NMAM INSTITUTE OF TECHNOLOGY, NITTE
(An Autonomous Institution affiliated to VTU, Belagavi)
VII Sem B.E. (CSE) Mid Semester Examinations - II, October 2017
14CS716 – MACHINE LEARNING

Duration: 1 Hour

Max. Marks: 20

*Note: Answer any **One** full question from **each Unit**.*

Unit – I

Marks   BT*

1. a) Consider the below dataset.

Tid	Refund	Marital Status	Taxable Income	Class
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

    i.   Compute entropy for the entire dataset.
    ii.   Compute the information gain for Refund attribute.   7   L*4

  b) Give the equation for computing posterior probability and class conditional probability in Bayesian classifier assuming input x is of single dimension and y∈ {c1, c1, ..., ck}

    3   L2

2. a) For the dataset given in Q 1(a), compute the class conditional probabilities for all the possible values of Refund and Marital status attribute with both class labels yes and no.   7   L4
  b) Write the steps for K-nearest neighbor algorithm.   3   L2

Unit – II

3. a) Consider the set of five points A1(2,10), A2(2,5), A3(8,4), A4(5,8) and A5(7,5). Assuming the initial cluster means as, A1 and A2, find the cluster grouping after first iteration.   7   L4
  b) List the different methods of combining classifiers.   3   L1

4. a) Consider the set of five points A1(2,10), A2(2,5), A3(8,4), A4(5,8) and A5(7,5) Apply hierarchical clustering to group the points.   7   L5
  b) What is cross validation?   3   L2

BT* Bloom's Taxonomy, L* Level

****