Shahjalal University of Science and Technology
Institute of Information and Communication Technology (IICT)
Software Engineering
4th Year 2nd Semester Final Examination' Dec 2022 (Session: 2017-18)
Course Code: SWE 427 Credits: 3 Course Title: Machine Learning
Time: 3 hrs Total Marks: 60

Group A
[Answer all the questions]

1. **Answer any FIVE** 5x1=5

a) What is underfitting?

b) K-means automatically adjust the number of clusters - true or false?

c) What is decision boundary?

d) What is model bias?

e) What is the "learning rate" of an algorithm?

f) What is a feed forward network?

g) What is a confusion matrix?

2. **Answer any FOUR** 4x2.5=1

a) Is a model with higher or lower variance better? Explain. 2.5

b) Calculate the gradient descent formula for the following Polynomial Regression model: 2.5

$$h(x) = a_1 x_1 + a_2 x_1^2 + a_3 x_2$$

c) What is loss function? Give a few examples of loss function. 2.5

d) Derive the expression of Naive Bayes classifier from Conditional Probability. 2.5

e) What is Regularization? Explain in brief. 2.5

f) What is overfitting in model development? How can it be avoided? 2.5

3. **Answer any THREE** 3x5=15

a) Does gradient descent require a convex cost function? Can we use Mean Squared Error for calculating gradient descent of Logistic Regression to converge to the global optima? If not, why? 5

b) Plot the following data points: (0,0,Red), (0,1,Red), (0,2,Red), (1,0,Red), (1,1,Red), (1,3,Green), (2,3,Green), (2,4,Green), (3,2,Green) and find the optimal separating line between classes. Point out the support vectors and calculate the margin. 5

c)

Type	Long	Not Long	Sweet	Not Sweet	Yellow	Not Yellow	Total
Banana	400	100	350	150	450	50	500
Orange	0	300	150	150	300	0	300
Other	100	100	150	50	50	150	200
Total	500	500	650	350	800	200	1000

Let's say you are given a fruit that is : Long, Sweet and Yellow, can you predict what fruit it is? (1)

d) Suppose there is a population of army personnel of two regions A : B = 1 : 20. Twelve percent of the people from region A and thirty percent of the people from region B get a chance in the UN peace program every year. Now, if we choose a person who got a chance in the UN peace program, what are the chances that he is from region A?

Group B

[Answer all the questions]

4. Answer any FIVE **5x1=5**

a) What is the Sensitivity of a ML model?

b) What is MinMax scaling?

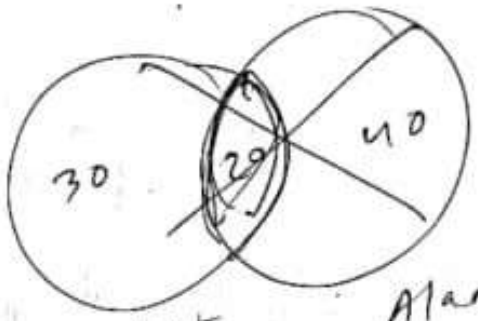
c) What is "Kernel Trick"?

d) What is activation function?

e) What is MSE?

f) What is a Generative learning algorithm?

g) When the feature space is large, overfitting is more likely - true or false?



bucket + Alarm

20

60

5. Answer any FOUR **4x2.5=10**

a) What is the F1 score of ML models? How is it used in selecting the best model? 2.5

b) Why do we need dimensionality reduction? What are its drawbacks? 2.5

c) What is propagation? Explain in brief. 2.5

d) What are the limitations of PCA? 2.5

e) In a group of 100 sports car buyers, 40 bought alarm systems, 30 purchased bucket seats, and 20 purchased an alarm system and bucket seats. If a car buyer chosen at random bought an alarm system, what is the probability they also bought bucket seats? 2.5

f) What is cross validation? Why is it used? 2.5

6. Answer any THREE **3x5=1**

a) What are support vectors? How do support vectors help to find the optimal margin of a model? 1+

b) Apply K-means clustering and Manhattan distance to the cluster {3,4,6,12,18,22,23,32,33,35} into 2 clusters. Suppose the initial centroids are 9 and 22. Show all the steps in a chart. Draw a line with all the 10 points and show the final cluster and respective center. Use at best 3 iterations if it does not converge.

c) What is RBF? Briefly explain how this function transforms the data into higher dimensions and reduces computation.

	Target	Prediction
1	Spam	Not Spam
2	Spam	Spam
3	Not Spam	Not Spam
4	Spam	Spam
5	Not Spam	Spam
6	Not Spam	Not Spam
7	Not Spam	Not Spam
8	Spam	Spam
9	Spam	Not Spam
10	Not Spam	Not Spam

Calculate the confusion matrix from the table above and then find Accuracy, Sensitivity, Specificity and Precision from the matrix.