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?

Ans:. Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, uncovering key insights within data mining projects. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase, requiring them to assist in the identification of the most relevant business questions and subsequently the data to answer them.Supervised machine learning

Supervised learning, also known as supervised machine learning, is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately. As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately. This occurs as part of the cross validation process to ensure that the model avoids overfitting or underfitting. Supervised learning helps organizations solve for a variety of real-world problems at scale, such as classifying spam in a separate folder from your inbox. Some methods used in supervised learning include neural networks, naïve bayes, linear regression, logistic regression, random forest, support vector machine (SVM), and more.

Unsupervised machine learning

Unsupervised learning, also known as unsupervised machine learning, uses machine learning algorithms to analyze and cluster unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention. Its ability to discover similarities and differences in information make it the ideal solution for exploratory data analysis, cross-selling strategies, customer segmentation, image and pattern recognition. It's also used to reduce the number of features in a model through the process of dimensionality reduction; principal component analysis (PCA) and singular value decomposition (SVD) are two common approaches for this. Other algorithms used in unsupervised learning include neural networks, k-means clustering, probabilistic clustering methods, and more.

2. Describe the process of human learning: i. Under the supervision of experts

IL With the assistance of experts in an indirect manner iii. Self-education.?

Ans:. Learning is the process of acquiring new understanding, knowledge, behaviors, skills, values, attitudes, and preferences. The ability to learn is possessed by humans, animals, and some machines; there is also evidence for some kind of learning in certain plants.Supervised learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y).Human-guided machine learning is a type of supervised learning, which uses a set of human-

labeled training data to develop a model. In supervised learning, the algorithm learns a set of inputs along with corresponding correct outputs.Meaning : Human learning is a process of acquiring knowledge. Our behavior, skills, values and ethics are acquired when we process information through our minds and learn. Human learning may occur as part of education, personal development or any other informal/formal trainingSelf-learning is an approach to learning where the individual makes the effort to identify their own learning needs, set learning goals, find the necessary resources, and evaluate their own knowledge

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

Ans:.

Types of Machine Learning

Machine learning is a subset of AI, which enables the machine to automatically learn from data, improve performance from past experiences, and make predictions. Machine learning contains a set of algorithms that work on a huge amount of data. Data is fed to these algorithms to train them, and on the basis of training, they build the model & perform a specific task.

Types of Machine Learning

These ML algorithms help to solve different business problems like Regression, Classification, Forecasting, Clustering, and Associations, etc.Based on the methods and way of learning, machine learning is divided into mainly four types, which are:

Supervised Machine Learning

Unsupervised Machine Learning

Semi-Supervised Machine Learning

Reinforcement Learning. Supervised Machine Learning

Let’s understand supervised learning with an example. Suppose we have an input dataset of cats and dog images. So, first, we will provide the training to the machine to understand the images, such as the shape & size of the tail of cat and dog, Shape of eyes, colour, height (dogs are taller, cats are smaller), etc. After completion of training, we input the picture of a cat and ask the machine to identify the object and predict the output. Now, the machine is well trained, so it will check all the features of the object, such as height, shape, colour, eyes, ears, tail, etc., and find that it’s a cat. So, it will put it in the Cat category. This is the process of how the machine identifies the objects in Supervised Learning.. Unsupervised Machine Learning

Unsupervised learning is different from the Supervised learning technique; as its name suggests, there is no need for supervision. It means, in unsupervised machine learning, the machine is trained using the unlabeled dataset, and the machine predicts the output without any supervision.

In unsupervised learning, the models are trained with the data that is neither classified nor labelled, and the model acts on that data without any supervision.. Semi-Supervised Learning

Semi-Supervised learning is a type of Machine Learning algorithm that lies between Supervised and Unsupervised machine learning. It represents the intermediate ground between Supervised (With Labelled training data) and Unsupervised learning (with no labelled training data) algorithms and uses the combination of labelled and unlabeled datasets during the training period.

Although Semi-supervised learning is the middle ground between supervised and unsupervised learning and operates on the data that consists of a few labels, it mostly consists of unlabeled data. As labels are costly, but for corporate purposes, they may have few labels. It is completely different from supervised and unsupervised learning as they are based on the presence & absence of labels.

To overcome the drawbacks of supervised learning and unsupervised learning algorithms, the concept of Semi-supervised learning is introduced. The main aim of semi-supervised learning is to effectively use all the available data, rather than only labelled data like in supervised learning. Initially, similar data is clustered along with an unsupervised learning algorithm, and further, it helps to label the unlabeled data into labelled data. It is because labelled data is a comparatively more expensive acquisition than unlabeled data. Reinforcement Learning

Reinforcement learning works on a feedback-based process, in which an AI agent (A software component) automatically explore its surrounding by hitting & trail, taking action, learning from experiences, and improving its performance. Agent gets rewarded for each good action and get punished for each bad action; hence the goal of reinforcement learning agent is to maximize the rewards.

In reinforcement learning, there is no labelled data like supervised learning, and agents learn from their experiences only.

The reinforcement learning process is similar to a human being; for example, a child learns various things by experiences in his day-to-day life. An example of reinforcement learning is to play a game, where the Game is the environment, moves of an agent at each step define states, and the goal of the agent is to get a high score. Agent receives feedback in terms of punishment and rewards.

4. Examine the various forms of machine learning.?

Ans:.

Supervised Learning

Supervised learning is one of the most basic types of machine learning. In this type, the machine learning algorithm is trained on labeled data. Even though the data needs to be labeled accurately for this method to work, supervised learning is extremely powerful when used in the right circumstances.

In supervised learning, the ML algorithm is given a small training dataset to work with. This training dataset is a smaller part of the bigger dataset and serves to give the algorithm a basic idea of the problem, solution, and data points to be dealt with. The training dataset is also very similar to the final dataset in its characteristics and provides the algorithm with the labeled parameters required for the problem.

The algorithm then finds relationships between the parameters given, essentially establishing a cause and effect relationship between the variables in the dataset. At the end of the training, the algorithm has an idea of how the data works and the relationship between the input and the output.

This solution is then deployed for use with the final dataset, which it learns from in the same way as the training dataset.Unsupervised Learning

Unsupervised machine learning holds the advantage of being able to work with unlabeled data. This means that human labor is not required to make the dataset machine-readable, allowing much larger datasets to be worked on by the program.

In supervised learning, the labels allow the algorithm to find the exact nature of the relationship between any two data points. However, unsupervised learning does not have labels to work off of, resulting in the creation of hidden structures. Relationships between data points are perceived by the algorithm in an abstract manner, with no input required from human beings.

The creation of these hidden structures is what makes unsupervised learning algorithms versatile. Instead of a defined and set problem statement, unsupervised learning algorithms can adapt to the data by dynamically changing hidden structures. This offers more post-deployment development than supervised learning algorithms.Reinforcement learning directly takes inspiration from how human beings learn from data in their lives. It features an algorithm that improves upon itself and learns from new situations using a trial-and-error method. Favorable outputs are encouraged or ‘reinforced’, and non-favorable outputs are discouraged or ‘punished’.

Based on the psychological concept of conditioning, reinforcement learning works by putting the algorithm in a work environment with an interpreter and a reward system. In every iteration of the algorithm, the output result is given to the interpreter, which decides whether the outcome is favorable or not.

In case of the program finding the correct solution, the interpreter reinforces the solution by providing a reward to the algorithm. If the outcome is not favorable, the algorithm is forced to reiterate until it finds a better result. In most cases, the reward system is directly tied to the effectiveness of the result.

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.?

Ans:. Well posed learning problems

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

Certain examples that efficiently defines the well-posed learning problem are –

1. To better filter emails as spam or not

Task – Classifying emails as spam or not

Performance Measure – The fraction of emails accurately classified as spam or not spam

Experience – Observing you label emails as spam or not spam

2. A checkers learning problem

Task – Playing checkers game

Performance Measure – percent of games won against opposer

Experience – playing implementation games against itself

3. Handwriting Recognition Problem

Task – Acknowledging handwritten words within portrayal

Performance Measure – percent of words accurately classified

Experience – a directory of handwritten words with given classifications

4. A Robot Driving Problem

Task – driving on public four-lane highways using sight scanners

Performance Measure – average distance progressed before a fallacy

Experience – order of images and steering instructions noted down while observing a human driver

5. Fruit Prediction Problem

Task – forecasting different fruits for recognition

Performance Measure – able to predict maximum variety of fruits

Experience – training machine with the largest datasets of fruits images

6. Face Recognition Problem

Task – predicting different types of faces

Performance Measure – able to predict maximum types of faces

Experience – training machine with maximum amount of datasets of different face images

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

Ans:. .We are always amazed at how machine learning has made such an impact on our lives. There is no doubt that ML will completely change the face of various industries, as well as job profiles. While it offers a promising future, there are some inherent problems at the heart of ML and AI advancements that put these technologies at a disadvantage. While it can solve a plethora of challenges, there are a few tasks which ML fails to answer. We are listing five such problems in this article.

1. Reasoning Power

One area where ML has not mastered successfully is reasoning power, a distinctly human trait. Algorithms available today are mainly oriented towards specific use-cases and are narrowed down when it comes to applicability. They cannot think as to why a particular method is happening that way or 'introspect' their own outcomes.

2. Contextual Limitation

If we consider the area of natural language processing (NLP), text and speech information are the means to understand languages by NLP algorithms. They may learn letters, words, sentences or even the syntax, but where they fall back is the context of the language. Algorithms do not understand the context of the language used. A classic example for this would be the "Chinese room" argument given by philosopher John Searle, which says that computer programs or algorithms grasp the idea merely by 'symbols' rather than the context given. (You can find the complete information on Chinese room here).

So, ML does not have an overall idea of the situation. It is limited by umnemonic interpretations rather than thinking to see what is actually going on.

3. Scalability

Although we see ML implementations being deployed on a significant basis, it all depends on data as well as its scalability. Data is growing at an enormous rate and has many forms which largely affects the scalability of an ML project. Algorithms cannot do much about this unless they are updated constantly for new changes to handle data. This is where ML regularly requires human intervention in terms of scalability and remains unsolved mostly.

In addition, growing data has to be dealt the right way if shared on an ML platform which again needs examination through knowledge and intuition apparently lacked by current ML.

4. Regulatory Restriction For Data In ML

ML usually need considerable amounts (in fact, massive) of data in stages such as training, cross-validation etc. Sometimes, data includes private as well as general information. This is where it gets complicated. Most tech companies have privatise TV TVd data and these data are the ones which are actually useful for ML applications. But, there comes the risk of the wrong usage of data, especially in critical areas such as medical research, health insurance etc.,

Even though data are anonymised at times, it has the possibility of being vulnerable. Hence this is the reason regulatory rules are imposed heavily when it comes to using private data.

5. Internal Working Of Deep Learning

This sub-field of ML is actually responsible for today's AI growth. What was once just a theory has appeared to be the most powerful aspect of ML. Deep Learning (DL) now powers applications such as voice recognition, image recognition and so on through artificial neural networks.

But, the internal working of DL is still unknown and yet to be solved. Advanced DL algorithms still baffle researchers in terms of its working and efficiency. Millions of neurons that form the neural networks in DL increase abstraction at every level, which cannot be comprehended at all. This is why deep learning is dubbed a 'black box' since its internal agenda is unknown.

Conclusion

All of these problems are very challenging for computer scientists and researchers to solve. The reason for this is uncertainty. If researchers aim at more groundwork related to ML rather than improving this field, we might have an answer to the unsolved problems listed here. After all, ML should be realised apart from being utilitarian

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

Ans: 10 Machine Learning Methods that Every Data Scientist Should Know 1.Regression.

2.Classification.

3.Clustering.

4.Dimensionality Reduction.

5.Ensemble Methods.

6.Neural Nets and Deep Learning

7.Transfer Learning.

8.Reinforcement Learning

9.Natural Language Processing

10.Word Embeddings Clustering With clustering methods, we get into the category of unsupervised ML because their goal is to group or cluster observations that have similar characteristics. Clustering methods don't use output information for training, but instead let the algorithm define the output. In clustering methods, we can only use visualizations to inspect the quality of the solution. The most popular clustering method is K-Means, where "K" represents the number of clusters that the user chooses to create. (Note that there are various techniques for choosing the value of K, such as the elbow method.) Roughly, what K-Means does with the data points: Randomly chooses K centers within the data. Assigns each data point to the closest of the randomly created centers. Re-computes the center of each cluster. If centers don't change (or change very little), the process is finished. Otherwise, we return to step 2. (To prevent ending up in an infinite loop if the centers continue to change, set a maximum number of iterations in advance.) The next plot applies K-Means to a data set of buildings. Each column in the plot indicates the efficiency for each building. The four measurements are related to air

conditioning, plugged-in equipment (microwaves, refrigerators, etc…), domestic gas, and heating gas. We chose K=2 for clustering, which makes it easy to interpret one of the clusters as the group of efficient buildings and the other cluster as the group of inefficient buildings. To the left you see the location of the buildings and to right you see two of the four dimensions we used as inputs: plugged-in equipment and heating gas. Dimensionality Reduction As the name suggests, we use dimensionality reduction to remove the least important information (sometime redundant columns) from a data set. In practice, I often see data sets with hundreds or even thousands of columns (also called features), so reducing the total number is vital. For instance, images can include thousands of pixels, not all of which matter to your analysis. Or when testing microchips within the manufacturing process, you might have thousands of measurements and tests applied to every chip, many of which provide redundant information. In these cases, you need dimensionality reduction algorithms to make the data set manageable. The most popular dimensionality reduction method is Principal Component Analysis (PCA), which reduces the dimension of the feature space by finding new vectors that maximize the linear variation of the data. PCA can reduce the dimension of the data dramatically and without losing too much information when the linear correlations of the data are strong. (And in fact you can also measure the actual extent of the information loss and adjust accordingly.) Another popular method is t-Stochastic Neighbor Embedding (t-SNE), which does non-linear dimensionality reduction. People typically use t-SNE for data visualization, but you can also use it for machine learning tasks like reducing the feature space and clustering, to mention just a few.

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

Ans:.

In Supervised Learning, a machine is trained using 'labeled' data. Datasets are said to be labeled when they contain both input and output parameters. In other words, the data has already been tagged with the correct answer.

So, the technique mimics a classroom environment where a student learns in the presence of a supervisor or teacher. On the other hand, unsupervised learning algorithms let the models discover information and learn on their own.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. . 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.. Naïve 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. 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.

6. 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. 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.?

Ans:

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.Difference between Supervised and Unsupervised Learning

Supervised and Unsupervised learning are the two techniques of machine learning. But both the techniques are used in different scenarios and with different datasets. Below the explanation of both learning methods along with their difference table is given.

The main differences between Supervised and Unsupervised learning are given below:

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 Depth.?

Ans:. In Deep Neural Networks the depth refers to how deep the network is but in this context, the depth is used for visual recognition and it translates to the 3rd dimension of an image. In this case you have an image, and the size of this input is 32x32x3 which is (width, height, depth) .was taking a look at Convolutional Neural Network from CS231n Convolutional Neural Networks for Visual Recognition. In Convolutional Neural Network, the neurons are arranged in 3 dimensions(height, width, depth). I am having trouble with the depth of the CNN. I can't visualize what it is.

In the link they said The CONV layer's parameters consist of a set of learnable filters. Every filter is small spatially (along width and height), but extends through the full depth of the input volume.

In Deep Neural Networks the depth refers to how deep the network is but in this context, the depth is used for visual recognition and it translates to the 3rd dimension of an image. In this case you have an image, and the size of this input is 32x32x3 which is (width, height, depth) .

In-depth guide to machine learning in the enterprise

Home Machine learning platforms Business software machine learning

DEFINITION

Machine learning

Ed Burns, Executive Editor

What is machine learning?

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

Recommendation engines are a common use case for machine learning. Other popular uses include fraud detection, spam filtering, malware threat detection, business process automation (BPA) and Predictive maintenance.

Machine learning is important because it gives enterprises a view of trends in customer behavior and business operational patterns, as well as supports the development of new products. Many of today's leading companies, such as Facebook, Google and Uber, make machine learning a central part of their operations. Machine learning has become a significant competitive differentiator for many companies.

Classical machine learning is often categorized by how an algorithm learns to become more accurate in its predictions. There are four basic approaches:supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. The type of algorithm data scientists choose to use depends on what type of data they want to predict.