**Homework 1**

**Name: Megha Godwal**

**UID: U01695574**

**Problem 1.** (20 Points; 2.5 Points each) Discuss whether or not each of the following activities is a data mining task. If the answer is yes, then also specify which one of the following categories it will belong to:  
(i) classification   
(ii) association analysis  
(iii) clustering   
(iv) regression or   
(v) anomaly detection   
  
(a) Sorting a student database based on student identification numbers.  
(b) By looking at a CT scan, a doctor wants to identify if a patient has cancer or not. There are a lot of labeled CT scans that the doctor will use for making the decision.  
(c) An image analyst obtains some new images and wants to automatically detect the number of distinct objects in the image. He does not have any prior information about these objects.  
(d) Predicting the outcomes of tossing a (fair) pair of dice.  
(e) Predicting the future stock price of a company using historical records.  
(f) In an Internet search engine company, there is a need to find potential users who will click a particular advertisement on the webpage.  
(g) Monitoring the heart rate of a patient for abnormalities.  
(h) Extracting the frequencies of a sound wave.

**Answer 1-**

1. Sorting a student database based on student identification numbers.

**Ans:** No

1. By looking at a CT scan, a doctor wants to identify if a patient has cancer or not. There are a lot of labeled CT scans that the doctor will use for making the decision.

**Ans:** Yes, it belongs to clustering.

1. An image analyst obtains some new images and wants to automatically detect the number of distinct objects in the image. He does not have any prior information about these objects.

**Ans:** Yes, it belongs to classification.

1. Predicting the outcomes of tossing a (fair) pair of dice.

**Ans:** No

1. Predicting the future stock price of a company using historical records.

**Ans:** Yes, it belongs to regression.

1. In an Internet search engine company, there is a need to find potential users who will click a particular advertisement on the webpage.

**Ans:** Yes, it belongs to association analysis.

1. Monitoring the heart rate of a patient for abnormalities.

**Ans:** Yes, it belongs to anomaly detection.

1. Extracting the frequencies of a sound wave.

**Ans:** No.

**Problem 2.**(10 Points; 2 Points each) Classify the following attributes as binary, discrete, or continuous. Also, classify them as qualitative (nominal or ordinal) or quantitative (interval or ratio). Some cases may have more than one interpretation, so briefly indicate your reasoning if you think there may be some ambiguity.  
For example: Age in years. Answer: Discrete, quantitative, ratio  
(a) Time in terms of AM or PM.  
(b) Angles as measured in degrees between 0 and 360.  
(c) Bronze, Silver, and Gold medals as awarded at the Olympics.  
(d) ISBN numbers for books.  
(e) Military rank.

**Answer 2-**

1. Time in terms of AM or PM.

**Ans:** Binary, Qualitative, Ordinal

1. Angles as measured in degrees between 0 and 360.

**Ans:** Continuous, Quantitative, Ratio

1. Bronze, Silver, and Gold medals as awarded at the Olympics.

**Ans:** Discrete, Qualitative, Ordinal

1. ISBN numbers for books.

**Ans:** Discrete, Qualitative, Nominal

1. Military rank.

**Ans:** Discrete, Qualitative, Ordinal

f) In an Internet search engine company, there is a need to find potential users who will

click a particular advertisement on the webpage.

Yes.

It belongs to Association Analysis

g) Monitoring the heart rate of a patient for abnormalities.

Yes.

It belongs to Anomaly Detection.

h) Extracting the frequencies of a sound wave.

No.

Q.2) Classify the following attributes as binary, discrete, or continuous. Also, classify them

as qualitative (nominal or ordinal) or quantitative (interval or ratio). Some cases may

have more than one interpretation, so briefly indicate your reasoning if you think there

may be some ambiguity.

For example: Age in years. Answer: Discrete, quantitative, ratio

a) Time in terms of AM or PM.

b) Angles as measured in degrees between 0 and 360.

c) Bronze, Silver, and Gold medals as awarded at the Olympics.

d) ISBN numbers for books.

e) Military rank.

Ans.)

a) Time in terms of AM or PM.

Binary, qualitative, ordinal.

b) Angles as measured in degrees between 0 and 360.

Continuous, quantitative, ratio.

c) Bronze, Silver, and Gold medals as awarded at the Olympics.

Discrete, qualitative, ordinal.

d) ISBN numbers for books.

Discrete, qualitative, nominal.

e) Military rank.

Discrete, qualitative,