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 the process by which we enable machines to learn from data and it’s past experiences based on the patterns it observes in the data without programming it explicitly.

The 2 machine learning applications in the real world are:

1. Image recognition: here a machine can identify various objects in an image, people in the particular or assigning usernames to photos on social media and etc.
2. Medical diagnosis: Pathology and oncology utilize machine learning to recognize cancerous tissues.

2. Describe the process of human learning:

i. Under the supervision of experts

Answer: Process of learning of humans under the supervision of experts is when we start studying in college under a particular stream of our choosing and we have no prior idea about the same. This is how learning under supervision of experts work.

ii. With the assistance of experts in an indirect manner

Answer: Suppose start our career in any field after completing our college degree and starting working in a particular field, we are expected to complete a job with a minimal supervision of our immediate reporting manager or person in charge. This way of learning work is called assistance of experts in an indirect manner.

iii. Self-education

Answer: Here as per the name itself suggests, learning starts from complete scratch and we have to learn things on our own by studying things and performing them to observe.

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

Answer: Various types of machine learning are:

1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning.

4. Examine the various forms of machine learning.

Answer: Following are the different forms of machine learning present today:

1. Chatbots present on various websites.
2. Recommendor systems used in Netflix,Amazon Prime,Youtube for user interface and interaction.

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 well posed learning problem for machines is when a machine can perform some task after gaining some experience having some relevant performance measure to check if the task done by machine is correct or not as per our acceptance. Any learning task can be segregated in to well posed learning problem if it has 3 traits:

1. Task
2. Performance measure
3. Experience

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

Answer: At the moment according to me , machine learning cannot solve all the problems. It is due to following reasons:

1. Reasoning Power: Mostly oriented towards used case and donot work in general.
2. Contextual limitation: when we perform NLP tasks, ML models do take in speeches and text but they donot take in is context.
3. Scalability: ML models need to be retrained and changed constantly even after deployment as the data fed to it will be constantly changing and it may result in bad performance of the model . It requires human intervention and if nothing is done, it will definitely produce bad results.
4. Internal working of Deep learning: No one knows actually how are the nodes and featurization part of DL actually work. It is yet to resolved and it is risky to implement something which we are not completely aware of.

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

Answer: Various approaches used in solving machine learning problems is collection of relevant and good data, it’s cleaning, performing EDA on the data to understand the underlying pattern and do feature engineering accordingly, design ML models based on use case and measure it’s accuracy on train and unknown new data known as test data and then deploy it accordingly on any cloud platform like AWS,GCP,Azure, Heroku depending on the need.

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

Answer: Various forms of supervised learning are:

1. Logistic regression: Classification tasks like if a student will pass an exam or not.
2. Linear regression: Prediction of house price or people’s income
3. Naïve bayes : Spam or ham filter
4. K Means clustering: Charactering people based on their watch history on OTT platforms.
5. Support Vector machines: Classification tasks like if a student will pass an exam or not.
6. Recommendor system: Suggestions given to users based on their watch history for almost same content on Youtube,Netflix.

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

Answer: Following are the differences between supervised and unsupervised ML:

Supervised:

Algorithms are trained using labelled data , computational complexity is not that much and most of the time, accuracy is good.

Classification tasks are mostly done right by algorithms like Logistic regression, SVMs etc.

Unsupervised:

Algorithms are trained on data which are not labelled, high computation required and accuracy is not that great too at times.

Clustering tasks done by K means clustering take a lot of time for grouping one or multiple points in a suitable group.

10. Describe the machine learning process in depth.

Answer:

1. First major challenge in ML process path is collection of valid and good quality data, it also maybe the case that we have very small quantity of data due to which training of the model will show overfitting.
2. After the sufficient data collection is done, we go for EDA step to analyse the data distribution amongst the various features in dataset. This is mostly done with the help of numpy , seaborn, plotly libraries and after observing the patterns we apply feature engineering techniques accordingly.
3. Feature engineering is done so that we can make our data distribution of different features as normal distribution if they are skewed or they show any different distribution by any chance as it is important for us to have a normal/Gaussian distribution of data for our models to work correctly.
4. Once we are till previous step, we go for different ML algorithms accordingly based on the use and based on performance metric to check.
5. After modelling we check our model performance on test/unknown data and if performs sufficiently well, we go for deployment on Cloud services where we generate an API for accessing the model via frontend and get the correct results as per model’s prediction.

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)

Answer:

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

MATLAB’s function name and signatures are very trivial and memorable which helps in easy recalling and implementation. The desktop environment is fairly tuned to accommodate to user’s requirements and helps you explore more in a very less amount of time.

1. Linear regression:

Linear regression is one of the simplest machine learning algorithms which find the best fit line between values of the real dependent features and values predicted by our machine learning model so that the error is minimized completely near to 0 and the numerical value predicted by model will be an accurate one.

11. Make a comparison between:-

1. Generalization and abstraction

Answer: While abstraction reduces complexity by hiding irrelevant detail, generalization reduces complexity by replacing multiple entities which perform similar functions with a single construct.

2. Learning that is guided and unsupervised

Answer: Explained above in question 2 , point number 2.

1. Regression and classification

Answer: Regression is the ML operation for prediction for continuous numerical value whereas classification is process of segregating one particular user into one class of user from the available list of classes.

In ML , regression can be mostly be performed by Linear regression but classification can be done via logistic regression, SVMs, decision trees, random forests etc.