Yurii Oleksandr Bibik

Dr. Kang

CS 422

26 February 2023

Homework 2

My code implements from the scratch the k-nearest neighbors algorithm for image recognition. It uses pandas library to load the MNIST training and test data from CSV files, numpy library to perform mathematical computations, and sklearn.preprocessing library to standardize the data using z-score normalization. The algorithm uses cosine similarity as the distance metric to find the k nearest neighbors for each test sample. The value of k is set to 20 in this implementation, purely empirically. After identifying the k nearest neighbors, the algorithm decides the majority class among them and assigns that class to the test sample. Finally, the overall accuracy is calculated by comparing the predicted labels with the true labels of the test samples, and the answer is outputted to the console.

Code Execution:

To execute the code you need Python 3.11.2. You can run the program with the following command *python3 homework2.py.* If you want to change the k value, then you need to open the file and edit the line *k = 20* and replace 20 with your k values.