Annexure 41.8

Centre for Machine Learning and Intelligence
Multidisciplinary Course
Machine Learning for Biochemistry, Biotechnology

Semester: III **Hours of Instructions / Week: 2**
25MAIM01 **No. of Credits: 2**

Course Objectives:

1. To understand the fundamentals of Machine Learning concepts
2. To familiarize working with Google Colab environment
3. To learn how to design and develop Biochemistry related applications in Google Colab

Unit I: Fundamental Concepts of Machine Learning **- 6 Hrs**

Machine Learning (ML) - Supervised Vs Unsupervised Learning - Reinforcement Learning - Regression - Classification – (Clustering)*

Unit II: Basics of Google Colaboratory **- 6 Hrs**

Introduction to Google Colab - Advantages of Google Colab - Features of Google Colab - GPU and TPU Facility in Google Colab - Colabpro - Installation of Google Colab - Starting up with New Colab Notebook - Setting the Notebook Name - Adding Cells and Entering Code in Google Colab - Executing Code in Google Colab - Modifying Cell Order and Deleting Cell in Google Colab - Saving and Sharing Notebook in Google Drive - Arithmetic Operations

Unit III: Working with Google Colaboratory **- 6 Hrs**

Loading File System: Uploading Files from your Local File System - Downloading Files to your Local File System - Mounting Google Drive - Loading Image in Google Colab: Accessing Images from Google Drive - Accessing Images from Computer - Loading Imageset in Google Colab- Displaying a Single and Multiple Images in Google Colab - Loading Data in Google Colab:Accessing Data from Google Drive - Accessing Data from Computer –(Loading Data from GoogleColab - Displaying Data using Colab)*

Unit IV: Data Importing, Exporting and Visualization **- 6 Hrs**

Importing and Displaying Data from Kaggle - Importing and Displaying Data from Github-Importing and Extracting Zip Files in Google Colab - Saving and Sharing Notebook in Github -Exporting Code and Dataset to Github and Kaggle - Cloning Git Repository in Google Colab -Charting: Line Plotting - Bar Plotting - Scatter Plotting - Histogram - Pie Chart –(Fill_between and alphas - Sub Plotting)* - 3D Graph - Saving Charts in Different Formats.

Unit – V: Forecasting Oxidative Coupling of Methane using Machine Learning - 6 Hrs

Package Installation: Numpy - Pandas - Matplotlib - scikit-learn - TensorFlow. Data Accessing: Reading excel file using read_excel(). Data Preprocessing: Sorting and concatenating data using Pivot(). Unsupervised Learning: K-Mean Clustering - PCA. Supervised Learning: Support Vector Machine (SVM) –(Random Forests)*.

*** Indicates Self - Study Component**

Total Hours: 30

Reference Books:

1. Chris Albon. (2019). "*Machine learning with python cookbook*".2nd edn. USA: O"Reil y Media, Inc.
2. John Paul Mueller, and Luca Massaron. (2019). "*Python for data science for dummies*",2nd edn. New Jersey: John Wiley& Sons, Inc.
3. Palkovits, S. (2020) "*A primer about machine learning in catalysis - a tutorial with code*", ChemCatChem, 12(16), pp. 3995-4008.
4. Google Colab (2020) "*Google Colaboratory*" Available at: https://colab.research.google.com/notebooks (Accessed 10 November 2020)

Course Outcomes

CO1: Enhance the knowledge about machine learning

CO2: Able to apply Machine Learning concepts for Biochemistry related application

CO3: Write programmes in Google Colab

CO4: Experiment Biochemistry applications using Google Colab

CO5: Explain how datasets can be imported and used in Google Colab