1. What is the concept of human learning? Please give two examples.

🡪Human learning is the form of learning which requires higher order mental processes like thinking, reasoning, intelligence, etc. We learn different concepts from childhood. For example: When we see a dog and attach the term 'dog', we learn that the word dog refers to a particular animal.

2. What different forms of human learning are there? Are there any machine learning equivalents?

🡪 Learning through Association - Classical Conditioning.

Learning through consequences – Operant Conditioning.

Learning through observation – Modelling/Observational Learning.

Different Forms of ML are as follows:

Artificial Intelligence Learning Theories, Machine Learning, Reinforcement Learning, Supervised Learning, Unsupervised Learning.

ML equivalents like Linear regression, decision trees, random forest and support vector machines are some commonly used techniques that are actually examples of supervised learning.

3. What is machine learning, and how does it work? What are the key responsibilities of machine learning?

🡪 Machine learning is a branch of Artificial intelligence (AI) that teaches computers on how to think in a similar way to how humans do, like by Learning and improving upon past experiences.

It works by exploring data and identifying patterns, and involves minimal human intervention.

Roles and responsibilities of a machine learning engineer are:

Designing ML systems.

Researching and implementing ML algorithms and tools, Selecting appropriate data sets.

Picking appropriate data representation methods, identifying differences in data distribution that affects model performance, Verifying data quality.

4. Define the terms "penalty" and "reward" in the context of reinforcement learning.

🡪 A Reinforcement Learning Algorithm, which may also be referred to as an agent, learns by interacting with its environment. The agent receives rewards by performing correctly and penalties for performing incorrectly. The agent learns without intervention from a human by maximizing its reward and minimizing its penalty.

5. Explain the term "learning as a search"?

🡪 Learning can be viewed as a search through the space of all sentences in a concept description language for a sentence that best describes the data. Alternatively, it can be viewed as a search through all hypotheses in a hypothesis space. Concept learning can be viewed as the task of searching through a large space of hypotheses implicitly defined by the hypothesis representation. The goal of this search is to find the hypothesis that best fits the training examples.

6. What are the various goals of machine learning? What is the relationship between these and human learning?

🡪 The Goal of machine learning, closely coupled with the goal of AI, is to achieve a thorough understanding about the nature of learning process (both human learning and other forms of learning), about the computational aspects of learning behaviours, and to implant the learning capability in computer systems. The Goals of Machine Learning in simple words, is to understand the nature of human and other forms of learning, and to build learning capability in computers. To be more specific, there are three aspects of the goals of ML.

1.To make the computers smarter, more intelligent. The more direct objective in this aspect is to develop systems (programs) for specific practical learning tasks in application domains.

2.To develop computational models of human learning process and perform computer simulations. The study in this aspect is also called cognitive modelling.

3.To explore new learning methods and develop general learning algorithms independent of application.

Humans have the ability to learn, however with the progress in artificial intelligence, machine learning has become a resource which can augment or even replace human learning. Learning does not happen all at once, but it builds upon and is shaped by previous knowledge. Humans acquire knowledge through experience either directly or shared by others. Machines acquire knowledge through experience shared in the form of past data. We have the terms, Knowledge, Skill, and Memory being used to define intelligence.

7. Illustrate the various elements of machine learning using a real-life illustration.

🡪 The Various elements of the the Machine Learning are:

Data

Task

Model

Loss Function

Learning Algorithm

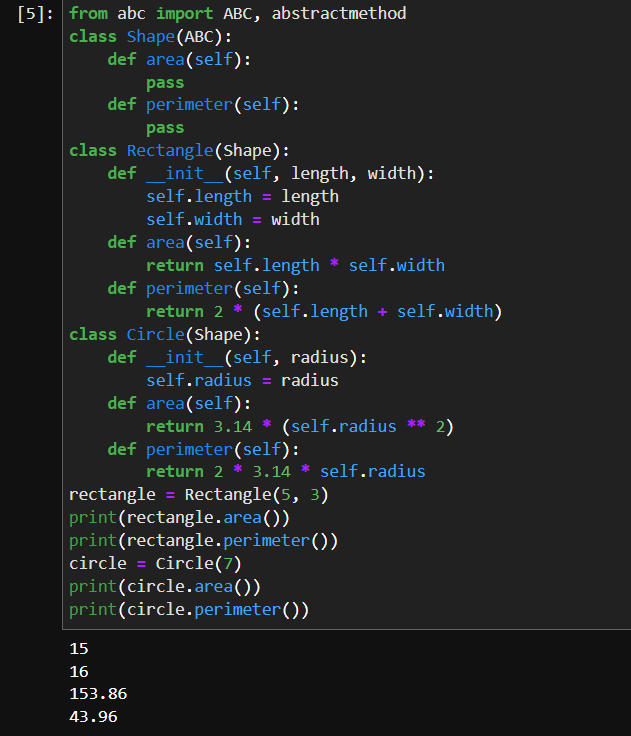
Evaluation

8. Provide an example of the abstraction method.

🡪 we define an abstract base class Shape that has two abstract methods: area() and perimeter(). These methods don't have any implementation; they are just placeholders that must be overridden by any concrete subclass that inherits from the Shape class.

We then create two concrete subclasses, Rectangle and Circle, that inherit from the Shape class. These subclasses provide their own implementations of the area() and perimeter() methods specific to rectangles and circles.

By using the abstraction method, we ensure that any class inheriting from the Shape class must implement the area() and perimeter() methods. This allows us to create objects of different shapes while enforcing the presence of these common methods.



9. What is the concept of generalization? What function does it play in the machine learning process?

🡪 Generalization refers to your model's ability to adapt properly to new, previously unseen data, drawn from the same distribution as the one used to create the model.

An example is when we train a model to classify between dogs and cats. If the model is provided with dogs images dataset with only two breeds, it may obtain a good performance. But, it possibly gets a low classification score when it is tested by other breeds of dogs as well. This issue can result to classify an actual dog image as a cat from the unseen dataset. Therefore, data diversity is very important factor in order to make a good prediction. In the sample above, the model may obtain 85% performance score when it is tested by only two dog breeds and gains 70% if trained by all breeds. However, the first possibly gets a very low score (e.g. 45%) if it is evaluated by an unseen dataset with all breed dogs. This for the latter can be unchanged given than it has been trained by high data diversity including all possible breeds.

It should be taken into account that data diversity is not the only point to care in order to have a generalized model. It can be resulted by nature of a machine learning algorithm, or by poor hyper-parameter configuration. In this post we explain all determinant factors. There are some methods (regularization) to apply during model training to ensure about generalization. But before, we explain bias and variance as well as underfitting and overfitting.

10. What is classification, exactly? What are the main distinctions between classification and regression?

🡪 In Machine Learning, Classification refers to a predictive modelling problem where a class label is predicted for a given example of input data. Classification is the task of predicting a discrete class label. Whereas Regression is the task of predicting a continuous quantity.

Classification predictive modelling is the task of approximating a mapping function (f) from input variables (X) to discrete output variables (y).

The output variables are often called labels or categories. The mapping function predicts the class or category for a given observation.

For example, an email of text can be classified as belonging to one of two classes: “spam“ and “not spam“.

-A classification problem requires that examples be classified into one of two or more classes.

-A classification can have real-valued or discrete input variables.

-A problem with two classes is often called a two-class or binary classification problem.

-A problem with more than two classes is often called a multi-class classification problem.

-A problem where an example is assigned multiple classes is called a multi-label classification problem.

It is common for classification models to predict a continuous value as the probability of a given example belonging to each output class. The probabilities can be interpreted as the likelihood or confidence of a given example belonging to each class. A predicted probability can be converted into a class value by selecting the class label that has the highest probability.

Regression predictive modelling is the task of approximating a mapping function (f) from input variables (X) to a continuous output variable (y).

A continuous output variable is a real-value, such as an integer or floating point value. These are often quantities, such as amounts and sizes.

For example, a house may be predicted to sell for a specific dollar value, perhaps in the range of

200,000.

-A regression problem requires the prediction of a quantity.

-A regression can have real valued or discrete input variables.

-A problem with multiple input variables is often called a multivariate regression problem.

-A regression problem where input variables are ordered by time is called a time series forecasting problem.

11. What is regression, and how does it work? Give an example of a real-world problem that was solved using regression.

🡪 Regression is a Supervised Machine Learning technique which is used to predict continuous values. The ultimate goal of a regression algorithm is to plot a best-fit line or a curve between the data.

The three main metrics that are used for evaluating the trained regression model are Variance, Bias and Error.

A simple linear regression real life example could mean you finding a relationship between the revenue and temperature, with a sample size for revenue as the dependent variable. In case of multiple variable regression, you can find the relationship between temperature, pricing and number of workers to the revenue.

12. Describe the clustering mechanism in detail.

🡪 Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups. In simple words, the aim is to segregate groups with similar traits and assign them into clusters. Clustering is the process of making a group of abstract objects into classes of similar objects.

-A cluster of data objects can be treated as one group.

-While doing cluster analysis, we first partition the set of data into groups based on data similarity and then assign the labels to the groups.

-The main advantage of clustering over classification is that, it is adaptable to changes and helps single out useful features that distinguish different groups.

Applications of Cluster Analysis

Clustering analysis is broadly used in many applications such as market research, pattern recognition, data analysis, and image processing.

Clustering can also help marketers discover distinct groups in their customer base. And they can characterize their customer groups based on the purchasing patterns.

In the field of biology, it can be used to derive plant and animal taxonomies, categorize genes with similar functionalities and gain insight into structures inherent to populations.

Clustering also helps in identification of areas of similar land use in an earth observation database. It also helps in the identification of groups of houses in a city according to house type, value, and geographic location.

Clustering also helps in classifying documents on the web for information discovery.

Clustering is also used in outlier detection applications such as detection of credit card fraud.

As a data mining function, cluster analysis serves as a tool to gain insight into the distribution of data to observe characteristics of each cluster.

Clustering methods can be classified into the following categories

Partitioning Method

Hierarchical Method

Density-based Method

Grid-Based Method

Model-Based Method Constraint-based Method

13. Make brief observations on two of the following topics:

i. Machine learning algorithms are used

ii. Studying under supervision

iii. Studying without supervision

iv. Reinforcement learning is a form of learning based on positive reinforcement.

**Machine learning algorithms are used**: At its Most basic, Machine Learning uses programmed algorithms that receive and analyse input data to predict output values within an acceptable range. As new data is fed to these algorithms, they learn and optimise their operations to improve performance, developing intelligence over time.

**Studying Under Supervision**: In machine learning, there are two important categories- Supervised and Unsupervised learning. Supervised learning, an algorithm learns from a training dataset. We know the correct answers or desired output, the algorithm makes predictions using the given dataset and is corrected by the “supervisor”.