Lab 3 Exercise – Optimise it!

Justin Ugwudike (jknu1g19@soton.ac.uk)

February 28, 2023

**1.1 Question 1**

The arrival points using different optimisers in the function shown in Appendix A:

**Text

Description automatically generated**

Figure 1 – The different arrival points of three different optimisers.

The function in Appendix A also produces loss plots showing the values of the Rastrigin functions at each optimisation step. These plots a shown below.

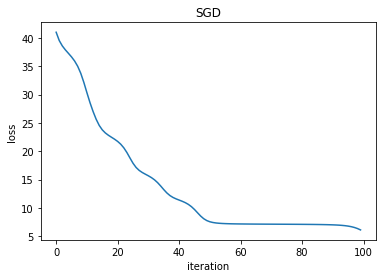


Figure 2 – Loss plot when using stochastic gradient descent (SGD) optimiser.

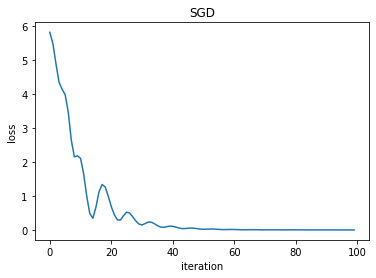


Figure 3 – Loss plot when using SGD optimiser with momentum.

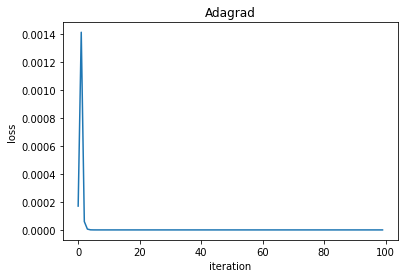


Figure 4 – Loss plot when using Adagrad optimiser.

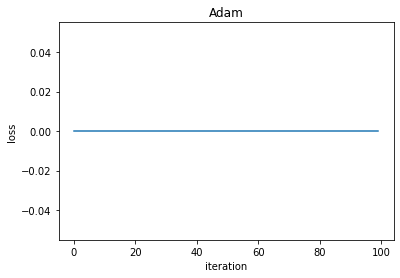


Figure 5 – Loss plot when using Adam optimiser.

The Adam optimiser seems to work the best because by inspecting Figure 5 after the first step (iteration 0) the loss is already approximately zero (at the global minimum of the Rastrigin function).

**1.2 Question 2**

Before evaluating the models, I expect the Adam optimiser to have higher accuracies than SGD. On top of this, I expect the decreasing learning rates to lower the accuracy in both cases. The variance of both optimisers should be low. The different accuracies of the SVM models are shown in Appendix C. The results were produced using the code in Appendix B.

By inspecting the results in Appendix C, I can see that using a learning rate of 0.01 achieves the highest accuracies. This could be because the higher learning rate allows the models to escape local minima.

Using lower learning rates with the optimiser can lead to higher accuracy on the rare occasion that the model does not get stuck in a local minimum.

**2 Appendix**

Appendix A – The optimise() function used to optimise the Rastrigin function with different optimisers and parameters. The Rastrigin function is also defined below.

Text

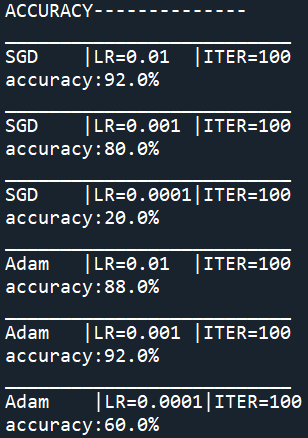
Description automatically generated

Appendix B – The optimise\_svm() function used to train Soft-margin Linear SVMs.

Text

Description automatically generated

Appendix C – Three sets of accuracies and variances for two optimisers with three different learning rates.

Text

Description automatically generatedA screenshot of a computer

Description automatically generated with low confidence