**Assignment 3 – Decision Trees**

**Question 1.** Consider the following data for purchasing automobiles:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Example | Purchase $ | Maintenance $ | # of Doors | Trunk Size | Safety | Acceptability |
| 1 | very high | very high | 2 | small | low | unacceptable |
| 2 | very high | medium | 2 | medium | high | unacceptable |
| 3 | very high | low | 2 | small | low | unacceptable |
| 4 | medium | high | 4 | small | low | unacceptable |
| 5 | low | medium | 4 | large | low | unacceptable |
| 6 | very high | medium | 2 | medium | high | acceptable |
| 7 | high | high | 2 | medium | high | acceptable |
| 8 | high | medium | 4 | medium | medium | acceptable |
| 9 | medium | very high | 4 | small | high | acceptable |
| 10 | low | low | 4 | small | medium | acceptable |

Where each attribute is described as:

Purchase $ = {very high, high, medium, low}

Maintenance $ = {very high, high, medium, low}

# of Doors = {2, 4}

Trunk Size = {large, medium, small}

Safety = {high, medium, low}

Acceptability = {acceptable, unacceptable}

Using **Acceptability** as the target attribute, construct a **decision tree** for the data. Use the ID3 algorithm as described in the notes.

**Question 2.** Consider the following census data:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Example | Max. Education | Marital Status | Race | Gender | Country of Birth | Income |
| 1 | High School | Divorced | White | Male | US | <=50K |
| 2 | High School | Married | Black | Male | US | <=50K |
| 3 | Undergraduate | Married | Black | Female | Other | <=50K |
| 4 | Graduate | Married | White | Female | US | <=50K |
| 5 | High School | Married | Black | Female | Other | <=50K |
| 6 | High School | Married | White | Male | US | >50K |
| 7 | Graduate | Never Married | White | Female | US | >50K |
| 8 | Undergraduate | Married | White | Male | US | >50K |
| 9 | Undergraduate | Married | Black | Male | US | >50K |
| 10 | Undergraduate | Married | Other | Male | Other | >50K |

Where each attribute is described as:

Max. Education = {High School, Undergraduate, Graduate}

Marital Status = {Never Married, Married, Divorced}

Race = {White, Black, Other}

Gender = {Male, Female}

Country of Birth = {US, Other}

Income = {<=50K, >50K}

Using **Income** as the target attribute, construct a **decision tree** for the data. Use the ID3 algorithm as described in the notes.