Name: Nazmul Islam

ID: 232113

Ans. to the Q. no. 1

a) Yes. This activity is a data mining task. It falls under the category of clustering.

b) Yes. This activity is a data mining task. It falls under the category of classification

c) Yes. This activity is a data mining task. It falls under prediction task.

d) No. This activity is not a data mining task. Monitoring the heart rate of a patient for abnormalities is a real-time prediction or detection task. Usually data mining involves with large datasets to find out pattern. So, This activity is not a data mining task.

Ans: to the Q.no. 2

a) Loan default prediction:

Goal: The goal of loan prediction is to identify potential customers who are likely to default on their loan payment.

Approach:

1) Analyze the historical loan data
2) Apply various classification algorithm on loan data
3) ~~Build~~ Build a model based on machine leonning.
4) Consider the attributes of data: credit history, income, employment status, loan amount.
5) Preprocess the data with feature selection
6) Train the model with qualityful ~~data~~ and balanced data.

b) Credit card fraud detection:

Goal: The goal of credit card fraud detection is to identify and prevent fraudlent transaction and protect both card holders and financial institutions.

Another Q. no. 3

The Three data quality problem can be:

1) Noise
2) Duplicate data
3) Missing values

Noise: For objects noise is an extraneous object. Noise refers to modification of original values. It can misleading the data mining task. Train with noise decrese the accuracy of the model. It also the noise the problem of overfitting, complexity, in accurate decision making.

Duplicate data: A data set include data objects that are duplicates or almost duplicates, of one another. To detect and eliminate such duplicates, two main issues are addressed. Deduplication can help with this process.

Missing values: Missing values also decrease the model performance. To handle this we can eliminate data object or estimate the missing values.

Ans. to the Q.no. 4

Temperature can be considered as either an interval or ratio, depending on the scale being used.

In Celcius (C°) or Fahrenheit (°F) temperature is measured on an interval scale.

On the other hand, in kelvin (K) scale, temperature is measured on a ratio scale.

The distiction between interval and ratio scales lies in the presence or absence of a true zero point. In an interval scale, ratios between values are not meaningful since zero does not indicate the absence of the measured attributes. In a ratio scale, ratios have meaningful interpretation because zero represent the absence of the attributes.

Thats why, in kelvin scale, temperature can be used ratio on the other hand, in celcious or fahrenheit scales, temperature can be used as an interval.