CSEN3261	MACHINE LEARNING AND ITS APPLICATIONS	L	T	P	S	J	C
		3	0	2	0	0	4
Pre-requisite	MATH2291 : Linear Algebra						
Co-requisite	None						
Preferable exposure	None						

Course Description:

Machine Learning is a flourishing subject in Computer Science which devises models that can automatically learn from data and detect patterns from data. The applications of machine learning are diverse ranging from self- driven cars to disaster management systems. With easy availability of data from different devices and measurements, machine learning techniques become imperative in analysing trends hidden in the data. This course focuses on the major tasks of machine learning that can robustly address data that is non-linear, noisy as well as high-dimensional in nature.

Course Educational Objectives:

- To introduce various key paradigms of machine learning approaches
- To familiarize with mathematical relationships across various machine learning algorithms
- To understand various key approaches in supervised learning
- To understand various key approaches in unsupervised learning
- To illustrate the concept of the neural network

UNIT 1 Machine Learning Fundamentals 9 hours, P - 6 hours

Machine Learning Fundamentals: Use of Machine Learning, Types of machine learning systems, Machine learning challenges, Testing and validating, working with real data, Obtaining the data, Visualizing the data, Data preparation, Training and fine tuning the model.

UNIT 2 Supervised Learning 9 hours, P - 6 hours

Supervised Learning: Classification, training a binary classifier, performance measures, multiclass classification, error analysis, multi label classification, multi output classification. Linear Regression, Gradient Descent, Polynomial Regression, learning curves, regularized linear models, logistic regression.

UNIT 3 Unsupervised Learning 9 hours, P - 6 hours

Unsupervised Learning: Clustering, K-Means, using clustering for image segmentation, Semi-supervised learning, DBSCAN, other clustering algorithms.

Gaussian Mixtures, anomaly detection, selecting number of clusters, Bayesian Gaussian Mixture Models, anomaly and novelty detection algorithms.

UNIT 4 Dimensionality Reduction& Ensemble Learning 9 hours, P - 6 hours

Dimensionality Reduction: The curse of dimensionality, main approaches for dimensionality reduction, PCA, Non Negative Matrix Factorization.

Ensemble Learning: voting classifiers, bagging, random patches and random spaces, random forests, boosting, stacking.

UNIT 5 Neural Networks & Deep Neural Networks 9 hours,P - 6 hours

Neural Networks: From biological to artificial neurons, implementing MLPs with Keras, fine tuning neural network hyper parameters.

Deep Neural Networks: Vanishing/Exploding Gradients Problem, avoiding overfitting through regularization, Dropout Regularization.

List of Experiments:

1. Write a python program to characterize the density functions

2. Write a python program to model statistically the feature space using distribution functions

3. Write a python program to understand the distribution functions (Normal, Binomial, Poissonetc)

4. Write a python program to estimate co variance matrix and its properties

5. Write a python program to visualize the changes of distribution as changes in parameters

(mean vector, covariance matrix)

6. Write a python program for perceptron learning and test the linear separability

7. Write a python program for Bayesian classification and analyze the decision boundaries by

varying the means and covariance matrices

8. Write a python program to classify the given data using maximum likelihood

Estimation. Write a program to solve Robot traversal problem (Understanding Means End

Analysis)

9. Write a phyton program to understand Morkov Chians and Monto Calro methods. Write a

program to implement Hangman game

10. Write a python program to Decision trees
11. Write a python program to build a Bayesian network for given data set
12. Write a python program to understand Kernel methods.
13. Write a program to implement a linear regression problem
14. Write a program to implement kNN neighbour problem
15. Write a program to implement logistic regression

TextBooks:

1. Aurelion Geron, Hands-on Machine Learning with Scikit-Learn, Keras, and Tensor Flow: Concepts, Tools and Techniques to build Intelligent Systems, 2/e, O'Reilly Media, 2019.(Chapters 1,3,4,5)

2. Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep learning, MIT press, 2016 (Chapter2)

References:

1. Tom M. Mitchell, "Machine Learning" First Edition by Tata McGraw- Hill Education.
2. Ethem Alpaydin,"Introduction to Machine Learning " 2nd Edition, The MIT Press, 2009
3. Christopher M. Bishop, "Pattern Recognition and Machine Learning" By Springer, 2007.
4. Mevi P. Murphy, "Machine Learning: A Probabilistic Perspective" by The MIT Press, 2012.
5. https://www.coursera.org/learn/uol-machine-learning-for-all
6. https://www.coursera.org/learn/introduction-to-machine-learning-supervised-learning
7. https://www.coursera.org/learn/machine-learning-with-python? (Clustering Week -4))
8. coursera.org/learn/mixture-models?specialization=bayesian-statistics (GMM & Bayesian GMM)
9. 5.https://www.coursera.org/learn/ibm-unsupervised-machine-learning(Curse Dimensionality, PCA)
10. https://www.coursera.org/learn/supervised-machine-learning-classification
11. https://www.coursera.org/learn/introduction-to-deep-learning-with-keras(Week1 &Week 2Part-1)
12. https://www.coursera.org/learn/deep-neural-network(Week 1- part 2)

Course Outcomes:

After successful completion of the course the student will be able to:

1. To formulate the different machine learning problems
2. Apply various learning approaches on real time problems using Classification
3. Apply various learning approaches on real time problems using Regression
4. Apply various learning approaches on real time problems using Clustering
5. Construct the neural networks for classification problems

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	3	3	2	2	1			2	1	1	2	3	2	1
CO2	3	3	3	2	2	1			2	1	1	2	3	2	2
CO3	2	3	3	2	2	1			2	1	1	2	3	2	2
CO4	1	2	2	2	2	1			2	1	1	2	3	2	2
CO5	2	3	3	3	3	1			2	1	2	2	3	2	2

Note: 1 - Low Correlation 2 - Medium Correlation 3 - High Correlation

APPROVED IN:

BOS : 06-09-2021 **ACADEMIC COUNCIL: 22nd AC (01-04-2022)**

SDG No. & Statement:

SDGs: 3, 6, 11

SDG:3 Good Health and Well-being

Statement: Machine Learning has the potential to personalize healthcare monitoring, diagnosis and treatment for the individual in the community and at home. It puts consumers in control of health and well-being.

SDG:6 Clean Water and Sanitation

Statement: Machine Learning will help to resolve challenges related to clean water and sanitation. It is helping utilities and municipalities to better manage their water and wastewater systems to ensure a clean and sanitized water supply.

SDG:11 Sustainable Cities and Communities

Statement: Machine Learning enable smart urban solutions brings multiple benefits, including more efficient energy, water and waste management, reduced pollution, noise and traffic congestions

SDG Justification: