Machine Learning Assignment - 1

February 5, 2024

0.1 Q1- Explain the following with an example?

1) Artificial Intelligence : Artificial Intelligence (AI) refers to the ability of a machine or computer program to perform tasks that typically require human intelligence, such as learning, problem-solving, decision-making, perception, and natural language understanding. AI can be implemented in various fields such as gaming, robotics, healthcare, and finance. Example: An AI-powered virtual assistant, such as Apple's Siri or Amazon's Alexa, can understand voice commands and respond to questions, play music, set reminders, and control smart home devices, among other things.

2) Machine Learning: Machine Learning (ML) is a subset of AI that involves training algorithms on a large dataset to learn from patterns and make predictions without being explicitly programmed. ML algorithms can be supervised, unsupervised, or semi-supervised, and they can be used for tasks such as image and speech recognition, natural language processing, and recommendation systems. Example: A spam filter that learns to distinguish between spam and legitimate emails based on patterns in the content, sender, and subject line.

3) Deep Learning: Deep Learning (DL) is a subset of ML that uses artificial neural networks with multiple layers to learn from data and make predictions. DL algorithms can be used for tasks such as image and speech recognition, natural language processing, and autonomous driving. Example: A DL algorithm that can recognize images of different breeds of dogs by learning from millions of labeled images of dogs with different features, such as fur color, snout length, and ear shape.

0.2 Q2- What is supervised learning? List some examples of supervised learning.

Ans: Supervised learning is a type of machine learning where an algorithm is trained on a labeled dataset, meaning the input data has corresponding output data. The goal is to use this training to make accurate predictions or classifications on new, unlabeled data. Examples include image recognition, spam filtering, and language translation.

0.3 Q3- What is unsupervised learning? List some examples of unsupervised learning.

Ans: Unsupervised learning is a type of machine learning where the model learns to identify patterns and relationships in data without being given any specific guidance or labels. Examples of unsupervised learning include clustering similar data points, dimensionality reduction, and anomaly detection.

0.4 Q4- What is the diffference between AI, ML, DL, and DS?

Ans: * AI (Artificial Intelligence) is a broad field that encompasses various techniques and approaches to make machines perform tasks that typically require human intelligence, such as perception, reasoning, and learning. * ML (Machine Learning) is a subset of AI that involves training algorithms to make predictions or decisions based on data. The focus of ML is on developing algorithms that can learn and improve from experience without being explicitly programmed. * DL (Deep Learning) is a subset of ML that uses artificial neural networks to learn from large amounts of data. DL algorithms are designed to automatically extract complex features from raw data and perform tasks such as image recognition, natural language processing, and speech recognition. * DS (Data Science) is a field that involves extracting insights and knowledge from data using various techniques such as statistics, machine learning, and data visualization. Data scientists use statistical and computational methods to extract meaningful insights and patterns from data and use them to make informed business decisions.

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

Ans: * Supervised learning: In supervised learning, the machine learning algorithm is trained on a labeled dataset, where each example in the dataset is associated with a specific label or output. The algorithm learns to predict the output for new, unseen examples based on the input features provided. Examples of supervised learning include image classification, sentiment analysis, and speech recognition. * Unsupervised learning: In unsupervised learning, the machine learning algorithm is trained on an unlabeled dataset, where the data has no specific labels or outputs. The algorithm learns to identify patterns and relationships in the data without any specific guidance. Examples of unsupervised learning include clustering, dimensionality reduction, and anomaly detection. * Semi-supervised learning: In semi-supervised learning, the machine learning algorithm is trained on a combination of labeled and unlabeled data. The algorithm learns to use the labeled examples to make predictions and the unlabeled examples to identify patterns and relationships in the data. Semi-supervised learning is useful when labeled data is scarce or expensive to obtain. Examples of semi-supervised learning include image and speech recognition. In summary, supervised learning requires labeled data, unsupervised learning does not require labels, and semi-supervised learning combines both labeled and unlabeled data.

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

Ans: Train, test, and validation split is a technique used in machine learning to divide a dataset into three subsets: a training set, a validation set, and a testing set. The training set is used to train the model, the validation set is used to evaluate the performance of the model during training, and the testing set is used to evaluate the final performance of the model after training. This split is important because it helps to prevent overfitting, where the model memorizes the training data and performs poorly on new data.

0.7 Q7- How can unsupervised learning be used in anormaly detection?

Ans: Unsupervised learning can be used in anomaly detection by training a model to learn the normal patterns in a dataset and then flagging any data that does not fit those patterns as an anomaly. The model does not need to be explicitly told what an anomaly is, but instead, it

identifies patterns that deviate from the norm on its own. This makes unsupervised learning useful for detecting novel or unknown anomalies in data.

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

Ans: * Supervised Learning Algorithms: - Linear Regression - Logistic Regression - Decision Tree - Random Forest - Support Vector Machines (SVM) - K-Nearest Neighbors (KNN) - Naive Bayes

- Unsupervised Learning Algorithms:

  - K-Means Clustering
  - Hierarchical Clustering - PCA (Principal Component Analysis)
  - Apriori Algorithm
  - Anomaly Detection
  - Isolation Forest
  - Autoencoders.