Question Paper

Exam Date & Time: 20-Apr-2018 (10:00 AM - 01:00 PM)

MANIPAL ACADEMY OF HIGHER EDUCATION

SCHOOL OF INFORMATION SCIENCES (SOIS)
SECOND SEMESTER ME- (BIG DATA AND DATA ANALYTICS)
DEGREE EXAMINATION- APRIL 2018
Friday, 20 April 2018
Time : 10:00am to 1:00pm
Machine Learning [BDA 602]

Marks: 100 **Duration: 180 mins.**

Answer all the questions.

1) A) Write the steps involved in choosing the Training Experience in designing a machine learning systems. (6)

B) How will you choose the Target Function in designing a machine learning system? Explain with example (4)

2) Write FIND-S algorithm to find a Maximally Specific Hypothesis and obtain the hypothesis space search by Find-S for the given training examples. (10)

(3+7 Marks)

Example	*Sky*	*AirTemp*	*Humidity*	*Wind*	*Water*	*Forecast*	*EnjoySport*
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

3) Define inductive bias. Explain the inductive bias of the Candidate-elimination algorithm? (8)

(4+4 Marks)

4) Obtain the entropy and information gain to find the root node of the decision tree for the following training examples (12)

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

5) A) Defining the true error of a hypothesis h with respect to target concept c and instance distribution D. With suitable diagram explain the error of a hypothesis h with respect to target concept (7)

B) Define PAC Learning (3)

6) State and prove the ϵ-Exhausted Version Space theorem to determine the number of training examples required to reduce this probability of failure below some desired level δ. (10)

7) A) Define Bayes theorem. (3)

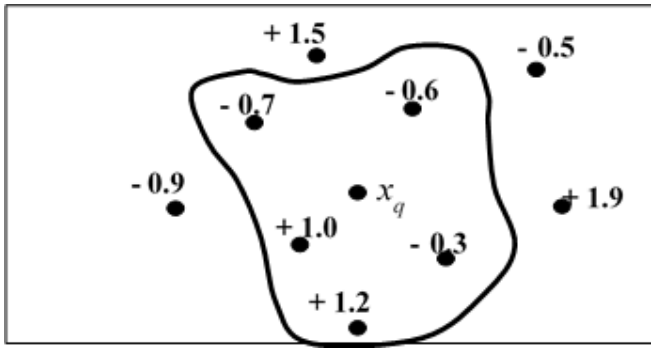
B) Find a patient's probability of having liver disease if they are an alcoholic for the following data: The past data tells you that 10% of patients entering your clinic have liver disease and five percent of the clinic's patients are alcoholics. Among those patients diagnosed with liver disease, 7% are alcoholics. If the patient is an alcoholic, what is the chances of having liver disease? (7)

8) A) Data set given in the table is for a company produce tissues (used by biological labs). Company objective is to predict how well their products are accepted by the clients. They conducted a survey with their clients to find the acceptance levels of the products. As shown in figure, Type- 1 and 2 are not well accepted whereas Type- 3 and 4 are well accepted. (6 Marks) (6)

Name	Acid Durability	Strength	Class
Type-1	7	7	Bad
Type-2	7	4	Bad
Type-3	3	4	Good
Type-4	1	4	Good

B) In the diagram below the numbers next to the + and - signs refer to the values taken by a real-valued target function. Calculate the (4)

value predicted for the target function at the query instance xq by the 5-Nearest Neighbour Learning.



9) Explain the structure of single layer and multilayer artificial neural network models with suitable diagrams (5)

A)

B) Describe the architecture models of feed forward and Recurrent Neural Network models with suitable diagrams. (5)

10) Describe the architecture of McCulloch-Pitts Neuron Model with suitable diagram. Also realize the output of logical NOT function using McCulloch-Pitts neuron model. (10)

-----End-----