

CS 432/532: Web Science

Spring 2017

Assignment 10

Michael Micros

Instructor: Michael L. Nelson

Old Dominion University
Norfolk, Virginia

May 1, 2017

Problem 1: Computing kNN for Dr. Nelson's blogs

Having substituted the euclidean distance with the cosine, the Knn function provided in the “Programming Collective Intelligence” textbook was used on the blogdata collected in Assignment 8. The blog data are located in “blogdata.txt”.

```
f = open("blogdata.txt", "r")

for line in f:
    a = line.strip('\n').split('\t');
    b = a.pop(0)
    vecs[b] = a

print len(vecs)

fm = 'F-Measure'
ws = 'Web Science and Digital Libraries Research Group'

a = vecs[fm]
temp = vecs.values()
temp.pop(vecs.keys().index(fm))

a = knnestimate(temp,a,k=5)

k = [1, 2, 5, 10, 20]
print "\n-----F-Measure kNN-----"
for i in k:
    print "\n---k = "+str(i)
    for j in range(i):
        b = a[j][1]
        print vecs.keys()[b]

a = vecs[ws]
temp = vecs.values()
temp.pop(vecs.keys().index(ws))

a = knnestimate(temp,a,k=5)

k = [1, 2, 5, 10, 20]
print "\n-----WS-DL kNN-----"
for i in k:
    print "\n---k = "+str(i)
    for j in range(i):
        b = a[j][1]
        print vecs.keys()[b]
```

Figure 1: Code for calculating kNN for F-Measure and WS-DL.

```
-----F-Measure kNN-----  
  
---k = 1  
Katy @ Sommerville Special School  
  
---k = 2  
Katy @ Sommerville Special School  
My Kitchen's Aroma  
  
---k = 5  
Katy @ Sommerville Special School  
My Kitchen's Aroma  
Sisterhood of the Traveling Recipes  
Agita makeup  
Very Hungry People  
  
---k = 10  
Katy @ Sommerville Special School  
My Kitchen's Aroma  
Sisterhood of the Traveling Recipes  
Agita makeup  
Very Hungry People  
Try This Yourself  
Best Recepis  
Rebecca's Amazing Creations  
The Wicked Truth  
What's Cooking!  
  
---k = 20  
Katy @ Sommerville Special School  
My Kitchen's Aroma  
Sisterhood of the Traveling Recipes  
Agita makeup  
Very Hungry People  
Try This Yourself  
Best Recepis  
Rebecca's Amazing Creations  
The Wicked Truth  
What's Cooking!  
Desirable Recipes  
Color and Spices  
Dining In Austin  
Mom Tested, Kid Approved  
Nibbling Gypsy  
Mom's cooking  
Angie's Recipes  
Excellent Kitchen  
Top Class Food  
Teacher's recipes
```

Figure 2: kNN results for <http://f-measure.blogspot.com/>.

```
----WS-DL kNN----  
  
---k = 1  
Sisterhood of the Traveling Recipes  
  
---k = 2  
Sisterhood of the Traveling Recipes  
LandOfTheLeI  
  
---k = 5  
Sisterhood of the Traveling Recipes  
LandOfTheLeI  
Very Hungry People  
My Little Corner of the World  
Piecey Num Nums  
  
---k = 10  
Sisterhood of the Traveling Recipes  
LandOfTheLeI  
Very Hungry People  
My Little Corner of the World  
Piecey Num Nums  
What's Cooking!  
Shape Your Life  
Mom Tested, Kid Approved  
Little Kitchen  
Recipe for Fitness  
  
---k = 20  
Sisterhood of the Traveling Recipes  
LandOfTheLeI  
Very Hungry People  
My Little Corner of the World  
Piecey Num Nums  
What's Cooking!  
Shape Your Life  
Mom Tested, Kid Approved  
Little Kitchen  
Recipe for Fitness  
"Agreed, Number One"  
Try This Yourself  
Dining In Austin  
Mom's cooking  
Barron County Retired Educators  
Nibbling Gypsy  
Angie's Recipes . Taste Of Home  
sense-a(Ha!)-bility  
This A.M.  
geeks.rate(food)  
mike@canus:~/CS432/asst10$
```

Figure 3: kNN results for <http://ws-dl.blogspot.com/>.

Problem 2: LIBSVM