7 PHASES IN DATA SCIENCE LIFE CYCLE

Data Science is a mixture of various tools, algorithms and machine learning principles with the aim of identifying hidden patterns or insights from our data which helps us to make improvised decisions.

1. PROBLEM DEFINITION:

One of the most important aspects of any data science project.
Any data science and machine learning project typically begins with the problem definition.

2. DATA COLLECTION:

Once the problem is clearly defined now we need to collect the data needed for our data science project.
We need to collect as much as relevant data as possible.

3. DATA PREPARATION:

The data collected is rarely available in the desired form.
We need to preprocess the acquired data to make it into suitable form.

4. DATA ANALYSIS:

Explore the data to find some interesting patterns and relationships.There are several machine learning algorithms that help us model the data.

5. FEATURE ENGINEERING:

This phase requires expertise and domain knowledge as it involves transforming raw data into more informative features.

6. MODEL BUILDING:

There are several machine learning algorithms that help us model the data.
Experiment with different machine learning algorithms to find out which works best for our data.

7. COMMUNICATE THE RESULTS:

Findings are of little value if it can be interpreted only by the one who made the analysis.
Communicate the results visually using the most suitable visualization techniques for effective interpretation by stakeholders.