		DOMAIN		
CUTM	**MLCU2000**	**Data Science and Machine Learning**	**26**	2+9+15
	CUML2010	ML for Predictive Analysis	4	0+2+2
	CUML2011	ML for Image Analytics	4	0+2+2
	CUML 2009	Mathematics for ML	3	2+1+0
	CUML 2008	IoT Analytics	4	0+2+2
	CUML 2012	Digital video Processing	3	0+2+1
	CUML2004	ML for Hyperspectral imaging (Elective)	6	0+4+2
	CUML2005	Internship	4	0+0+4
	CUML2006	Project	4	0+0+4

Course Outline

Code	Course Title	T-P-Pj (Credit)	Prerequisite
MLCU2000	**Data Science and Machine Learning**	2-9-15	

Objective

- *Understand the scope, stages, applications, effects and challenges of ML.*
- *Understand the mathematical relationships within and across ML algorithms and the paradigms of supervised and unsupervised learning.*

Course outcome

- *Ability to Create and incorporate ML solutions in their respective fields of study.*
- *Get skill to design and implement various machine learning algorithms in a range of real-world applications.*
- *Ability to design product/ publish article/ file patent.*
- *Able to get jobs in AI/ML field*

Course Content

ML for Predictive Analysis (0+2+2)

https://careerfoundry.com/en/blog/data-analytics/regression-vs-classification/

Project/Task: (Choose two projects, Prediction and Classification)

Time Series Analysis

Health Care System

Concept Required:

3.1 Data pre-processing: -

- Accessing / collecting the datasets from different online repository.
- Missing values handling, noise reduction, finding Correlation between features, outlier elimination.
- Label Encoding / Encoding the categorical data
- Splitting the dataset
- Data Normalization

3.2 Learning Algorithms: -

- Supervised Learning Algorithms
- Unsupervised Learning Algorithms

3.3 Feature extraction and selection: -

- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)
- Different Feature Selection Techniques / Algorithms

3.4 Model building: -

- Regression (Linear, Polynomial, multiple, logistic), Decision Tree, Random Forest.
- Artificial Neural Network (Feed Forward Neural Network, Gradient Descent, Back Propagation Neural Network).
- Convolutional Neural Network
- Other Pretrained Models

3.5 Performance measures: -

- Perdition: Root Mean Square Error (RMSE), Average Percentage Error (APE), Mean Average Percentage Error (MAPE).
- Classification: Confusion Matrix (TN, TP, FP, FN), Sensitivity, Specificity, Gmean, F-score, Overall Accuracy, (Receiver Operating Characteristic) ROC Curve. Area under Curve (AUC)

3.6 Reading and Writing Research Articles

ML for Image Analytics (0-2-2)

Project/Task: (Choose one among six Tasks)

Detection of optometry diseases using retinal fundus imaging.

1. Diabetic Retinopathy
2. Glaucoma
3. Cataract

Detection of various diseases using X-ray imaging.

1. Covid19

Leaf disease classification using RGB images.

1. Tomato leaf
2. Potato leaf

Concept Required:

Image Pre-processing:-

- Accessing individual pixels using matrix concept
- Image resize, grey scale conversion, Colour channel splitting
- Histogram equalisation (CLACH).

Image Feature Extraction: -

- Edge detection (Sobel, Canny), Morphological operations
- Image segmentation, Image Thresholding, Binary conversion
- Cluster based segmentation
- Feature extraction based on size, shape and colour
- Feature extraction using predefined functions: SIFT, SURF, STAR, ORB.
- Feature Extraction using convolutional neural network (CNN).

Creation of Feature Matrix by combining Extracted Features: -

- Matrix flattening, Horizontal stacking, Vertical stacking, padding.
- Splitting the feature matrix (training/testing) and labelling.

Classification algorithms: -

- Support vector machine (SVM)
- Different kernels of SVM (linear, polynomial, radial basis function).
- Gradient Boosting (GB)

- Multi-layer Perceptron (MLP), Deep Learning.

Mathematics for ML (2+1+0)

When Models Meet Data:-

- Data, Models, and Learning
- Empirical Risk Minimization
- Parameter Estimation
- Probabilistic Modeling and Inference
- Directed Graphical Models
- Model Selection

Linear Regression:-

- Problem Formulation
- Parameter Estimation
- Bayesian Linear Regression
- Maximum Likelihood as Orthogonal Projection

Dimensionality Reduction with Principal Component Analysis:-

- Problem Setting
- Maximum Variance Perspective
- Projection Perspective
- Eigenvector Computation and Low-Rank Approximations
- PCA in High Dimensions
- Key Steps of PCA in Practice
- Latent Variable Perspective

Density Estimation with Gaussian Mixture Models:-

- Gaussian Mixture Model
- Parameter Learning via Maximum Likelihood
- EM Algorithm
- Latent-Variable Perspective

Classification with Support Vector Machines:-

- Separating Hyperplanes
- Primal Support Vector Machine
- Dual Support Vector Machine
- Kernels
- Numerical Solution

Practice:

- Curve Fitting in Python.

- Exploratory Data Analysis in Python.
- Kernel Density Estimation in Python.
- Probability Distribution Function Plotting in Python.
- Cumulative Distribution Function Plotting in Python.
- Dimensionality Reduction and Feature Extraction in Python.

References:

1. Mathematics for Machine Learning by Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong.
2. https://youtube.com/playlist?list=PLLy_2iUCG87D1CXFxE-SxCFZUiJzQ3IvE

IoT Analytics (0+2+2)

Defining IoT Analytics and Challenges

IoT

Benefits of Deploying IoT

End to End IoT architecture

IoT challenges

IoT Protocols

1 Wireless Protocol

Connectivity Protocols (when Power is Limited)

Bluetooth Low Energy (BLE)

Zigbee

LoRaWAN

NFC

2 Connectivity Protocols (when Power is Not a problem)

Wifi

3 Data Communication Protocol

MQTT

Web-Socket

HTTP

2 Sensors

Types of Sensors based on communication-I2C, SPI

Types of Sensors based on Application

3 Overview of 32 -bit Controller

ESP8266

ESP32

Raspberry Pi

4 AWS IoT for Cloud

AWS IoT Core services

AWS IoT Analytics services

AWS DynamoDB Services

5 Thingspeak for IoT

Getting and posting Data to IoT Cloud using ESP devices

Posting Data to IoT Cloud using Raspberry Pi

6 ThingWorx for Industrial IoT

Building Dashboard on Thingworx platform

Binding the senor value to the dashboard

Text Book:

1. Minteer, Andrew. *Analytics for the Internet of Things (IoT)*. Packt Publishing Ltd, 2017.

Reference Books:

2. Geng, Hwaiyu, ed. *Internet of things and data analytics handbook*. John Wiley & Sons, 2017.

Digital video Processing (0+2+1)

UNIT 1:

Fundamentals of Video Processing: Digital Video Acquisition, Principles of Color Video, Video Camera, Video Display, **Analog Vs Digital Video**: Progressive Vs Interlaced scans,

Signal, Bandwidth Characterization of a Digital Video Signal.

Practice:

- Read and play video files
- Extract frames from video files
- Combine frames to create a video file

UNIT 2:

Fourier Analysis of Digital Video Signals: Spatial and Temporal resolution, Fourier Analysis of Digital Video Signals, **Spatial-Temporal Sampling:** Temporal Frequency Response and Flicker Perception. Spatial Frequency Response, Spatiotemporal Frequency Response, Smooth Pursuit Eye Movement

Prctice:

- Applying fourier transformation on video
- Time domain analysis
- Frequency domain Analysis

UNIT 3:

Digital Video Formats: Significance of Video Formatting, Data rate and bandwidth trade-off, **File Formats:** MP4, MOV, WMV, AVCHD, FLV, AVI, WebM, MKV

Digital Video Compression Standards: Digital Video Compression Metrics, Digital Video Storage Precisions, Significance of Video compression, **Video Compression Codec's:** Motion JPEG, JPEG 2000, H.264/MPEG-4 AVC, VP8, HEVC, H.265 High Efficiency Video Codec.

Prctice:

- Conversion of video files from one format to another.
- Using Motion JPEG Codec
- Using MPEG-4 Codec
- Using H.265 Codec

UNIT 4:

Digital Video Editing Basics: Video Editing Types- Online, Offline, Linear, Non-linear, Assemble, Insert, Rough-cut, Video Shot Transition Effects: Cut, Fade, Wipe, Dissolve, B-roll, Video Shot Boundary Detection Methods: pixel differences, statistical differences, histogram comparisons, edge differences and motion vectors. Video Shot Detection Performance Metrics: ROC Curves, Recall, Precision, F-Measure

Practice:

- Video Shot Detection using pixel Difference
- Video Shot Detection using Histogram based methods
- Video Shot Detection using Edge based methods
- Video Shot Detection using Motion Vectors

Project List

1) Creating a VIDEO object detection system
2) Vehicle detection in Videos using OpenCV and Python
3) Detecting faces in live camera feed with identification of the person.

TEXT BOOK:

1. Rafael C Gonzalez and Richard E Woods, "Digital Image Processing'', Pearson Education, 3rd Edition, 2009.
2. Handbook of Image and Video processing - Al Bovik (Alan C Bovik), Academic Press,

REFERENCE BOOK:

1. Fundamentals of Digital Image Processing", Anil K. Jain, PHI, 1995.
2. "Digital Image Processing", William. K.Pratt, Wiley Interscience, 2nd Ed, 1991.

Ml for Hyperspectral imaging (0-4-2)

Project/Task: (Choose one among four Tasks)
Agriculture

1. Crop yield prediction.
2. Crop quality prediction
3. Soil health monitoring

Mining

1. Iron ore quality prediction

Concept Required:

Introduction to Remote Sensing: -

- Multi-Spectral Imagery (MSI)
- Hyperspectral Imagery (HSI)

Scientific Principles:

- Physics of imaging spectroscopy
- Electromagnetic propagation
- Sensor physics
- Atmospheric Corrections.

Hyperspectral Concepts and System Trade-offs:-

- Signal-to-Noise ratio (SNR)
- Spectral resolution, sampling.

HSI Data Processing Techniques:-

- Spectral angle mapping
- Principal Component Analysis (PCA)
- Minimum Noise Fraction (MNF)
- Spectral feature fitting.

Classification Techniques:-

- Support Vector Machine (SVM)
- Partial Least Squares Regression (PLSR)
- Neural Network
- Deep learning and CNN

Clustering Techniques:-

- K-mean clustering

Project (0-0-4)

Internship (0-0-4)

Course Outline Prepared by: Prof. (Dr.) Sujata Chakravarty