

Machine Learning and Intelligent Systems

Project 2: The Perceptron

REPORT

Model Training

The Perceptron model is a simple single layer neural network used for binary classification problems. It makes predictions by computing a linear activation function on the input features, passing that through a step function to generate a 0 or 1 prediction, and updating the weights and bias when there is a misclassification.

The Perceptron model was trained using the perceptron learning algorithm with a form of stochastic gradient descent (SGD). The training data was split into train (**60%**), validation (**20%**), and test (**20%**) sets. The model was trained on the training set by looping through each sample for **1000** epochs and updating the weights and bias on misclassifications.

A learning rate of **0.01** was used with SGD to control the size of the weight updates. The model fitting was monitored on the validation set to tune hyperparameters and select the best model.

Validation & Test Set Usage

After training, the model achieved **100%** accuracy on the validation set. This indicates it learned the patterns in the validation data successfully without overfitting. So we did not need to use hyperparameter tuning on the model. Usually, the validation set would have been used to tune hyperparameters and select the best model with the best accuracy.

Finally, the trained model was evaluated on the test set, achieving **98.6%** accuracy. Since the test set is completely held out during training, it gives an unbiased estimate of how well the model can generalize to new data.

The high test accuracy indicates the model was trained effectively and can accurately classify new data similar to what it was trained on. The validation set helped select model hyperparameters for good generalization.

It is important to note that the UCI ML hand-written digits dataset the model was trained on is relatively small so it may not perform as well on larger datasets.