

(DA-Lab)
MS. SHWETA TIWARI
April 19, 2022

DA-Lab

DATA ANALYTICS-LAB

By SHWETA TIWARI

Data Preprocessing in R

Topics: Machine Learning

Data preprocessing is the initial phase of Machine Learning where data is prepared for machine learning models. This part is crucial and needs to be performed properly and systematically. If not, we will end up building models that are not accurate for their purpose.

We will learn all the steps that are involved in the data preprocessing stage.

Prerequisites

To follow this file along - the reader should have the following:

1. RStudio installed on your computer.
2. Package packages ('caTools') installed.

Steps in data preprocessing

- Steps in Data Preprocessing
 - Step 1: Importing the Dataset
 - Step 2: Handling the Missing Data
- Step 3: Encoding Categorical Data.
 - Output
- Step 4: Splitting the Dataset into the Training and Test sets
 - Training set
 - Test set
- Step 5: Feature Scaling
 - training set
 - test set