Question 1. Classification with Decision tree and Bayes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Age | Income | Student | Credit | Target  Buys Computer |
| 1 | Youth | High | No | Fair | No |
| 2 | Youth | High | No | Excellent | No |
| 3 | Middle | High | No | Fair | Yes |
| 4 | Senior | Medium | No | Fair | Yes |
| 5 | Senior | Low | Yes | Fair | Yes |
| 6 | Senior | Low | Yes | Excellent | No |
| 7 | Middle | Low | Yes | Excellent | Yes |
| 8 | Youth | Medium | No | Fair | No |
| 9 | Youth | Low | Yes | Fair | Yes |
| 10 | Senior | Medium | Yes | Fair | Yes |
| 11 | Youth | Medium | Yes | Excellent | Yes |
| 12 | Middle | Medium | No | Excellent | Yes |
| 13 | Middle | High | Yes | Fair | Yes |
| 14 | Senior | Medium | No | Excellent | No |

Which class the following object is belonging to?

Age=youth, Income=medium, Student=yes, Credit\_rating=fair

**Giải:**

* **Thuật toán decision tree**

Value (Age)= Youth, Middle, Senior

Value (Income)=High, Medium,Low

Value(Student)= Yes,No

Value (Credit\_rating)= Fair, Excellent

Vậy chọn Age làm root

Youth

Middle

Senior

Age

D4,D5,D6,D10,D14

[3+,2-]

D3,D7,D12,D13

[4+,0-]

D1,D2,D8,D9,D11

[2+,3-]

Yes

Vậy chọn Student làm root tiếp theo sau Young

No

Yes

No

Yes

Yes

Student

Yes

Age

Youth

Middle

Senior

Vậy chọn Credit rating là gốc tiếp theo sau Senior

Age

Youth

Middle

Senior

Student

Yes

Yes

No

Yes

No

Vậy Age=youth, Income=medium, Student=yes, Credit\_rating=fair là Yes

* **Thuật toán NB**

|  |  |  |
| --- | --- | --- |
| Age | Yes | No |
| Youth | 2/9 | 3/5 |
| Middle | 4/9 | 0/5 |
| Senior | 3/9 | 2/5 |

|  |  |  |
| --- | --- | --- |
| Income | Yes | No |
| High | 2/9 | 2/5 |
| Medium | 4/9 | 2/5 |
| Low | 3/9 | 1/5 |

|  |  |  |
| --- | --- | --- |
| Student | Yes | No |
| Yes | 6/9 | 1/5 |
| No | 3/9 | 4/5 |

|  |  |  |
| --- | --- | --- |
| Credit rating | Yes | No |
| Fair | 6/9 | 2/5 |
| Excellent | 3/9 | 3/5 |

Age=youth, Income=medium, Student=yes, Credit\_rating=fair:

Vậy Age=youth, Income=medium, Student=yes, Credit\_rating=fair là Yes

Question 2. Classification with Decision tree and Bayes.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Object | Body | Nationality | Marital status | Group |
| O1 | Small | Germany | Single | A |
| O2 | Big | France | Single | A |
| O3 | Big | Germany | Single | A |
| O4 | Small | Italy | Single | B |
| O5 | Big | Germany | Maried | B |
| O6 | Big | Italy | Single | B |
| O7 | Big | Italy | Maried | B |
| O8 | Small | Germany | Maried | B |

1. Find the classification rules of the above decision table with:

* The condition attribute set {Body, Nationality, Marital status}
* Classification attribute {Group}

1. If the following 2 objects are belong to any class? If so, which class are they belong to?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| O9 | Big | Italy | Single | ? |
| O10 | Small | Italy | Maried | ? |

**Giải**



Rule 1:

If Body = Small, Nationality = Germany, Marital status = Single, then Group = A.

Rule 2:

If Body = Big, Nationality = France, Marital status = Single, then Group = A.

Rule 3:

If Body = Big, Nationality = Germany, Marital status = Single, then Group = A.

Rule 4:

If Body = Small, Nationality = Italy, Marital status = Single, then Group = B.

Rule 5:

If Body = Big, Nationality = Germany, Marital status = Married, then Group = B.

Rule 6:

If Body = Big, Nationality = Italy, Marital status = Single, then Group = B.

Rule 7:

If Body = Big, Nationality = Italy, Marital status = Married, then Group = B.

Rule 8:

If Body = Small, Nationality = Germany, Marital status = Married, then Group = B.

b.

**Giải thuật :**

Áp dụng các Rules vào O9 (Big, Italy, Single):

Rule 6 phù hợp: O9 thuộc nhóm B.

Áp dụng các rules vào O10 (Small, Italy, Married):

Quy tắc 4,8 phù hợp: O10 thuộc nhóm B.

**Giải thuật NB**

|  |  |  |
| --- | --- | --- |
| Body | A | B |
| Small | 1/3 | 2/5 |
| Big | 2/3 | 3/5 |

|  |  |  |
| --- | --- | --- |
| Nationaly | A | B |
| Germany | 2/3 | 2/5 |
| France | 1/3 | 0 |
| Italy | 0 | 3/5 |

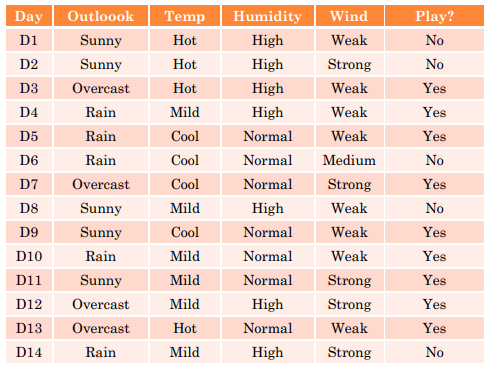
|  |  |  |
| --- | --- | --- |
| Marital status | A | B |
| Single | 3/3 | 2/5 |
| Maried | 0 | 3/5 |

O9 (Big Italy Single)

O10 (Small Italy Maried)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| O9 | Big | Italy | Single | B |
| O10 | Small | Italy | Maried | B |

**Question 3. Given data set**

****

If X= (O = Sunny, Temp = Hot, Humidity = Normal, Wind = Weak),

Which answer should be? Yes or No

**Giải**

* **Thuật toán Decision Tree**

Value(Outlook)= Sunny, Overcast, Rain

Value(Temp)=Hot,Mid,Cool

Value(Humidity)=High,Normal

Value (Windy)= True, False

Vậy chọn Outlook làm gốc

D3,D7,D12,D13

[4+,0-]

D1,D2,D8,D9,D11

[2+,3-]

Outlook

Sunny

Overcast

Rain

Yes

Vậy chọn Humidity là gốc tiếp theo sau Sunny

Yes

Outlook

Sunny

Overcast

Rain

Humidity

Normal

High

Yes

No

Vậy chọn Wind là gốc tiếp theo sau Rain

No

Yes

Strong

Weak

Wind

Outlook

Sunny

Overcast

Rain

Yes

Humidity

Normal

High

Yes

No

Nếu X= (O = Sunny, Temp = Hot, Humidity = Normal, Wind = Weak) = Yes

* **Thuật toán NB**

|  |  |  |
| --- | --- | --- |
| Outlook | Yes | No |
| Sunny | 2/9 | 3/5 |
| Overcast | 4/9 | 0/5 |
| Rain | 3/9 | 2/5 |

|  |  |  |
| --- | --- | --- |
| Temp | Yes | No |
| Hot | 2/9 | 2/5 |
| Mid | 4/9 | 2/5 |
| Cool | 3/9 | 1/5 |

|  |  |  |
| --- | --- | --- |
| Humidity | Yes | No |
| Normal | 6/9 | 1/5 |
| High | 3/9 | 4/5 |

|  |  |  |
| --- | --- | --- |
| Windy | Yes | No |
| Weak | 6/9 | 2/5 |
| Strong | 3/9 | 3/5 |

Nếu X= (O = Sunny, Temp = Hot, Humidity = Normal, Wind = Weak)

Nếu X= (O = Sunny, Temp = Hot, Humidity = Normal, Wind = Weak) = Yes