CSCD 429-040 HW 2

Blake Chalpin

5 November 2021

Project Structure

Running The Program

NOTE: Before running any of the following commands below, make sure that the current directory is `<path to this zip file>/src`

Using Conda

Using Conda, run the following commands to create the environment to run the program in (this will handle the Python version and packages):

```
$conda env create -f environment.yml
$conda activate cscd429-hw2-blake-chalpin
```

Once the Conda environment is created and activated, run the program with the command:

```
$python main.py
```

Without Conda

If you do not have Conda installed, the Python packages may still be installed using the following command:

```
$pip install -r requirements.txt
```

Note: Python 3.8 is used for this project

Once the Python packages are installed, run the program with the command:

```
$python main.py
```

Methods Used

Performance

Accuracy

The overall accuracy of the 3 Nearest Neighbors (KNN with $k = 3$) classification model using the entire training data set to predict our test set is **45.779%**.

Results File

The results of our prediction are stored in the "results file" name `results.txt`