**Task: Preprocess the data to make the data ready for Machine Learning model creation**

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model.

When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So for this, we use data preprocessing task.

A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data preprocessing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

It involves below steps:

* Getting the dataset
* Importing libraries
* Importing datasets
* Finding Missing Data
* Encoding Categorical Data
* Splitting dataset into training and test set
* Feature scaling

**Resources:**

1. <https://www.v7labs.com/blog/data-preprocessing-guide>

2. <https://www.geeksforgeeks.org/data-preprocessing-machine-learning-python/>

3. <https://towardsdatascience.com/an-introduction-to-preprocessing-data-for-machine-learning-8325427f07ab>

4. <https://www.section.io/engineering-education/data-preprocessing-python/>

5. <https://www.kaggle.com/code/ajay1216/practical-guide-on-data-preprocessing-in-python>

6. <https://data-flair.training/blogs/python-ml-data-preprocessing/>