Breast Cancer Detection

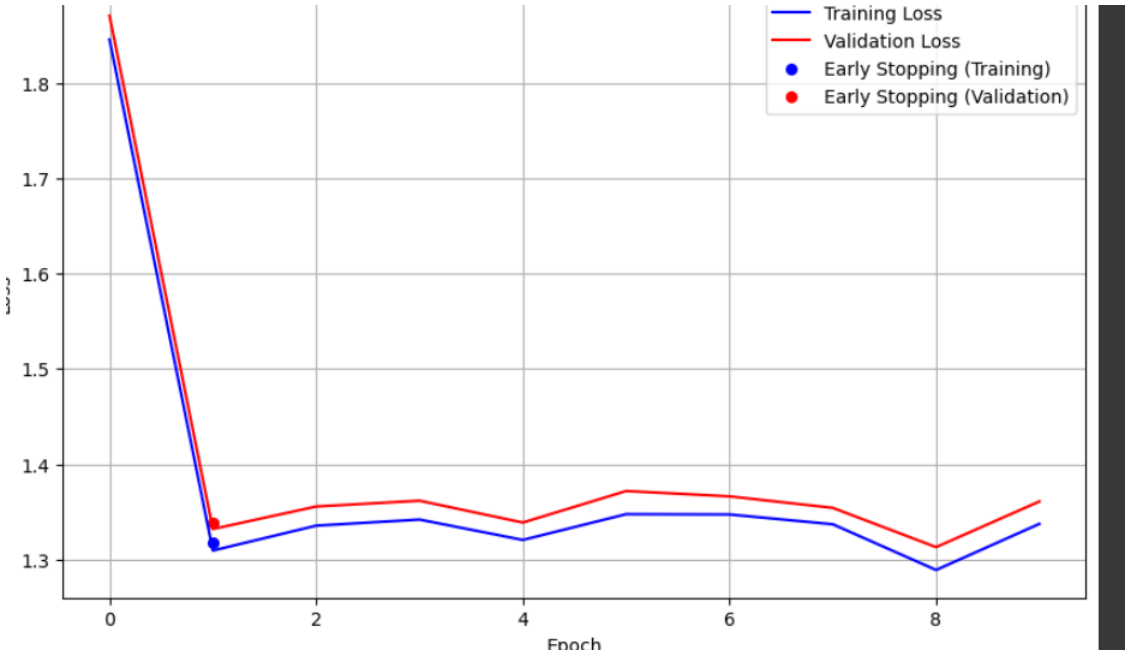
Akash Santhanakrishnan

November 13, 2024

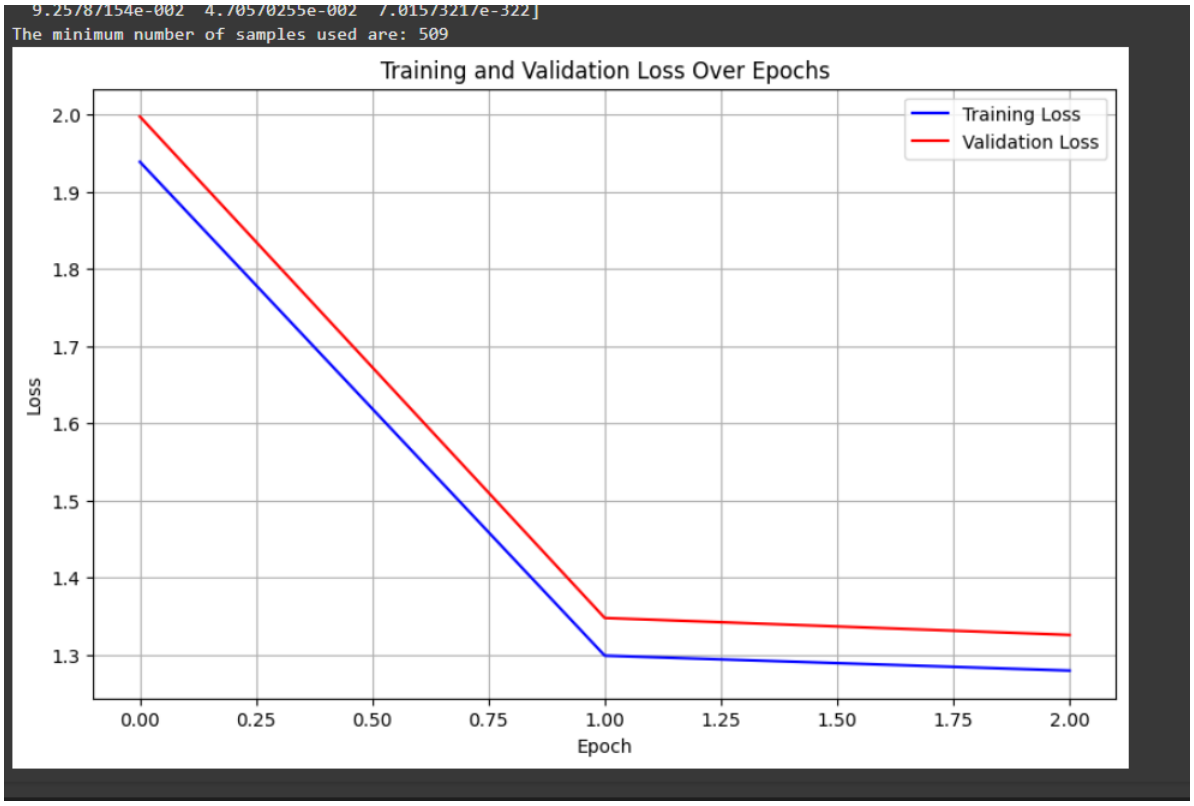
Results

Experiement 1

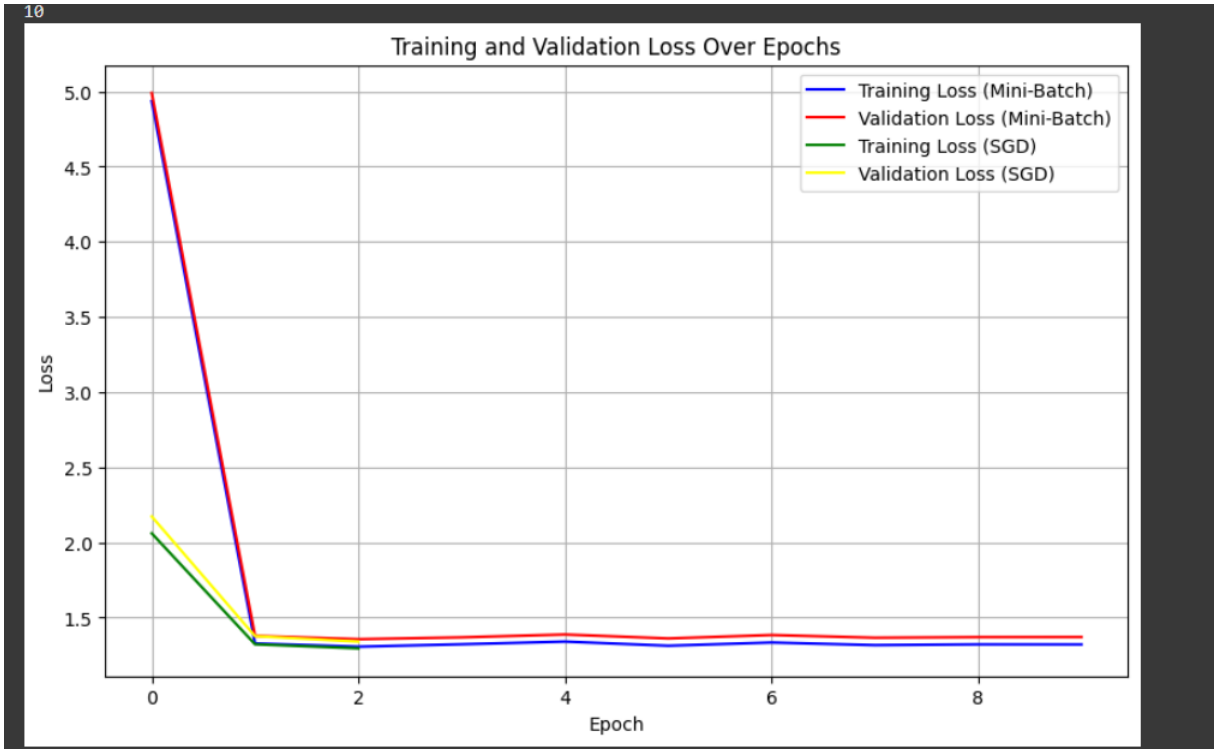
Without Early Stopping Break

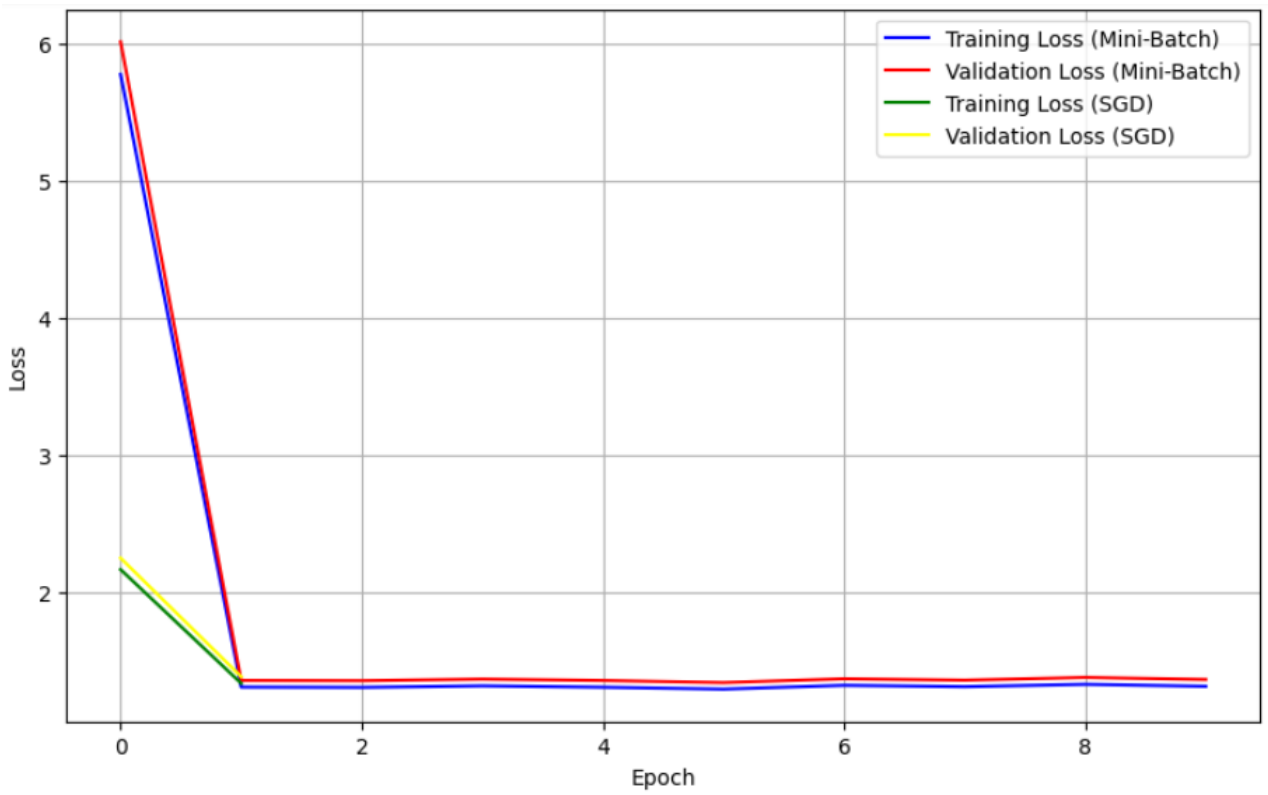


With Early Stopping Break

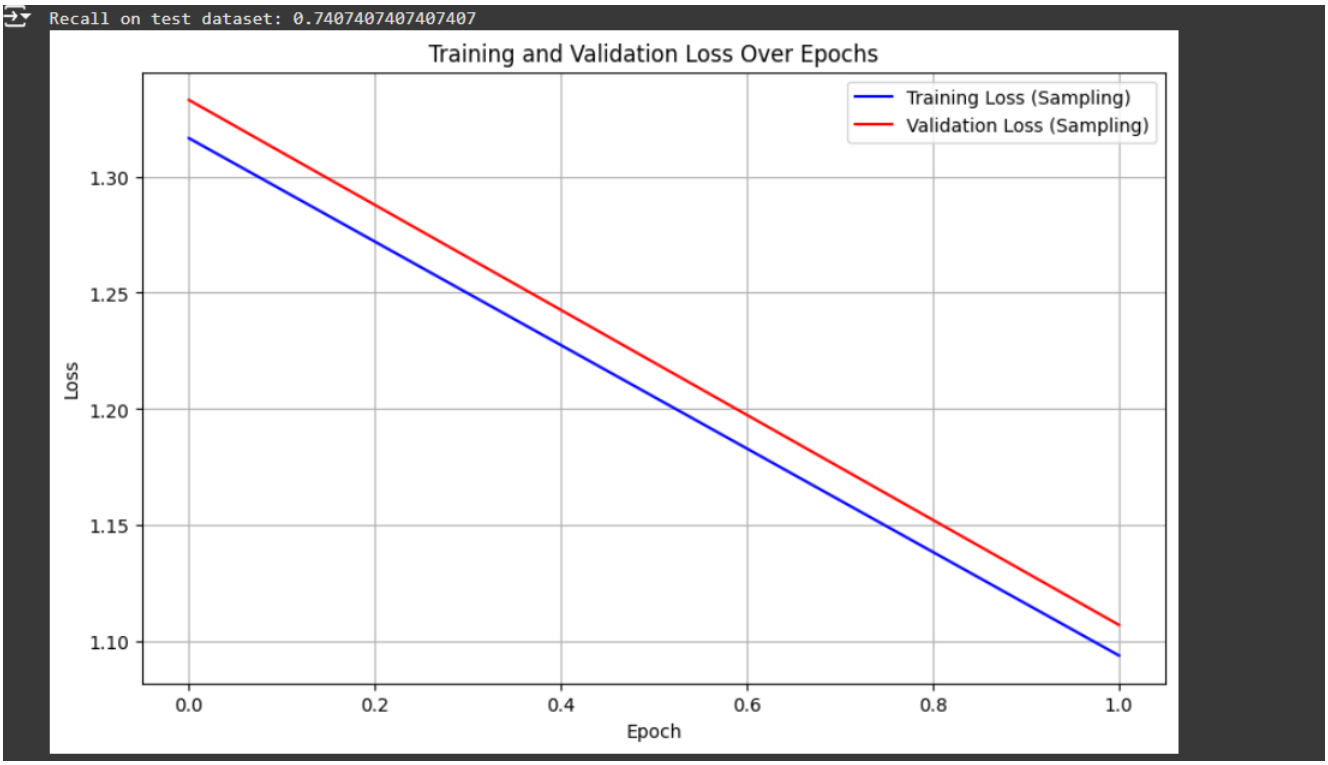


Experiment 2





Experiment 3



Summary

For the first experiment, after working with multiple different values I was able to find an optimal model with parameters $C = 1, \lambda = 0.0035, \text{epochs} = 2500, \text{threshold} = 1e^{-4}$. Notice that I added two different graphs one with early stopping and one with to show that algorithm does stop once the training and validation plateau. Also notice that the training loss is lower than the validation which makes sense as weights of the model are computed using the training set. The overall accuracy on the test set generally sits above 0.90.

For the 2nd experiment, I used parameters $C = 1, \lambda = 0.0085, \text{epochs} = 2500$. I increased the learning rate and this is due to the gradients being processed in batches, the algorithm automatically converges fast so when I use a lower learning rate the difference between validation and training isn't distinguishable. SGD and Mini-batch generally seem to have the same convergence region possibly because the batch-size is 5 which is closer to standard SGD than batch GD, meaning the weight updates and losses will be similar except for the initial epochs. Overall, mini-batch performed either the same or slightly better than the SGD algorithm, which may be because sometimes the early stopping happens too early before reaching convergence which prevents the model from having the optimal weights.

For the final experiment, The general trend was losses decreasing for both validation and training as we continue to train the classifier with the best samples (lowest hinge loss). The minimum number of samples needed to defined the classifier would be around 300. Overall, the metrics performed worse than the other 2 when you look at every single sample which suggests that the model over fits at some point but if you add the performance threshold, the metrics are around the same as SGD.