Machine Learning for Engineering Applications
Homework #3
Fall 2023

Due: **10/6/2023 by 11:59 PM – submit your zip file to CANVAS**

Dataset

- Dataset info and download: https://archive.ics.uci.edu/dataset/480/2+4+ghz+indoor+channel+measurements

Part 1

- Use a ***scikit-learn logical regression model*** to classify all targets
- Use a pipeline, GridSearch, and cross-validation to setup, tune, train, and validate the model. *Thigs to consider: Standardization vs. Normalization, PCA vs. LDA, etc*.
- Print out the best parameters, training & test accuracy values, include the precision, recall, and F1-score to a text file.
- Plot the classification outcome of the **best model**, the training and validation accuracy learning curves of the **best model**, training and test confusion matrices of the **best model**, and the ROC AUC graph of the **best model**, save them as images for submission.

Part 2

- Use a ***scikit-learn perceptron model*** to classify all targets.
- Same requirements from Part 1

Part 3

- Use a ***scikit-learn decision tree model*** to classify all targets.
- Same requirements from Part 1

Part 4

- Use Ensemble Learning using the models from Parts 1, 2, and 3.
- Implement the voting technique for classification
- Print the accuracy, precision, recall and F1-score for the ensembled model
- Plot the classification outcome and save it as an image for submission.

Submit for each part:

- Python file (.py)
- Best classification plot image (.png or .jpg)
- Best accuracy learning curves image (.png or .jpg) [Not for Part 4]
- ROC AUC plot image (.png or .jpg) [Not for Part 4]
- Text file with all the metrics and hyperparameters values (.txt) [Results matter now]
- The SLURM/Batch file used for the LEAP cluster execution.
- The .OUT file from the job submitted

Rubric per Part and Overall:

- Header in the code (5pts)
- Comments in the code (10pts)
- Running code without errors (20pts)
- Executes requirements & produces required output information (65pts)

Template:

```
#*****************************************************************************
# Damian Valles
# ML - HW#1
# Filename:  hw1-perceptron.py
# Due:  Sept. 6, 2023
#
# Objective:
# To demonstrate header and comment expectations for assignments for the semester.
#*****************************************************************************
#Importing all required libraries
from sklearn import datasets
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import Perceptron

#Downloading the dataset, creating dataframe
iris = datasets.load_iris()

#Separating data & target information
X = iris.data[:, [2, 3]]
y = iris.target

#Data split for training and testing
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,
random_state=1, stratify=y)

#Scaling training data
sc = StandardScaler()
sc.fit(X_train)
X_train_std = sc.transform(X_train)

#Creating perceptron with hyperparameters
ppn = Perceptron(max_iter=40, eta0=0.01, shuffle=True)

#This is training the model
ppn.fit(X_train_std, y_train)

#Scaling test data
sc.fit(X_test)
X_test_std = sc.transform(X_test)

#Testing the model data
y_pred = ppn.predict(X_test_std)
```