SKILLCONNECT
Innovating Engineering!

Week 1: Advanced Challenge

Agenda: Developing end-to-end projects
Last Date for submission: Sunday, 12th April, 2020

Topics Covered: Model training, Algorithms from scratch, Flask server

Problem Statement

Develop a Google Colab notebook with well documented code for the following topics. Each topic must be shown in a separate section using “text” cells in the colab.

1. Training following Classification Algorithms
    a. Linear Classifiers: Logistic Regression, Naive Bayes Classifier
    b. Nearest Neighbor
    c. Support Vector Machines
    d. Decision Trees
    e. Boosted Trees
    f. Random Forest
    g. Neural Networks

2. KMeans Clustering Algorithm
3. Linear Regression

** At least two models should be written from scratch in python. Rest can be imported using the Sklearn library, better if you try coding them by yourselves.

Dataset: (Better if Choose your own dataset, we’ll get a variety of different projects)

1. Use Iris dataset : http://archive.ics.uci.edu/ml/machine-learning-databases/iris/
1. Use NSL KDD dataset: https://github.com/defcom17/NSL_KDD
2. General dataset link: http://archive.ics.uci.edu/ml/datasets.php

*** It would be great if you perform training on multiple datasets.*
*** Create separate Colab files for each training Dataset. Make a separate section at top for data downloading and reading into Colab part.*
*** Prefer using the “wget” command from within colab for testers to easily download your dataset.*

SKILLCONNECT
Innovating Engineering!

Results:

The github repository should contain following list of files

1. Trained models
2. Google Colab or Python Code used for developing and training the model
3. Flask server
4. UI for uploading sample input (test data)
5. UI page for showing and comparing results of different models
6. Refer to the README file mentioned in point 4 in the reference section for the expected UI part.

Resource Reference

1. Types of classification algorithms in Machine Learning - Theory regarding Classifiers
2. Linear Regression on Boston Housing Dataset
3. https://github.com/shubham99bisht/Machine_Learning_Tutorials - Sample codes from scratch
4. For flask server and UI pages: https://github.com/shubham99bisht/Intrusion-Detection-using-ML