**Department of Electrical Engineering**

**Faculty Member:** LE Munadi Sial **Date:** 22-Nov-2023

**Semester:** 7th **Group:**

# CS471 Machine Learning

**Lab 10: Decision Trees, K-NNs and Support Vector Machines**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **PLO4 - CLO4** | **PLO4 -CLO4** | **PLO5 -CLO5** | **PLO8 -CLO6** | **PLO9 -CLO7** |
| **Name** | **Reg. No** | **Viva /Quiz / Lab Performance** | **Analysis of data in Lab Report** | **Modern Tool Usage** | **Ethics** | **Individual and Team Work** |
|  |  | **5 Marks** | **5 Marks** | **5 Marks** | **5 Marks** | **5 Marks** |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

## Introduction

This laboratory exercise will focus on the Scikit Learn (or SKLearn) library for machine learning implementations in python. Scikit Learn contains many useful functions for fitting models using various machine learning techniques such as linear regression, logistic regression, decision trees, support vector machines, k-means clustering, anomaly detection and more.

## Objectives

The following are the main objectives of this lab:

* Implement Decision Trees using Scikit learn
* Implement K-Nearest Neighbors using Scikit learn
* Implement K-Nearest Neighbors from scratch
* Implement Support Vector Machines using Scikit learn

## Lab Conduct

* Respect faculty and peers through speech and actions
* The lab faculty will be available to assist the students. In case some aspect of the lab experiment is not understood, the students are advised to seek help from the faculty.
* In the tasks, there are commented lines such as #YOUR CODE STARTS HERE# where you have to provide the code. You must put the code/screenshot/plot between the #START and #END parts of these commented lines. Do NOT remove the commented lines.
* Use the tab key to provide the indentation in python.
* When you provide the code in the report, keep the font size at 12

**Theory**

Scikit Learn is a python library that contains a wide arsenal of functions pertaining to machine learning. It also contains its own datasets for trying out the machine learning algorithms. Scikit learns API interface can be divided into three types: estimator, predictor and transformer. The estimators are used to fit the model in accordance with some algorithm. The predictors use the fitted model to make prediction on test features. The transformers are used for the conversion of data.

A brief summary of the relevant keywords and functions in python is provided below:

**print()** output text on console

**input()** get input from user on console

**range()**  create a sequence of numbers

**len()** gives the number of characters in a string

**if** contains code that executes depending on a logical condition

**else** connects with **if** and **elif**, executes when conditions are not met

**elif** equivalent to **else if**

**while** loops code as long as a condition is true

**for** loops code through a sequence of items in an iterable object

**break** exit loop immediately

**continue** jump to the next iteration of the loop

**def** used to define a function

**pd.read\_csv** import csv file as a dataframe

**df.to\_csv** export dataframe as a csv file

**Lab Task 1 – Decision Trees \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Download a dataset containing at least 4 feature columns and a label column containing discrete data. Use functions from Sci-kit learn to train a model using decision trees. Try the following feature combinations:

* 2 features combination
* 3 features combination
* 4 features combination

Display the trained trees and also use the trees to make the predictions for all of the 3 combinations. Provide the code and all of the relevant screenshots of your work.

***### TASK 1 CODE STARTS HERE ###***

*### TASK 1 CODE ENDS HERE ###*

***### TASK 1 SCREENSHOT STARTS HERE ###***

*### TASK 1 SCREENSHOT ENDS HERE ###*

**Lab Task 2 – K-Nearest Neighbors \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Download a dataset containing at least 3 feature columns and a label column containing discrete data. Write a python program from scratch that uses KNNs to predict the class of an example for the following cases:

* 2 features combination
* 3 features combination

Make a scatter plot showing the predictions in both cases.

***### TASK 2 CODE STARTS HERE ###***

*### TASK 2 CODE ENDS HERE ###*

***### TASK 2 SCREENSHOT STARTS HERE ###***

*### TASK 2 SCREENSHOT ENDS HERE ###*

**Lab Task 3 – K-Nearest Neighbors Part 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Download a dataset containing at least 4 feature columns and a label column containing discrete data. Use functions from Sci-kit learn to predict the class of an example using KNNs for the following cases:

* 2 features combination
* 3 features combination
* 4 features combination

Make a scatter plot showing the predictions in both cases.

***### TASK 3 CODE STARTS HERE ###***

*### TASK 3 CODE ENDS HERE ###*

***### TASK 3 SCREENSHOT STARTS HERE ###***

*### TASK 3 SCREENSHOT ENDS HERE ###*

**Lab Task 4 – Support Vector Machines \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Download a dataset containing at least 4 feature columns and a label column containing discrete data. Use functions from Sci-kit learn to predict the class of an example using Support Vector Machines for the following cases:

* 2 features combination
* 3 features combination

Make a scatter plot showing the predictions in both cases.

***### TASK 4 CODE STARTS HERE ###***

*### TASK 4 CODE ENDS HERE ###*

***### TASK 4 SCREENSHOT STARTS HERE ###***

*### TASK 4 SCREENSHOT ENDS HERE ###*