IT-641 Deep Learning

Lab 1

1. Introduction

Machine Learning Pipeline

During this Lab Session we shall revise classification and regression tasks using standard machine learning algorithms. Moreover we shall also try and define a machine learning pipeline that shall help us develop more complex deep learning algorithms in the future:

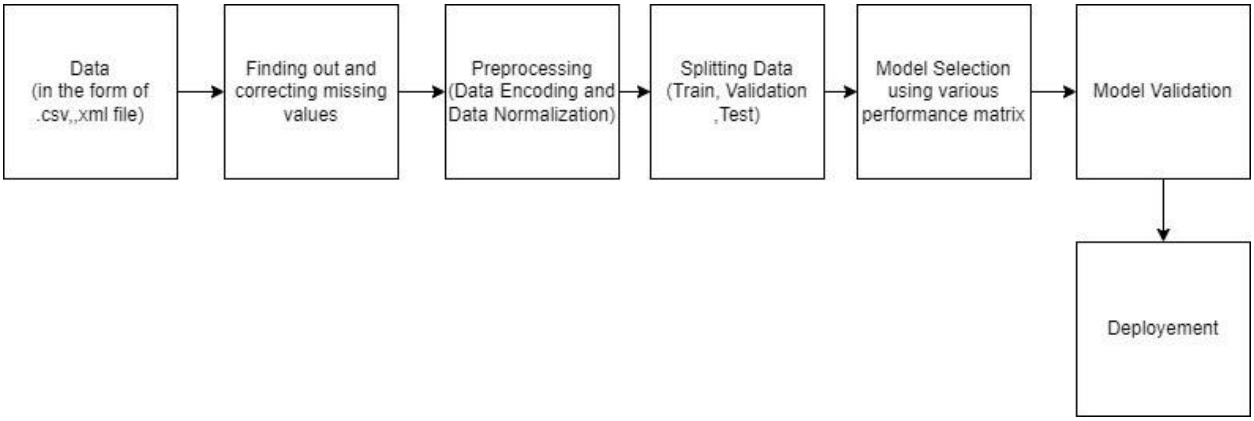

Figure 1(Machine Learning Flowchart)

2. Datasets

1. **User dataset**

   This dataset contains information of users from the company's database. It contains information about UserID, Gender, Age, EstimatedSalary,Purchased.

2. **Pima Indians Diabetes Database**

   This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. The datasets consist of several medical predictor variables and one target variable, Outcome. Predictor variables include the number of pregnancies the patient has had, their BMI, insulin level, age, and so on.

3. **50_Startups**

   This dataset has data collected from New York, California and Florida about 50 business Startups. The variables used in the dataset are Profit, R&D spending, Administration Spending, and Marketing Spending.

3. Tasks

For each of the above given datasets

1. Load Data and check if the data has missing value
2. Identify which features need to be encoded and encode them
3. Identify which features to normalize and normalize them
4. Identify whether the given task is of classification of regression
5. Split the data into train set (75%) validation set (10%) and test set (15%)
6. Fit the data into 2 models of your choice

4. Reference Code

https://colab.research.google.com/drive/1lEqUTriS66KBbn1V648NooaqIsTB3MvE?usp=sharing

5.Submission Details

1. The assignment can be submitted as IPython Notebook as well as PDF
2. The submitted file must be of format:

   "STUDENTID_FIRSTNAME_LASTNAME.ipynb" or
   "STUDENTID_FIRSTNAME_LASTNAME.pdf"