A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

ML.

Name: Piyusha Rajendra Supe.
Roll No: 23C0315. [BE. Computer B].

Machine Learning. Assignment.

Attempt Q1 any three UNIT - 05.

Q1] With reference to clustering, explain the issue of "Optimization of clusters"?

→ ◦ Clustering aims to group data points such that objects within the same cluster are highly similar, while those in different clusters are dissimilar.

• The optimization of clusters refers to the process of finding the best possible division of data into clusters that maximize intracluster similarity and minimize inter cluster similarity.

◦ It involves determining the optimal cluster centroids or representatives that minimize a cost or objective function (eg. minimizing the sum of squared distances in K means).

◦ This optimization problem is challenging because it is NP Hard, meaning the number of possible cluster assignments grows exponentially with the dataset size.

A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

- Various algorithms (like K means, K-medoids, DBSCAN, and Heirarchical clustering use heuristics or iterative refinement to approximate an optimal clustering solution.
- The process also includes evaluating the quality of clusters using metrics such as Silhoutte Score, davies-Bouldin index, or Dunn Index. to decide the best structure.
- A major issue is that the optimization result often depends on initialization, distance metric, and outliers, leading to potentially suboptimal or unstable clusters.

Q2) Explain how a cluster is formed in the density based clustering algorithm.

→
- In density based clustering (like DBSCAN) clusters are formed based on regions of high data density separated by regions of low density.
- The algorithm defines two parameters ε epsilon - the maximum distance between two points to be considered neighbours and MinPts, the minimum number of points required to form a dense region.
- It begins by selecting an arbitrary point if it has at least MinPts neighbours within ε it is classified as a core point.
- All points within the ε-neighbourhood of a core point are included in same cluster.

- The process expands by recursively including density reachable points i.e. points that can be reached through a chain of neighbouring core points.
- Points not belonging to any dense region (not density reachable from any core point) are labelled as noise or outliers.
- As a result, clusters form naturally as dense regions and the algorithm can discover arbitrary shaped clusters and handle noise effectively.

Q3] How would you choose the number of clusters when designing a K-medoid clustering algorithm?

→
- The K-medoid algorithm requires the number of clusters (k) as an input, so selecting an appropriate K is crucial for meaningful clustering.
- One common method is the Elbow method, where the total cost (sum of dissimilarities between points and their medoids) is plotted against different K values, the elbow point indicates an optimal trade off between cost reduction and cluster compactness.
- The silhoutte coefficient can also be used, measuring how similar a point is to its own cluster compared to other clusters, the K with highest avg silhoutte score is ideal.

A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

- The gap statistic contains the total within cluster variation for different K values with that expected under a null reference distribution of the data choosing the K with largest gap.
- Domain knowledge or businesses context can also guide choice of K, especially when a certain number of meaningful categories is expected.
- A trial and error approach may be applied, testing different K values and evaluating stability and interpretability of the results.

UNIT - 06)
Q4]. What are the building blocks of a neural network? Elaborate?

→ • The basic building block of a neural network is the artificial neuron, inspired by biological neurons.
• A neural network is composed of the following essential components.

- Input layer: Accepts input features or data into the network.
- Weights (w): Numerical parameters that control the influence of input signals; they are learnt during training.
- Summation function - Computes the weighted sum of inputs.

A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

- Activation function: Introduces non linearity to the network (eg. sigmoid, ReLU, Tanh allowing it to learn complex relationships)
- Hidden Layers - Intermediate Layers between input and output that capture patterns and heirarchical representations.
- Output layer - Produces final results (eg. classification labels, regression values).
- Bias (b) - A constant added to the weighted sum to improve model flexibility.
- Learning Algorithm - Adjusts weights based on error feedback (eg. Backpropogation with gradient descent).
- These interconnected neurons collectively form feedforward or recurrent architectures capable of approximating non linear mappings between inputs and outputs.

Q5] Describe characteristics of back propagation algorithm.

→ - The back propagation algorithm is a supervised learning method used for training multilayer neural networks.
- It is based on the gradient descent optimization principle. weights are adjusted in proportion to the negative gradient of the error function.
- It operates in two main phases -

(1) Forward pass - Inputs are propagated through the network to generate an output.

A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

- Backward pass - The output error is propogated backward computing partial derivatives of the loss function with respect to each weight.
- The algorithm minimizes the mean squared error or other loss functions by iteratively updating weights.
- It requires differentiable activation functions such as sigmoid, Tanh, or ReLU.
- Key characteristics include supervised learning, error correction capability, layer wise weight adjustment and convergence through iterative optimization.
- While effective it may suffer from issues like slow convergence, local minima and vanishing or exploding gradients. particularly in deep networks.

Q6] Write a short note on Reurrent neural network and convolutional neural network. RNN and CNN.

→ 1. Reurrent Neural Network (RNN) -

- RNNs are specialized neural networks designed for sequential or time series data, where current output depends on previous computations.
- They include feedback loops that allow information to persist over time, making them suitable for tasks like speech recognition, language modeling and text generation.

- The main element is the hidden state, which carries forward learned information across time steps.
- Variants like LSTM (Long short Term memory) and GRU (Gated recurrent unit) address problems like vanishing gradients by introducing gates for better control of information flow.

2] Convolutional Neural Network -

- CNNs are primarily used for image and spatial data analysis, leveraging local connectivity and weight sharing.
- The main components include convolutional layers, pooling layers and fully connected layers.
- The convolutional layers apply filters (kernels) to extract spatial features like edges, textures and shapes.
- Pooling layers reduce dimensionality and help acheive translation invariance.
- CNNs excel in computer vision tasks like image classification, object detection and facial recognition
- Their architecture efficiently captures heirarchical patterns from low level features (edges) to high level abstractions (objects).