Course Objective

In this course, I will introduce the essential machine learning algorithms, including Linear Regression, Logistic Regression, Decision Tree, Support Vector Machine, Naive Bayes Classifier, Artificial Neural Networks, Frequent Itemset mining, K-Means, Hierarchical clustering, Dimension Reduction, Ensemble Algorithm, and deep learning algorithms such as CNN. This course will use the Python Environment for Data Science to demonstrate the algorithms with some example datasets. The Python Data Science environment consists of Python, Numpy, Pandas, SciPy, Matplotlib, Scikit-Learn, Keras, and Tensorflow. We will use the Jupyter Notebook and Spyder as IDEs for data mining. In this course, you will learn the most useful Machine Learning (AI) techniques, which are very useful for your jobs and research.

Outline

1. Introduction

2. Python for machine learning

3. Regression, Logistic Regression

4. Classification: Decision Tree, Naïve Bayes

5. Clustering

6. Text Mining (NLP) (Guest lecturer)

7. Association rules (Frequent itemset mining)

8. Classification: Support Vector Machine

9. Neural Networks.

10. Deep Learning: CNN

Text Books

1. Data Mining: Concepts and Techniques, Jiawei Han and Micheline Kamber

2. Introduction to data mining, Pan-ning Tan, Mich. Steinbach, and Vipin Kumar

3. Data Mining: Methods and Models---by Daniel T. Larose

4. Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-learn, and TensorFlow2

5. 打下最紮實的AI基礎: Scikit-learn 一步一腳印， 黃永昌編著

Prerequisite

Calculus (differential) and some Statistics Background

Grading Policy

Midterm 40%, Final 40%, Class attendance 20%.