**Machine Learning Questions**

1. What is Data?
2. The set of observations and information.
3. What is Data
4. Data Science is a process of extracting knowledge and insights the data by using scientific methods. It is a multidisciplinary approach that combines principles and practices from the fields of mathematics, statistics, artificial intelligence, and computer engineering to analyze large amounts of data.
5. What is Data Processing?
6. Data Processing occurs when the data is collected and translated into useable information. This is the process of cleaning the data, so that the data gives good result.
7. What Libraries Utilized to work with Data preprocessing?
8. Numpy and Pandas are the libraries to work with Data preprocessing.
9. What are roles of Data Scientist?
10. 1. Collecting the data.

2. Analyzing the current data.

3. Selecting proper model for the data.

4. Checking the accuracy of the model.

1. What is Machine Learning?
2. The capability of AI systems to learn by extracting patterns form data is known as Machine Learning. ML is about creating and implementing algorithms that let machine receive data and use these data to make predictions, analysis pattern and giving recommendations.
3. Why Machine Learning?
4. Machine Learning is to handle the unlimited amount of data the ML helps in analyze the data easily and quickly.
5. Purpose of Machine Learning?
6. The main purpose of Machine Learning is to predict the future based on the past data.
7. Machine Learning current examples?
8. 1. Suggestion in YouTube based on your viewed videos.

2. Detecting future salary based on the past salaries.

1. How many types of Machine Learning models?
2. There are 3 types.
3. 1. Supervised Learning
4. 2. Unsupervised Learning
5. 3. Reinforcement Learning
6. On what based machine Learning model is selected, brief?
7. The model is selected based on the data and the expected output of the data.
8. Describe Difference between Supervised, Unsupervised and Reinforcement Learning.

|  |  |  |  |
| --- | --- | --- | --- |
| S.No | Supervised Learning | Unsupervised Learning | Reinforcement Learning |
| 1. | Learns by using Labelled Data. | Trained using unlabelled data without any guidance. | Works on interacting with the environment. |
| 2. | Calculate outcomes. | Discover underlying patterns. | Learn a series of action. |

1. Difference between Machine Learning and Deep Learning?
2. Machine Learning and deep Learning are both types of AI. In short, Machine Learning is a part of AI that can predict the future. Deep Learning is a subset of machine Learning that uses Artificial Neural Networks to mimic the learning process of the human brain.
3. What is Scikit Learn?
4. Scikit Learn is a machine learning library for python that provides simple and efficient tools for data analysis and modeling, including various algorithms for classification, regression, clustering, and more. It is widely used for its user-friendly interface and compatibility with other scientific computing libraries.
5. List out Data Preprocessing Steps?
6. Steps for Data preprocessing are:
7. Identify your source.
8. Remove Duplicate entries
9. Handle missing value
10. Standardize Formats
11. Validate & verify
12. In Dataset describe Independent Variables and Dependent Variables?

|  |  |  |
| --- | --- | --- |
| S.No | Independent Variables | Dependent Variables |
| 1. | The independent variables does not depend on other variables. | The Dependent variables depends on independent variables. |
| 2. | The universal declaration is ‘ X ’. | The universal declaration is ‘ y ’. |

1. What is the default ratio for Train/test set?
2. The default percentage taken for train is 70% and for test it is 30%.
3. Brief steps involved in Machine Learning Model Approach?
4. Machine Learning Pipeline:

Step-1:- First we need to understand the data.

Step-2:- We apply data clearing process.

Step-3:- Split the data into train data as 70% (Default) but configurable, and test into 30% (Default) configurable.

Step-4:- Train data (X\_train, y\_train)

Step-5:- Training the model

Once the model gets trained, model will gain knowledge, this knowledge is called as model\_fit.

Step-6:- We use test data

• X\_test data used to predict y\_predict data of model\_fit.

• Now compare (y\_predict with y\_test), we get performance of machine.

1. What is Over\_fit?
2. When we train model with huge amount of data, then it is considered as Over\_fit.
3. What is Under\_fit?
4. When we train model with less amount of data, then it is considered as under\_fit.