15 March Assignment

March 19, 2023

1 Q1 - Explain the following with an example

1) Artifical Intelligence
2) Machine Learning
3) Deep Learning

2 Answer :- 1) Artificial Intelligence

Artificial intelligence refers to the simulation of human intelligence in machines that are programmed to perform tasks that would normally require human intelligence, such as visual perception, speech recognition, decision-making, and language translation

AI has many applications in various fields, including healthcare, finance, transportation, and entertainment. Some examples of AI-powered technologies include virtual assistants, self-driving cars, and facial recognition systems.

Some AI are designed in such a way that they can make decisions on their own without any human interaction. Such AI systems can be both beneficial and dangerous to us, but this depends on a number of factors

3 2) Machine Learning

Machine learning is a subfield of artificial intelligence (AI) that involves the development of algorithms and statistical models that allow computer systems to learn from data and improve their performance on a specific task over time, without being explicitly programmed.

In machine learning, the computer system is given a dataset and is trained to recognize patterns and make predictions based on that data. As the system processes more data, it adjusts its algorithms and models to improve its accuracy and performance.

There are mainly 3 types of Machine Learning 1) Supervised Machine Learning 2) Unsupervised Machine Learning 3) Reinforcement Learning

4 3) Deep Learning

Deep learning is a subfield of machine learning that uses artificial neural networks, which are modeled after the structure and function of the human brain, to learn from large amounts of data. These neural networks consist of layers of interconnected nodes or "neurons" that process and

analyze data, and the connections between these neurons are weighted to determine their impact on the final output.

Some of famous Deep Learning Architectures include :- 1) Convolutional Neural Networks (CNNs) 2) Recurrent Neural Networks (RNNs)

and there are many more

5 Q2) What is supervised Machine Learning ? Give Some Examples of Supervised Machine Learning ?

6 Answer

Supervised machine learning is a type of machine learning where the algorithm is trained using labeled data, meaning the data comes with known input-output pairs. The algorithm learns to recognize patterns in the input data and predict the correct output based on that learning. The goal is to generalize this learning to new, unseen data.

In supervised machine learning, the input data is called the "features", and the desired output is called the "label" or "target". The algorithm is trained to find the relationship between the features and the label.

Further there are 2 types of problems we solve by Supervised Machine Learning is 1) Classification Problems 2) Regression Problems

Example of Classification :- Given an image, the algorithm is trained to identify the objects or people in the image and assign a label to it. For example, a machine learning algorithm could be trained to classify images of animals as either cats or dogs.

Example of Regression :- For example, a machine learning algorithm could be trained to predict the price of a house based on its size, location, and other features.

In regression we have a continuous range of values whereas in classification set of output values if fixed as discussed like giving grade to students on the basis of their marks, here grades are limited but the example of price of House can be anything between a range depending on various conditions

7 Q3) What is unsupervised Machine Learning ? Give Some Examples of unsupervised Machine Learning ?

8 Answer

Unsupervised machine learning is a type of machine learning where the algorithm is trained using unlabeled data, meaning the data does not come with known input-output pairs. The goal of unsupervised learning is to find patterns and relationships within the data without prior knowledge of what the output should be.

The algorithm does this by identifying similarities and differences between data points, and clustering them accordingly.

Some examples of unsupervised machine learning include: 1) Clustering : the algorithm is trained to group similar data points together into clusters. For example, a machine learning algorithm could be trained to group customers based on their purchasing behavior

2) Dimensionality reduction - Given a high-dimensional dataset, the algorithm is trained to reduce the dimensionality of the data while preserving its structure

Example, a retail company could use clustering algorithms to identify groups of customers who are likely to be interested in certain products, and then target them with relevant marketing campaigns.

Also, a bank could use anomaly detection algorithms to identify unusual transactions that deviate from the typical spending patterns of a customer.

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

10 Answer

AI (Artificial Intelligence), ML (Machine Learning), DL (Deep Learning), and DS (Data Science) are related but distinct concepts in the field of computer science. Here are the key differences between them:

1) AI is a broad field that encompasses the development of intelligent systems that can perform tasks that normally require human intelligence. ML, DL, and DS are all subsets of AI.

2) ML is a method of teaching computers to learn from data, without being explicitly programmed. ML algorithms use statistical techniques to learn patterns in data and make predictions or decisions based on those patterns.

3) DL is a type of ML that uses neural networks to learn from data. DL algorithms are designed to mimic the structure and function of the human brain, allowing them to learn complex patterns and relationships in data.

4) DS is a multidisciplinary field that involves using techniques from statistics, ML, and other fields to extract insights from data. DS involves everything from data acquisition and cleaning, to analysis, visualization, and interpretation.

11 Q5) What is main difference between Supervised, Supervised and Semi Supervised ML ?

12 Answer

In supervised learning, the training dataset consists of labeled examples, where the desired output or "label" is provided along with the input data. The goal of supervised learning is to learn a mapping from input to output based on the labeled examples, whereas in unsupervised learning, the training dataset consists of unlabeled data, where the goal is to learn patterns or structure in the data without any specific labels. Unsupervised learning algorithms are often used for exploratory data analysis, feature extraction as discussed and in semi-supervised learning, the training dataset consists of both labeled and unlabeled data. The goal is to use the labeled data to guide the learning process and improve the accuracy of the model

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

14 Answer

In machine learning, it is common practice to split a dataset into three subsets: training set, validation set, and test set. Here is an explanation of each subset and its importance:

1) Training set: The training set is a subset of the dataset used to train the machine learning model. It contains input data and the corresponding output or "label" for each data point. The model learns from this data by adjusting its internal parameters to minimize the error between the predicted output and the true output.

2) Validation set: The validation set is a subset of the dataset used to evaluate the performance of the model during training. It is also used to increase the accuracy of our model

3) Test set: The test set is a subset of the dataset used to evaluate the final performance of the model after training is complete. It is used to estimate the generalization error of the model, which is the error rate when the model is applied to new, unseen data

In summary, train-test-validation split is a crucial step in machine learning to ensure that the model is trained, validated, and tested on different subsets of the data.

Training Set is important because It helps the model to learn from the input data and the corresponding output or "label" for each data point.

Validation set is important because it is used to evaluate the performance of the model during training and adjust the hyperparameters to improve performance and helps in increasing accuracy

Test set is important because it is used to evaluate the final performance of the model and estimate its generalization error. It provides an unbiased estimate of the performance of the model on new, unseen data.

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

16 Answer

Unsupervised machine learning algorithms can be used to identify unusual patterns or behaviors that may indicate fraud or other malicious activities. For example, a bank could use anomaly detection algorithms to identify unusual transactions that deviate from the typical spending patterns of a customer.

17 Q8- List down some commonly used supervised and unsupervised machine learning algorithms ?

18 Answer

Here are some most commonly used Machine Learning Algorithms

Supervised Machine Learning Algorithms :- 1) Linear Regression 2) Logistic Regression 3) Decision Trees 4) Random Forest 5) Naive Bayes

Unsupervised Machine Learning Algorithms :- 1) K-Means Clustering 2) Hierarchical Clustering 3) DBSCAN Clustering 4) Principal Component Analysis (PCA) 5) Independent Component Analysis (ICA)

[ ]: