CS-3120

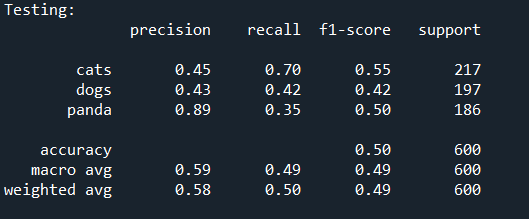
Dr. Feng Jiang

Xiaosong Wang

Mar 15, 2020

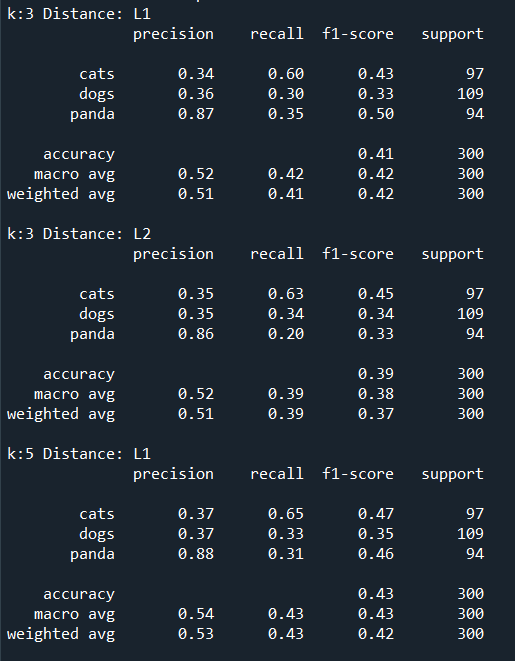
HW3 Report

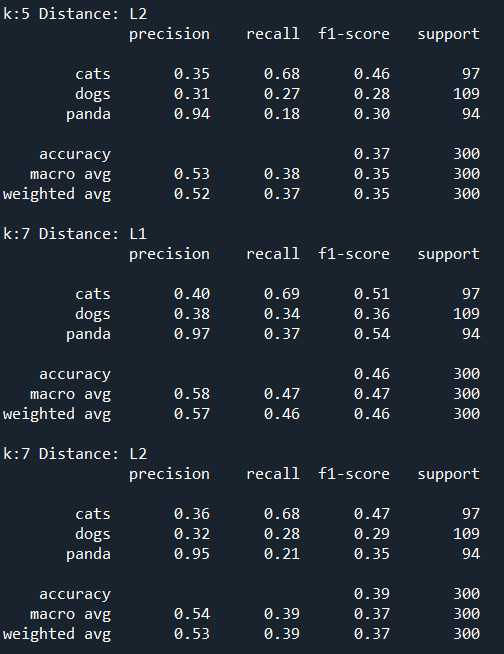
Evaluate performance on the test set:

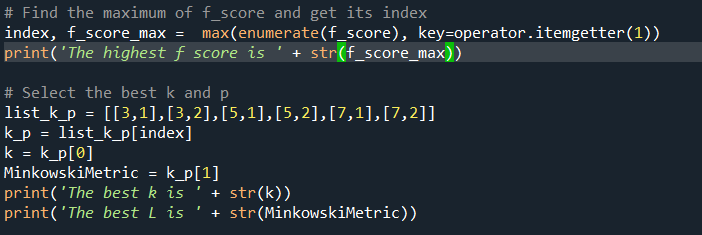


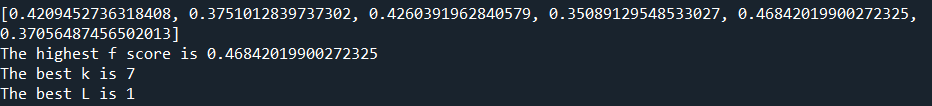
Choosing the best k and distance:

First calculate the performance for the 6 pairs of different k and different distance by passing in the training and validation dataset.

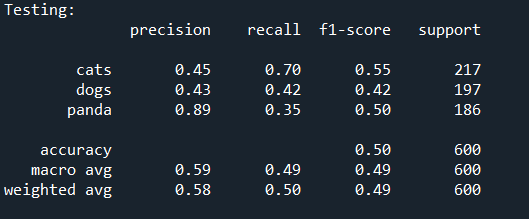




Then I put all the f\_score in a list, then use the max method to find the highest f\_score, and the model with the highest f score is the model we want to use in the testing part.



Finally, use the best model(which has the best performance in k and L)to do the testing



Reference:

The structure of my code came from the slides and video you uploads.

I studied more about KNeighborsClassifier with:

<https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html#sklearn.neighbors.KNeighborsClassifier>

How to print out f\_score:

<https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html>

How to find the maximum number and its index in a list:

<https://stackoverflow.com/questions/6193498/pythonic-way-to-find-maximum-value-and-its-index-in-a-list>