

1.Explain the term machine learning, and how does it work? Explain two machine learning applications in the business world. What are some of the ethical concerns that machine learning applications could raise?

Solution - Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Example of machine learning would be:

Machine learning's ability to identify patterns and discover anomalies that deviate from these patterns make it a great tool for spotting fraudulent banking activities.

Machine learning tools not only assist businesses in setting prices. They also help in delivering the right products and services to the right regions at the right time via predictive inventory planning and customer segmentation.

Ethical concerns machine learning application can raise are lack of transparency of AI tools: AI decisions are not always intelligible to humans. AI is not neutral: AI-based decisions are susceptible to inaccuracies, discriminatory outcomes, embedded or inserted bias. Surveillance practices for data gathering and privacy of court users

2. Describe the process of human learning:

i. Under the supervision of experts

Solution – Human Learning under the supervision of experts provides a clear roadmap and defines what kind of output is required at every stage of a task.

ii. With the assistance of experts in an indirect manner

Solution – Human Learning with the assistance of experts in an indirect manner helps a human to learn with a basic understanding of the end goal but does not have constant idea of what is to be achieved at each stage of a task.

iii. Self-education

Solution - Self-education enables one to interpret one's own ideas about a task and design a self-thought road map which can be used to achieve a task without the guidance or supervision of any external factors.

3. Provide a few examples of various types of machine learning.

Solution -

Supervised Machine Learning

Unsupervised Machine Learning

Semi-Supervised Machine Learning

Reinforcement Learning

4. Examine the various forms of machine learning.

Solution -

Supervised Machine Learning: is defined by its use of labelled datasets to train algorithms that to classify data or predict outcomes accurately.

Unsupervised Machine Learning: uses machine learning algorithms to analyse and cluster unlabelled datasets

Semi-Supervised Machine Learning: the type of machine learning that uses a combination of a small amount of labelled data and a large amount of unlabelled data to train models

Reinforcement Learning: a machine learning training method based on rewarding desired behaviours and/or punishing undesired ones.

5. Can you explain what a well-posed learning problem is? Explain the main characteristics that must be present to identify a learning problem properly.

Solution - Well Posed Learning Problem – A computer program is said to learn from experience E in context to some task T and some performance measure P , if its performance on T , as was measured by P , upgrades with experience E .

Any problem can be segregated as well-posed learning problem if it has three traits:

- Task
- Performance Measure
- Experience

6. Is machine learning capable of solving all problems? Give a detailed explanation of your answer.

Solution - Although machine learning is being used in every industry and helps organizations make more informed and data-driven choices that are more effective than classical methodologies, it still has so many problems that cannot be ignored.

Inadequate Training Data: The major issue that comes while using machine learning algorithms is the lack of quality as well as quantity of data.

Overfitting: Whenever a machine learning model is trained with a huge amount of data, it starts capturing noise and inaccurate data into the training data set. It negatively affects the performance of the model.

Underfitting: Underfitting is just the opposite of overfitting. Whenever a machine learning model is trained with fewer amounts of data, and as a result, it provides incomplete and inaccurate data and destroys the accuracy of the machine learning model.

Slow implementations and results: However, machine learning models are highly efficient in producing accurate results but are time-consuming.

7. What are the various methods and technologies for solving machine learning problems? Any two of them should be defined in detail.

Solution – Regression, Classification, Clustering, Dimensionality Reduction, Ensemble Methods are the different methods and techniques for solving machine learning problems.

Regression methods fall within the category of supervised ML. They help to predict or explain a particular numerical value based on a set of prior data, for example predicting the price of a property based on previous pricing data for similar properties.

Classification

Another class of supervised ML, classification methods predict or explain a class value. For example, they can help predict whether an online customer will buy a product. The output can be yes or no: buyer or not buyer. But classification methods aren't limited to two classes.

8. Can you explain the various forms of supervised learning? Explain each one with an example application.

Solution –

1. Regression

Moreover, it is a type of supervised learning that learns from labelled data sets to predict continuous output for different data in an algorithm.

There two types of Regression

Linear regression - It is used to identify the relationship between two variables, typically used for making future predictions.

Logistic regression - Logistic regression is used when the dependent variable is categorical or has binary outputs like 'yes' or 'no'.

2. Naive Bayes

A Naive Bayes algorithm is used for large datasets. The approach works on the fundamental that every programme in the algorithm works independently. This means that the presence of one feature will not impact the other.

3. Classification

It is a type of supervised learning algorithm that accurately assigns data into different categories or classes. It recognizes specific entities and analyses them to conclude where those entities must be categorized.

4. Neutral networks

This type of supervised learning algorithm is used to group or categorize raw data. In addition, it is used for finding a pattern or interpreting sensory data. However, the algorithm requires numerous amounts of computation resources. As a result, its uses are constrained.

5. Ensemble Method

A random forest algorithm is often called an ensemble method because it combines different supervised learning methods to conclude. Moreover, it uses many decision trees to output the classification of individual trees. As a result, it is widely used across industries.

9. What is the difference between supervised and unsupervised learning? With a sample application in each region, explain the differences.

Solution –

Supervised learning algorithms are trained using labelled data while Unsupervised learning algorithms are trained using unlabelled data.

Supervised learning model takes direct feedback to check if it is predicting correct output or not while Unsupervised learning model does not take any feedback.

Supervised learning model predicts the output while Unsupervised learning model finds the hidden patterns in data.

In supervised learning, input data is provided to the model along with the output while In unsupervised learning, only input data is provided to the model.

10. Describe the machine learning process in depth.

Solution –

Machine learning life cycle involves seven major steps, which are given below:

Gathering Data - Data Gathering is the first step of the machine learning life cycle. The goal of this step is to identify and obtain all data-related problems.

Data preparation - After collecting the data, we need to prepare it for further steps. Data preparation is a step where we put our data into a suitable place and prepare it to use in our machine learning training.

Data Wrangling - Data wrangling is the process of cleaning and converting raw data into a useable format.

Analyse Data - The aim of this step is to build a machine learning model to analyze the data using various analytical techniques and review the outcome.

Train the model - In this step we train our model to improve its performance for better outcome of the problem.

Test the model - In this step, we check for the accuracy of our model by providing a test dataset to it.

Deployment - The last step of machine learning life cycle is deployment, where we deploy the model in the real-world system.

a. Make brief notes on any two of the following:

MATLAB is one of the most widely used programming languages.

ii. Deep learning applications in healthcare

Solution - heart disease, cancer, and brain tumours are diagnosed using medical imaging procedures such as MRI scans, CT scans, and ECG. As a result, deep learning assists doctors in better analysing diseases and providing the best treatment to patients.

iii. Study of the market basket

iv. Linear regression (simple)

Solution - Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

11. Make a comparison between: -

1. Generalization and abstraction

Solution - Abstraction: It involves the translation of data into broader representations.

Generalization: It uses abstracted data to form a basis for action.

2. Learning that is guided and unsupervised

Solution - The main distinction between the two approaches is the use of labelled datasets. To put it simply, supervised learning uses labelled input and output data, while an unsupervised learning algorithm does not.

3. Regression and classification

Solution - The key distinction between Classification vs Regression algorithms is Regression algorithms are used to determine continuous values such as price, income, age, etc. and Classification algorithms are used to forecast or classify the distinct values such as Real or False, Male or Female, Spam or Not Spam, etc.