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. A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

Two machine learning applications in business world are google recomendation engine, Stock market predictor.

The following list enumerates all the ethical issues that were identified from the case studies are:

a)Cost to innovation

b)Harm to physical integrity.

c)Lack of access to public services.

d)Lack of trust.

e)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

Ans. i) Someone learns any topic and when he/she does something wrong the expert guide him/her the right thing.

ii) Expert doesn't help directly but assist some queries.

iii) Learn by ownself without taking help from others.

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

Ans. There are three types of machine learning:

i) Supervised Learning: Stock Market Prediction.

ii) Unsupervised Learning: Customer Segmentation

iii) Reinforcement Learning: Gaming Industry.

4. Examine the various forms of machine learning.

Ans. 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.

Ans. 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 main characteristics that must be present to identify a learning problem properly are: a) Task b) Performance Measure c) Experience.

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

Ans. 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.

Nowadays, hyperbole about machine learning and artificial intelligence is ubiquitous. This is perhaps rightly so, given the potential for this field is massive. The number of AI consulting agencies has soared in the past few years, and, according to a report from Indeed, the number of jobs related to AI ballooned by 100% between 2015 and 2018.

As of December 2018, Forbes found that 47% of business had at least one AI capability in their business process, and a report by Deloitte projects that a penetration rate of enterprise software with AI built-in, and cloud-based AI development services, will reach an estimated 87 and 83 percent respectively. These numbers are impressive — if you are planning to change careers anytime soon, AI seems like a pretty good bet.

So it all seems great right? Companies are happy and, presumably, consumers are also happy — otherwise, the companies would not be using AI.

It is great, and I am a huge fan of machine learning and AI. However, there are times when using machine learning is just unnecessary, does not make sense, and other times when its implementation can get you into difficulties.

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.

Limitation 4 — Misapplication

Related to the second limitation discussed previously, there is purported to be a “crisis of machine learning in academic research” whereby people blindly use machine learning to try and analyze systems that are either deterministic or stochastic in nature.

For reasons discussed in limitation two, applying machine learning on deterministic systems will succeed, but the algorithm which not be learning the relationship between the two variables, and will not know when it is violating physical laws. We simply gave some inputs and outputs to the system and told it to learn the relationship — like someone translating word for word out of a dictionary, the algorithm will only appear to have a facile grasp of the underlying physics.

For stochastic (random) systems, things are a little less obvious. The crisis of machine learning for random systems manifests itself in two ways:

P-hacking

Scope of the analysis

P-hacking

When one has access to large data, which may have hundreds, thousands, or even millions of variables, it is not too difficult to find a statistically significant result (given that the level of statistical significance needed for most scientific research is p < 0.05). This often leads to spurious correlations being found that are usually obtained by p-hacking (looking through mountains of data until a correlation showing statistically significant results is found). These are not true correlations and are just responding to the noise in the measurements.

This has resulted in individuals ‘fishing’ for statistically significant correlations through large data sets, and masquerading these as true correlations. Sometimes, this is an innocent mistake (in which case the scientist should be better trained), but other times, it is done to increase the number of papers a researcher has published — even in the world of academia, competition is strong and people will do anything to improve their metrics.

Scope of the Analysis

There are inherent differences in the scope of the analysis for machine learning as compared with statistical modeling — statistical modeling is inherently confirmatory, and machine learning is inherently exploratory.

We can consider confirmatory analysis and models to be the kind of thing that someone does in a Ph.D. program or in a research field. Imagine you are working with an advisor and trying to develop a theoretical framework to study some real-world system. This system has a set of pre-defined features that it is influenced by, and, after carefully designing experiments and developing hypotheses you are able to run tests to determine the validity of your hypotheses.

Exploratory, on the other hand, lacks a number of qualities associated with the confirmatory analysis. In fact, in the case of truly massive amounts of data and information, the confirmatory approaches completely break down due to the sheer volume of data. In other words, it simply is not possible to carefully lay out a finite set of testable hypotheses in the presence of hundreds, much less thousands, much less millions of features.

Therefore and, again, broadly speaking, machine learning algorithms and approaches are best suited for exploratory predictive modeling and classification with massive amounts of data and computationally complex features. Some will contend that they can be used on “small” data but why would one do so when classic, multivariate statistical methods are so much more informative?

ML is a field which, in large part, addresses issues derived from information technology, computer science, and so on, these can be both theoretical and applied problems. As such, it is related to fields such as physics, mathematics, probability, and statistics but ML is really a field unto itself, a field which is unencumbered by the concerns raised in the other disciplines. Many of the solutions ML experts and practitioners come up with are painfully mistaken…but they get the job done.

Limitation 5 — Interpretability

Interpretability is one of the primary problems with machine learning. An AI consultancy firm trying to pitch to a firm that only uses traditional statistical methods can be stopped dead if they do not see the model as interpretable. If you cannot convince your client that you understand how the algorithm came to the decision it did, how likely are they to trust you and your expertise?

As bluntly stated in “Business Data Mining — a machine learning perspective”:

“A business manager is more likely to accept the [machine learning method] recommendations if the results are explained in business terms”

These models as such can be rendered powerless unless they can be interpreted, and the process of human interpretation follows rules that go well beyond technical prowess. For this reason, interpretability is a paramount quality that machine learning methods should aim to achieve if they are to be applied in practice.

The blossoming -omics sciences (genomics, proteomics, metabolomics and the like), in particular, have become the main target for machine learning researchers precisely because of their dependence on large and non-trivial databases. However, they suffer from the lack of interpretability of their methods, despite their apparent success.

Summary and Peter Voss’ List

While it is undeniable that AI has opened up a wealth of promising opportunities, it has also led to the emergence of a mindset that can be best described as “AI solutionism”. This is the philosophy that, given enough data, machine learning algorithms can solve all of humanity’s problems.

As I hope I have made clear in this article, there are limitations that, at least for the time being, prevent that from being the case. A neural network can never tell us how to be a good person, and, at least for now, do not understand Newton’s laws of motion or Einstein’s theory of relativity. There are also fundamental limitations grounded in the underlying theory of machine learning, called computational learning theory, which are primarily statistical limitations. We have also discussed issues associated with the scope of the analysis and the dangers of p-hacking, which can lead to spurious conclusions. There are also issues with the interpretability of results, which can negatively impact businesses that are unable to convince clients and investors that their methods are accurate and reliable.

Whilst in this article I have covered very broadly some of the most important limitations of AI, to finish, I will outline a list published in an article by Peter Voss in October 2016, outlining a more comprehensive list on the limitations of AI. Whilst current mainstream techniques can be very powerful in narrow domains, they will typically have some or all of a list of constraints that he sets out and which I’ll quote in full here:

Each narrow application needs to be specially trained

Require large amounts of hand-crafted, structured training data

Learning must generally be supervised: Training data must be tagged

Require lengthy offline/ batch training

Do not learn incrementally or interactively, in real-time

Poor transfer learning ability, reusability of modules, and integration

Systems are opaque, making them very hard to debug

Performance cannot be audited or guaranteed at the ‘long tail’

They encode correlation, not causation or ontological relationships

Do not encode entities or spatial relationships between entities

Only handle very narrow aspects of natural language

Not well suited for high-level, symbolic reasoning or planning

All that being said, machine learning and artificial intelligence will continue to revolutionize industry and will only become more prevalent in the coming years. Whilst I recommend you utilize machine learning and AI to their fullest extent, I also recommend that you remember the limitations of the tools you use — after all, nothing is perfect.

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

Ans. The various methods and technologies for solving machine learning problems are:

i) Regression: Regression is a statistical method used in finance, investing, and other disciplines that attempts to determine the strength and character of the relationship between one dependent variable (usually denoted by Y) and a series of other variables (known as independent variables).

ii) Classification: In machine learning, classification refers to a predictive modeling problem where a class label is predicted for a given example of input data. Examples of classification problems include: Given an example, classify if it is spam or not.

iii) Clustering.

iv) Dimensionality Reduction.

v) Ensemble Methods.

vi) Neural Nets and Deep Learning.

vii) Transfer Learning.

viii) Reinforcement Learning.

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

Ans. Different Types of Supervised Learning

i) Regression: Regression: Regression is a statistical method used in finance, investing, and other disciplines that attempts to determine the strength and character of the relationship between one dependent variable (usually denoted by Y) and a series of other variables (known as independent variables).

ii) Classification: In machine learning, classification refers to a predictive modeling problem where a class label is predicted for a given example of input data. Examples of classification problems include: Given an example, classify if it is spam or not.

iii) Ensemble Models.

iv) Neural Networks.

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

Ans. Supervised learning: Supervised learning is the learning of the model where with input variable ( say, x) and an output variable (say, Y) and an algorithm to map the input to the output.

That is, Y = f(X)

Why supervised learning?

The basic aim is to approximate the mapping function(mentioned above) so well that when there is a new input data (x) then the corresponding output variable can be predicted.

It is called supervised learning because the process of an learning(from the training dataset) can be thought of as a teacher who is supervising the entire learning process. Thus, the “learning algorithm” iteratively makes predictions on the training data and is corrected by the “teacher”, and the learning stops when the algorithm achieves an acceptable level of performance(or the desired accuracy).

Example of Supervised Learning

Suppose there is a basket which is filled with some fresh fruits, the task is to arrange the same type of fruits at one place.

Also, suppose that the fruits are apple, banana, cherry, grape.

Suppose one already knows from their previous work (or experience) that, the shape of each and every fruit present in the basket so, it is easy for them to arrange the same type of fruits in one place.

Here, the previous work is called as training data in Data Mining terminology. So, it learns the things from the training data. This is because it has a response variable which says y that if some fruit has so and so features then it is grape, and similarly for each and every fruit.

This type of information is deciphered from the data that is used to train the model.

This type of learning is called Supervised Learning.

Such problems are listed under classical Classification Tasks.

Unsupervised Learning: Unsupervised learning is where only the input data (say, X) is present and no corresponding output variable is there.

Why Unsupervised Learning?

The main aim of Unsupervised learning is to model the distribution in the data in order to learn more about the data.

It is called so, because there is no correct answer and there is no such teacher(unlike supervised learning). Algorithms are left to their own devises to discover and present the interesting structure in the data.

Example of Unsupervised Learning

Again, Suppose there is a basket and it is filled with some fresh fruits. The task is to arrange the same type of fruits at one place.

This time there is no information about those fruits beforehand, its the first time that the fruits are being seen or discovered

So how to group similar fruits without any prior knowledge about those.

First, any physical characteristic of a particular fruit is selected. Suppose color.

Then the fruits are arranged on the basis of the color. The groups will be something as shown below:

RED COLOR GROUP: apples & cherry fruits.

GREEN COLOR GROUP: bananas & grapes.

So now, take another physical character say, size, so now the groups will be something like this.

RED COLOR AND BIG SIZE: apple.

RED COLOR AND SMALL SIZE: cherry fruits.

GREEN COLOR AND BIG SIZE: bananas.

GREEN COLOR AND SMALL SIZE: grapes.

The job is done!

Here, there is no need to know or learn anything beforehand. That means, no train data and no response variable. This type of learning is known as Unsupervised Learning.

10. Describe the machine learning process in depth.

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

i.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)

Ans. i) MATLAB is a high-level language and interactive environment for numerical computation, visualization, and programming. ... It is a high-level programming language that can communicate with its cousins, e.g., FORTRAN and C. You can produce sound and animate graphics. Millions of engineers and scientists worldwide use MATLAB for a range of applications, in industry and academia, including deep learning and machine learning, signal processing and communications, image and video processing, control systems, test and measurement, computational finance, and computational biology. 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. In contrast, Python is a general-purpose programming language.

ii) 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.

iii) 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.

iv) Linear regression quantifies the relationship between one or more predictor variable(s) and one outcome variable. For example, it can be used to quantify the relative impacts of age, gender, and diet (the predictor variables) on height (the outcome variable).

11. Make a comparison between:-

1. Generalization and abstraction

2. Learning that is guided and unsupervised

3. Regression and classification

Ans. i) Abstraction is finding common between uncommon and generalisation is transferring that common to new similar concept. Abstraction and generalisation are the tool for formation of concepts and categoration in mind.Better at generalisation and abstraction indicate superior thinking and intelligence.

ii) It is easy to get confused when comparing supervised vs unsupervised learning since both of them are essentially two techniques of machine learning. In other words, supervised and unsupervised learning describes two different ways in which machine learning algorithms can learn from data and make predictions. The most fundamental difference between them is that supervised learning algorithms already know the output, while unsupervised algorithms don’t. Let’s dig a little deeper to understand the differences and core functionalities of these two.

Supervised vs Unsupervised Learning: What is the difference?

Before we go into the differences between supervised and unsupervised learning, let’s discuss what these terms actually mean:

Supervised learning

In supervised machine learning, you train the machine with the help of labeled data in order to predict outcomes of unforeseen data. By analyzing both desired inputs and outputs, the algorithm has to determine the right method to arrive at those particular outputs.

While we already know the answer to the problem, the algorithm’s responsibility is to recognize the data patterns involved and make accurate predictions about unavailable, unseen, and future data based on those patterns. In this case, the programmer is responsible for correcting the algorithm in order to achieve a high accuracy level.

Supervised learning algorithms can be divided into two main categories:

a. Classification

The algorithm is tasked with determining which category the given data belongs to, based on the previous values or data. If the algorithm tries to label the input data into two distinct categories, then it is called binary classification. In case there are more than two categories, then it is referred to as multiclass classification.

An example of classification in supervised learning is determining whether a customer is likely to default in paying their loan or not. Another example is email spam detection where the algorithm has to determine whether the email is spam or not.

b. Regression:

In regression unsupervised learning, the algorithm needs to determine a real or continuous output, like age or weight of a person. It can also be used to predict future stock prices of a company. An example of the regression algorithm is using an equipment’s performance history as input to determine when the next malfunction will occur and schedule maintenance accordingly.

Unsupervised learning

In unsupervised learning, the machine does not need any supervision or training of any kind. The algorithm is responsible for learning on its own by determining and adapting according to the characteristics of the input data. It also uses unlabelled data to detect patterns, identify information structure, and discover valuable insights on its own. While it allows you to perform more complex processes, as compared to supervised learning, it is not as accurate as its counterpart.

The main goal of unsupervised learning is to analyze and identify the innate structure of the dataset.

The two main categories of unsupervised machine learning include:

a. Clustering

Clustering unsupervised algorithms are mainly used to categorize input data into different clusters or groups based on the pattern of the data. Since there are no previously known groups, the algorithm has to first segment data according to the similarities and dissimilarities and then divide the data into different categories. For instance, clustering can be used in manufacturing to detect any anomalies in production equipment and find the root cause behind the malfunctions.

b. Association

These algorithms are mainly used to discover relationships in the distribution of the input data. Association algorithms can be used for determining consumer behavior and target users accordingly for maximum conversions.

Supervised vs Unsupervised Learning: Key Differences

Let’s understand what are the key differences between supervised and unsupervised learning.

Supervisvised vs Unsupervised Learning: Data available

Supervised learning: Both input and output data is available

Unsupervised learning: Only unlabeled input data is available

Supervised vs Unsupervised Learning: Goal

Supervised learning: The main goal is to understand the relationship between input and output data, and predict future data accordingly.

Unsupervised learning: The main goal is to identify the underlying structure and hidden pattern present in the input data.

Supervised vs Unsupervised Learning: Feedback

Supervised learning: It takes direct feedback from the programmer to check if the predictions are correct or not.

Unsupervised learning: It does not take any feedback.

Supervised vs Unsupervised Learning: Complexity and accuracy

Supervised learning: While it is comparatively less complex, it provides a higher accuracy rate:

Unsupervised learning: It more complex than supervised learning and the accuracy levels are also relatively less

Supervised vs Unsupervised Learning: Use cases

Supervised learning: It is often used for speech recognition, image recognition, financial analysis, forecasting, and training neural networks.

Unsupervised learning: It is mainly used to pre-process data or to pre-train supervised learning algorithms.

Adopting, learning, and executing machine learning starts by understanding the key differences between supervised vs unsupervised learning.

iii) Classification predictive modeling problems are different from regression predictive modeling problems.

Classification is the task of predicting a discrete class label.

Regression is the task of predicting a continuous quantity.

There is some overlap between the algorithms for classification and regression; for example:

A classification algorithm may predict a continuous value, but the continuous value is in the form of a probability for a class label.

A regression algorithm may predict a discrete value, but the discrete value in the form of an integer quantity.

Some algorithms can be used for both classification and regression with small modifications, such as decision trees and artificial neural networks. Some algorithms cannot, or cannot easily be used for both problem types, such as linear regression for regression predictive modeling and logistic regression for classification predictive modeling.

Importantly, the way that we evaluate classification and regression predictions varies and does not overlap, for example:

Classification predictions can be evaluated using accuracy, whereas regression predictions cannot.

Regression predictions can be evaluated using root mean squared error, whereas classification predictions cannot.