**ML ASSESMENT**

**PW SKILLS**

Q1) Explain the following with an example:

a) Artificial Intelligence

b) Machine Learning

c) Deep Learning

- Artificial Intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think and act like humans. For example, AI can be used in virtual assistants like Siri or Alexa, which can understand voice commands and perform tasks.

- Machine Learning (ML) is a subset of AI that enables machines to learn from data without being explicitly programmed. An example of machine learning is email spam detection, where algorithms learn to classify emails as spam or not based on past examples.

- Deep Learning (DL) is a subset of machine learning that uses artificial neural networks with multiple layers to learn from large amounts of data. An example of deep learning is image recognition, where deep neural networks can identify objects in images with high accuracy.

Q2) What is supervised learning? List some examples of supervised learning.

- Supervised learning is a type of machine learning where the model is trained on a labeled dataset, meaning each input data point is paired with the correct output. The model learns to make predictions or decisions based on this labeled data. Examples include:

- Spam email detection

- Handwritten digit recognition

- Predicting house prices based on features like location, size, etc.

Q3) What is unsupervised learning? List some examples of unsupervised learning.

- Unsupervised learning is a type of machine learning where the model is trained on unlabeled data, and the algorithm learns to find patterns or structure in the input data. Examples include:

- Clustering similar documents together

- Customer segmentation based on purchasing behavior

- Anomaly detection in network traffic

Q4) What is the difference between AI, ML, DL, and DS?

- AI (Artificial Intelligence) is the broader concept of machines being able to carry out tasks in a way that we would consider "smart."

- ML (Machine Learning) is a subset of AI that involves the development of algorithms that can learn from and make predictions or decisions based on data.

- DL (Deep Learning) is a subset of ML that uses neural networks with many layers to learn representations of data.

- DS (Data Science) is a multidisciplinary field that uses scientific methods, algorithms, processes, and systems to extract knowledge and insights from structured and unstructured data.

Q5) What are the main differences between supervised, unsupervised, and semi-supervised learning?

- Supervised learning requires labeled data for training, unsupervised learning uses unlabeled data, and semi-supervised learning uses a combination of labeled and unlabeled data.

Q6) What is train, test, and validation split? Explain the importance of each term.

- Train, test, and validation split is a common practice in machine learning where the dataset is divided into three subsets:

- Training set: Used to train the model.

- Validation set: Used to tune hyperparameters and avoid overfitting.

- Test set: Used to evaluate the final performance of the model on unseen data.

The importance of each term:

- Training set helps the model learn patterns from data.

- Validation set helps in tuning model hyperparameters and assessing model performance during training.

- Test set provides an unbiased evaluation of the model's performance on unseen data.

Q7) How can unsupervised learning be used in anomaly detection?

- Unsupervised learning can be used in anomaly detection by identifying patterns or outliers in data that deviate from normal behavior. For example, clustering algorithms like K-means can group similar data points together, and data points that are significantly different from the clusters may be flagged as anomalies.

Q8) List down some commonly used supervised learning algorithms and unsupervised learning algorithms.

Commonly used supervised learning algorithms:

- Linear Regression

- Logistic Regression

- Decision Trees

- Random Forest

- Support Vector Machines (SVM)

- Naive Bayes

- k-Nearest Neighbors (k-NN)

Commonly used unsupervised learning algorithms:

- K-means Clustering

- Hierarchical Clustering

- DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

- Principal Component Analysis (PCA)

- t-Distributed Stochastic Neighbor Embedding (t-SNE)

- Autoencoders

These algorithms are used for various tasks in machine learning, depending on the nature of the data and the problem at hand.