22/09/2020

ASSIGNMENT 1

1) What is machine learning and list its current applications.

2) Find out the features of the following learning Problems. i) Weather prediction problem ii) Handwritten recognition problem iii) Tumor detection problem.

3) Discuss the following with respect to the checkers learning problem

   i) Choosing the Training Experience

   ii) Choosing the target function

   iii) Choosing the function approximation algorithm

4)

Perform the steps of the Find S and Candidate Elimination Algorithm on the Training data given in the table below and show the values of set G and S after each record and show the final results.

Example No	Newspaper Name	Circulation	Publishing city	Truth Value
1	The News	Pakistan	Lahore	1
2	Jang	Punjab	Karachi	1
3	The Nation	Punjab	Lahore	0
4	Jang	Pakistan	Islamabad	0
5	The Nation	Sindh	Lahore	1

This table has three attributes Newspaper Name, Circulation and publishing city that can contain the values,

Newspaper Name = [The News, The Nation, Jang]
Circulation = [Pakistan, Punjab, Sindh]
Publishing city = [Lahore, Karachi, Islamabad]

5) Apply the Find S and Candidate Elimination algorithm in constructing the possible consistent hypotheses

EXAMPLE	COLOR	TOUGHNESS	FUNGUS	APPEARANCE	POISONOUS
1.	GREEN	HARD	NO	WRINKELD	YES
2.	GREEN	HARD	YES	SMOOTH	NO
3.	BROWN	SOFT	NO	WRINKLED	NO
4.	ORANGE	HARD	NO	WRINKLED	YES
5.	GREEN	SOFT	YES	SMOOTH	YES
6.	GREEN	HARD	YES	WRINKLED	YES
7.	ORANGE	HARD	NO	WRINKLED	YES

ATTRIBUTES ON WHICH THE CONCEPT DEPENDS ON — CONCEPT

6) Consider the following set of training examples to train a machine to predict whether or not an office contains a recycling bin.

	STATUS	**FLOOR**	**DEPT.**	**OFFICE SIZE**	**RECYCLING BIN?**
1	faculty	four	cs	medium	Yes
2.	faculty	four	ee	medium	Yes
3.	student	four	cs	small	No
4.	faculty	five	cs	medium	Yes

Give a sequence of S and G boundary sets computed by the CANDIDATE-ELIMINATION algorithm if it is given the sequence of examples above in the order in which they appear on the table.

7) Comment on the issues in machine learning

8) What are the limitations of Find-S, write an algorithm to overcome the issues.

9) Write and explain the Decision tree for the following transactions

Tid	Refund	Marital Status	Taxable Income	**Cheat**
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**

10) For the transactions shown in the table compute the following
(i) Entropy of the collection of transaction records with respect to classification
(ii) What are the information gain of the a1 and a2 related to the transactions given

Instance	1	2	3	4	5	6	7	8	9
a1	T	T	T	F	F	F	F	T	F
a2	T	T	F	F	T	T	F	F	T
Target class	+	+	-	+	-	-	-	+	-

11) Construct the decision tree for the following data using ID3 algorithm

DAY	A1	A2	A3	classification
1	TRUE	HOT	HIGH	NO
2	TRUE	HOT	HIGH	NO
3	FALSE	HOT	HIGH	YES
4	FALSE	COOL	NORMAL	YES
5	FALSE	COOL	NORMAL	YES
6	TRUE	COOL	HIGH	NO
7	TRUE	HOT	HIGH	NO
8	TRUE	HOT	NORMAL	YES
9	FALSE	COOL	NORMAL	YES
10	FALSE	COOL	HIGH	NO

12) Give the Decision tree representation for the following Boolean function.

i) A **XOR** B

ii) [A ∧ B] ∨ [C ∧ D]