ASSIGNMENT 2

NAME: FALAQ KHAN

BATCH:Machine Learning and AI
Batch A3

1. **Data science** is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data and apply knowledge and actionable insights from data across a broad range of application domains. Data science is related to data mining, machine learning and big data.

   The field of data science is growing as technology advances and big data collection and analysis techniques become more sophisticated.

Uses: Data science can identify patterns, permitting the making of inferences and predictions, from seemingly unstructured or unrelated data. Tech companies that collect user data can use techniques to turn what's collected into sources of useful or profitable information.

- 2 STEPS TO CREATE DATA SCIENCE PROJECT:
- a. Choose a dataset
- b. Choose an IDE
- c. List down the activities clearly
- **Data ingestion –** It is a process of reading the data into a data-frame
- **Data cleaning –** It is the process of identifying and removing the anomalies in the dataset
- **Data Transformation –** It involves changing the data type of the columns, creating derived columns or removing duplicate data to name a few
- **Exploratory data analysis –** Perform multi-variate analysis on the datasets to find hidden insights and patterns in them
- **Model building –** Try and test all possible models on the dataset before you choose the right one based on business/technical constraints. During this phase, you can try some bagging or boosting techniques as well
- **Model evaluation –** In this phase, we test if our model is good enough to get a predicted result. We measure accuracy, specificity, sensitivity, or adjusted R-square depending on the model that we have used
- d. Take up the tasks one by one
- e. Prepare a summary
- f. Share it on open source platforms

3. IDENTITY OPERATORS:

is : returns true if both the variables are the same objects

is not: returns true if both the variables are not the same objects.

LOGICAL OPERATORS:

and :returns true if both statements are true

or : returns true of one of the statement is true

not: reverse the result, returns false if the result is true

COMPARSION OPERATORS:

< : checks if the left value is lesser than that on the right

> : checks whether the left value is greater than the one on the right.

<= : returns True only if the value on the left is either less than or equal to that on the right of the operator.

>= : returns True only if the value on the left is greater than or equal to that on the right.

== : returns True if the values on either side of the operator are equal.

!= : True if the values on either side of the operator are unequal.