**Department of Electrical Engineering**

**Faculty Member:** LE Munadi Sial **Date:** 20-Sep-2023

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

# CS471 Machine Learing

**Lab 2: Data Structures, NumPy Arrays and SciPy Functions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **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 is focused on the introduction of data structures native in python particularly Lists and Dictionaries which are very commonly used. This lab also introduces the importing of modules that are used for machine learning tasks. In this lab, the NumPy and SciPy libraries will be introduced which are very important to the field of Machine Learning.

## Objectives

The following are the main objectives of this lab:

* Implement data structures such as lists and dictionaries in python
* Create, alter and loop through lists
* Use slicing to access range of items in a list
* Utilize various list methods such as append, insert, extend, remove, pop etc
* Create and implement a dictionary
* Create Numpy arrays and perform matrix operations and broadcasting
* Use Scipy for minimization, scarce matrices and iterpolation

## 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 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
* Upon completing the lab, you must delete the manual from the lab computer

**Theory**

Data structures are an important part of python. The 4 main data structures are lists, tuples, sets and dictionaries. Lists and dictionaries are the most commonly used for machine learning tasks. The *import* keyword is used to load modules and libraries. In machine learning, there are many popular libraries. The most basic of these is the NumPy library which provides an optimized array implementation for very fast matrix computations necessary for machine learning. The SciPy library provides numerous functions for scientific computations.

A brief summary of the list functions in python is provided below:

**append(I)** append item I to the end of the list

**insert(i, I)** insert item I at i position of the list

**extend(L)** extend/concatenate a second list L

**remove(I)** remove a specified item I from a list

**pop(i)** remove item at specific index i in the list

**count(I)** return total number of a specific item I from a list

**index(I)** return index of first occurrence of a specific item I

**reverse** reverse the items of the list

**Lab Task 1 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [1]**

Create a 1-D list containing the characters of the name of any one person in your group. Loop through the list and display each character on a new line.

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

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

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

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

**Lab Task 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [1]**

Write a program that repeatedly prompts the user for input. The user will keep entering numbers which are added to a list. Each time a number is added to the list, it must be placed in such a way that the list items are always in ascending order. Each time a number is input, the list is to be printed showing the newly added number. This continues until the word “done” is input at which point the prompts will stop. The final list is then displayed. Do NOT use any inbuilt sorting function for this task.

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

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

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

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

**Lab Task 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [1]**

Create a list with the sequence 1, 2, 3… 20. Then using the slice operation (:) on this list, print the following sub-lists:

5, 6, 7… 20

1, 2, 3… 12

7, 8, 9 … 16

4, 5

11, 12, 13, 14

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

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

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

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

**Lab Task 4 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [1]**

In this task, you will make use of dictionaries. Write a program that first prompts the user to input five strings which will be the keys of the dictionary. Then, the program must prompt the user to input the values of the respective keys. When entering the values, the user must be shown the key whose value is being input. Once all values are entered, display the dictionary.

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

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

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

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

**Lab Task 5 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [1]**

Import the NumPy Library. Use the np.array function to define a 4x5 array with elements of your choice. Ensure that the elements are numbers (not strings). Then, perform the following:

* Print the array
* Print element(4,4)
* Print rows 2 and 3 via slicing
* Print the central 3x3 elements in a matrix
* Compute the sum of the matrix elements
* Compute the sum of the matrix elements along axis 0
* Compute the sum of the matrix elements along axis 1
* Compute the mean of the matrix elements
* Compute the standard deviation of the matrix elements

***### TASK 5 CODE STARTS HERE ###***

*### TASK 5 CODE ENDS HERE ###*

***### TASK 5 SCREENSHOTS START HERE ###***

*### TASK 5 SCREENSHOTS END HERE ###*

**Lab Task 6 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [1]**

Use the np.array function to define two matrices of size 3x3. Place numerical elements of your choice in the matrices. Write code to perform the following:

* Print the arrays
* Compute the sum of the matrices
* Compute the difference of the matrices
* Compute the element-wise product of the matrices
* Compute the element-wise division of the matrices
* Compute the matrix multiplication of the matrices

***### TASK 6 CODE STARTS HERE ###***

*### TASK 6 CODE ENDS HERE ###*

***### TASK 6 SCREENSHOTS START HERE ###***

*### TASK 6 SCREENSHOTS END HERE ###*

**Lab Task 7 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [1]**

In this task, you will use various functions of the SciPy library that are commonly used in machine learning. Import the various modules from SciPy Library:

**from scipy.optimize import root**

**from scipy.optimize import minimize**

**from scipy.sparse import csr\_matrix**

**from scipy.interpolate import interp1d**

1. Use the root function to determine the roots of the equation 3x2 + 2x – 10.
2. In machine learning, it is very often required to find the argument that minimizes a complex equation with the given data. Use the minimize function to determine the roots of the equation x2 - 20x + 45.
3. In machine learning, sometimes there are matrices in which most of the elements are zero. In such cases, it is more convenient to store them as sparse matrices which holds information of the non-zero elements. In this task, create a 3x10 sparse matrix (A) with elements of your choice. Ensure that about 2/3 of the elements are zero. Then, print the matrix information using csr\_matrix(A), csr\_matrix(A).data and csr\_matrix(A).count\_nonzero().
4. In this task, you will perform interpolation. Create two lists x and y. The list x contains elements 1,2,3… 10. The list y contains the elements for y = 2x + 1. Use the interp\_func = interp1d(x, y) to get the interpolating function. Then, use interp\_func(val) to get any 3 interpolated values.

***### TASK 7 CODE STARTS HERE ###***

*### TASK 7 CODE ENDS HERE ###*

***### TASK 7 SCREENSHOTS START HERE ###***

*### TASK 7 SCREENSHOTS END HERE ###*