Reg. No.										

MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

(A constituent unit of MAHE, Manipal)

II SEMESTER M.TECH. (COMPUTER NETWORKING AND ENGINEERING / SOFTWARE ENGINEERING)

END SEMESTER EXAMINATIONS, APRIL/MAY 2019

SUBJECT: ELECTIVE I MACHINE LEARNING [ICT 5239]

REVISED CREDIT SYSTEM

(29/04/2019)

Time: 3 Hours MAX. MARKS: 50

Instructions to Candidates:

- Answer ALL the questions.
- Missing data, if any, may be suitably assumed.

1A.	Given a training set comprising of a m by n matrix X and a m-dimensional vector $\overline{y}$ containing all the target values from the training set. Assuming Least Squares Regression, find in closed form the value of θ that minimizes Least Squares Cost function $J(\theta)$	5
1B.	Consider the Bernoulli distribution parameterized by ϕ given by $$p(y; \phi) = \phi^y(1-\phi)^{y-1}$$ Show that the Bernoulli distribution is an exponential family distribution. Explicitly specify $b(y)$, η, $T(y)$ and $a(\eta)$. Also write ϕ in terms of η.	3
1C.	"Locally weighted linear regression is a non-parametric algorithm". Justify the statement.	2
2A.	Consider a classification problem in which the response variable y can take on any one of the k values, so $y \in \{1, 2, \ldots, k\}$. Derive a Generalized Linear Model (GLM) for modelling this type of multinomial data.	5
2B.	Justify the need for feature selection. Discuss the various methods for feature selection.	3
2C.	"GDA and logistic regression will, in general, give different decision boundaries when trained on the same dataset". Which is better? Discuss.	2
3A.	Describe the following cross-validation techniques. For each technique write its suitability condition. (i) Hol d-out (ii) K-fold (iii)Leave-one out	5
3B.	Write the k-means algorithm. Also write the distortion function J and explain the reasons for J not to converge to a global minima and how to overcome that.	3
3C.	Briefly discuss the various types of inherent ambiguities associated with Independent Component Analysis(ICA).	2

4A. Suppose you have an estimation problem in which you are given a training set $\{x(i), i=1,....m\}$ consisting of m independent examples. It is required to fit the parameters of a model $p(x,z)$ where z is a latent variable. The likelihood is given by 5

$$l(\theta) = \sum_{i=1}^{m} \log \sum_{z} p(x, z; \theta)$$

Explicit finding of maximum likelihood estimate of the parameters θ may be difficult. With all the necessary steps, show how Expectation Maximization algorithm can be applied for the given problem.

4B. Given the unit vector u and $x^{(i)}$ the points in the dataset, how would you maximize the variance of the projection $x^{(i)}$ onto u. Obtain the relation for the principal eigen vector of the covariance matrix. Specify why is it a dimensionality reduction algorithm. 3

4C. The factor analysis model is defined as follows: 2

$$z \sim \mathcal{N}(0, I)$$
$$\varepsilon \sim \mathcal{N}(0, \Psi)$$
$$x = \mu + \Lambda z + \varepsilon$$

where ε and z are independent. The random variables z and x have a joint Gaussian distribution

$$\begin{bmatrix} z \\ x \end{bmatrix} \sim \mathcal{N}(\mu_{zx}, \Sigma).$$

Find μ_{zx} and Σ.

5A. Starting with the following optimization problem for SVM, 5

$$minimize_{\gamma,\omega,b} \; \frac{1}{2} \|\omega\|^2$$
$$s.t \; y^{(i)}(\omega^T x^{(i)} + b) \geq 1, i = 1, ..., m$$

Solve for ω and b using primal and dual problem approach.

5B. Describe the Markov Decision Processes(MDP). Define *policy* and *optimal value function* with respect to MDP. 3

5C. Write the algorithm for value iteration and policy iteration for finite state MDP. 2