

6. Data Science – Machine Learning – Life Cycle

Contents

1. Machine learning life cycle	2
1.1. Data Collection.....	3
1.2. Data Preparation.....	3
1.3. Data Wrangling	3
1.4. Train the model.....	3
1.5. Test the model	4
1.6. Model deployment	4

6. Data Science – Machine Learning – Life Cycle

1. Machine learning life cycle

- ✓ There are mainly six steps involved here,
 - Data collection
 - Data preparation
 - Data Wrangling
 - Train the model
 - Test the model
 - Model deployment

1.1. Data Collection

- ✓ Data Gathering is the first step in machine learning life cycle.
- ✓ In this step, we need to identify the different data sources.
- ✓ Data can be collected from different sources such as files, database, web & etc.

1.2. Data Preparation

- ✓ During data preparation we should understand about the data format, quality of data & etc.
- ✓ A better understanding of data leads to an effective outcome.
- ✓ We will find correlations, general trends, and outliers.

1.3. Data Wrangling

- ✓ Data wrangling is the process of cleaning and converting raw data into a useable format.
- ✓ It is the process of,
 - Cleaning the data
 - Missing Values
 - Duplicate data
 - Invalid data
 - Selecting the variable to use,
- ✓ It is not necessary that data we have collected is always of our use as some of the data may not be useful.

1.4. Train the model

- ✓ During training, the model can learn different patterns and rules.
- ✓ We need to use train dataset to train the model.

1.5. Test the model

- ✓ Once model trained then testing is required to check the accuracy.
- ✓ We need to use test dataset to test the model.

1.6. Model deployment

- ✓ Once model trained and tested, if model is producing good results then we can deploy the model in the real time.