Washington State University
School of Electrical Engineering and Computer Science
Fall 2018

CptS 440/540 Artificial Intelligence
Homework 10 Solution
Due: November 15, 2018 (11:59pm)

General Instructions: Put your answers to the following problems into a PDF document and submit as an attachment under Content → Homework 10 for the course CptS 440 Pullman (all sections of CptS 440 and 540 are merged under the CptS 440 Pullman section) on the Blackboard Learn system by the above deadline. Note that you may submit multiple times, but we will only grade the most recent entry submitted before the above deadline.

For this homework we will be applying two different learning methods to the following 10 examples. Each example describes weather conditions in terms of the features: Sky ∈ {clear, cloudy, overcast}, Sea ∈ {blue, gray}, and Sun ∈ {true, false}. The class Sail ∈ {yes, no} indicates whether or not we should go sailing in these conditions.

Sky	**Sea**	**Sun**	**Sail**
clear	blue	false	yes
clear	gray	true	yes
clear	gray	false	no
cloudy	blue	true	yes
cloudy	blue	false	yes
cloudy	gray	true	yes
cloudy	gray	false	no
overcast	blue	true	yes
overcast	blue	false	yes
overcast	gray	false	no

1. *Naïve Bayes*. Suppose we want to classify the new instance <Sky=overcast, Sea=gray, Sun=true> using the naïve Bayes learning method. Compute the following. Show your work.
    a. Compute the prior probabilities P(Sail=yes) and P(Sail=no).
    b. Compute P(Sky | Sail) for all values of Sky ∈ {clear, cloudy, overcast} and Sail ∈ {yes, no}.
    c. Compute P(Sea | Sail) for all values of Sea ∈ {blue, gray} and Sail ∈ {yes, no}.
    d. Compute P(Sun | Sail) for all values of Sun ∈ {true, false} and Sail ∈ {yes, no}.
    e. Compute P(Sail=yes | Sky=overcast, Sea=gray, Sun=true) and P(Sail=no | Sky=overcast, Sea=gray, Sun=true).
    f. Which class would Naïve Bayes choose for the new instance?

Solution:

a. P(Sail=yes) = 7/10 = 0.7
   P(Sail=no) = 3/10 = 0.3

b. P(Sky=clear | Sail=yes) = P(Sky=clear, Sail=yes) / P(Sail=yes) = (2/10) / (7/10) = 2/7 = 0.29
P(Sky=cloudy | Sail=yes) = P(Sky=cloudy, Sail=yes) / P(Sail=yes) = (3/10) / (7/10) = 3/7 = 0.43
P(Sky=overcast | Sail=yes) = P(Sky=overcast, Sail=yes) / P(Sail=yes) = (2/10) / (7/10) = 2/7 = 0.29
P(Sky=clear | Sail=no) = P(Sky=clear, Sail=no) / P(Sail=no) = (1/10) / (3/10) = 1/3 = 0.33
P(Sky=cloudy | Sail=no) = P(Sky=cloudy, Sail=no) / P(Sail=no) = (1/10) / (3/10) = 1/3 = 0.33
P(Sky=overcast | Sail=no) = P(Sky=overcast, Sail=no) / P(Sail=no) = (1/10) / (3/10) = 1/3 = 0.33

c. P(Sea=blue | Sail=yes) = P(Sea=blue, Sail=yes) / P(Sail=yes) = (5/10) / (7/10) = 5/7 = 0.71
P(Sea=gray | Sail=yes) = P(Sea=gray, Sail=yes) / P(Sail=yes) = (2/10) / (7/10) = 2/7 = 0.29
P(Sea=blue | Sail=no) = P(Sea=blue, Sail=no) / P(Sail=no) = (0/10) / (3/10) = 0/3 = 0
Since equals 0, we add 1 to numerator and |values(Sea)| = |{blue,gray}| = 2 to denominator:
P(Sea=blue | Sail=no) = (0 + 1) / (3 + 2) = 1/5 = 0.2
P(Sea=gray | Sail=no) = P(Sea=gray, Sail=no) / P(Sail=no) = (3/10) / (3/10) = 3/3 = 1.0

d. P(Sun=true | Sail=yes) = P(Sun=true, Sail=yes) / P(Sail=yes) = (4/10) / (7/10) = 4/7 = 0.57
P(Sun=false | Sail=yes) = P(Sun=false, Sail=yes) / P(Sail=yes) = (3/10) / (7/10) = 3/7 = 0.43
P(Sun=true | Sail=no) = P(Sun=true, Sail=no) / P(Sail=no) = (0/10) / (3/10) = 0/3 = 0
Since equals 0, we add 1 to numerator and |values(Sun)| = |{true,false}| = 2 to denominator:
P(Sun=true | Sail=no) = (0 + 1) / (3 + 2) = 1/5 = 0.2
P(Sun=false | Sail=no) = P(Sun=false, Sail=no) / P(Sail=no) = (3/10) / (3/10) = 3/3 = 1.0

e. P(Sail=yes | Sky=overcast, Sea=gray, Sun=true)
= P(Sky=overcast,Sea=gray,Sun=true | Sail=yes) * P(Sail=yes) / P(Sky=overcast,Sea=gray, Sun=true)
= α P(Sail=yes) * P(Sky=overcast | Sail=yes) * P(Sea=gray | Sail=yes) * P(Sun=true | Sail=yes)
= α (0.7)(0.29)(0.29)(0.57) = α (0.034)
P(Sail=no | Sky=overcast, Sea=gray, Sun=true)
= P(Sky=overcast,Sea=gray,Sun=true | Sail=no) * P(Sail=no) / P(Sky=overcast,Sea=gray, Sun=true)
= α P(Sail=no) * P(Sky=overcast | Sail=no) * P(Sea=gray | Sail=no) * P(Sun=true | Sail=no)
= α (0.3)(0.33)(1.0)(0.2) = α (0.020)
P(Sail | Sky=overcast, Sea=gray, Sun=true) = <0.63,0.37>

f. Since P(Sail=yes | Sky=overcast,Sea=gray,Sun=true) > P(Sail=no | Sky=overcast,Sea=gray,Sun=true) then naïve Bayes would classify the new instance as Sail=yes.

2. *Nearest Neighbor*. Suppose we want to classify the new instance <Sky=overcast, Sea=gray, Sun=true> using the nearest neighbor learning method.
   a. We will be using Euclidean distance, so we must first convert the data to numeric values. Translate the examples and instance according to the mapping: clear → 0, cloudy → 1, overcast → 2, blue → 0, gray → 1, false → 0, true → 1. Show a new table of examples using this mapping.
   b. Show the Euclidean distance between the new instance (expressed using the mapping from part (a)) and the 10 training examples in your table from part (a).
   c. Indicate which examples would be used by 3-nearest-neighbor to classify this new instance. Which class would 3-nearest-neighbor choose for this new instance?
   d. Indicate which examples would be used by 7-nearest-neighbor to classify this new instance. Which class would 7-nearest-neighbor choose for this new instance?

Solution:

a. Table below. New instance would be <Sky=2, Sea=1, Sun=1>.

Sky	Sea	Sun	Sail	Distance
0	0	0	yes	2.45
0	1	1	yes	2.00
0	1	0	no	2.24
1	0	1	yes	1.41**
1	0	0	yes	1.73**
1	1	1	yes	1.00*
1	1	0	no	1.41**
2	0	1	yes	1.00*
2	0	0	yes	1.41**
2	1	0	no	1.00*

b. Distances shown in table above.

c. Examples with one asterisk (*) are the 3 nearest neighbors, which would classify the new instance as Sail=yes.

d. Examples with one (*) or two (**) asterisks are the 7 nearest neighbors, which would classify the new instance also as Sail=yes.