CE362: MACHINE LEARNING

Credits and Hours:

Teaching Scheme	Theory	Practical	Tutorial	Total	Credit
Hours/week	3	2	-	5	**4**
Marks	100	50	-	150	

Pre-requisite courses:

- Probability, linear algebra, calculus and programming language

Outline of the Course:

Sr. No.	Title of the unit	Minimum number of hours
1.	Data cleaning and Preprocessing	06
2.	Supervised Machine Learning	14
3.	Unsupervised Machine Learning	11
4.	Ensemble learning	04
5.	Deep Neural Networks	10
	Total hours (Theory) :	45
	Total hours (Lab) :	30
	Total hours :	75

Detailed Syllabus:

1.	**Data Cleaning and Pre-processing**	**06 Hours**	**13%**
	Handling Missing data, Data Exploration and Visualization, Handling Outliers, Data Integration and Aggregation, Feature selection: Filter Methods, Wrapper Methods, Embedded Methods, Data Transformation: Feature scaling and normalization, Encoding categorical variables: one-hot encoding, label encoding, Data transformation for skewed distributions		
2.	**Supervised Machine Learning**	**14 Hours**	**31%**
	What is supervised ML, Linear Regression for univariate and Multivariate data, Cost function, Gradient Descent, Logistic regression, Under fitting and Over fitting, Support Vector		

	Machine, Decision Tree, Random Forest, Artificial Neural Network architecture, Activation functions, Forward pass in ANN, Back propogation in ANN, Model Evaluation techniques.		
3.	**Unsupervised Machine Learning**	**11 Hours**	**25%**
	What is unsupervised ML, Distance Techniques: Euclidean Distance, Manhattan Distance, Minkowski Distance, Cosine Similarity, Hamming Distance. Different clustering algorithm: K-means, Hierarchical Clustering, DBSCAN, Evaluate clustering algorithm, Evaluation Techniques: Evaluate Clustering metrics - Silhouette Score, Calinski-Harabasz Index, Davies-Bouldin Index, Dunn Index. Dimensionality reduction: Principal Component Analysis(PCA)		
5.	**Ensemble learning**	**04 Hours**	**09%**
	Gradient Boosting, XGBoost (Extreme Gradient Boosting), AdaBoost (Adaptive Boosting)		
5.	**Deep Neural Networks**	**10 Hours**	**22%**
	Introduction to Convolutional Neural Networks, Convolutional Neural Network Layers, CNN Architectures and Variants, Introduction to Recurrent Neural: Understanding sequential data and its challenges, Anatomy of a Recurrent Neural Network (RNN), vanishing gradient problem Networks, Long Short-Term Memory (LSTM) Networks.		

Course Outcome (COs):

At the end of the course, the students will be able to

CO1	Addressing data quality issues, effectively cleaning and preprocessing raw datasets, and preparing them for machine learning algorithm.
CO2	Capable of building predictive models for classification and regression tasks, and skilled in selecting, training, and evaluating machine learning algorithms on labeled data

CO3	Solid understanding of unsupervised learning methods, including clustering and dimensionality reduction
CO4	Proficient in implementing and optimizing gradient boosting models for both regression and classification tasks
CO5	Understand the principles and architecture of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) Networks.

Course Articulation Matrix:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	3	-	1	1	-	-	-	-	-	-	-	-	-	-
CO2	3	-	3	2	-	-	-	-	-	-	-	-	1	-
CO3	3	-	3	2	-	-	-	-	-	-	-	-	1	-
CO4	3	-	3	2	-	-	-	-	-	-	-	-	1	-
CO5	3	-	3	2	-	-	-	-	-	-	-	-	1	-

Enter correlation levels 1, 2 or 3 as defined below:

1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High)

If there is no correlation, put "-"

Recommended Study Material:

- **Text book:**
  1. Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow, 2nd Edition, Aurelien Géron, OReilly
  2. Machine Learning Yearning, Andrew Ng, GitHub; eBook (Draft, 2018); eBook (MIT Licensed)
  3. Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville, The MIT Press (17 December 2016)

- **Reference book:**

1. Christopher M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006.
2. Richard O. Duda, Peter E. Hart & David G. Stork, "Pattern Classification. Second Edition", Wiley & Sons, 2001.
3. Trevor Hastie, Robert Tibshirani and Jerome Friedman, "The elements of statistical learning", Springer, 2001.

❖ **Web material:**

1. https://www.youtube.com/watch?v=foHSmB48rY&list=PLKvX2d3IUq586Ic9gIhZj6ubpWV-OJfl4
2. https://www.youtube.com/watch?v=CS4cs9xVecg&list=PLkDaE6sCZn6EcXTbcX1uRg2_u4xOEky0
3. https://www.youtube.com/watch?v=UzxYlbK2c7E
4. https://www.youtube.com/playlist?list=PLAwxTw4SYaPkQXg8TkVdIvYv4HfLG7SiH
5. Recommended online course:
6. Coursera:https://www.coursera.org/specializations/machine-learning-introduction
7. Coursera: https://www.coursera.org/specializations/deep-learning

❖ **Software:**

1. Anaconda
2. GoogleColab