

Eric Gerner
18 June 2025

What I have done:


This assignment aimed to implement a k-Nearest Neighbors (KNN) classifier from scratch to perform digit classification using the scikit-learn digits dataset. The dataset contains 8x8 images of handwritten digits (0–9). I explored different distance metrics and values of k to optimize model performance, then evaluated the model on a separate test set. I implemented two distance comparison functions, Euclidean and Manhattan distances. Both were tested to see which produced better accuracy

Best Accuracy:

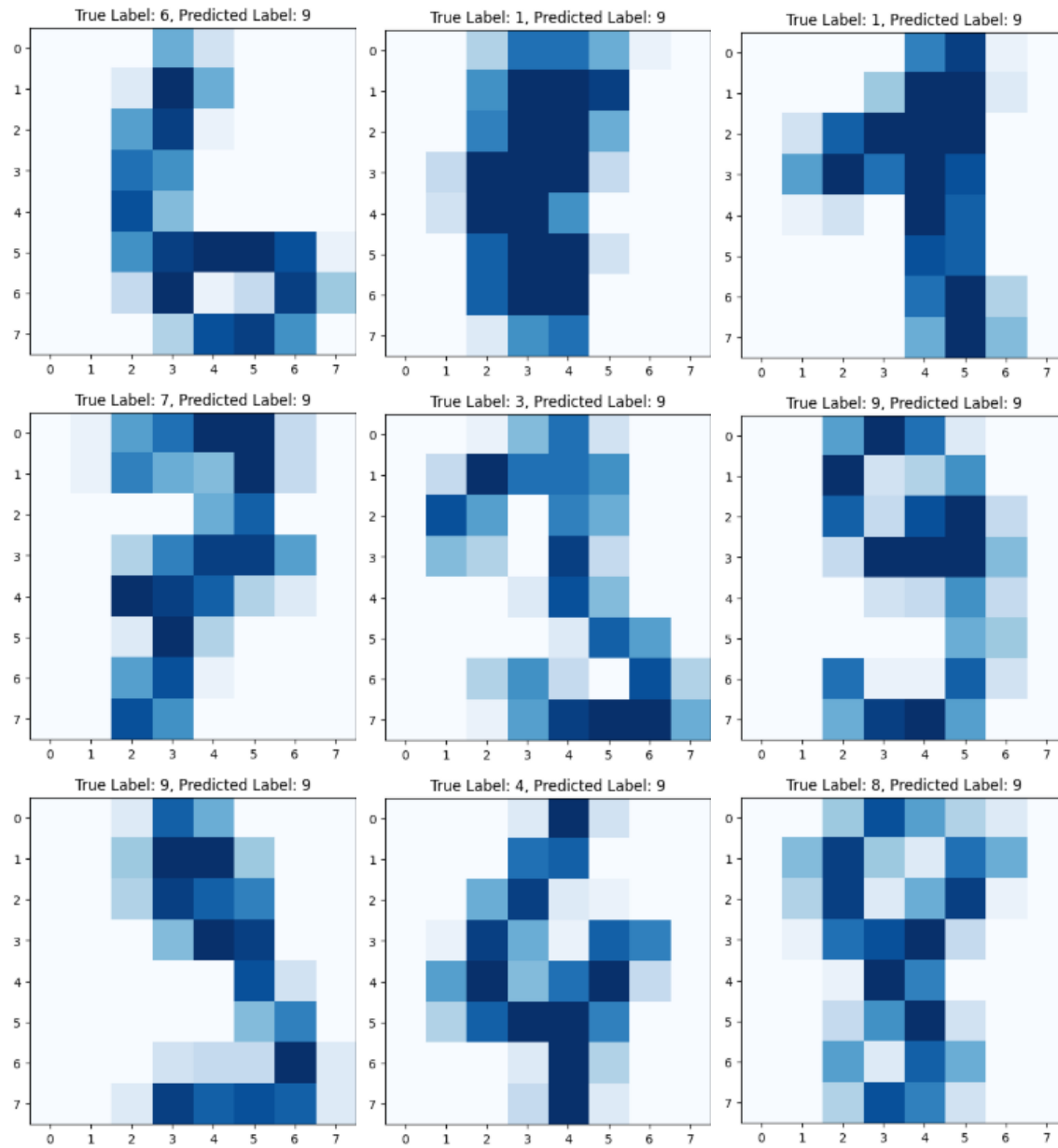
I evaluated several values of k: K = 1, 3, 5, 7, 9

For each k-value, I tested both distance metrics using the val set to determine the optimal combination. The best one I found was K = 1 with an accuracy of 0.988929889298893 in Euclidean and K = 7 with an accuracy of 0.988929889298893 in Manhattan.

Code:  KNN.ipynb

Prediction visuals:  Below are 10 randomly selected samples from the test set with true/predicted labels. The images were made using matplotlib.pyplot

Eric Gerner
18 June 2025



Eric Gerner
18 June 2025

