Report

1. Check list of functionalities
   f1: implemented in function Part1
   f2: implemented in function Part2
   f3: implemented in function Part3
   f4: implemented in function Part3
   f5: implemented in the rest of the file
2. About running my program
   Directly run file.py to run the model. Only numpy, sklearn and joblib needed for my program as dependency.
3. Details of my program
   Part 1:
   The dataset is split into training set and test set in proportion as 3:1.
   About Part 3, my kNN model:
   My algorithm is traditional k-nearest neighborhood. At first construct a matrix "diff", whose (i,j) entry is the distance of i-th train data and j-th test data. Then using function numpy.argpartition to get the k smallest distance for each column, i.e. each test data. Then for each test data, counting appearance for each class, and choose the class which has maximum appearance.
   Part 4
   About error for both models:
   Training error of library model: 1.3%
   Test error of library model: 4.4%
   Training error of library model: 1.2%
   Test error of library model: 4.4%
   Part 5:
   The query function support for query of single index.
4. Running time of my program
   The running time of the whole file is less than 30 second.