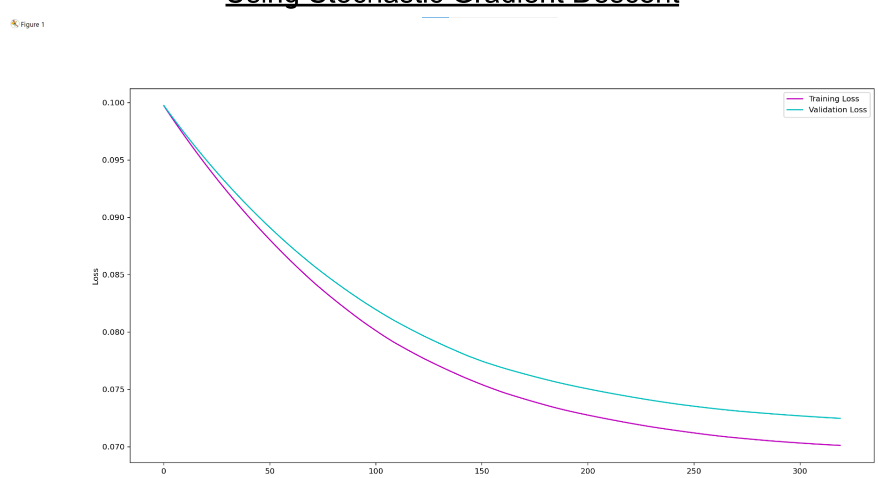
Assignment III

John Wu
Nov 10, 2024
WU103 # 400312093

Part I: Stochastic Gradient Descent (SGD) Training

In Part 1, a custom Support Vector Machine (SVM) model was trained using Stochastic Gradient Descent (SGD). This training approach updates the model's weights one sample at a time, enabling faster convergence with frequent weight updates. Training employed early stopping based on the convergence of training loss, and the model stopped at 319 epochs out of the maximum allowed 1000. Evaluation on the validation set showed an accuracy of 91.2%, precision of 86.8%, and recall of 91.7%, indicating that SGD effectively optimized the model parameters while controlling overfitting through early stopping

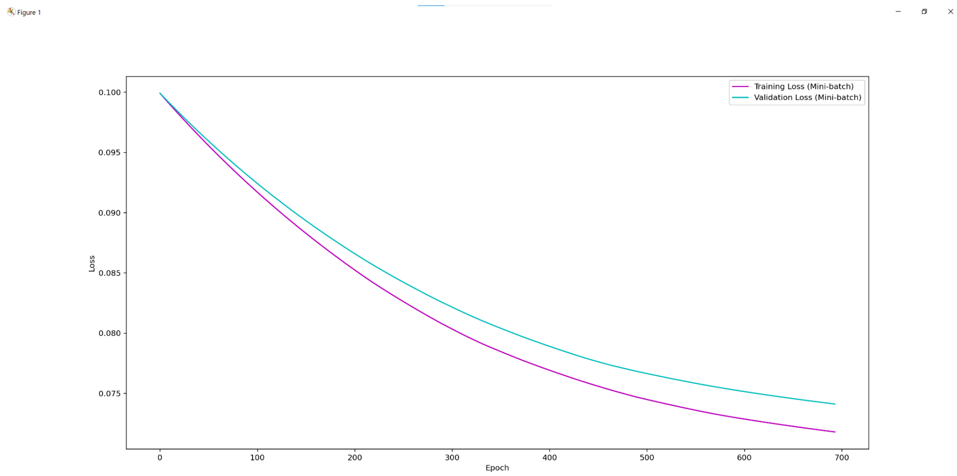
Using Stochastic Gradient Descent

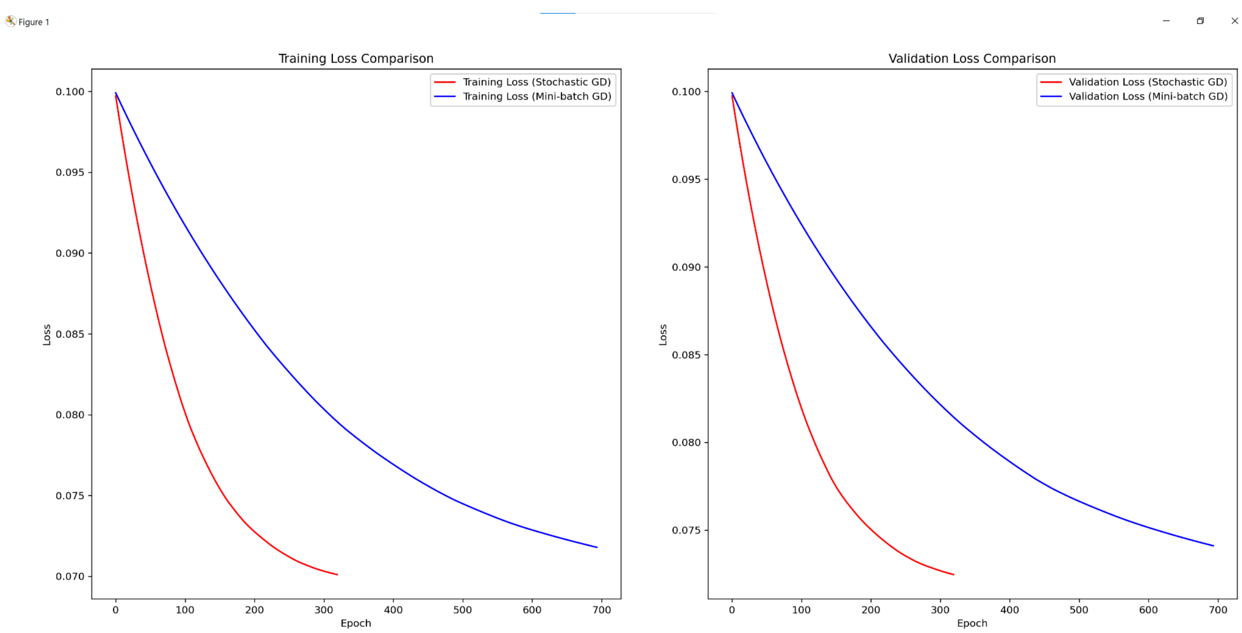


Part II: Mini-Batch Gradient Descent Training

Part 2 implemented a mini-batch gradient descent approach, where the model was trained with mini-batches of size 32. Mini-batch gradient descent combines the benefits of batch and stochastic methods, balancing computational efficiency with frequent updates. This method allowed the model to converge in 694 epochs, displaying comparable performance to SGD with accuracy, precision, and recall metrics similar to those in Part 1. The training and validation losses were logged and plotted, revealing a gradual decline in loss value. In comparison SGD may perform better and converge faster, though may be more computationally expensive.

Using mini-batch train

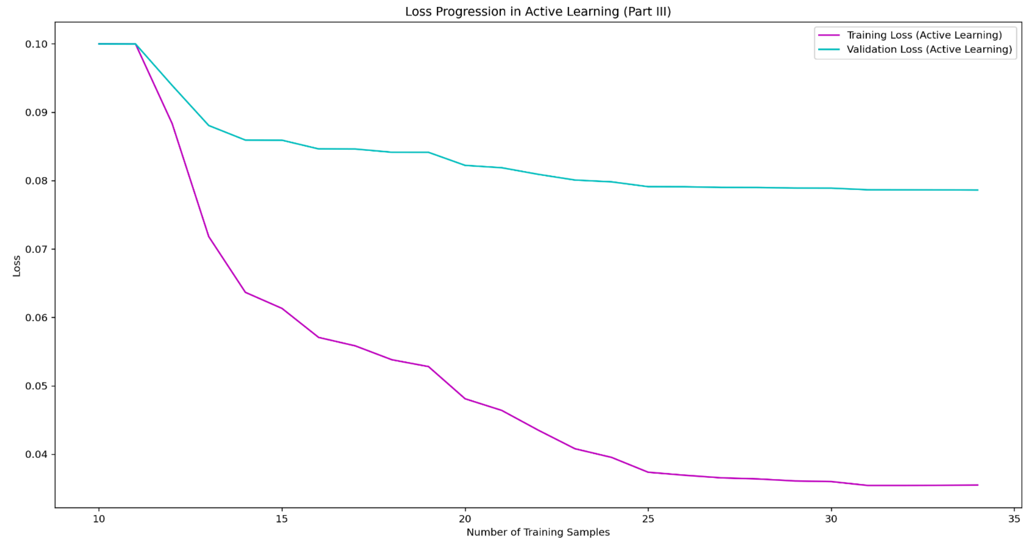




Part III: Active Learning with Uncertainty Sampling

In Part 3, an active learning approach with least-loss sampling was used. In this method, the model iteratively retrains by selecting and adding the sample from an unlabeled pool that exhibits the least loss with the current model. This approach helps the model focus on easy-to-learn samples first, building a strong foundation before gradually incorporating more challenging data. Training terminated based on early stopping criteria across several iterations, achieving a high test accuracy of 95.6%, with precision and recall at 93.2% and 95.4%, respectively.

Figure 1


Loss Progression in Active Learning (Part III)
Training Loss (Active Learning)
Validation Loss (Active Learning)
0.10
0.09
0.08
0.07
0.06
0.05
0.04
Loss
10
15
20
25
30
35
Number of Training Samples