Yurii Oleksandr Bibik

Dr. Kang

CS 422

12 April 2023

Homework 4

In this assignment, I was tasked with comparing the performance of Support Vector Machines (SVMs) with three different kernels—linear, polynomial, and radial basis function (RBF)—on the MNIST dataset. To evaluate the performance, I used the Python package scikit-learn and implemented a 5-fold cross-validation approach. First, I imported the necessary libraries such as pandas, NumPy, and scikit-learn. Then, I loaded the dataset and separated the feature matrix (X) and the target vector (y). After preparing the data, I initialized three SVM models with different kernels: linear, polynomial, and RBF. To perform 5-fold cross-validation, I used the KFold class from scikit-learn, making sure to shuffle the data to ensure a good distribution of samples in each fold. For each fold, I trained the three SVM models on the training set and predicted the labels for the test set. I calculated the accuracy for each model by comparing the predicted labels with the actual labels in the test set. See the table below for the results:

A picture containing text, scoreboard, clapperboard

Description automatically generated