5. Content based ← Collaborative based

Requires only user behaviour data

Item features are used for group similar items

Requires both user & item data

Item features are not required

2 marks : U-5:

1. ANN :

Artificial Neural Network is modelling technique inspired by human nervous system that allows learning by example ~~by~~ from representative data that describes decision process.

2. Zero padding :

Most common padding value is Zero Padding which involves adding Zeros to the borders of input feature map.

3. Strides :

The no. of pixels forming to the input matrix is called strides.

4. CNN Archi :

* i/p layer * convolutional layer
* Max pooling layer * dense layer
* o/p layer

2 marks : U-4 :

1. Application of Clustering :
   * Market segmentation * image segmentation
   * Social network analysis
   * Search result grouping
   * Medical Imaging

2. Merge clusters methods :
   * Partitioning method
   * Hierarchical Method
   * Density based method

3. * Soft clustering
   * One datapoint belong to one cluster only.

   * Hard clustering
   * If provided datapoint belonging to each of pre-defined number of clusters

4. Parameters in DBSCAN :
   * ε (eps)
   * (minPts) minimum no. of points required to form dense region

ML ; CCET-2;

2 marks ; U-3;

1. How does the decision trees make decision:

They clearly layout the problem so that all options can be challenged.

2. Bagging or Ensemble of Decision trees:

Decision tree ensembles are commonly used in wide range of applications. The Different trees in ensemble can processed in parallel during tree inference.

3. When do you stop splitting decision:

Stopping rules determine when to stop Spliting specific branches of tree. Minimum record in parent branch prevents a split if the no. of records in node to be split is less that specified value.

4. Leaf nodes:

Leaf nodes are nodes of tree that have no additional nodes coming off them.

2 marks

1) How decision trees are making decision

A decision tree starts with a point (node) and starts splitting in two or more directions. Each branch offers different possible outcomes until final outcome is achieved

2) Intro of Ensemble of decision tree

Decision tree ensembles are commonly used in a wide range of applications becoming algorithms for decision tree based classifiers. Different trees in an ensemble can be processed in parallel during tree inference.

3) When do yo stop splitting decision tree

Decision tree stops splitting when the number of records in the node to be split is less than the specified value

4) Interpret on kernel trick

kernel trick allows us to project data from a training set which isn't linearly seperable into higher dimensional space where it becomes linearly seperable

5) Entropy and Information gain

↳ Entropy measures impurity in data

↳ Information gain measures reduction in impurity in data

6) Why support vectors important

The use of support vector ensures that only a subset of data points determine the decision boundary

1) Leaf node

Leaf node are nodes of the tree that have no additional nodes coming off them.

2) Applications of Clustering techniques

* Medical imaging
* Image segmentation
* Anomaly detection
* market segmentation
* Google photos

3) Demerits of k-means algorithm

* Time complexity
* Sensitivity
* Cannot handle noisy data & outliers
* Defining number of cluster

4) Methods to merge cluster in Agglomerative clustering

* Single link
* Complete link
* Average link

5) Parameters required to do clustering in DBSCAN algorithm

* Epsilon
* Midpoint

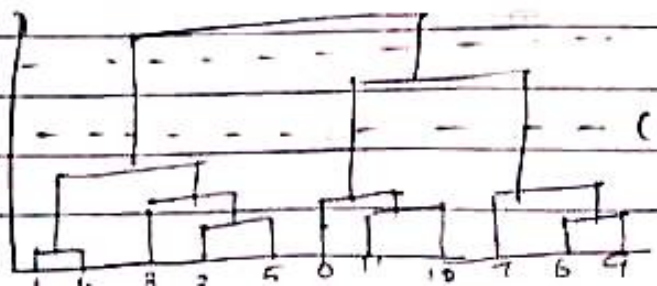
6) Collaborative filtering

* Uses past interactions to recommend new items
* Item features are not required

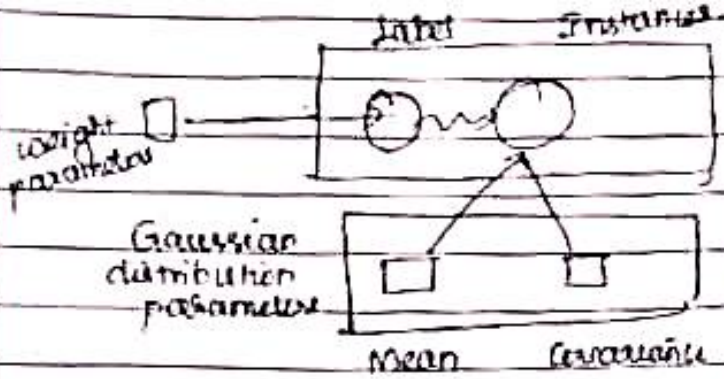
Content based filtering

* Uses ML algorithms to predict & recommend new items
* Are required

7) Dendrogram for bottom-up approach



21) Architecture of Gaussian mixture model



22) Hard clustering
→ Every node belong to only one cluster
→

→ Eg. k means

Soft clustering
→ Every node may belong to several clusters
→

→ Eg: HMMs.

23) ANN.

14) List the parameters used in DBSCAN algorithm

Ans:-

15) Different scales of K-means clustering algorithm
- Standard Scale
- Robust scale
- Min max Scale
- Normalizer

16) Challenges in unsupervised learning
- Lack of labelled data
- Determining no. of clusters
- Scaling and normalization
- Noise & outliers

17) Steps involved in K-means clustering algorithm
- Assigning each data point to the closest cluster centroid
- Setting each cluster centroid to the mean of data points assigned to it

18) How choice of number of clusters affect K-mean result
- K-means is non-deterministic
- The choice of initial cluster have impact on final cluster formation

19) How DBSCAN classify datapoints into different groups
- eps: If the distance between two points are less than equal to eps, they are considered as neighbours
- minpoints: Minimum number of points to form a dense layer

20) Idea behind Gaussian Mixture model

It assumes that all the data points are generated from the mixture of finite no. of gaussian distributions with unknown parameters