ME4 Machine Learning Coursework Assignment

Justin Kek - 01273515

Dataset 1

Started with Sequential Neural Network with three layers (1 input, 1 output, 1 hidden), read that one hidden layer should be sufficient for most problems. Hidden layer had 4 nodes (read that it should be between the number of input and output nodes) with ReLU. Output layer used softmax with 2 nodes to generate probabilities for each output (hit, miss). Dataset separated into training/test using 80/20 split. Optimised using Stochastic Gradient Descent, but initially faced issue where accuracy was not improving. Tried changing to AdaM, which worked, but read that it does not generalise well. Realised the SGD was getting stuck in a local minima, so decreased learning rate to 0.001. Tried different batch sizes (32, 64, …) and epoch sizes. Settled for batch size of 32 to ensure convergence (at computational cost). Reduced epoch size from 2000 as it plateaus in accuracy after 500. Accuracy achieved at this point was approximately 75%, but scaling the features to ensure all features are evenly minimised increased the accuracy to 82%.

Dataset 2

Used the base model from Dataset 1, but got a low accuracy of 50%. Added additional hidden layer with 4 nodes and increased epoch but showed no improvement. Realised that the learning rate was too slow, and training was stopping before reaching the optimal model parameters. Increased SGD learning rate to 0.01, and accuracy improved to 70%. Read that increasing hidden nodes can improve accuracy, so tried 12 (less than twice of input nodes) and 6 (two-thirds the size of input nodes + output nodes) for the first and second hidden layer respectively. Accuracy of predictions improved to 80%.