Homework assignment: would you survive the Titanic?

Fabio Zamboni

3/04/23

1 Model Training

1.1 Which is a good value for the learning rate?

A good possible learning rate choice is 0.001

1.2 How many iterations are required to converge?

Set the learning rate as 0.001, the number of iteration required to converge are 2.75ML

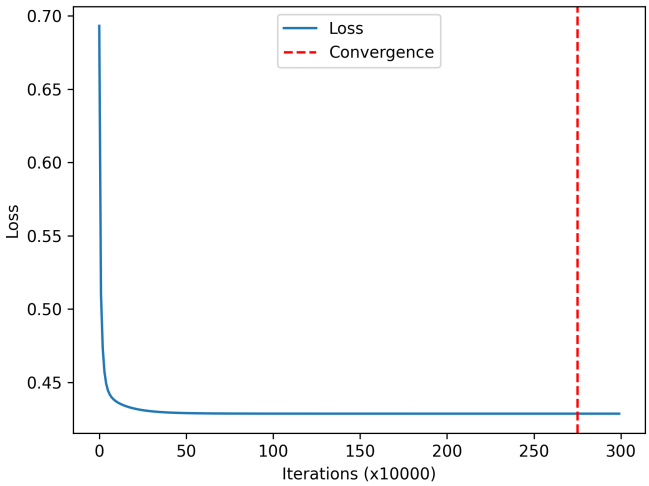

Figure 1: Loss Convergence

2 Model Analysis

2.1 What would be your probability to survive?

Unfortunately, no, I would not have survived or at least the chances of survival for me were very low: 13.08%. The probability has been calculated providing the following input to the model:

$$X = [3, 0, 23, 0, 0, 200]$$

2.2 What is the training accuracy of the trained model?

The trained model accuracy is about 0.811, in percent 81.1%

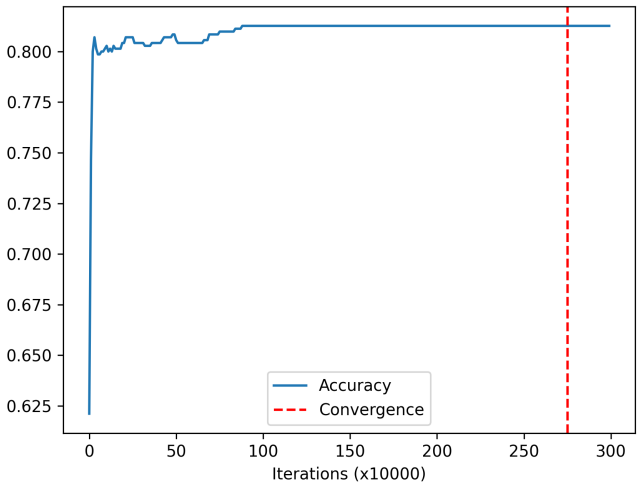

Figure 2: Accuracy

2.3 Looking at the learned weights, how the individual features influence the probability of surviving?

The weights vector w is the following:

$$w = \ [-1.297,\ 2.806,\ -0.047,\ -0.336,\ -0.089,\ 0.0007]$$

as we can see some weights are dominant, in order to have a more clear view we can look at the following figure, Fig.(3)

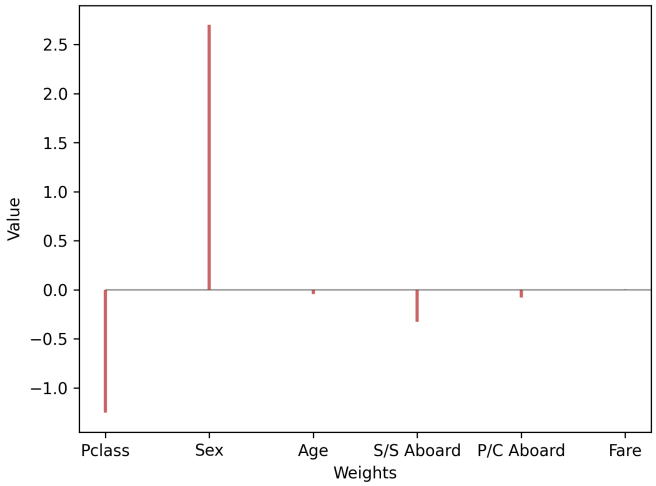

Figure 3: Weights

We can see that there are 2 main weights, the *Sex* that gives a positive contribute than the *Pclass* that gives a negative contribute. We can also see a discrete negative contribute from the *S/S Aboard*. The remaining weights contribute is really little.

2.4 What kind of passengers was most likely to survive? And what kind to die?

Looking at the Fig.(3) we can find the type passenger who was most likely to survive. In particular the passenger type that was most likely to survive is a first class young female without siblings, spouses, parents, children aboard. Instead a third class male with a high number of siblings, spouses, parents, children aboard have the worst probability to survive.

2.5 Scatter plot distribution of the two classes in the plane defined by the two most influential features.

Looking at the scatterplot we can confirm our last hypothesis.

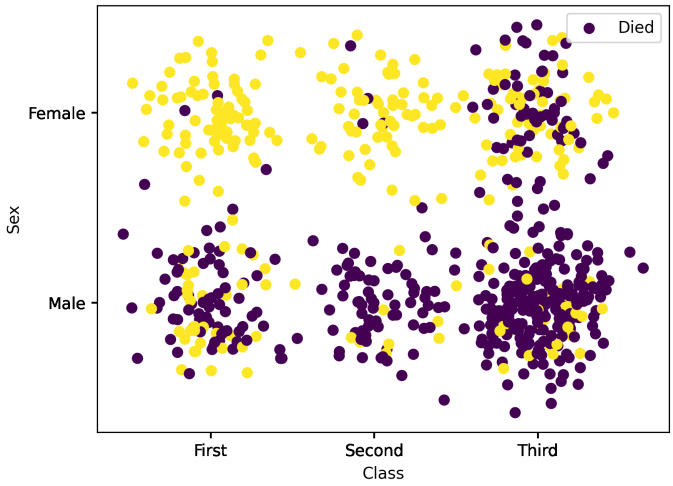

Figure 4: Sex vs Class

3 Model Evaluation

3.1 What is the test accuracy of the model?

The model accuracy on the test set is 0.7966, in percent 79.66%

3.2 Is the model overfitting or underfitting the training set?

Looking at the Fig.(6) we can see that the two curve looks good and there is no sign of either overfitting or underfitting.

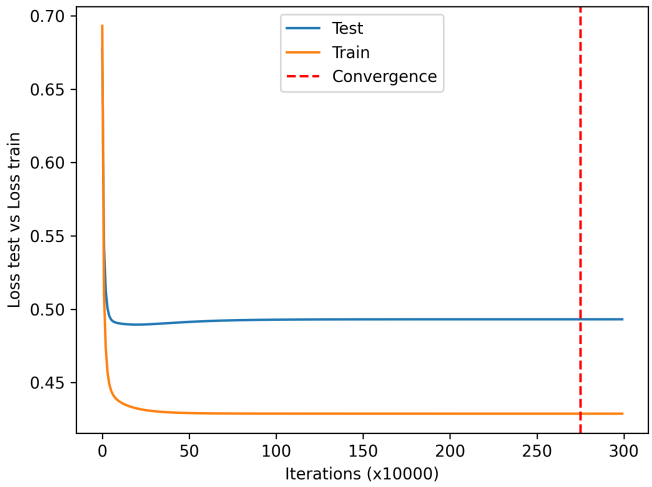

Figure 5: Loss Test vs Loss Train

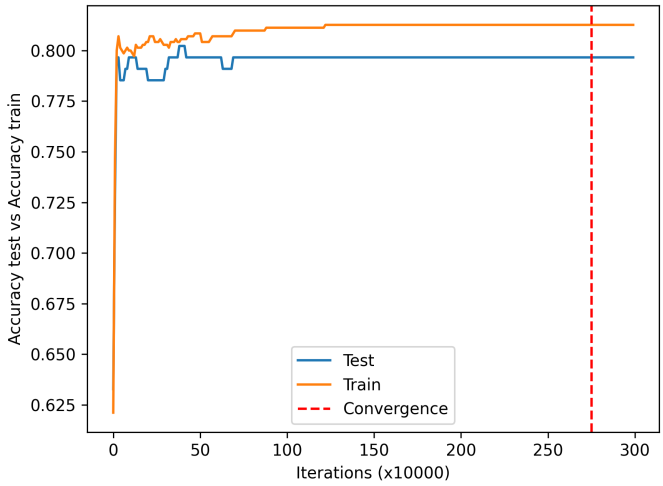

Figure 6: Accuracy Test vs Accuracy Train

3.3 How can you increase the performance of the model?

In order to increase the performance of the model we can try to use regularization.
Using $\lambda = 0.001$ we have a small improvement but only in the train accuracy, in fact we have the new train accuracy equal to 0.813, in percent 81.3%, the test accuracy remain perfectly the same.

I affirm that this report is the result of my own work and that I did not share any part of it with anyone else except the teacher.