TYCS-Semester VI

DATA SCIENCE

1) _____ type of data is measurable like width, length etc
a) numerical   b) categorical   c)nominal   d) ordinal

2) _____ type of data have only two values.
a) ordinal   b) nominal   c) binary   d) continuous

3) Facebook, twitter are _____type of data.
a) structured   b) unstructured c) semistructue   d) none

4) _____type of data are processed, stored and available in specific format.
a) structured   b) unstructured c) semistructue   d) none

5) Identify correct option for assembly language
a) contains strings of 0's and 1's
b) contain English language and mathematical symbol
c) contain alphanumeric code
d) all

5) List processing and program in logic LISP, Prolog are used for
a) window based application   b) solving mathematical and statistical problem
c) handling files   d) string manipulation, search pattern

6) Pick out incorrect option for IDE
a) It is open source and commercial   b) It contain editor, debugger with GUI interface
c) Approach to analyse the data   d) all

7) Bar chart, Histogram, Box plot, Scatter plot are _____technique
a) data analysis   b) data visualization   c) data integration   d) data collection

8) Scatter plot is
a) 2D data visualization method   b) representation of data in circle with slices
c) distribution of data over continuous interval   d) data representation with bars

9) _____plot indicates values like minimum, first quartile, median, third quartile and maximum.
a) Histogram   b) barchart   c) scatter plot   d) boxplot

10) CSV files, Oracle, Cloud system are the examples of_____
a) data collection   b) data visualization   c) data sources   d) data

11) _____is a statistical method which shows peak of frequency distribution
a) skewness   b) kurtosis   c) box plot   d) outliers

12) _____ is a process of transforming data from row form to more usable format.
a) data extraction   b) data cleaning   c) data wangling   d) data management

13) Observation method comes under_____type of data collection method.
a) primary   b) secondary   c) ternary d) all

14) Secondary data collection method indicates
a) data collection from questionnaire   b) collection of data with interview
c) data collected from laboratory   d) data collection from historical and statistical documents

15) Data cleaning is a process of
a) removing unwanted observations   b) by handling missing data   c) by fixing errors   d) all

16) Which is incorrect extraction method?
a) full extraction   b) internal extraction c) physical extraction d) incremental extraction

17) _____is data analysis method.
a) data collection   b) data mining   c) data curation   d) data wrangling

18) Select incorrect data model
a) Relational   b) ER   c) Use case d) hierarchical

19) Data curation is _____
a) extraction of data from various sources   b) is query language   c) collection of data from various sources and integration into various repositories   d) is a cloud service

20) Dplyr, Tider are _____
a) data wrangling method   b) query language c) clustering method   d) regression method

21) An example of semi-structured data is _____
 a) XML   b) SQL   c) Spreadsheet   d) CSV files

22) Nosql ,XML,JSON are examples of_____ system
a) Structured   b) semi-structured c) Unstructured   d) None

23) XML is ___(Select incorrect statement)

a) tool for storing and transporting data   b) markup language   c) data stored in text format
d) unstructured data system

24)_____is used for selecting child node or parent node in XML document.
a) Xpath   b) Xquery   c) JSON   d) both A & b

25)Information collected from digital survillence, satellite images are in _____
format
a) structured   b) semi-structured   c) unstructured   d) all

26) Web Crawling is _____
a) tool for storing and transporting data
b) regression method
c) browses WWW in systematic and automatic manner
d) process of handling missing data

27) Which is incorrect Web crawler
a) focused   b) incremental   c) distributed   d) None

28) _____is crawler which can run on network of workstation
a) focused   b) incremental   c)parallel   d) distributed

29) _____term indicate getting data from web.
a) curation   b) scrapping   c) crawling   d) modelling

30) Data for research, E-commerce website , Marketing and sales campaigns are examples of
a) web scraping   b) web crawling   c) data extraction   d) distributed system

31) _____is giving permission to do something
a) Authentication   b) authorization   c) both a & b   d) none

32) GitHub is_____
a) code hosting and collaborative platform   b) data managing tool
c) data testing tool   d) web application development tool

33) _____system mostly uses distributed system.
a) small scale system   b) medium scale system   c) large scale system   d) all

34) Nosql database is_____
a) non-relational database   b) do not define fixed data structure   c) both a & b
d) heterogeneous distributed system

35) _____is cross platform, document related platform, works on collection and documentation.
a) HBase   b) NoSQL   c) MongoDB   d) JSON

36) In MongoDB ,document is present in _____format.
 a) doc   b) JSON   c) CSV   d) doc

37) HBase is _____
a) distributed column oriented database built on Hadoop platform
b) cloud platform
c) cross platform database system
d) database which is not based on relational database

38) _____is cloud platform service where service providers provides services like databases, queues, workflow engines to their customers.
a) Iaas   b) Paas   c) Saas   d) Mass

39) _____is a is machine learning method and resampling procedure used to evaluate model for limited data.
a) regression   b) sampling   c) cross validation   d) lasso regression

40) Process of converting data from one format to other format is known as_____
a) data transformation   b) data migration   c) data segregation   d) data management

41) ETL tool is stand for
a) Extract, Transform, Leak   b) Export, Transport, Load
c) Extract, Transport, Load   d) extract, Transform ,Load

42) Feature Extraction is based on _____
a) Regression   b) clustering   c) Decision trees   d) PCA

43) Binning method, clustering, combined computer and human inspection are the techniques of_____
a) smoothing   b) dimension reduction   c) supervised learning   d) none

44) _____system involves building machine model based on labelled samples.
a) PCA   b) unsupervised learning   c) supervised learning   d) Hypothesis

45) In Regression model indicates
a) calculate prediction of future
b) relationship between dependent and independent variables
c) shows relationship between classes
d) none

46) Identify Incorrect Regression model
a) simple   b) multiple   c) Lasso   d) binary

47) _____is consists of nodes and links ,nodes further divided into decision node, leaf node.
a) decision tree   b) regression tree   c) both a & b   d) classification

48) Trend, Seasonality, Cycles are the components of_____
a) time series analysis   b) decision trees   c) logistic regression   d)clustering

49) _____learning technique there is no trainer to teach model.
a) Classification   b) supervised   c) unsupervised   d)   all

50) Identify which wrong statement for k-means is clustering
a) an exploratory analysis system
b) implements non-hierarchical method of grouping of objects together
c) determines centroid using Euclidian distance
d) it has combined solution or decision