Assignment - 3

Q.1. Define Cluster Analysis outlier Analysis :-

1] Cluster Analysis :

Cluster analysis is a data analytics technique is use to group similar data point into Cluster So that item is set the same Cluster are more Similar.

2] Oulier Analysis.

The process of identifying data points ar that significantly diffe From the rest of the datuset. often indicuting anomalies, error or rare event.

Q.2. write the Formulas For support and confidence.

1] support.

$$Support (x \rightarrow y) = \frac{Number\ of\ transaction\ which\ contain\ the\ item\ set\ x\ and\ y}{Total\ numbe\ of\ transaction.}$$

whereas,

$$Support(x) = \frac{Number\ of\ transaction\ which\ Contain\ the\ item\ Set\ \&\ x}{Total\ number\ of\ transaction.}$$

2] Confidence :-

The Confidence of rule $x \to y$ is defined as :

$$Confidence\ (x \to y) = \frac{Suppor\ (x \to y)}{Support(x)}$$

Q.3. Explain ~~FRE~~ Frequent itemset, Closed itemset and Association Rule.

1] Frequent Item Set.

A Frequent itemset typically refer to a set of items that frequently appear together in a transactional data set.

maximal frequent Itemset (~~MFE~~ MFI) is a Frequent item set For which none of its imediate Supersets are Frequent

2] Closed Item SET

The closed Frequent Item SET (CFI) is a Frequent Item set For which is non of its immediate Superset has same support as that of the intemset.

3] Association Rule

An 6 association Rule is that implies a strong relationship beth items in a transactional dataset

It is generaly written as
A => B.

Assignment - 4

Q.1. list the layer of Social media Aanalytics.

=> layer of Social media analytics

1] Text
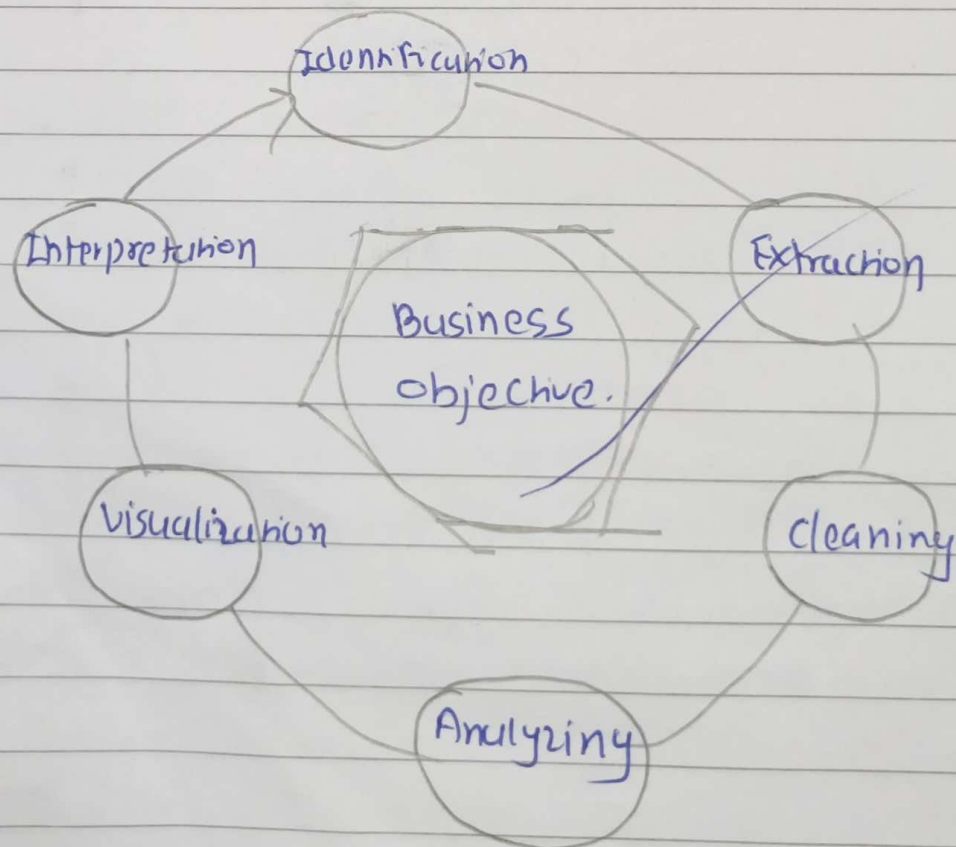2] Networks
3] Actions
4] Hyperlinks
5] mobile
6] Location
7] Search Engine.

Q.2. Draw the diagram For Social mediat analytics life cycle.

=>

Q.3. Define a link prediction and tokenization

1] linkprediction.

A data mining and mach learning task that aims to predi the likehood of a Future or missin connection (ink) betn two nodes in a network graph, base on the current structure and proporti of network.

2] tokenization.

The First part of any NLP process is simpley breaking a pie of text into its constituent parts (usually words) this process is called tokenization.