Reg. No.										

MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

A Constituent Institution of Manipal University

VI SEMESTER B.TECH. MAKEUP EXAMINATIONS, JUNE 2017

SUBJECT:OPEN ELECTIVE II- MACHINE LEARNING [ICT 3285]

REVISED CREDIT SYSTEM
(24 / 06 /2017)

Time: 3 Hours MAX. MARKS: 50

Instructions to Candidates:

- Answer ALL questions.
- Missing data may be suitable assumed.

1A Write and describe the practical usage of the following dimensionality reduction algorithms in machine learning: **5**
(i) Principal Component Analysis (PCA)
(ii) Independent Component Analysis (ICA)

1B Compare and contrast supervised and unsupervised learning with an example for each. **3**

1C Describe any two major strengths and weaknesses of partition based algorithms. **2**

2A Describe the working principles of AdaBoost algorithm for a classification problem. Illustrate the working of the algorithm for a two-class classification problem. **5**

2B Give two examples in any area of interest to you where regression analysis can be used as a data analytic tool to answer some questions of interest. For each example: **3**
(i) What is the question of interest?
(ii) Identify the response and the predictor variables.
(iii) Classify each of the variables as either quantitative or qualitative.
(iv) Which type of regression can be used to analyze the data?

2C Mention any four major characteristics of a Hidden Markov Model (HMM). **2**

3A Describe the following: **5**
(i) Bootstrapping and its usage in Classification Algorithms.
(ii) Bagging Algorithm.

3B The data for a road resurfacing project is given in Table Q.3B. In this we can assume to have a unit of association: The connection between a particular cost and mileage is that they are based on the same project. **3**

Table Q. 3B

Cost yi (in $1000)	6.0	14.0	10.0	15.0	26.0
Mileage xi (in miles)	1.0	3.0	4.0	5.0	7.0

(i) Draw a scatter plot for the given data.
(ii) Derive the regression line equation from the scatter plot.
(iii) Find the least squares estimates of the slope and intercept of the regression equation thus formulated.

3C Describe the Q-Learning algorithm. 2

4A Describe the Expectation-Maximization (EM) algorithm for unsupervised learning. 5

4B Describe the algorithm for generic ensemble framework for supervised learning. 3

4C Describe the concept of Overfitting in the context of model complexity: 2
(i) Overfitting due to lack of representative samples; &
(ii) Overfitting due to presence of noise.

5A Describe the following: 5
(i) Convex Optimization Problems.
(ii) Gaussian Processes.
(iii) Feature Selection Algorithms.

5B Compare and contrast the techniques of "Sampling with Replacement" and "Sampling without Replacement" giving one relevant example for each. 3

5C Explain the role of Kernel functions in solving Non-Linear Support Vector Machine Problems. 2