Module 2

Decision Tree Learning

<table>
<tr><td>1</td><td>Explain Decision Tree Learning and also explain the problem characteristics for which decision tree learning is suitable.</td><td>4 Marks</td></tr>
<tr><td>2</td><td>Explain the concepts of entropy and information gain.</td><td>6 Marks</td></tr>
<tr><td>3</td><td>Explain ID3 algorithm for decision tree learning with an illustrative example.</td><td>6Marks</td></tr>
<tr><td>4</td><td>Explain which attribute is chosen as the best classifier in ID3 algorithm.</td><td>6 Marks</td></tr>
<tr><td>5</td><td>Explain the construction of decision tree for the below given dataset using ID3 algorithm:
<table>
<tr><th>outlook</th><th>temperature</th><th>humidity</th><th>wind</th><th>answer</th></tr>
<tr><td>Sunny</td><td>hot</td><td>high</td><td>weak</td><td>no</td></tr>
<tr><td>sunny</td><td>hot</td><td>high</td><td>strong</td><td>no</td></tr>
<tr><td>overcast</td><td>hot</td><td>high</td><td>weak</td><td>yes</td></tr>
<tr><td>Rain</td><td>mild</td><td>high</td><td>weak</td><td>yes</td></tr>
<tr><td>Rain</td><td>cool</td><td>normal</td><td>weak</td><td>yes</td></tr>
<tr><td>Rain</td><td>cool</td><td>normal</td><td>strong</td><td>no</td></tr>
<tr><td>overcast</td><td>cool</td><td>normal</td><td>strong</td><td>yes</td></tr>
<tr><td>Sunny</td><td>mild</td><td>high</td><td>weak</td><td>no</td></tr>
<tr><td>Sunny</td><td>cool</td><td>normal</td><td>weak</td><td>yes</td></tr>
<tr><td>Rain</td><td>mild</td><td>normal</td><td>weak</td><td>yes</td></tr>
<tr><td>Sunny</td><td>mild</td><td>normal</td><td>strong</td><td>yes</td></tr>
<tr><td>overcast</td><td>mild</td><td>high</td><td>strong</td><td>yes</td></tr>
<tr><td>overcast</td><td>hot</td><td>normal</td><td>weak</td><td>yes</td></tr>
<tr><td>Rain</td><td>mild</td><td>high</td><td>strong</td><td>no</td></tr>
</table></td><td>8 Marks</td></tr>
<tr><td>6</td><td>Consider the following dataset:
<table>
<tr><th>No.</th><th>Student</th><th>First last year?</th><th>Male?</th><th>Works hard?</th><th>Drinks?</th><th>First this year?</th></tr>
<tr><td>1</td><td>Richard</td><td>yes</td><td>yes</td><td>No</td><td>yes</td><td>yes</td></tr>
<tr><td>2</td><td>Alan</td><td>yes</td><td>yes</td><td>Yes</td><td>no</td><td>yes</td></tr>
<tr><td>3</td><td>Alison</td><td>no</td><td>no</td><td>Yes</td><td>no</td><td>yes</td></tr>
<tr><td>4</td><td>Jeff</td><td>no</td><td>yes</td><td>No</td><td>yes</td><td>no</td></tr>
<tr><td>5</td><td>Gail</td><td>yes</td><td>no</td><td>Yes</td><td>yes</td><td>yes</td></tr>
<tr><td>6</td><td>Simon</td><td>no</td><td>yes</td><td>Yes</td><td>yes</td><td>no</td></tr>
</table></td><td>8 Marks</td></tr>
</table>

<table>
<tr><td></td><td>1. What is the entropy of this collection of training examples with respect to the target function classification?
2. Construct a minimal decision tree for this dataset. Show your work.</td><td></td></tr>
<tr><td>7</td><td>1. Use your decision tree from the previous part to classify the following instances:

<table><tr><th>No.</th><th>Student</th><th>First last year?</th><th>Male?</th><th>Works hard?</th><th>Drinks?</th><th>First this year?</th></tr><tr><td>7</td><td>Matthew</td><td>no</td><td>yes</td><td>No</td><td>yes</td><td>??</td></tr><tr><td>8</td><td>Mary</td><td>no</td><td>no</td><td>Yes</td><td>yes</td><td>??</td></tr></table></td><td>8 Marks</td></tr>
<tr><td>8</td><td>Give decision trees to represent the following Boolean function:
a. A ∧ ~B
b. A V (B ∧ C)
c. A XOR B
d. (A ∧ B) V (C ∧ D)</td><td>8 Marks</td></tr>
</table>