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?

Machine learning is a form of artificial intelligence (AI) that teaches computers to think in a similar way to how humans do: Learning and improving upon past experiences. It works by exploring data and identifying patterns, and involves minimal human intervention.

* Real-time chatbot agents
* Decision support
* Customer recommendation engines
* Customer churn modelling
* Dynamic pricing tactics
* Market research and customer segmentation
* Fraud detection

When people can look at the source code for a program, they can see how it makes decisions. Another issue is the use of machine learning training data and possible biases. There have been several instances of racial and other biases making it into machine learning programs unintentionally

2. Describe the process of human learning:

i. Under the supervision of experts

ii. With the assistance of experts in an indirect manner

iii. Self-education

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

There are perhaps 14 types of learning that you must be familiar with as a machine learning practitioner; they are:

Learning Problems

1. Supervised Learning

2. Unsupervised Learning

3. Reinforcement Learning

Hybrid Learning Problems

4. Semi-Supervised Learning

5. Self-Supervised Learning

6. Multi-Instance Learning

Statistical Inference

7. Inductive Learning

8. Deductive Inference

9. Transudative Learning

Learning Techniques

10. Multi-Task Learning

11. Active Learning

12. Online Learning

13. Transfer Learning

14. Ensemble Learning

4. Examine the various forms of machine learning.

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.

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

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

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

Different Types of Supervised Learning

1. Regression

In regression, a single output value is produced using training data. This value is a probabilistic interpretation, which is ascertained after considering the strength of correlation among the input variables. For example, regression can help predict the price of a house based on its locality, size, etc.

In logistic regression, the output has discrete values based on a set of independent variables. This method can flounder when dealing with non-linear and multiple decision boundaries. Also, it is not flexible enough to capture complex relationships in datasets.

2. Classification

It involves grouping the data into classes. If you are thinking of extending credit to a person, you can use classification to determine whether or not a person would be a loan defaulter. When the supervised learning algorithm labels input data into two distinct classes, it is called binary classification. Multiple classifications means categorizing data into more than two classes.

3. Naive Bayesian Model

The Bayesian model of classification is used for large finite datasets. It is a method of assigning class labels using a direct acyclic graph. The graph comprises one parent node and multiple children nodes. And each child node is assumed to be independent and separate from the parent.

Decision Trees

A decision tree is a flowchart-like model that contains conditional control statements, comprising decisions and their probable consequences. The output relates to the labelling of unforeseen data.

In the tree representation, the leaf nodes correspond to class labels, and the internal nodes represent the attributes. A decision tree can be used to solve problems with discrete attributes as well as boolean functions. Some of the notable decision tree algorithms are ID3 and CART.

4. Random Forest Model

The random forest model is an ensemble method. It operates by constructing a multitude of decision trees and outputs a classification of the individual trees. Suppose you want to predict which undergraduate students will perform well in GMAT – a test taken for admission into graduate management programs. A random forest model would accomplish the task, given the demographic and educational factors of a set of students who have previously taken the test.

5. Neural Networks

This algorithm is designed to cluster raw input, recognize patterns, or interpret sensory data. Despite their multiple advantages, neural networks require significant computational resources. It can get complicated to fit a neural network when there are thousands of observations. It is also called the ‘black-box’ algorithm as interpreting the logic behind their predictions can be challenging.

6. Support Vector Machines

Support Vector Machine (SVM) is a supervised learning algorithm developed in the year 1990. It draws from the statistical learning theory developed by Vap Nick.

SVM separates hyperplanes, which makes it a discriminative classifier. The output is produced in the form of an optimal hyperplane that categorizes new examples. SVMs are closely connected to the kernel framework and used in diverse fields. Some examples include bioinformatics, pattern recognition, and multimedia information retrieval.

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

|  |  |
| --- | --- |
| **Supervised Learning** | **Unsupervised Learning** |
| Supervised learning algorithms are trained using labeled data. | Unsupervised learning algorithms are trained using unlabeled data. |
| Supervised learning model takes direct feedback to check if it is predicting correct output or not. | Unsupervised learning model does not take any feedback. |
| Supervised learning model predicts the output. | Unsupervised learning model finds the hidden patterns in data. |
| In supervised learning, input data is provided to the model along with the output. | In unsupervised learning, only input data is provided to the model. |
| The goal of supervised learning is to train the model so that it can predict the output when it is given new data. | The goal of unsupervised learning is to find the hidden patterns and useful insights from the unknown dataset. |
| Supervised learning needs supervision to train the model. | Unsupervised learning does not need any supervision to train the model. |
| Supervised learning can be categorized in **Classification** and **Regression** problems. | Unsupervised Learning can be classified in **Clustering** and **Associations** problems. |
| Supervised learning can be used for those cases where we know the input as well as corresponding outputs. | Unsupervised learning can be used for those cases where we have only input data and no corresponding output data. |
| Supervised learning model produces an accurate result. | Unsupervised learning model may give less accurate result as compared to supervised learning. |
| Supervised learning is not close to true Artificial intelligence as in this, we first train the model for each data, and then only it can predict the correct output. | Unsupervised learning is more close to the true Artificial Intelligence as it learns similarly as a child learns daily routine things by his experiences. |
| It includes various algorithms such as Linear Regression, Logistic Regression, Support Vector Machine, Multi-class Classification, Decision tree, Bayesian Logic, etc. | It includes various algorithms such as Clustering, KNN, and Apriori algorithm. |

10. Describe the machine learning process in depth.

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

1. MATLAB is one of the most widely used programming languages.

MATLAB is the easiest and most productive computing environment for engineers and scientists. It includes the MATLAB language, the only top programming language dedicated to mathematical and technical computing. ... "With MATLAB, I can code and debug a new capability much faster than with other languages.

1. Deep learning applications in healthcare

Deep learning applications in healthcare have already been seen in medical imaging solutions, chatbots that can identify patterns in patient symptoms, deep learning algorithms that can identify specific types of cancer, and imaging solutions that use deep learning to identify rare diseases or specific types of diseases

Drug discovery, Medical imaging, Insurance fraud, Alzheimer's disease, Genome

1. Study of the market basket

Market basket analysis is a data mining technique used by retailers to increase sales by better understanding customer purchasing patterns. It involves analyzing large data sets, such as purchase history, to reveal product groupings, as well as products that are likely to be purchased together.

1. Linear regression (simple)

Linear Regression is a machine learning algorithm based on supervised learning. Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x (input) and y(output).

11. Make a comparison between:-

1. Generalization and abstraction

Abstraction is the process of removing details of objects. We step back from concrete objects to consider a number of objects with identical properties. So a concrete object can be looked at as a “superset” of a more abstract object.

A generalization, then, is the formulation of general concepts from specific instances by abstracting common properties. A concrete object can be looked at as a “subset” of a more generalized object.

2. Learning that is guided and unsupervised

The main distinction between the two approaches is the use of labeled datasets. To put it simply, supervised learning uses labeled input and output data, while an unsupervised learning algorithm does not. Unsupervised learning models, in contrast, work on their own to discover the inherent structure of unlabeled data

1. Regression and classification

Classification is the task of predicting a discrete class label. Regression is the task of predicting a continuous quantity.