NETAJI SUBHAS UNIVERSITY OF TECHNOLOGY

MACHINE LEARNING

KNN QUIZ

ADARSH KUMAR

2020UCO1663

COE-3

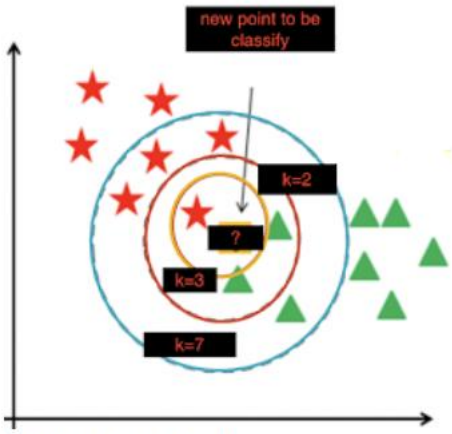
LINKS TO FILES

QUIZ	Quiz google form
STUDENT RESPONSES	Responses

QUESTIONS

QUESTION	ANSWER
How to avoid a clashes between two classes of data? * *Mark only one oval.* ◯ Choose a larger value of k to avoid any chances of the confusion. ◯ Choose a smaller value of k to avoid any chances of the confusion. ◯ Choose even value of k to avoid any chances of the confusion. ◯ Choose odd value of k to avoid any chances of the confusion.	Choose odd value of k to avoid any chances of the confusion.
A small value of k could lead to _____ as well as a big value of k can lead to _____. *Mark only one oval.* ◯ underfitting, underfitting ◯ overfitting, underfitting ◯ underfitting, overfitting ◯ overfitting, overfitting	overfitting, underfitting

Which of the following distance variables do we use in case of categorical variables in K-NN? *Mark only one oval.* ◯ Hamming distance ◯ Euclidean distance ◯ Manhattan distance ◯ Cosine distance	Hamming distance
Which of the following is true about Manhattan distance? * *Mark only one oval.* ◯ It can be used for continuous variables. ◯ It can be used for categorical variables. ◯ It can be used for categorical and continous variables.	It can be used for continuous variables.
What is the correct order of time taken by 1-NN, 2-NN, 3-NN for creation of model? *Mark only one oval.* ◯ 1-NN > 2-NN > 3-NN ◯ 3-NN > 2-NN > 1-NN ◯ 1-NN ~ 2-NN ~ 3-NN ◯ 2-NN > 3-NN > 1-NN	1-NN ~ 2-NN ~ 3-NN

What will be the predicted value of unknown class '?' on choosing different value of k and which value of k should be preferred ?  *Mark only one oval.* ◯ k =2 class predicted be red ◯ k=3 class predicted be green ◯ k=7 class predicted be red ◯ k=5 class predicted be green	k =2 class predicted be red
IN order to reduce noise in data which of the following option would you consider in k-NN? *Mark only one oval.* ◯ I will increase the value of k ◯ I will decrease the value of k ◯ Noise can not be dependent on value of k ◯ None of the above	I will increase the value of k
Considering there is no limit of the size of data points, what is the most appropriate measure to pre process our data points? * 1 point *Mark only one oval.* ◯ It is not necessary to pre process the data points ◯ We need to standardize the data points considering it follows a gaussian pattern ◯ We need to normalise the data points considering they may not follow the gaussian patter	We need to standardize the data points considering it follows a gaussian pattern

If a data point is classified by both binary as well as real number values then which distance variable to be used to predict an unknown data point? * 1 point

X	Y	Class
1	3.7	+
0	3.9	-
0	2.4	+
1	3.3	-
0	2.8	+
1	2.2	-

Mark only one oval.

- ( ) we first normalize the real number variable then compute its euclidean distance
- ( ) we first normalize the real number variable then compute its Hamming distance
- ( ) just find the euclidean distance without normalizing
- ( ) if both features x and y has equal priority then calculating either of the two distance would be correct

we first normalize the real number variable then compute its Hamming distance

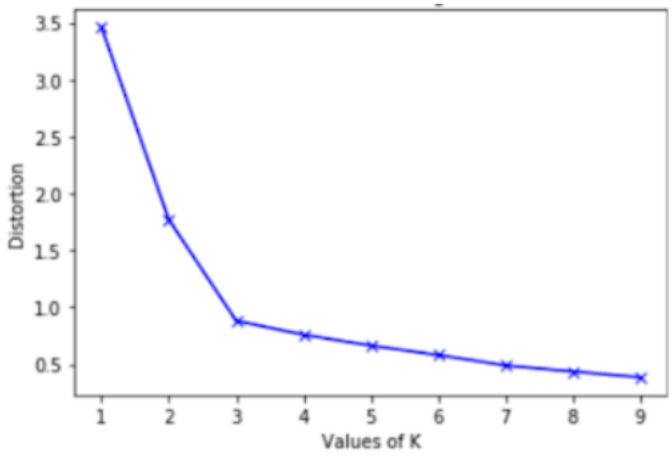
Which of the following will be true about k in k-NN in terms of variance and bais? * 1 point

Mark only one oval.

- ( ) When you increase the k the bias will be increases and variance increase
- ( ) When you decrease the k the bias will be increases and variance decrease
- ( ) When you decrease the k the variance will increases and bais will decrease
- ( ) value of k does not effect bais and varaince

When you decrease the k the variance will increases and bias will decrease

Which of the following statement is true about k-NN algorithm? 1. k-NN performs much better if all of the data have the same priority 2. k-NN works well with a small number of input data, but struggles when the number of data points is very large 3. k-NN makes no assumptions about the functional form of the problem being solved *Mark only one oval.* 1 and 2 1 and 3 Only 1 1, 2 and 3	1, 2 and 3
What is the time complexity of the *k*-NN algorithm? * **n = cardinality of the training set** **d = dimension of each sample** *Mark only one oval.* O(d + n) O(d) O(n*d) O(n)	O(n*d)

Using the Elbow Method, at which point is the value of K most suitable  *Mark only one oval.* ◯ 3 ◯ 2 ◯ 4 ◯ 1	3
In elbow method, if we choose the number of clusters equal to the data points, then the value of WCSS becomes _____, and that will be the _____ of the plot. *Mark only one oval.* ◯ Zero, Middle ◯ One, Endpoint ◯ Zero, Endpoint ◯ Infinite, Endpoint	Zero, Endpoint