Course number and name: **CS 07656: Machine Learning II**

Credits and contact hours: 3 credits / 3 contact hours

Instructor's or course coordinator's name: Shen-Shyang Ho

Instructional materials: 1. Lecture notes.

Catalog description: In this course, we will review and revisit some topics in Machine Learning 1. We will revisit the theory of generalization, support vector machine, neural network, and reinforcement learning with more advanced materials. New topics include ensemble learning, nonlinear dimensionality reduction, prediction interpretability and explainability, and other recent topics.

Prerequisites: Linear Algebra, Probability, Calculus II, or permission of instructor

Specific goals for the course:

1. **Generalization theory.** Learn how the theory of generalization explains the prediction performance of different machine learning approaches using VC dimensions.

2. **Solving overfitting.** Learn how the concept of regularization and validation are applied to overcome the overfitting problem.

3. **Multiple learners.** Learn how the use of multiple learners can improve on prediction performance and confidence

4. **Using non-standard data.** Learn to handle non-standard data or problem settings (sequential, graph, etc) using extensions of standard machine learning techniques.

5. **Neural networks.** Be able to appreciate and have basic understanding of recent advances in neural network technology (e. g. , Deep Learning, Residual Network, Recurrence Neural Network, Generative Adversarial Networks, etc.)

6. **Statistical Rigor.** Learn to perform statistically sound experiments when using machine learning approaches to solve problems.

7. **ML approaches.** Learn to decide what dimensionality reduction, unsupervised learning, pre-processing approaches to use on data before applying machine learning techniques

List of topics to be covered:

1. History of Machine Learning
2. Theory of Generalization
    a. Revisit materials in Machine Learning 1
    b. Structural Risk Minimization
    c. VC analysis of all discussed machine learning methods.
3. Support Vector Machines
    a. Revisit materials in Machine Learning 1
    b. SVM formulation for regression
    c. SVM formulation for imbalanced data
    d. Kernel Methods
        i. Mercer's Condition
        ii. Reproducing Kernel Hilbert Space
        iii. Application-driven kernels
4. Neural Network
    a. Revisit materials in Machine Learning 1
    b. Graph Neural Network
    c. Generative Adversarial Network
    d. Transformer & Attention
    e. Other Recent Development in Neural Network
5. Reinforcement Learning
    a. Revisit materials in Machine Learning 1
    b. Off-Policy and On-Policy Learning Algorithms
    c. Deep Reinforcement Learning
    d. Policy Gradient Methods – Actor-Critic Algorithm
    e. Other Recent Development in Reinforcement Learning
6. Dimensionality Reduction & Manifold Learning
    a. Kernel Principal Component Analysis
    b. Isomap
    c. Laplacian Eigenmap
    d. Locally Linear Embedding
7. Ensemble Learning
    a. ECOC classifier
    b. Random Forest
    c. Bagging
    d. Boosting
8. Interpretability & Explainability