

Mid-Semester Test
(EC-2 Regular)

Course No. : IS ZC464
Course Title : MACHINE LEARNING
Nature of Exam : Closed Book
Weightage : 30%
Duration : 2 Hours
Date of Exam : 30/09/2018 (FN)

No. of Pages	= 2
No. of Questions	= 5

Note:

1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

- Q.1. Explain the following concepts with respect to machine learning. [2 + 2 + 2 + 2 = 8]
- (i) Difference between prediction and regression
 - (ii) Difference between lazy and active learning
 - (iii) Highlight any two important ML achievements of recent past
 - (iv) What do we mean by the model and hypothesis in ML

- Q.2. Hidden Markov Model (HMM) presents us a setting where we do not have access to the actual state of the system instead we have some observations that are mostly probabilistic. Let there be a hotel in Shimla where you would be offered a free coffee on a cold day with probability 0.7 and on a hot day with 0.4. Also let the probability that for a hot day, next day would also be hot is 0.6 and on a cold day next day would also be cold is 0.7. Assume on a particular day (say day-1) when you checked in the hotel the probability of that day was cold be 0.7. You know for sure that when you stayed there for three consecutive days, you got coffee on day-1 and day-3 but not on day-2. What is the optimum day sequence in Shimla during your stay according to HMM. [Hint: it would be something like CCC, CHC, HHC, ...]. [5]

- Q.3. Consider a database having three attributed with possible values as show below [5]

Attribute	Possible Values
Sky	Sunny/Cloudy
AirTemp	Warm/Cold
Humidity	Normal/High

Apply **candidate elimination algorithm** to determine **version space** for the following training data. Clearly show general and specific boundary in each step during the training. (target variable is EnjoySport)

Sky	AirTemp	Humidity	EnjoySport
Sunny	Warm	Normal	+ve
Cloudy	Warm	Normal	+ve
Sunny	Cold	Normal	-ve
Sunny	Warm	High	+ve

- Q.4. Provide proof for the following statement (clearly state all the steps).
“A hypothesis that minimizes the sum of square distances corresponds to maximum likelihood hypothesis” [4]
- Q.5. Construct a decision tree out of the following training data for the classification of mango as ripe or not. [8]

Color	Weigh	Texture	Softness	Ripe?
Red	light	Rough	Very Soft	Y
Green	moderate	Rough	hard	Y
Yellow	heavy	Smooth	Very Soft	Y
Green	light	Smooth	Soft	N
Yellow	moderate	Smooth	Very Soft	N
Red	heavy	Smooth	Soft	Y
Green	light	Smooth	Very Soft	N
Yellow	moderate	Rough	Soft	Y
Red	moderate	Rough	Soft	Y
Green	heavy	Smooth	hard	N
Yellow	light	Rough	Very Soft	Y
Green	moderate	Rough	hard	Y
Yellow	heavy	Smooth	Soft	N
Red	light	Rough	hard	Y
