Machine Learning, 2022
Lab 1: Exploratory Data Analysis (EDA)

Objective:

Import the dataset and perform EDA such as number of data samples, number of features, number of classes, number of data samples per class, removing missing values, conversion to numbers, explore dimensionality, type the mean or average value, and using seaborn library to plot different graphs. Consider one of the datasets given below.

Public datasets for EDA

1.National Centers for Environmental Information**:** Dig into the world's largest provider of weather and climate data.

2.World Happiness Report 2021**:** What makes the world's happiest countries so happy?

3.NASA**:** If you're interested in space and earth science, see what you can find among the tens of thousands of public datasets made available by NASA.

4.US Census**:** Learn more about the people and economy of the United States with the latest census data from 2020.

5.FBI Crime Data Explorer (CDE)**:** Explore crime data collected by more than 18,000 law enforcement agencies.

6.World Health Organization COVID-19 Dashboard**:** Track the latest coronavirus numbers by country or WHO region.

7.Google Books Ngram**:** Download the raw data from the Google Books Ngram to explore phrase trends in books published from 1960 to 2015

Note:

Use Co lab to perform this experiment.

Deliverable:

You need to write document showing code that fulfills the objective requirement. You must write resulting inference for all intermediate steps, must create GitHub account and store your experiment write up there prior to next experiment.