KNN Algorithm

In this report, we discuss the implementation of the k-nearest neighbors (KNN) algorithm on a dataset of handwritten digits. The goal of this project was to train a KNN model that can accurately predict the digit label of an image of a handwritten digit.

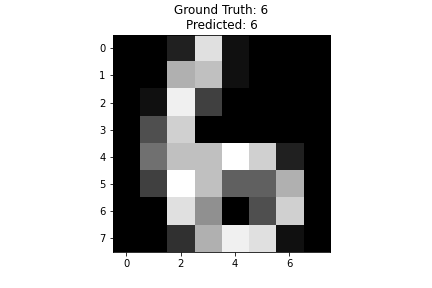
We first preprocessed the data by scaling the pixel values and splitting the data into train, dev, and test sets. We then experimented with different distance metrics (Euclidean distance, Manhattan distance, and cosine similarity) and different k values (ranging from 1 to 10) to find the best combination that resulted in the highest accuracy on the test data.

The best results were obtained when using a k value of 1 and the Euclidean distance metric, with an accuracy of 99% on the test data.

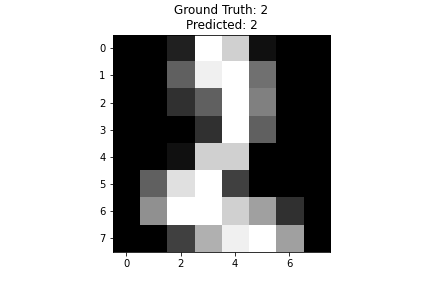
To visualize the predictions of the model, we randomly selected 10 samples from the test data and used the predict() method of the KNN model to obtain the predicted labels for these samples. We then plotted each sample along with its corresponding ground truth label and predicted label, using the imshow() function of the matplotlib library.

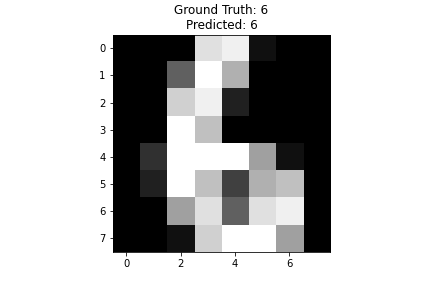


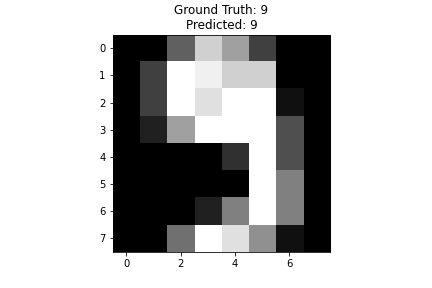




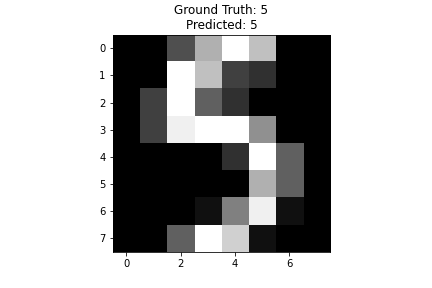








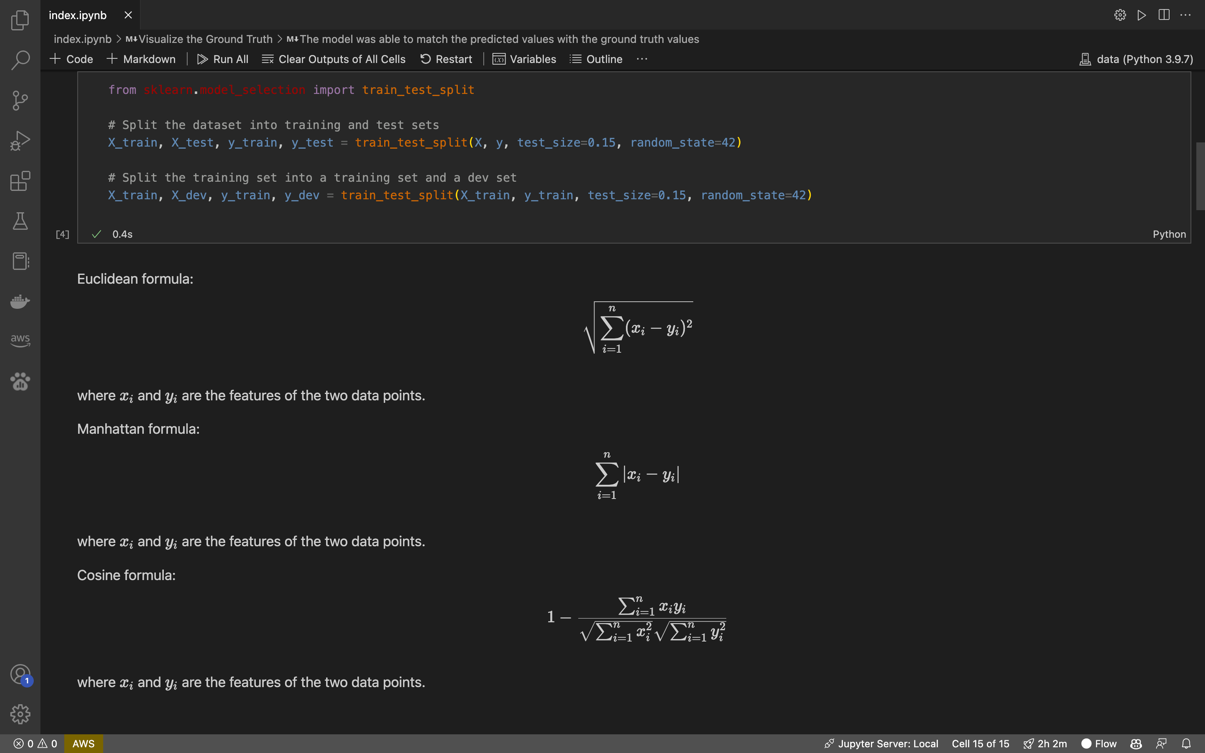


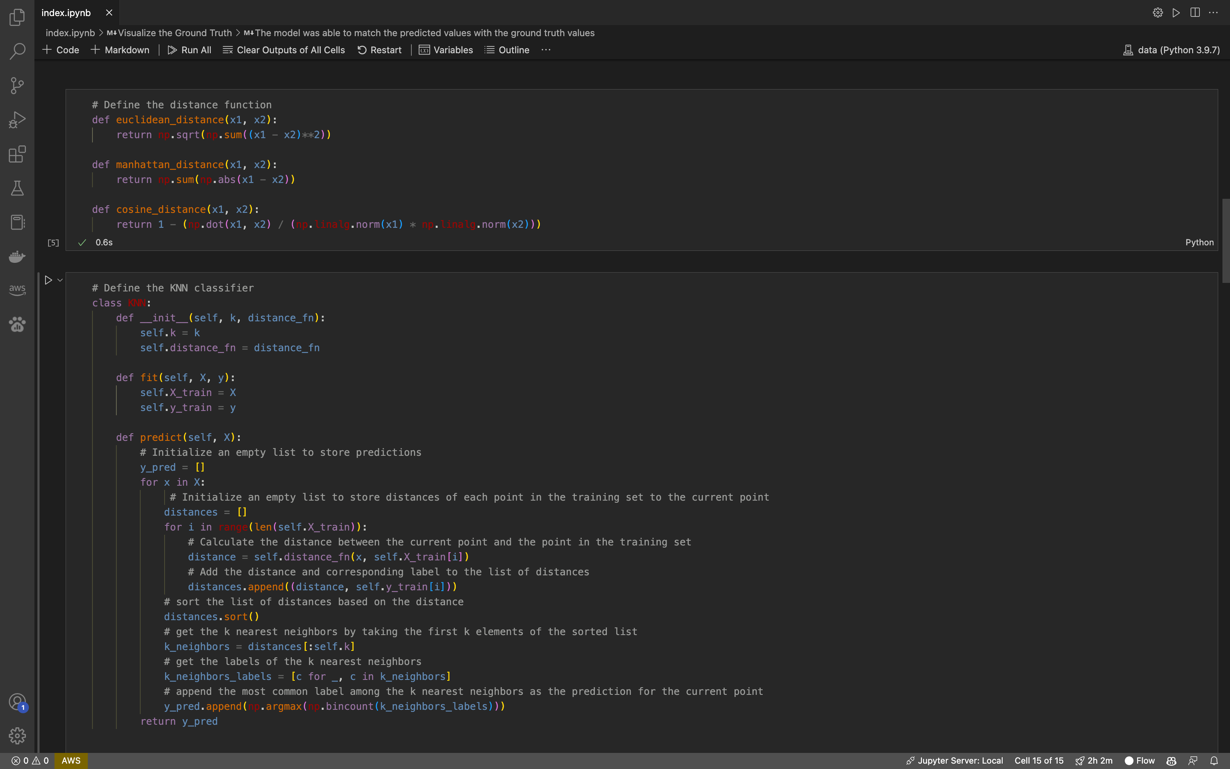


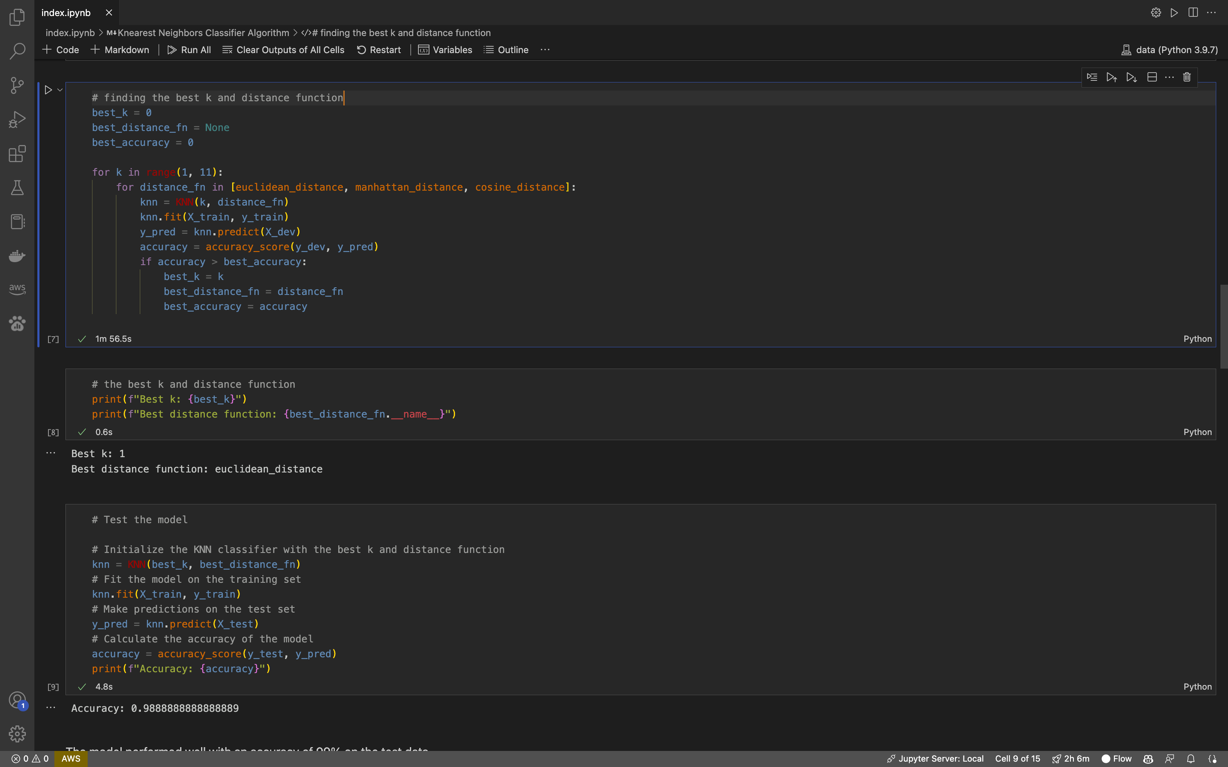


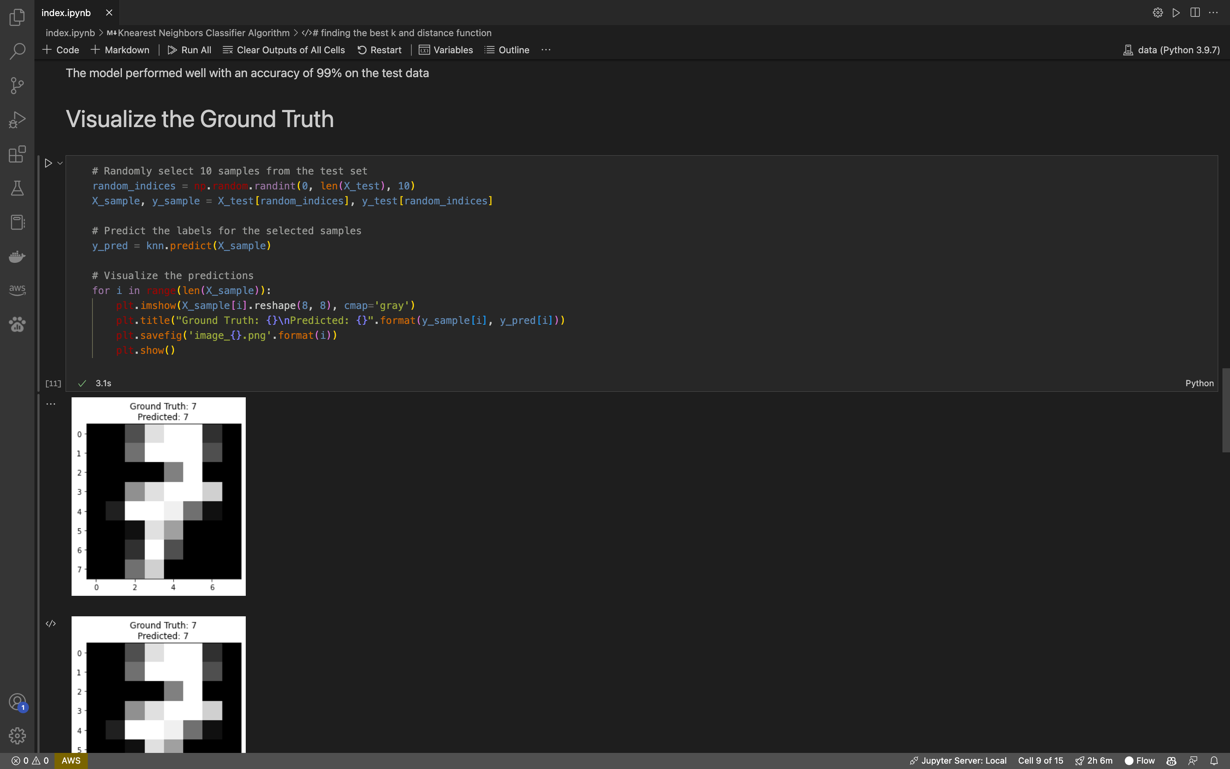
The codes used in the implementation are also included in the report as screenshots.











Overall, the KNN algorithm performed very well on this dataset, achieving a high accuracy rate and effectively distinguishing between different handwritten digits. The visualizations of the predictions further confirm the effectiveness of the model.