7th Sem B.Tech (CSE) Assignment-I – M.L

<table>
<tr><td rowspan="2">1</td><td>a</td><td>Discuss the two approaches to prevent over fitting the data.</td><td>7</td></tr>
<tr><td>b</td><td>Consider the following set of training examples.
<table>
<tr><th>Instance</th><th>Classification</th><th>A1</th><th>A2</th></tr>
<tr><td>1</td><td>1</td><td>1</td><td>1</td></tr>
<tr><td>2</td><td>1</td><td>1</td><td>1</td></tr>
<tr><td>3</td><td>0</td><td>1</td><td>0</td></tr>
<tr><td>4</td><td>1</td><td>0</td><td>0</td></tr>
<tr><td>5</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>6</td><td>0</td><td>0</td><td>1</td></tr>
</table>
i) What is the entropy of this collection of training examples with respect to the target function classification?
ii) What is the information gain of A2 relative to these training examples?</td><td>10</td></tr>
<tr><td rowspan="3">2</td><td>a</td><td>Define decision tree. Construct the decision tree to represent the following. Boolean functions:
i) A^~b ii) A v[B^C] iii) A XOR B</td><td>6</td></tr>
<tr><td>b</td><td>Write the ID3 algorithm.</td><td>5</td></tr>
<tr><td>c</td><td>What do you mean by gain and entropy? How it is used to build the decision tree.</td><td>6</td></tr>
<tr><td rowspan="2">3</td><td>a</td><td>Define perceptron. Explain the concepts of single perceptron with neat diagram.</td><td>7</td></tr>
<tr><td>b</td><td>Derive derivation of Backpropagation weight training rule.</td><td>10</td></tr>
</table>

7th Sem B.Tech - Assignment II . M.L

1	a	Write Gradient descent algorithm to train a linear unit along with the derivation of gradient descent rule.	7
	b	. Write derivation of Backpropagation rule considering unit j as output unit and unit j as hidden unit.	10
2	a	Derive derivation to show maximum likely hood hypothesis has least-squared error hypotheses.	7
	b	Classify the test data and {Red , SUV , Domestic } using Naïve Bayes classifier for the dataset show in Table Q8(b).	10

Color	Type	Origin	Stolen
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

3	a	Define i) Sample error ii) True error iii) Confidence intervals for discrete valued hypothesis.	7
	b	Last year, five randomly selected students took a math aptitude test before they began their statistics course. The Statitiscs Department has three questions. i) What linear regression equation best predicts statistics performance, based their aptitude scores? ii)If a student made on 80 on the aptitude test, what grade would we expect her to make in statistics? iii)How well does the regression equation fit the data?	10

Student	xi	yj
1	95	85
2	85	95
3	80	70
4	70	65
5	60	70