

CS412 Homework 3 Report

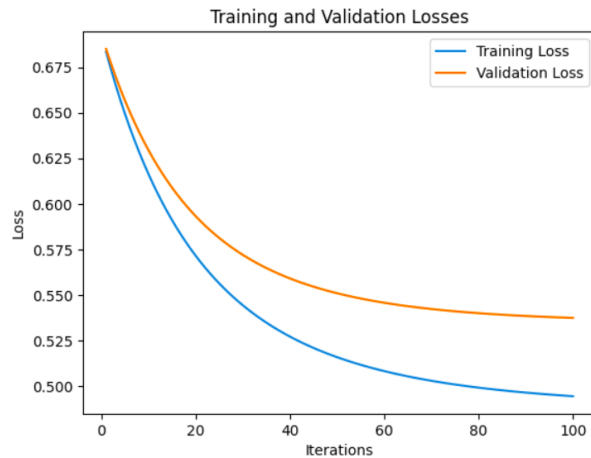
Ufuk Ulaş Tokat 28914

In this homework we had the titanic data. A 5-row sample of this data is shown below.

	Survived	Pclass	Sex	Age
709	1	3	2	29.699118
439	0	2	2	31.000000
840	0	3	2	20.000000
720	1	2	1	6.000000
39	1	3	1	14.000000

In **task 3** I plotted a loss graph with 0,1 learning rate and 100 iterations.

- **Learning rate** determines the step size at each iteration. A larger learning rate can make convergence to the minimum loss faster, but it can overshoot the minimum.
- **Number of iterations** determines how many times the entire data set is passed during training. A larger number of iterations can minimize the loss, but it can cause overfitting. It may work perfectly in training data but it won't work in real data. A lower number of iterations can be more stable.



We see in the graph that loss in the validation set is more than training data. If we further minimize the loss in the training set the gap between them may increase. Because it can overfit the training data.

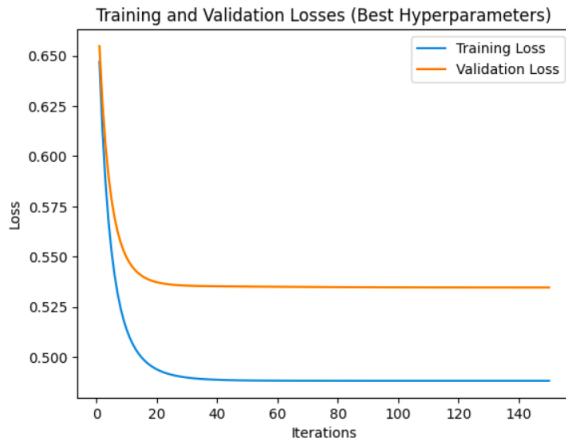
Also, we see that loss is minimized with respect to increase in the number of iterations.

In **task 4** I tried different **learning rates [0.01, 0.1, 0.5]**, and different **number of iterations [50, 100, 150]**. The validation losses according to these hyperparameters are as follows.

Learning Rate: 0.01, Num Iterations: 50, Validation Loss: 0.6570929512862813
Learning Rate: 0.01, Num Iterations: 100, Validation Loss: 0.6300067495652006
Learning Rate: 0.01, Num Iterations: 150, Validation Loss: 0.6095333324539743
Learning Rate: 0.1, Num Iterations: 50, Validation Loss: 0.551028498355107
Learning Rate: 0.1, Num Iterations: 100, Validation Loss: 0.537520415648391
Learning Rate: 0.1, Num Iterations: 150, Validation Loss: 0.5355700417328977
Learning Rate: 0.5, Num Iterations: 50, Validation Loss: 0.5351107144405552
Learning Rate: 0.5, Num Iterations: 100, Validation Loss: 0.5347181480969195

Learning Rate: 0.5, Num Iterations: 150, Validation Loss: 0.5346437900628989

I have a minimum validation error in **0,5 learning rate and 150 iterations** as expected because higher learning rate and higher number of iterations lower the loss. Then I plot the loss graph for these two hyperparameters.



The Resulting graph is very similar to the graph I found in task 3.

It can be seen that the convergence to the minimum loss is faster than task 3 due to the increase in learning rate.

Also, minimum loss is less than task 3 due to higher number of iterations.

In task 5 and 6 I trained the model with new hyperparameters and training + validation data. Then calculated the accuracy of the model which is **0.8324022346368715**.