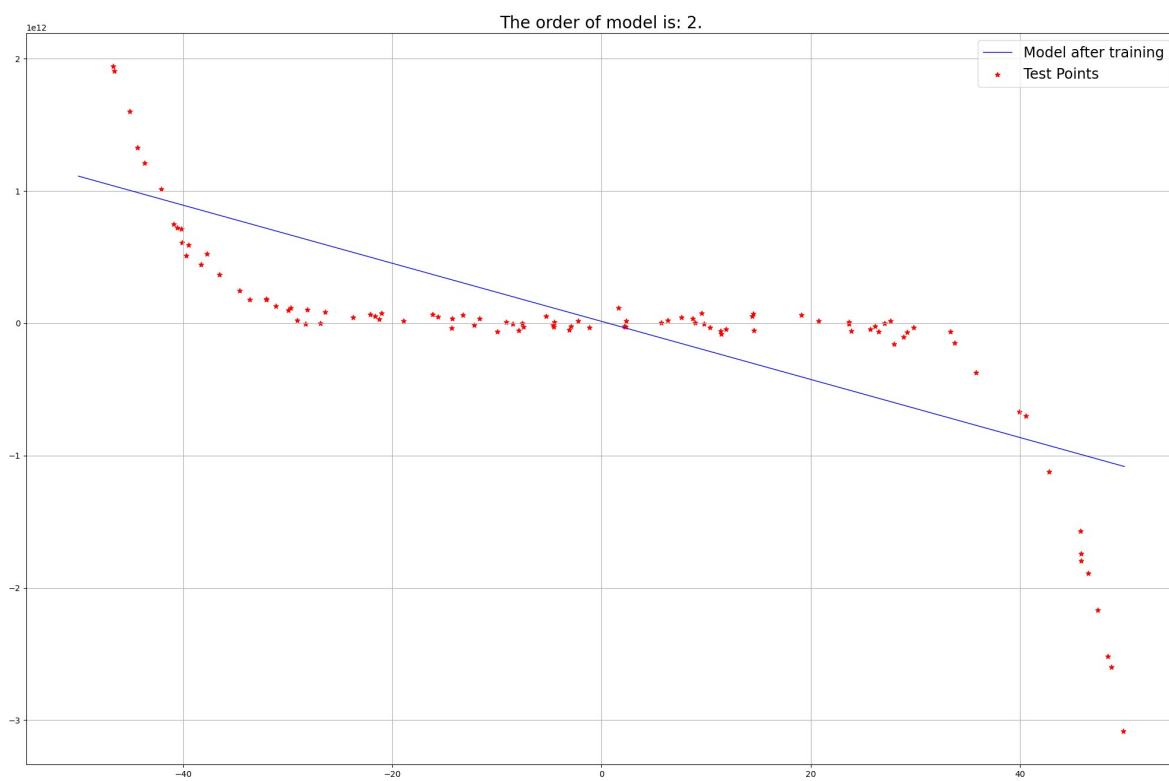
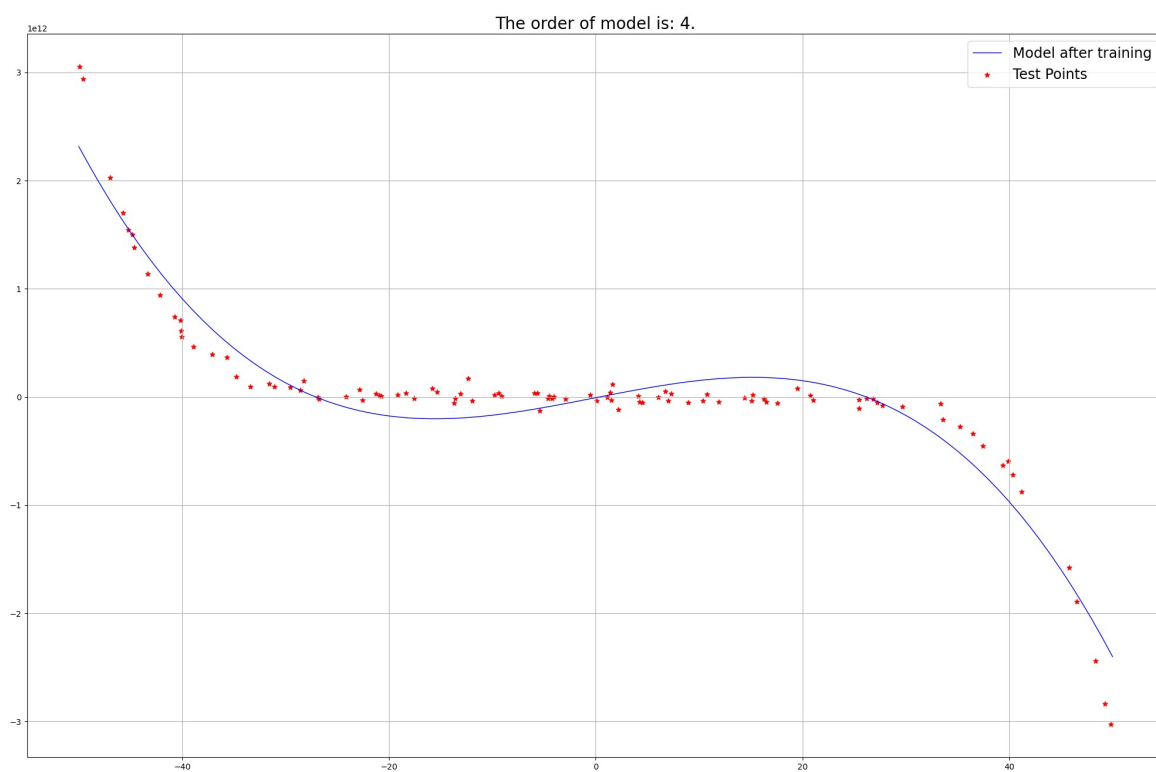Figure 1.2 Trained model with order = 2 overlaid on testing set

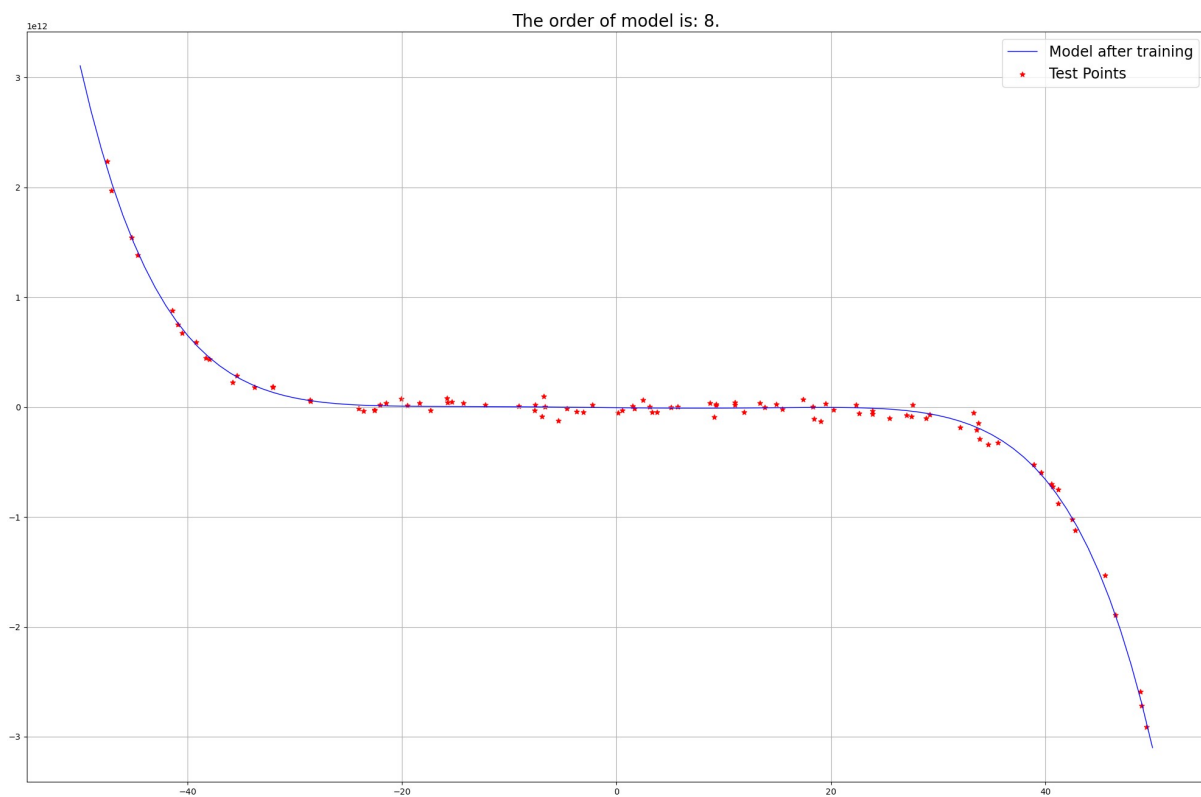Figure 1.3 Trained model with order = 4 overlaid on testing set

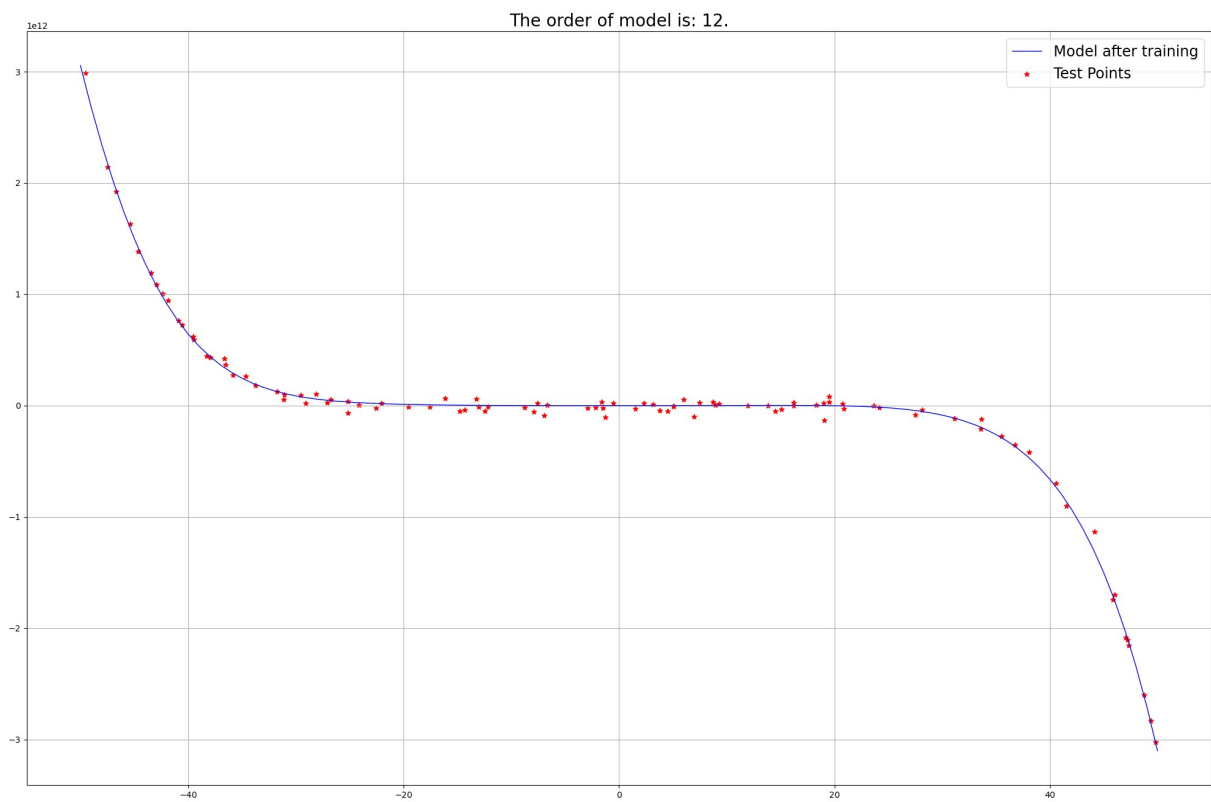Figure 1.4 Trained model with order = 8 overlaid on testing set

Figure 1.5 Trained model with order = 12 overlaid on testing set

Problem 4

Below are the results of the training process with different choices of parameters of the logistic regression model:
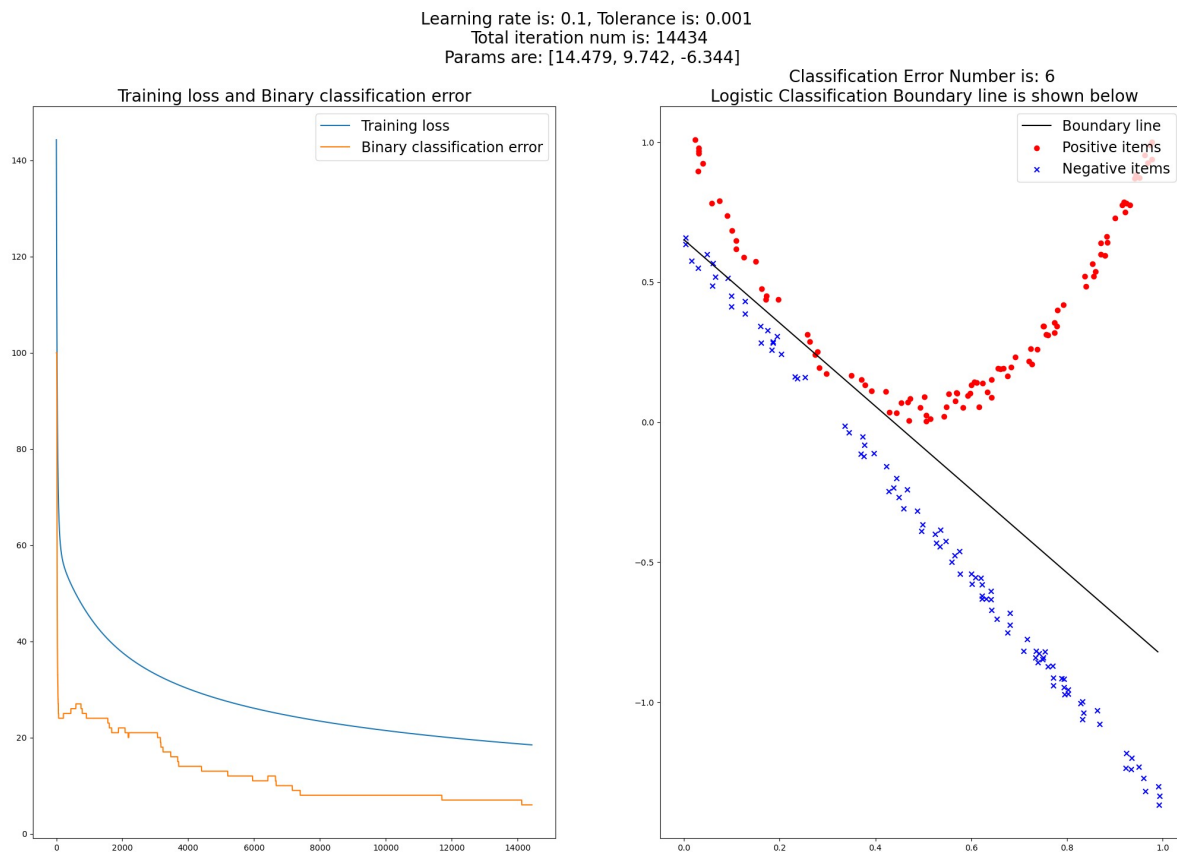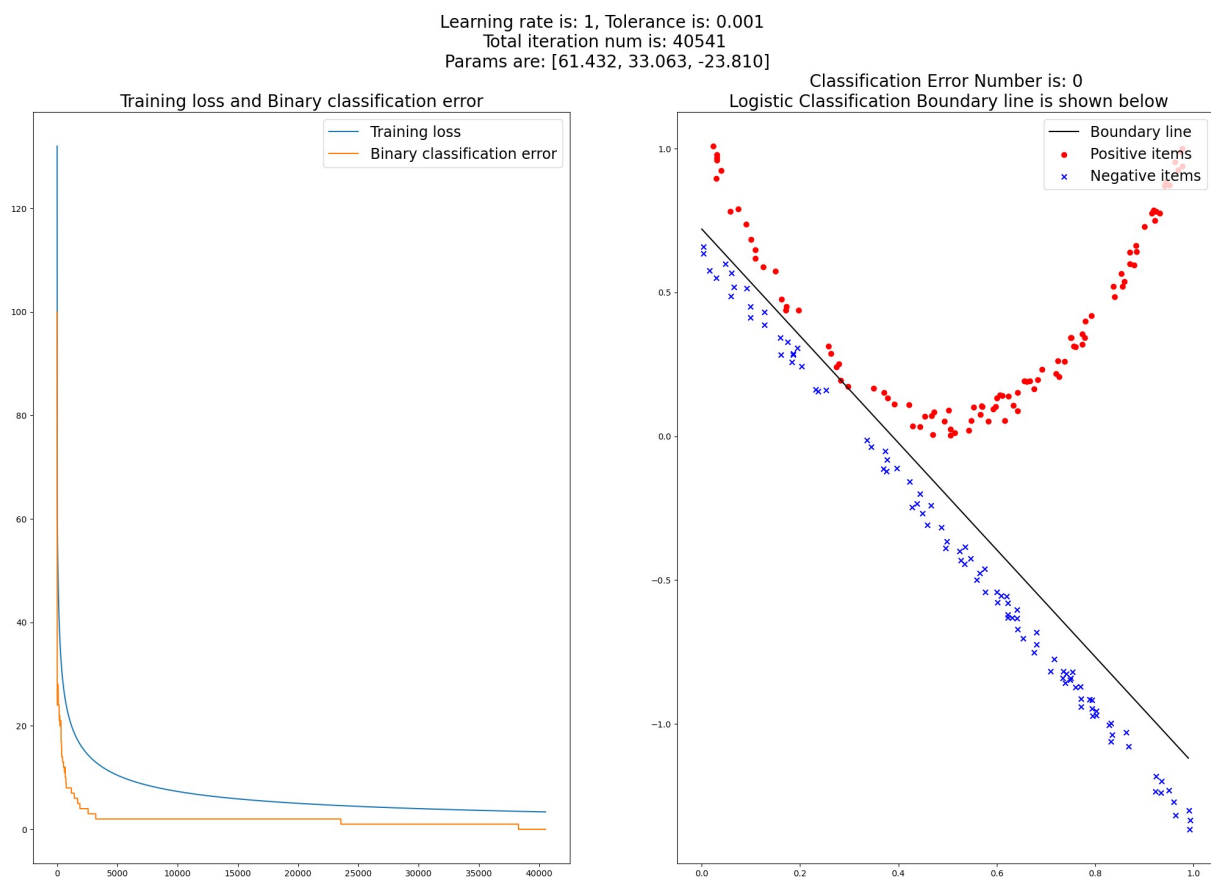

Figure 4.1 Training with learning rate 0.1 and tolerance 0.001

Figure 4.2 Training with learning rate 1 and tolerance 0.001
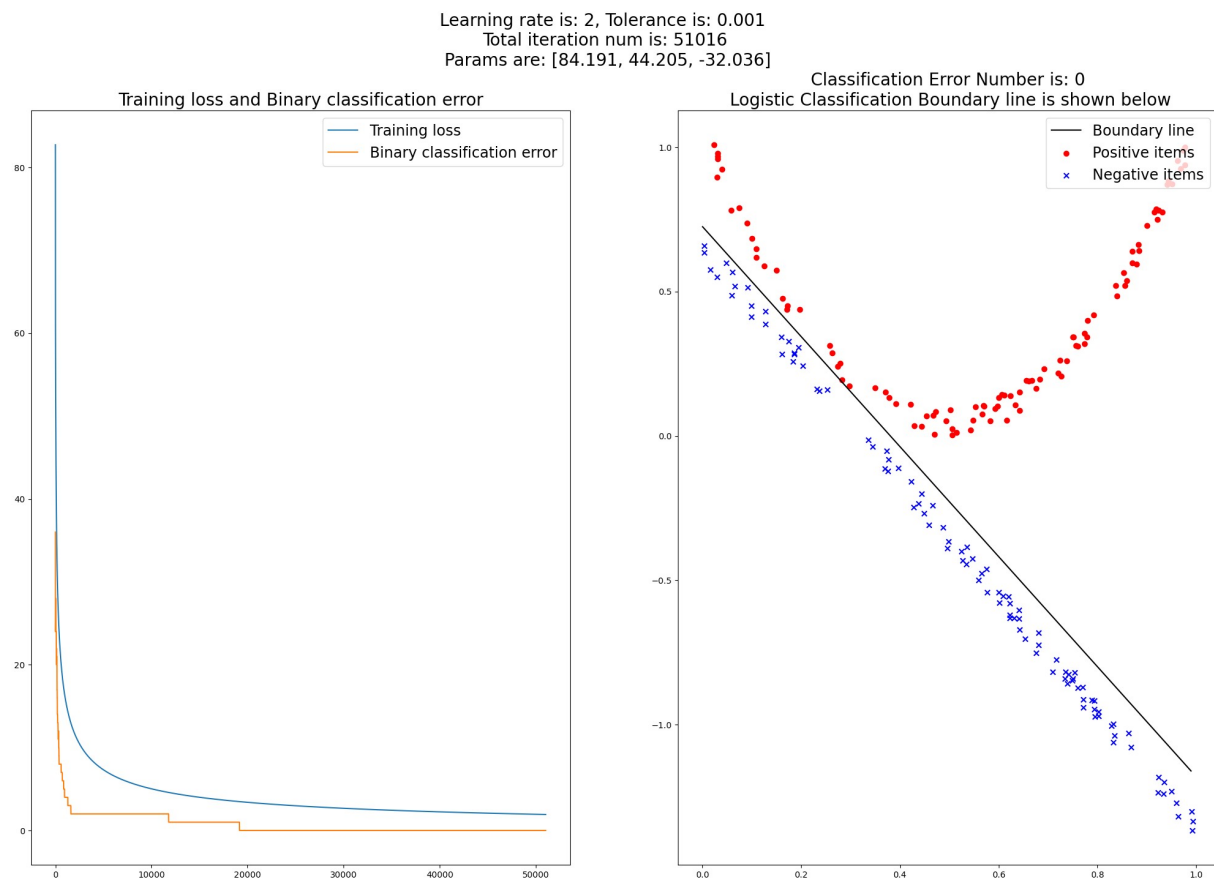
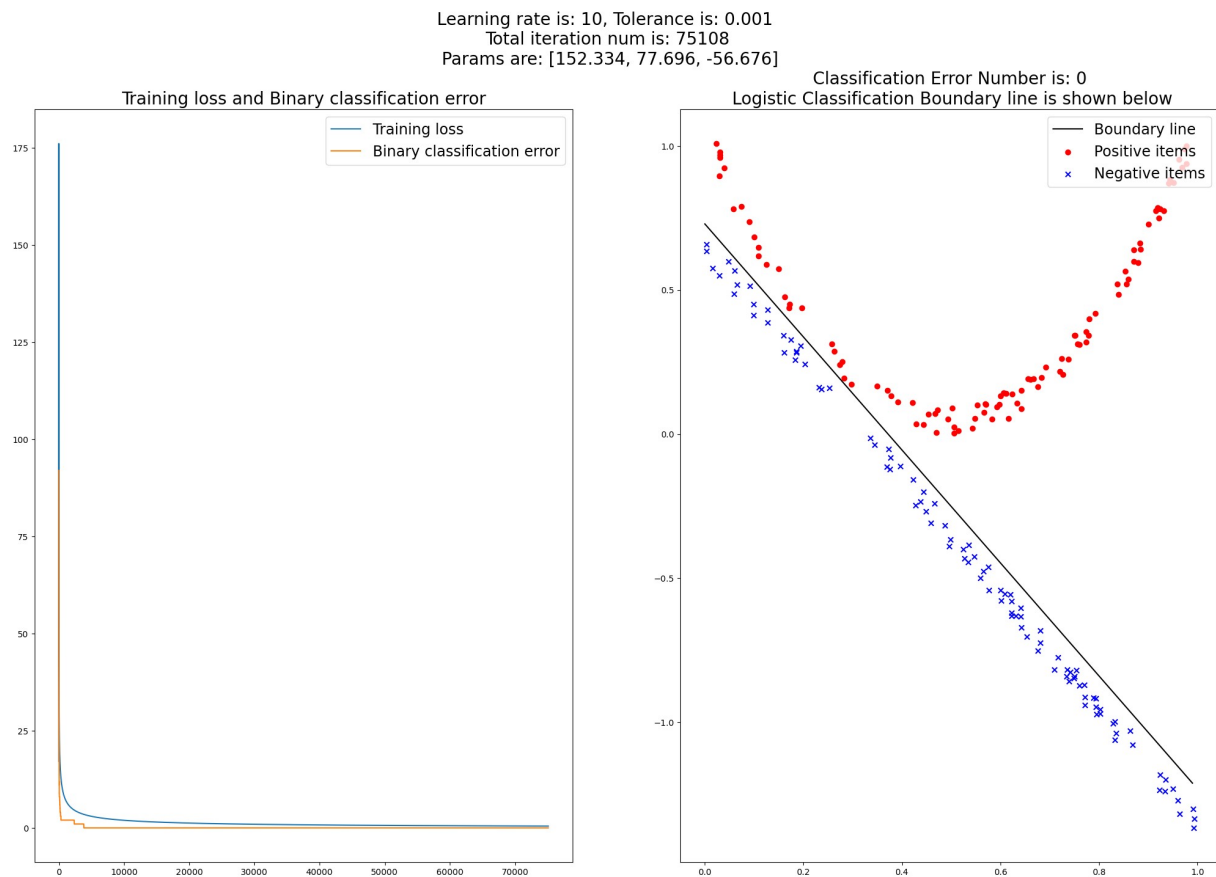Figure 4.3 Training with learning rate 2 and tolerance 0.001

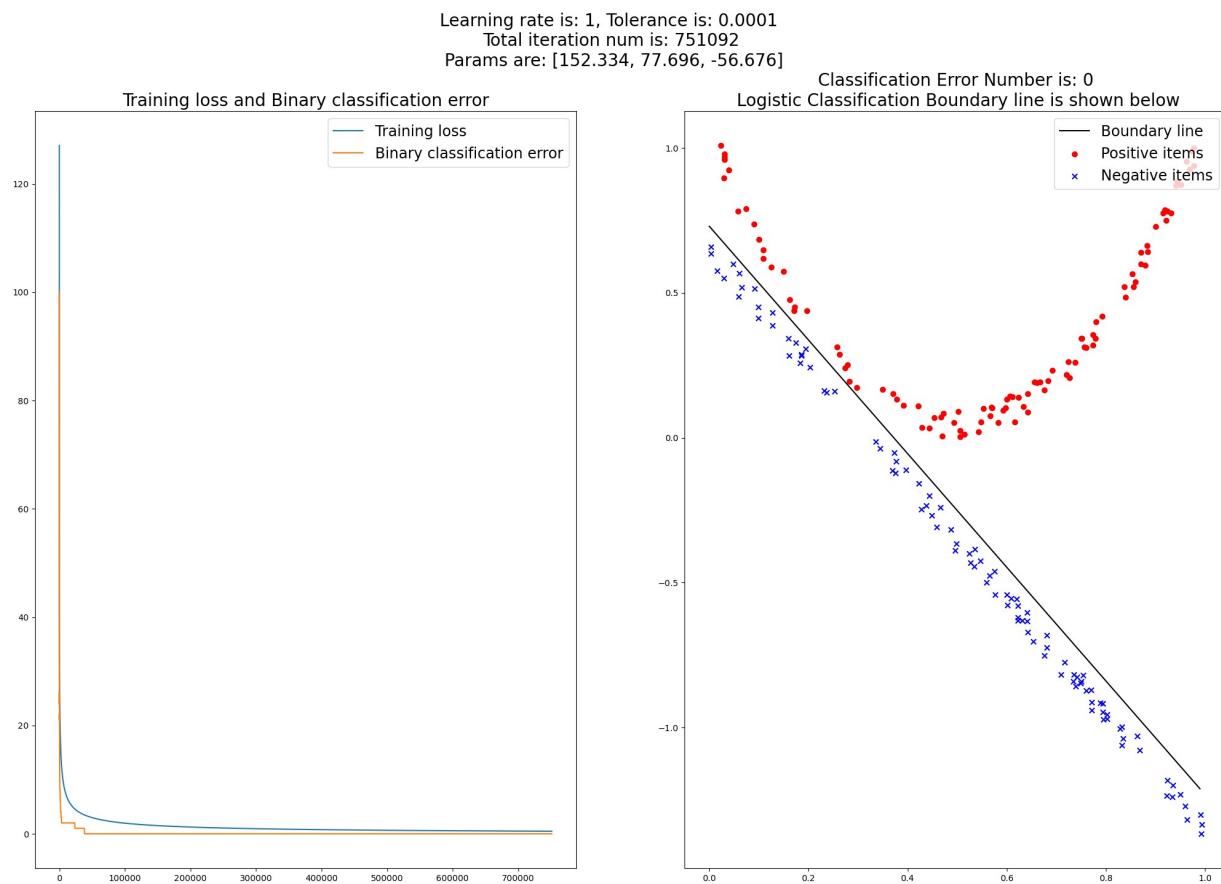Figure 4.4 Training with learning rate 10 and tolerance 0.001
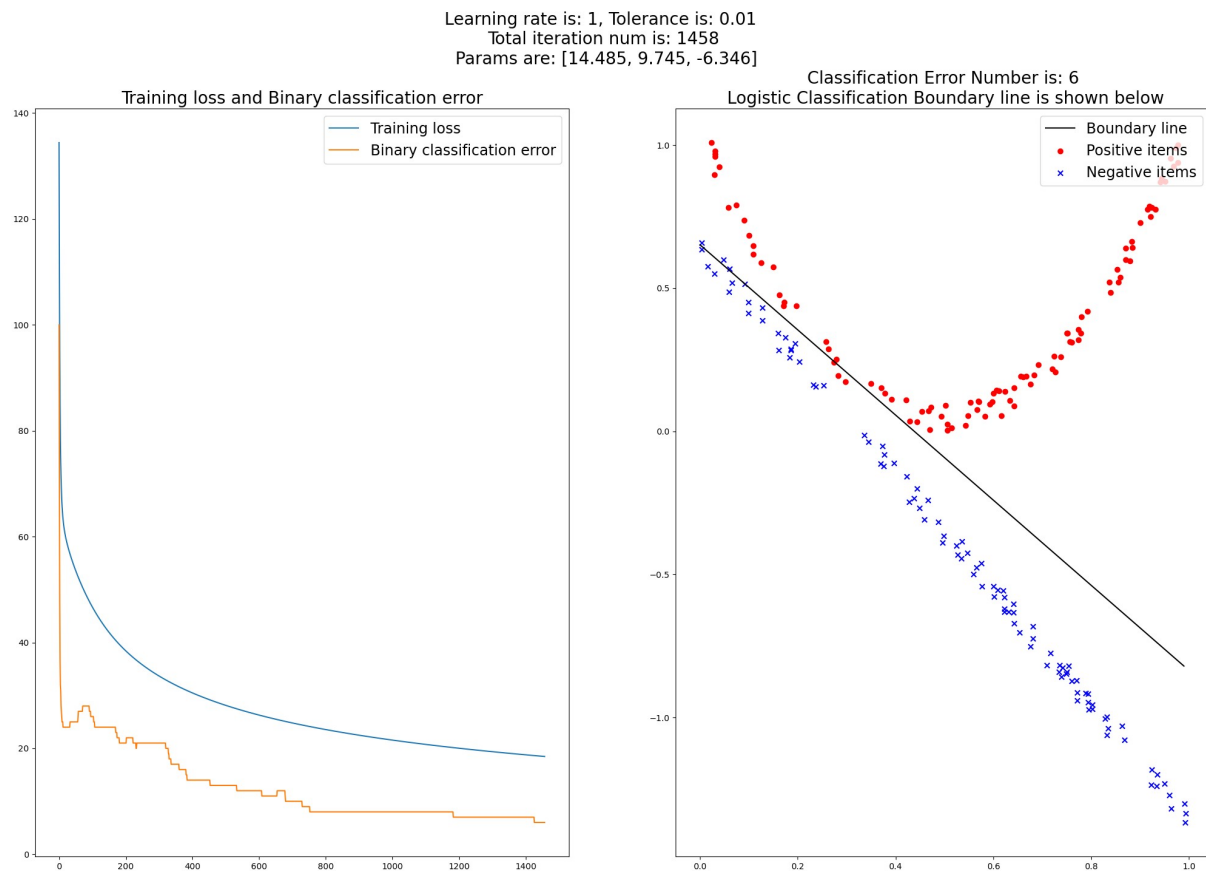
Figure 4.5 Training with learning rate 1 and tolerance 0.0001

Figure 4.6 Training with learning rate 1 and tolerance 0.01