MIT-WPU | MIT WORLD PEACE UNIVERSITY | Pune

COURSE STRUCTURE

Course Code	**CSE0PE22A**				
Course Category	**Program Elective**				
Course Title	**Applied Machine Learning**				
Teaching Scheme	**Lectures**	**Tutorials**	**Laboratory / Practical**	**Project**	**Total**
Weekly load hours	3	0	2	0	5
Credits	3	0	1	0	4
Assessment Schema Code	**TL3**				

Prerequisites: Basics of Machine Learning

Course Objectives:

1.To understand the basic concepts machine Learning and apply different dimensionality reduction techniques

2.To optimize the different linear methods of regression and classification

3.To optimize the different linear methods of regression and classification

4.To interpret the different supervised classification methods of support vector machine and tree based models

Course Outcomes:

After completion of this course students will be able to:

1. Understand the fundamental concepts of Machine Learning.
2. Understand the basic concepts of Feature Selection.
3. Use regression techniques to transform data.
4. Understand classification algorithms such as bayes algorithm and SVM.
5. Use of various Ensembling Techniques.

MIT-WPU, PUNE-38.
APPROVED BY
05 JUN 2021
ACADEMIC COUNCIL

Course Contents:

Unit 1: Introduction to Machine Learning

Introduction to Machine Learning, Examples of Machine Learning Applications, Learning Types Supervised Learning -Learning a Class from Examples, Vapnik-Chervonenkis (VC) Dimension, Probably Approximately Correct (PAC) Learning,Noise, Learning Multiple Classes, Regression, Model Selection and Generalization, Dimensions of a Supervised Machine Learning Algorithm

Dean
Faculty of Engineering and Technology

Unit 2: Feature Selection

Concept of Feature, Preprocessing of data: Normalization and Scaling, Standardization, Managing missing values, Introduction to Dimensionality Reduction, Principal Component Analysis (PCA),Feature Extraction: Kernel PCA,LDA

Introduction to various Feature Selection Techniques: Sequential Forward Selection, Sequential Backward Selection

Unit 3: Regression

Linear regression- Linear models, A bi-dimensional example, Linear Regression and higher dimensionality, Ridge, Lasso, Polynomial regression, Isotonic regression, Logistic regression-Linear classification, Logistic regression, Implementation and Optimizations, Stochastic gradient descendent algorithms. Finding the optimal hyper-parameters through grid search

Unit 4: Naïve Bayes and Support Vector Machine

Bayes Theorem, Naïve Bayes Classifiers, Naïve Bayes in Scikitlearn- Bernoulli Naïve Bayes, Multinomial Naïve Bayes, and Gaussian Naïve Bayes.

Support Vector Machine(SVM)- Linear Support Vector Machines, Scikit- learn implementation-Linear Classification, Kernel based classification, Non- linear Examples. Controlled Support Vector Machines

Unit 5: Decision Trees and Ensemble Learning

Decision Trees- Impurity measures, Feature Importance. Decision Tree Classification with Scikit-learn, Ensemble Learning-Random Forest, Gradient Tree Boosting, Voting Classifiers. Clustering Fundamentals- Basics, K-means: Finding optimal number of clusters, Introduction to Meta Classifier: Concepts of Weak and eager learner, Ensemble methods, Bagging: Random Forests, Boosting: XG,Adaboost,LGBM

Laboratory Exercises / Practical:

1. Apply data preprocessing techniques to make data suitable for machine learning.
2. Train the system using data set obtained from UCI ML repository. Use a partition of the same data set as a test set to determine accuracy using DecisionTree.
3. Train the system using data set obtained from UCI ML repository. Use a partition of the same data set as a test set to determine accuracy using RandomForest.
4. Train the system using data set obtained from UCI ML repository. Use a partition of the same data set as a test set to determine accuracy using Naïve Bayes.
5. Implement Find-S algorithm.
6. Train the system using data set obtained from UCI ML repository. Use a partition of the same data set as a test set to determine accuracy using SVM
7. Train the system using data set obtained from UCI ML repository. Use a partition of the same data set as a test set to determine accuracy using KNN classifier.
8. Train the system using data set obtained from UCI ML repository. Use a partition of the same

Dean
Faculty of Engineering and Technology

data set as a test set to determine accuracy using Kmeans clustering
9. Implement the ANN algorithm on a data set obtained from UCI ML repository
10. Apply PCA and SVD on a data set obtained from UCI ML repository

Learning Resources:

Text Books/ Reference Books:

1. 1. Ethem Alpaydin, " Introduction to Machine Learning", PHI 2nd Edition-2013, ISBN 978-0-262-01243-0
2. Peter Flach, "Machine Learning: The Art and Science of Algorithms that Make Sense of Data", Cambridge University Press, Edition 2012, ISBN-10: 1107422221; ISBN-13: 978-1107422223
3. Tom Mitchell "Machine Learning" McGraw Hill Publication, ISBN:0070428077 9780070428072
4. Nikhil Buduma, "Fundamentals of Deep Learning", O'REILLY publication, second edition 2017, ISBN: 1491925612

Supplemen Practical Tableautary Reading:

1. "Practical Machine Learning" by Sunila Gollapudi Publisher(s): Packt Publishing ISBN: 9781784399689
2. "Practical Machine Learning with Python" by Dipanjan Sarkar, Raghav Bali, Tushar Sharma , Apress, ISBN 1484232062, 9781484232064

Web Resources:

Weblinks:

1. https://www.geeksforgeeks.org/machine-learning/
2. https://www.javatpoint.com/machine-learning

MOOCs: Online courses for self learning

Courses by NPTEL and MIT Open Courseware etc

1.
2.

Pedagogy:

- Power point presentations
- Videos
- Demonstrations
- Systematic use of group work and project-based learning.

Dean
Faculty of Engineering and Technology