Course Name: **Applied Machine Learning**

Level: 5
Credits: 4.0
Type: Theory

Prerequisites
CS1101, CS3102 or equivalent

Syllabus

- Learning to use Python machine learning libraries
    - Writing and importing code in Python
    - Learning to use the 'scikit-learn' and 'TensorFlow' libraries
- Conducting classification tasks
    - A brief review of classification algorithms or classifiers
    - Recognizing hand-written digits using a support vector classifier
    - Conducting a comparative study of the widely used classifiers using in-silico datasets
    - Investigating the class boundaries learned by the linear discriminant analysis (LDA) versus that of the quadratic discriminant analysis (QDA)
    - Predicting whether the share market will move up or down using a naïve Bayes classifier and a k-nearest neighbours (KNN) classifier
    - Identifying the peak hours of bike rentals using linear regression and Poisson regression
- Performing feature extraction
    - Extracting features from text documents
    - Extracting features from image files
- Selecting linear models (with and without regularization) that best describe a given dataset
    - A brief review of linear model selection algorithms
    - Choosing the models that best describe how salaries of baseball players vary based on various statistics associated with the players' performance: Using least squares regression, partial least squares (PLS) regression, ridge regression, lasso regression, and principal components regression
- Utilizing tree-based algorithms for predication
    - A brief review of tree-based algorithms
    - Predicting housing prices with different types of decision trees (classification trees, regression trees) and random forests (with and without bagging and boosting)

- Employing support vector machines (SVMs) for classification tasks
    - A brief review of SVMs
    - Employing an SVM for one-class classification
    - Employing an SVM for two-class classification
    - Employing an SVM for multi-class classification
- Clustering
    - A brief review of clustering algorithms
    - Subgrouping cancer cell lines into cancer types using the k-means and hierarchical clustering algorithms
- Training and testing artificial neural networks (ANNs)
    - A brief review of ANNs
    - Deploying a single-layer ANN
    - Deploying a multi-layer ANN
- Modelling real-world systems with graphical models
    - A brief review of graphical models
    - Modelling a stock market using Bayesian networks
    - Denoising images using Markov random fields (MRFs)
    - Understanding speech with hidden Markov models (HMMs)
- Demonstrating a few state-of-the-art applications of machine learning
    - Application of large language models for legal text summarization
    - Application of word embedding for information retrieval
    - Classification of aerial photographs under scarcity of labelled samples
    - Prediction of 3D protein structures with AlphaFold

Weekly Plan

Week 1	Python recap, pip, scikit-learn, TensorFlow
Week 2-4	Support vector classifier, comparative study of classifiers, linear discriminant analysis, quadratic discriminant analysis, naïve Bayes, k-nearest neighbours, linear regression, Poisson regression, feature extraction from text, feature extraction from images
Week 5	Least squares regression, partial least squares regression, ridge regression, lasso regression, principal components regression
Week 6	Decision trees (classification trees, regression trees), random forests, bagging and boosting
Week 7	One-class, two-class, and multi-class classifications with support vector machines
Week 8	k-means clustering, hierarchical clustering algorithms
Week 9	Perceptron, single-layer and multi-layer artificial neural networks
Week 10	Bayesian networks, Markov random fields, hidden Markov models
Week 11	State-of-the-art applications
Week 12	Buffer

Textbook

- James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). An Introduction to Statistical Learning: With Applications in Python. Springer Nature.

Reference books

- Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
- Duda, R. O., & Hart, P. E. (2006). Pattern Classification. John Wiley & Sons.
- Marsland, S. (2011). Machine Learning: An Algorithmic Perspective. Chapman and Hall/CRC.

Reference Materials

- 'Machine Learning in Python' by the authors of the 'scikit-learn' Python library: https://scikit-learn.org/stable/index.html
- Books and videos available on the TensorFlow website: https://www.tensorflow.org/resources/learn-ml

Similar courses

- 'Machine Learning for Engineering and Science Applications' by IIT Madras: https://nptel.ac.in/courses/106106198
- 'Practical Machine Learning with Tensorflow' by IIT Madras and Google: https://nptel.ac.in/courses/106106213
- 'CS 5785 Applied Machine Learning' by the Cornell University: https://kuleshov-group.github.io/aml-website/
- 'Applied Machine Learning in Python' by the University of Michigan: https://www.coursera.org/learn/python-machine-learning