**B.E. COMPUTER SCIENCE AND ENGINEERING \FOURTH YEAR FIRST SEMESTER
SUPPLEMENTARY EXAM 2018**

Machine Learning

Time: Three hours   Full Marks: 100

Answer any FIVE questions.

1. a) Discuss gradient descent search algorithm for weight update of a linear unit perceptron model.

    b) What is the drawback of gradient descent search procedure?

    c) Justify with a suitable example the need of hidden layers in Artificial Neural Networks.

    10 + 5 +5 = 20 marks

2. a) Discuss KNN algorithm. Give a suitable numerical example to show how it works.

    b) Develop a **Naïve Bayes Classifier** (show the training phase by computing all necessary probabilities) using the following training data and predict label for the test instance: (Outlook=*Sunny,* Temperature=*Cool,* Humidity=*High,* Wind=*Strong*).

PlayTennis: training examples

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

10 + 10 = 20 marks

[Turn over

3. Describe *backpropagation algorithm* used for training Artificial Neural Networks. Give a numerical example to show how it works.   10 + 10=20 marks

4. a) Consider the following set of training examples and compute (a) the entropy of this collection of training examples with respect to the target function classification and (b) the information gain of a_2 relative to these training examples?

Instance	Classification	a1	a2
1	+	T	T
2	+	T	T
3	-	T	F
4	+	F	F
5	-	F	T
6	-	F	T

   b) Explain how continuous attributes are handled in decision tree learning. Give an example to illustrate the concept.

   c) Justify with an example the following statement
           "Noisy training data may lead to longer decision tree"   10 + 5 + 5 = 20 marks

5. a) Describe linear SVM.

    b) Consider the following confusion matrix depicting the performance of a classification model and compute (i) accuracy, (ii) error rate, (iii) class wise F-measure, (iv) overall F-measure

Predicted Actual	yes	no
yes	6954	46
no	412	2588

                                                        10 + (2+2+3+3)=20 marks

6. Write short notes on the following
    a) Unbiased Bayesian learning
    b) Impact of *Momentum* parameter while training Artificial Neural Networks
                                                        (10 x 2) = 20 marks

7. What are differences between supervised and unsupervised learning? Write the *K-means* clustering algorithm. Give a numerical example to show how the K-means algorithm works. Explain why this algorithm is named as "K-means"? Define error function for this algorithm.
                                                        4+5+5+3+3=20 marks