CI603 Data Mining

Tutorial 1

1. Discuss whether or not each of the following activities is a data mining task.

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

   b. Dividing the customers of a company according to their gender.

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

   d. Computing the total sales of a company.

   e. Monitoring seismic waves for earthquake activities.

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

   g. Extracting the frequencies of a sound wave.

   h. Identifying the key attributes of clients attracted by a mortgage offer.

   i. Increasing cross-selling to existing customers.

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

   k. Detecting fraudulent activity in telecommunication services.

   l. Classifying the current customers of a company according to their profitability.

2. Which statement is **NOT TRUE** regarding a data mining task?

   a. Clustering is a descriptive data mining task

   b. Classification is a predictive data mining task

   c. Regression is a descriptive data mining task

   d. Deviation detection is a predictive data mining task

3. Which data mining task can be used for predicting wind velocities as a function of temperature, humidity, air pressure, etc.?

   a. Cluster Analysis

   b. Regression

   c. Classification

   d. Sequential pattern discovery

4. Which of the following is not a data mining functionality?
    a. Characterisation and Discrimination

    b. Classification and regression

    c. Selection and interpretation

    d. Clustering and Analysis

5. Which data mining task is most suitable for a bank credit risk assessment for the granting of a mortgage?
    a. Neural network classification.

    b. Logistic regression.

    c. Clustering

    d. All of the above

    e. None of the above.

6. Which data mining task is most suitable for predicting the effect that current covid vaccines will have thirty years from now in the population under 12 who received the two doses in the UK.
    a. Logistic regression.

    b. Neural network classification.

    c. Clustering

    d. All of the above

    e. None of the above.

7. Which data mining task is most suitable for designing a supermarket layout to maximise sales?
    a. Classification.

    b. Clustering

    c. Affinity grouping.

    d. All of the above

    e. None of the above.