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?

Answer: Machine Learning is a branch of Artificial Intelligence focus on building application that learn fron data
and improve their accuracy over time without being programmed to do so.
Machine Learning is an application on it gives devices the abillity to learn from their exeprience and improve thier self without doing any coding.
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.
Machine Learning in finance helps in portfolio management, algorithmic trading, loan underwriting, and fraud detection. However, future applications of Machine Learning in finance will include Chatbots and other conversational interfaces for security, customer service, and sentiment analysis.
The following list enumerates all the ethical issues that were identified from the case studies and the Delphi study, totalling 39.
Cost to innovation.
Harm to physical integrity.
Lack of access to public services.
Lack of trust.
“Awakening” of AI.
Security problems.
Lack of quality data.
Disappearance of jobs.

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

Answer:(i)--->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. The training data used to create a machine learning model is assumed to be ground truth, meaning that its validity is not questioned–however, the model must still be tested for accuracy before it can be deployed. There are also subsets of supervised learning known as active learning, or semi-supervised learning, where the machine learning model is improved with each additional correction or piece of information collected. This is where humans come in.

```
Human-guided machine learning is a process whereby subject matter experts
accelerate the learning process by teaching the technology in real-time. For
example, if the machine learning model comes across a piece of data it is
uncertain about, a human can be asked to weigh in and give feedback. The model
then learns from this input, and uses it to make a more accurate prediction the
next time. Human-guided machine learning works from the bottom up by first
using algorithms to conduct the heavy lifting of identifying relationships
within the data, and engaging humans when necessary for training or validation.
This means that, inevitably, the amount of time a human needs to spend
performing a specific task will decrease as the machine learning accuracy
increases.

(ii)--->Guided learning is a term that refers to a process in which learners
initiate and advance their learning guided by more experienced partners and
socially derived sources, such as tools, text, and/or other artifacts. This
process of learning is seen as being distinct from didactic instruction
comprising unidirectional interaction between more and less knowledgeable
interlocutors, as in teaching. In contrast, the direction and process of guided
learning is premised more on the learners' intentions, capacity, and agency,
albeit being guided by social partners and norms and forms. This guidance will
likely take two forms: (1) close interpersonal interactions with more informed
partners (e.g., experts, teachers, parents) and (2) indirect guidance from
observing and interacting

(iii)--->Human learning begins before birth and continues until death as a
consequence of ongoing interactions between person and environment. The nature
and processes involved in learning are studied in many fields, including
educational psychology, neuropsychology, experimental psychology, and pedagogy.
```

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

Answer:Example of Supervised Learning Algorithms:

Linear Regression.

Logistic Regression.

Nearest Neighbor.

Gaussian Naive Bayes.

Decision Trees.

Support Vector Machine (SVM)

Random Forest.

4. Examine the various forms of machine learning.

Answer:These are three types of machine learning: supervised learning, unsupervised learning, and reinforcement 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.

Answer: A (machine learning) problem is well-posed if a solution to it exists, if that solution is unique, and if that solution depends on the data / experience but it is not sensitive to (reasonably small) changes in the data / experience.The number one problem facing Machine Learning is the lack of good data. While enhancing algorithms often consumes most of the time of developers in AI, data quality is essential for the algorithms to function as intended. The trick is to have a very clear idea about what does your problem intend to solve and how do you want to go about doing it. For example, if your problem is classification, then you should consider looking at various Machine Learning and Deep Learning Algorithms which will perform classification for you.To break it down, the three important components of a well-posed learning problem are, Task. Performance Measure. Experience.To break it down, the three important components of a well-posed learning problem are, Task. Performance Measure. Experience.

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

Answer:Most people reading this are likely familiar with machine learning and the relevant algorithms used to classify or predict outcomes based on data. However, it is important to understand that machine learning is not the answer to all problems. Given the usefulness of machine learning, it can be hard to accept that sometimes it is not the best solution to a problem.Machine learning, a subset of artificial intelligence, has revolutionalized the world as we know it in the past decade. The information explosion has resulted in the collection of massive amounts of data, especially by large companies such as Facebook and Google. This amount of data, coupled with the rapid development of processor power and computer parallelization, has now made it possible to obtain and study huge amounts of data with relative ease.

Limitation 1 — Ethics Machine learning, a subset of artificial intelligence, has revolutionalized the world as we know it in the past decade. The information explosion has resulted in the collection of massive amounts of data, especially by large companies such as Facebook and Google. This amount of data, coupled with the rapid development of processor power and computer parallelization, has now made it possible to obtain and study huge amounts of data with relative ease.

It is easy to understand why machine learning has had such a profound impact on the world, what is less clear is exactly what its capabilities are, and perhaps more importantly, what its limitations are. Yuval Noah Harari famously coined the term ‘dataism’, which refers to a putative new stage of civilization we are entering in which we trust algorithms and data more than our own judgment and logic.

Whilst you may find this idea laughable, remember the last time you went on vacation and followed the instructions of a GPS rather than your own judgment on a map — do you question the judgment of the GPS? People have literally driven into lakes because they blindly followed the instructions from their GPS.

The idea of trusting data and algorithms more than our own judgment has its pros and cons. Obviously, we benefit from these algorithms, otherwise, we wouldn't be using them in the first place. These algorithms allow us to automate processes by making informed judgments using available data. Sometimes, however, this means replacing someone's job with an algorithm, which comes with ethical ramifications. Additionally, who do we blame if something goes wrong?

The most commonly discussed case currently is self-driving cars — how do we choose how the vehicle should react in the event of a fatal collision? In the future will we have to select which ethical framework we want our self-driving car to follow when we are purchasing the vehicle?

If my self-driving car kills someone on the road, whose fault is it?

Whilst these are all fascinating questions, they are not the main purpose of this article. Clearly, however, machine learning cannot tell us anything about what normative values we should accept, i.e. how we should act in the world in a given situation. As David Hume famously said, one cannot 'derive an ought from an is'.

Limitation 2 — Deterministic Problems This is a limitation I personally have had to deal with. My field of expertise is environmental science, which relies heavily on computational modeling and using sensors/IoT devices.

Machine learning is incredibly powerful for sensors and can be used to help calibrate and correct sensors when connected to other sensors measuring environmental variables such as temperature, pressure, and humidity. The correlations between the signals from these sensors can be used to develop self-calibration procedures and this is a hot research topic in my research field of atmospheric chemistry.

However, things get a bit more interesting when it comes to computational modeling.

Running computer models that simulate global weather, emissions from the planet, and transport of these emissions is very computationally expensive. In fact, it is so computationally expensive, that a research-level simulation can take weeks even when running on a supercomputer.

Good examples of this are MM5 and WRF, which are numerical weather prediction models that are used for climate research and for giving you weather forecasts on the morning news. Wonder what weather forecasters do all day? Run and study these models.

Running weather models is fine, but now that we have machine learning, can we just use this instead to obtain our weather forecasts? Can we leverage data from satellites, weather stations, and use an elementary predictive algorithm to discern whether it is going to rain tomorrow?

The answer is, surprisingly, yes. If we have knowledge of the air pressures around a certain region, the levels of moisture in the air, wind speeds, and information about neighboring points and their own variables, it becomes possible to train, for example, a neural network. But at what cost?

Using a neural network with a thousand inputs to determine whether it will rain tomorrow in Boston is possible. However, utilizing a neural network misses the entire physics of the weather system.

Machine learning is stochastic, not deterministic.

A neural network does not understand Newton's second law, or that density cannot be negative — there are no physical constraints.

However, this may not be a limitation for long. There are multiple researchers looking at adding physical constraints to neural networks and other algorithms so that they can be used for purposes such as this.

Limitation 3 — Data This is the most obvious limitation. If you feed a model poorly, then it will only give you poor results. This can manifest itself in two ways: lack of data, and lack of good data.

Lack of Data

Many machine learning algorithms require large amounts of data before they begin to give useful results. A good example of this is a neural network. Neural networks are data-eating machines that require copious amounts of training data. The larger the architecture, the more data is needed to produce viable results. Reusing data is a bad idea, and data augmentation is useful to some extent, but having more data is always the preferred solution.

If you can get the data, then use it.

Lack of Good Data

Despite the appearance, this is not the same as the above comment. Let's imagine you think you can cheat by generating ten thousand fake data points to put in your neural network. What happens when you put it in?

It will train itself, and then when you come to test it on an unseen data set, it will not perform well. You had the data but the quality of the data was not up to scratch.

In the same way that having a lack of good features can cause your algorithm to perform poorly, having a lack of good ground truth data can also limit the capabilities of your model. No company is going to implement a machine learning model that performs worse than human-level error.

Similarly, applying a model that was trained on a set of data in one situation may not necessarily apply as well to a second situation. The best example of this I have found so far is in breast cancer prediction.

Mammography databases have a lot of images in them, but they suffer from one problem that has caused significant issues in recent years — almost all of the x-rays are from white women. This may not sound like a big deal, but actually, black women have been shown to be 42 percent more likely to die from breast cancer due to a wide range of factors that may include differences in detection and access to health care. Thus, training an algorithm primarily on white women adversely impacts black women in this case.

What is needed in this specific case is a larger number of x-rays of black patients in the training database, more features relevant to the cause of this 42 percent increased likelihood, and for the algorithm to be more equitable by stratifying the dataset along the relevant axes.

If you are skeptical of this or would like to know more, I recommend you look at this article.

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

Answer:machine learning algorithms have the ability to improve themselves through training. Today, ML algorithms are trained using three prominent methods. These are three types of machine learning: supervised learning, unsupervised learning, and reinforcement learning. 1.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. This means that supervised machine learning algorithms will continue to improve even after being deployed, discovering new patterns and relationships as it trains itself on new data.

2.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.

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

Answer:There are two types of Supervised Learning techniques: Regression and Classification. Classification separates the data, Regression fits the data

A. Classification: It is a Supervised Learning task where output is having defined labels(discrete value). For example in above Figure A, Output – Purchased has defined labels i.e. 0 or 1; 1 means the customer will purchase, and 0 means that the customer won't purchase. The goal here is to predict discrete values belonging to a particular class and evaluate them on the basis of accuracy. It can be either binary or multi-class classification. In binary classification, the model predicts either 0 or 1; yes or no but in the case of multi-class classification, the model predicts more than one class. Example: Gmail classifies mails in more than one class like social, promotions, updates, and forums.

B. Regression: It is a Supervised Learning task where output is having continuous value. For example in above Figure B, Output – Wind Speed is not having any discrete value but is continuous in a particular range. The goal here is to predict a value as much closer to the actual output value as our model can and then evaluation is done by calculating the error value. The smaller the error the greater the accuracy of our regression model.

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

Answer: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. To put it simply, supervised learning uses labeled input and output data, while an unsupervised learning algorithm does not. In supervised learning, the algorithm "learns" from the training dataset by iteratively making predictions on the data and adjusting for the correct answer.

SUPERVISED LEARNING UNSUPERVISED LEARNING

Real Time Uses off-line analysis --- Uses Real Time Analysis of Data

Number of Classes ---- Number of Classes are known Number of Classes are not known

Accuracy of Results Accurate and Reliable Results --- Moderate Accurate and Reliable Results

10. Describe the machine learning process in depth.

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

iii. Study of the market basket

iv. Linear regression (simple)

Answer:

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.

Why is machine learning important? 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.

What are the different types of machine learning? 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.

THIS ARTICLE IS PART OF

In-depth guide to machine learning in the enterprise Which also includes: Learn the business value of AI's various techniques 10 common uses for machine learning applications in business 6 ways to reduce different types of bias in machine learning DOWNLOAD 1 Download this entire guide for FREE now!

Supervised learning: In this type of machine learning, data scientists supply algorithms with labeled training data and define the variables they want the algorithm to assess for correlations. Both the input and the output of the algorithm is specified. Unsupervised learning: This type of machine learning involves algorithms that train on unlabeled data. The algorithm scans through data sets looking for any meaningful connection. The data that algorithms train on as well as the predictions or recommendations they output are predetermined. Semi-supervised learning: This approach to machine learning involves a mix of the two preceding types. Data scientists may feed an algorithm mostly labeled training data, but the model is free to explore the data on its own and develop its own understanding of the data set. Reinforcement learning: Data scientists typically use reinforcement learning to teach a machine to complete a multi-step process for which there are clearly defined rules. Data scientists program an algorithm to complete a task and give it positive or negative cues as it works out how to complete a task. But for the most part, the algorithm decides on its own what steps to take along the way. How does supervised machine learning work? Supervised machine learning requires the data scientist to train the algorithm with both labeled inputs and desired outputs. Supervised learning algorithms are good for the following tasks:

Binary classification: Dividing data into two categories. Multi-class classification: Choosing between more than two types of answers. Regression modeling: Predicting continuous values. Ensembling: Combining the predictions of multiple machine learning models to produce an accurate prediction. How does unsupervised machine learning work? Unsupervised machine learning algorithms do not require data to be labeled. They sift through unlabeled data to look for patterns that can be used to group data points into subsets. Most types of deep learning, including neural networks, are unsupervised algorithms. Unsupervised learning algorithms are good for the following tasks:

Clustering: Splitting the dataset into groups based on similarity. Anomaly detection: Identifying unusual data points in a data set. Association mining: Identifying sets of items in a data set that frequently occur together. Dimensionality reduction: Reducing the number of variables in a data set.

How does semi-supervised learning work? Semi-supervised learning works by data scientists feeding a small amount of labeled training data to an algorithm. From this, the algorithm learns the dimensions of the data set, which it can then apply to new, unlabeled data. The performance of algorithms typically improves when they train on labeled data sets. But labeling data can be time consuming and expensive. Semi-supervised learning strikes a middle ground between the performance of supervised learning and the efficiency of unsupervised learning. Some areas where semi-supervised learning is used include:

Machine translation: Teaching algorithms to translate language based on less than a full dictionary of words. Fraud detection: Identifying cases of fraud when you only have a few positive examples. Labelling data: Algorithms trained on small data sets can learn to apply data labels to larger sets automatically. How does reinforcement learning work? Reinforcement learning works by programming an algorithm with a distinct goal and a prescribed set of rules for accomplishing that goal. Data scientists also program the algorithm to seek positive rewards -- which it receives when it performs an action that is beneficial toward the ultimate goal -- and avoid punishments -- which it receives when it performs an action that gets it farther away from its ultimate goal. Reinforcement learning is often used in areas such as:

Robotics: Robots can learn to perform tasks the physical world using this technique. Video gameplay: Reinforcement learning has been used to teach bots to play a number of video games. Resource management: Given finite resources and a defined goal, reinforcement learning can help enterprises plan out how to allocate resources.

(1)---->What is deep learning in healthcare? Deep learning uses mathematical models that are designed to operate a lot like the human brain. The multiple layers of network and technology allow for computing capability that’s unprecedented, and the ability to sift through vast quantities of data that would previously have been lost, forgotten or missed. These deep learning networks can solve complex problems and tease out strands of insight from reams of data that abound within the healthcare profession. It’s a skillset that hasn’t gone unnoticed by the healthcare profession.

In 2018, IDC predicted that the worldwide market for cognitive and AI systems would reach US77.6 billion by 2022. Towards the end of 2019, IDC predicted it would reach $US97.9 billion by 2023 with a compound annual growth rate (CAGR) of 28.4%. The market is seeing steady growth thanks to the ubiquity of the technology and the potential it has in transforming multiple industries, not just healthcare.

Although, deep learning in healthcare remains a field bursting with possibility and remarkable innovation. Organizations have tapped into the power of the algorithm and the capability of AI and ML to create solutions that are ideally suited to the rigorous demands of the healthcare industry. Deep learning, a subfield of machine learning, advances nursing using neural networks for advanced pattern recognition, which has helped machine learning extend to new sources of data, including speech recognition and image analysismage result for ii. Deep learning applications in healthcare Deep learning provides the healthcare industry with the ability to analyze data at exceptional speeds without compromising on accuracy. It's not machine learning, nor is it AI, it's an elegant blend of both that uses a layered algorithmic architecture to sift through data at an astonishing rate. ***Heart disease, cancer, and brain tumors are diagnosed using medical imaging procedures such as MRI scans, CT scans, and ECG. As a result, deep learning assists doctors in better analyzing diseases and providing the best treatment to patients. (2)----> 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. A market basket is a selected mix of goods and services that tracks the performance of a specific market or segment. A popular market basket is the Consumer Price Index (CPI), which provides an estimate for inflation based on the average change of price paid for a specific basket of goods and services over time. Market baskets work by using a representative selection of goods and services to model the spending patterns of larger market segments. For example, in the broader category of medical care, the market basket may include prices for things like prescription drugs, medical supplies, physicians' services and eyeglasses. Market Basket Analysis is one of the key techniques used by large retailers to uncover associations between items. It works by looking for combinations of items that occur together frequently in transactions. To put it another way, it allows retailers to identify relationships between the items that people buy.

Association Rules are widely used to analyze retail basket or transaction data, and are intended to identify strong rules discovered in transaction data using measures of interestingness, based on the concept of strong rules.

An example of Association Rules

Assume there are 100 customers 10 of them bought milk, 8 bought butter and 6 bought both of them. bought milk => bought butter support = P(Milk & Butter) = 6/100 = 0.06 confidence = support/P(Butter) = 0.06/0.08 = 0.75 lift = confidence/P(Milk) = 0.75/0.10 = 7.5

11. Make a comparison between:-

1. Generalization and abstraction
2. Learning that is guided and unsupervised
3. Regression and classification

Answer:(1). Generalization and abstraction---->Abstraction aims at simplifying the description of an entity while generalization looks for common properties among these abstractions. Generalizations are clearly important and prevalent in many disciplines of study.

I understand that abstraction is about taking something more concrete and making it more abstract. That something may be either a data structure or a procedure. For example:

Data abstraction: A rectangle is an abstraction of a square. It concentrates on the fact a square has two pairs of opposite sides and it ignores the fact that adjacent sides of a square are equal. Procedural abstraction: The higher order function map is an abstraction of a procedure which performs some set of operations on a list of values to produce an entirely new list of values. It concentrates on the fact that the procedure loops through every item of the list in order to produce a new list and ignores the actual operations performed on every item of the list. So my question is this: how is abstraction any different from generalization? I'm looking for answers primarily related to functional programming. However if there are parallels in object-oriented programming then I would like to learn about those as well.

(2) Learning that is guided and unsupervised:-----> Unsupervised learning is the branch of machine learning that is aimed at learning patterns from data without labels. Supervised learning with millions of labels for image classification had driven the modern deep learning revolution in the past few years. Deep neural networks have exceeded human performance at this specific task. But the requirement of such large amounts of labeled data for these models makes one skeptical about the generalization of such intelligence to myriad tasks. While training a neural network to classify images of a cat, one might wonder : Do humans really need hundreds of images to differentiate a cat from an elephant.Or is there some underlying principle that can be rendered useful by a machine in its race to match human intelligence . Unsupervised learning unveils the potential of machine learning algorithms beyond empirical risk minimization and extend them to learning non-trivial representations of the data. At the core of such learning are two distinct principles - model-agnostic representation learning and model-guided inference. The goal of this thesis is to extend the present literature on unsupervised learning through design of novel unsupervised algorithms for clustering, information estimation and model-guided inference. Our journey starts with one of the simplest, yet most fundamental unsupervised learning problem, namely clustering. We explore how modern generative principles such as Generative Adversarial Learning (GAN) can be used to cluster diverse types of data. Even though auto-encoders had been used for clustering in the past, clustering using GANs was unexplored prior to this work. ClusterGAN modifies the vanilla GAN architecture to enable embedding of data in the latent space where cluster structure is revealed. It also improves the generation ability of vanilla GANs by segregating a complex multi-modal distribution into simpler components. Recently, information-theoretic quantities such as mutual information and cross-entropy have been used to regularize unsupervised representation learning and improve clustering. Estimation of such quantities is another fundamental problem in unsupervised learning, which is related to the broader statistical problem of estimating functionals of probability density. We design an estimator, CCMI, for mutual information estimation using classifier likelihood ratio in an unsupervised manner and demonstrate its suitability for high dimensional real-valued information estimation. The conditional variant of this quantity, conditional mutual information (CMI), is also estimated and applied to conditional independence testing. The above approaches to unsupervised learning do not assume any model for the data generation and learn it implicitly from data. However, in many real-world problems, one has domain knowledge about the data-generation process. Utilizing such domain knowledge can help to further reduce data complexity and abandon the need for deep learning models. It also imparts interpretability to the learning process. We apply such learning techniques to a specific phenomenon in genomics, known as segmental duplication. The problem can be formulated as either a (a) low-rank matrix completion or a (b) robust signed community detection based on suitable assumptions on the data. We design algorithms for resolving segmental duplication in genomes under these two formulations. Finally, we explore another application in natural language understanding where unsupervised and supervised approaches blend gracefully. This also illustrates a situation where labeled data could be difficult to obtain and an unsupervised solution may be used.

In [ ]: