Quiz No 2

Data Science Fall 2018 (Sep 18, 2018)

Name:_____	Roll#:_____

Q 1) Identify whether following tasks are data mining tasks or not? Explain each with valid reason. 5

(i) A company wants to sort his employees on the basis of employee salary to know highest salaried employees.

No, it is a simple database query.

(ii) Government of KPK wants to find common symptoms for dengue fever on the basis of historical data of patients.

Yes, You have to apply scientific method on data to discover patterns.

(iii) Monitoring the pulse rate of a patient.

No, this task can be completed by using a simple data storing and retrieval method.

(iv) Predicting future stock price of a company on the basis of historical data.

Yes, You have to apply scientific method on data to discover patterns.

(v) Computing revenue of a particular product.

No, You can calculate revenue by using simple queries of sql/other database query languages.

Q 2) what is difference between predictive data mining task and descriptive data mining task? Write 1 example of each 4

Predictive Data Mining tasks	Descriptive Data Mining tasks
Data mining task in which we use some variables to predict future values of other variables (target variables).	Data mining task in which we find patterns that summarize relationships in data.

Classification, Regression	Clustering

Q 3) Write down appropriate data mining task name against each data mining technique . 6

Data Mining Technique	Supervised/Unsupervised	Predictive/Descriptive
Anomaly Detection	Supervised	Predictive
Regression	Supervised	Predictive
Sequential pattern discovery	Unsupervised	Descriptive
Association rule discovery	Unsupervised	Descriptive
Classification	Supervised	Predictive
Clustering	Unsupervised	Descriptive

Q 4) Write down appropriate step name of CRISP-DM model against each of the following activities of a process model. 6

Process Model Activity	CRISP-DM Step
Write final report and presentation	Deployment
Build and verify models	Modeling
Data selection	Data Preparation
Verification of data quality	Data Understanding
Assessment of results	Evaluation
Set business success criteria	Business Understanding

Q 5) Write appropriate type against each of the following data sets 6

Data Set	Type of data set(Record/Graph/Ordered)
Google pages	Graph
Data matrix/Pattern matrix	Record
Spatial temperature data	Ordered
Structure of Ammonia(NH_3) molecule	Graph
Genomic sequence data	Ordered
Document-term matrix	Record

Q 6) Write appropriate types for each of the following attributes. 9

Attribute Name	Type of attribute(qualitative/quantitative)	Type of attribute(Nominal/Ordinal, Interval/Ratio)	Type of attribute(Discrete /Continuous)
Number of students	Quantitative	Ratio	Discrete
Blood pressure	Quantitative	Interval	Continuous
Temperature	Quantitative	Interval(C), Ratio(K)	Continuous
Occupation	Qualitative	Nominal	Discrete
Military rank	Qualitative	Ordinal	Discrete
Weight	Quantitative	Ratio	Continuous

Q7) What do you think about following techniques, where they will be used in practical life? 6

Data Mining Technique	Applications
Anomaly Detection	Fraud Detection, Network Intrusion
Regression	Stock market price prediction, Sales amount prediction, Income prediction on the basis of qualification
Sequential pattern discovery	Point of sales systems
Association rule discovery	Inventory management systems, Marketing and sales promotions
Classification	Customer churn prediction, Fraud detection
Clustering	Market segmentation, Document Clustering