CMSC 476/676

Spring 2021

Extra Credit Optional Homework 2

Name : Akarsh Kashamshetty

Roccio Feedback Questions

1. Vocabulary: {green, golf, putter, hue, color}

Original query: q0 = [1, 1, 0, 0, 0] = [1\*green, 1\*golf]

Relevant docs:

dr1 = [2, 1, 1, 0, 0] = [2\*green, 1\*golf, 1\*putter]

dr2 = [4, 1, 3, 0, 0] = [4\*green, 1\*golf, 3\*putter]

Non-relevant doc:

dn1 = [2, 0, 0, 2, 2] = [2\*green,2\*hue, 2\*color]

Assume alpha = 1 , beta = .5

What is q1?

A) [4,1,1,0,0]

B) [4,1,1,-1,-1]

C) [4,1,2,0,0]

D) [4,1,2,-1,-1]

E) None of the above

Answer for 1 is E

1. Vocabulary: {squid, recipe, calamari, caching, proxy, http}

Original query: q0 = [1, 1, 1, 0, 0,0] = [1\*squid, 1\*recipe, 1\*calamari]

Relevant docs:

dr1 = [2, 1, 1, 0, 0, 0] = [2\*squid, 1\*recipe, 1\*calamari]

dr2 = [4, 2, 3, 0, 0, 0] = [4\*squid, 2\*recipe, 3\*calamari]

Non-relevant docs:

dn1 = [2, 0, 0, 2, 2, 0] = [2\*squid,2\*caching, 2\*proxy]

dn2 = [2, 0, 0, 2, 0, 2] = [2\*squid,2\*caching, 2\*http]

Assume alpha = 1 , beta = .5

What is q1?

A) [3,1,1,0,0,0]

B) [3,1,1,-1,-1,0]

C) [3,2.5,3,-1,-.5,-.5]

D) [3,2.5,3,0,0,0]

E) None of the above

Answer for 2 is C

Naïve Bayes Questions

1. Suppose we have trained a generative model and get a new test example x. Suppose also that our model tells us that

p(x|y=0) = 0.02,

p(x|y=1) = 0.04, and

p(y=1) = p(y=0) = .5.

What is p(y=1|x)? Rephrasing the question, what is the probability that y is 1 (true), given x.

1. .77
2. .67

C) .91

D) .54

E) None of the above

Answer for 3 is B

1. Suppose we have trained another generative model to detect emails that are spam or ham using the presence of the words buy and deal. Our model tells us that:

p(buy|spam) = 20/25,

p(deal|spam) = 10/25,

p(spam) = 25/100,

p(buy|ham) = 5/75,

p(deal|ham) = 10/75, and

p(ham) = 75/100.

What is the probability that the email is spam if it has the words buy and deal? (answer will have 2 decimal places)

* 1. .77
  2. .67
  3. .92
  4. .54
  5. None of the above

Answer for 4 is C

Soundex Code Questions

1. What is the Soundex code for the words Yosemite and Yosemittee?
   1. Y253 and Y253
   2. Y532 and Y523
   3. Y532 and Y532
   4. Y123 and Y532
   5. None of the above

Answer for 5 is A

1. What is the Soundex code for the words Connecticut and Coneticut?
   1. C523 and C522
   2. C523 and C523
   3. C552 and C553
   4. C253 and C253
   5. None of the above

Answer for 6 is E