Course: Machine Learning Techniques
Week 7: Graded Questions Solutions

1. (1 point) Consider figure 1. Pink points represent class 1 while blue points represent class 2. What will be the label predicted by 3-NN classifier on the new green point? Use Euclidean distance as distance metric. Assume coordinates of all points are integers.

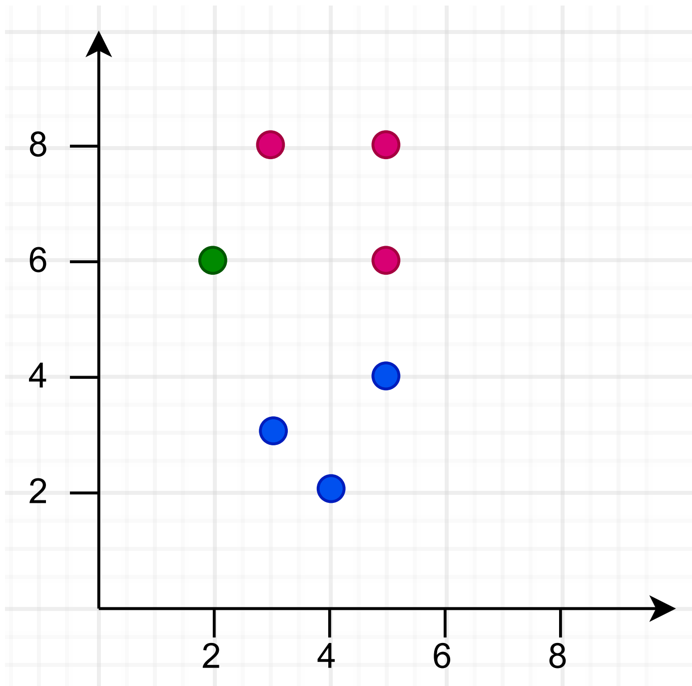

Figure 1:

A. Class 1

B. Class 2

C. can not be determined

Answer: A

Solution:

Point	Distance from new point	Class	Rank
(3,3)	$\sqrt{(3-2)^2+(3-6)^2} = \sqrt{10}$	blue	3
(4,2)	$\sqrt{(4-2)^2+(2-6)^2} = \sqrt{20}$	blue	5
(5,4)	$\sqrt{(5-2)^2+(4-6)^2} = \sqrt{13}$	blue	4
(5,6)	$\sqrt{(5-2)^2+(6-6)^2} = \sqrt{9}$	pink	2
(3,8)	$\sqrt{(3-2)^2+(8-6)^2} = \sqrt{5}$	pink	1
(5,8)	$\sqrt{(5-2)^2+(8-6)^2} = \sqrt{13}$	pink	4

Since the three nearest neighbours have labels (pink, pink, blue), hence the predicted label should be pink, i.e. class 1.

2. (1 point) What will be the class label for 4-NN classification? Use Euclidean distance as distance metric.

   A. Class 1

   B. Class 2

   C. can not be determined

Answer: B

Solution:

Looking at the table created in solution to previous question, there there two candidate points for 4^{th} neighbour. Hence the predicted label can not be computed deterministically.

3. (1 point) What will be the regression output label associated with the green point, if 3-NN regressor is used? x axis represents feature value and y axis represents corresponding label value. Use Euclidean distance as distance metric. Use mean strategy to compute label from the k nearest neighbours.

Answer: 6

Solution:

x	y	**Distance from new point**	**Rank**
3	3	$(3-2)=1$	1
4	2	$(4-2)=2$	3
5	4	$(5-2)=3$	4
5	6	$(5-2)=3$	4
3	8	$(3-2)=1$	1
5	8	$(5-2)=3$	4

Since the mean strategy is used, the predicted label should be $\dfrac{8+3+2}{3} = 4.33$

4. (1 point) A class of distributions is in the exponential family if it can be written in the form $p(y;\eta) = b(y)\exp(\eta^T T(y) - a(\eta))$. For Bernoulli distribution with mean ϕ

   A. $T(y) = -y$
   $a(\eta) = -\log(1-\phi)$
   $b(y) = 1$

   B. $T(y) = y$
   $a(\eta) = \log(1-\phi)$
   $b(y) = 1$

C. $T(y) = y$
$a(\eta) = -\log(1-\phi)$
$b(y) = 1$

D. $T(y) = -y$
$a(\eta) = \log(1-\phi)$
$b(y) = 1$

Answer: C

Solution: Refer to slides for this question.

5. (1 point) Given the following training set:

Point	**x**	y
A.	(-8,20)	0.4
B.	(-12, 11)	0.6
C.	(0, 26)	0.1
D.	(-5, 5)	1.0
E.	(-1, 22)	0.8
F.	(-7, 1)	-0.3

Then the target value of the test point (-10,0) obtained by performing 1-nearest neighbor regression using the Euclidean metric will be:

A. 0.4

B. 0.6

C. 0.1

D. 1.0

E. -0.3

F. 0.8

Answer: E

Solution:

Point	**x**	y	**Distance**
A.	(-8,20)	0.4	$\sqrt{(-8-(-10))^2+(20-0)^2} > 20$
B.	(-12, 11)	0.6	$\sqrt{(-12-(-10))^2+(11-0)^2} > 11$
C.	(0, 26)	0.1	$\sqrt{(0-(-10))^2+(26-0)^2} > 26$
D.	(-5, 5)	1.0	$\sqrt{(-5-(-10))^2+(5-0)^2} > 5$
E.	(-1, 22)	0.8	$\sqrt{(-1-(-10))^2+(22-0)^2} > 22$
F.	(-7, 1)	-0.3	$\sqrt{(-7-(-10))^2+(1-0)^2} = 3$

By observing from above table, the closest point is F, hence label F should be the correct answer, i.e. -0.3.

6. (1 point) For the following vectors, select the correct values of Euclidean and Manhattan distances respectively:
$x = [3, -45, 96, 56, -3.9]$, $y = [33, -102, 7, 0.7, -10]$

   A. 123.15, 237.4

   B. 100.88, 237.4

   C. 123.15, 237.6

   D. 100.88, 237.6

Answer: A

Solution: Euclidean Distance:

$\sqrt{(3-33)^2 + (-45-(-102))^2 + (96-7)^2 + (56-0.7)^2 + (-3.9-(-10))^2}$

$= \sqrt{900 + 3249 + 7921 + 3058.09 + 37.21}$

$= \sqrt{15165.3}$
$=123.147$

Manhattan Distance:

$|(3-33)| + |(-45-(-102))| + |(96-7)| + |(56-0.7)| + |(-3.9-(-10))|$
$= 30 + 57 + 89 + 55.3 + 6.1$
$= 237.4$

7. (1 point) Which of the following statements are true? (Multiple options may be correct.)

   A. kNN models with low values of k produce complex decision boundaries.

   B. kNN models with high values of k produce complex decision boundaries.

   C. kNN models with low values of k produce simple decision boundaries.

   D. kNN models with high values of k produce simple decision boundaries.

Answer: A, D

8. (1 point) Find the Manhattan distance between $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$

Answer: 2

Solution: Manhattan distance can be computed as following:

$= |1-2| + |0-0| + |1-0|$
$= 2$

Connected questions Q9-Q12 Consider the following vector $\begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}$. Use softmax function to answer the following

9. (1 point) Probability associated with first element of the vector is

   Answer: 0.088 (0.086 to 0.091)

10. (1 point) Probability associated with second element of the vector is

    Answer: 0.241 (0.238 to 0.244)

11. (1 point) Probability associated with third element of the vector is

    Answer: 0.012 (0.010 to 0.014)

12. (1 point) Probability associated with fourth element of the vector is

    Answer: 0.657 (0.637 to 0.667)

    Solution: Solution of connected questions Q9-12: Lets call the pertinent vector **a**.

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}$$

$$e^{\mathbf{a}} = \begin{bmatrix} 2.718 \\ 7.389 \\ 0.367 \\ 20.085 \end{bmatrix}$$

$$\frac{e^{\mathbf{a}}}{sum(e^{\mathbf{a}})} = \begin{bmatrix} 0.088 \\ 0.241 \\ 0.012 \\ 0.657 \end{bmatrix}$$

13. (2 points) Consider a 3-dimensional vector $\mathbf{v} = [2.1\ 4.8\ 3.5]$ of arbitrary real values. If you apply the $softmax$ function to this vector, which of the following is the output 3-dimensional vector of real values?

    A. [0.201 0.461 0.336]

    B. [0.886 0.0889 0.0242]

    C. [0.0502 0.7464 0.2034]

    D. [0.440 0.315 0.244]

Answer: C

Solution:

Using similar approach as in previous question:

$$\mathbf{v} = \begin{bmatrix} 2.1 \\ 4.8 \\ 3.5 \end{bmatrix}$$

$$e^{\mathbf{v}} = \begin{bmatrix} 8.166 \\ 121.510 \\ 33.115 \end{bmatrix}$$

$$\frac{e^{\mathbf{v}}}{sum(e^{\mathbf{v}})} = \begin{bmatrix} 0.050 \\ 0.746 \\ 0.203 \end{bmatrix}$$