

ESE 588: Pattern Recognition

Instructor: Alex Doboli, Ph.D.

Email: adoboli@ece.sunysb.edu

Office hours: Mon, Wed, 10-11AM, Light Eng. Room 261

Schedule: Spring Semester, Mon-Wed 8.30AM-9.50AM

Goals: Basic concepts of pattern recognition and data mining techniques are introduced, including statistical pattern recognition, syntactic pattern recognition, and graph matching. Topics on Bayes decision theory, parametric and nonparametric techniques, clustering techniques, formal languages, parsing algorithms, and graph-matching algorithms are covered.

Textbooks:

- J. Han, M. Kamber, J. Pei, “*Data Mining. Concepts and Techniques*”, Morgan Kaufman, second edition, 2012. (second edition can be also used)
- Published papers will be provided in class.

Prerequisites:

Stochastic processes and data structures. If prerequisites are not met, the instructor can grant permission to register for the course.

Topics:

1) *Introduction into Data Mining:*

- Data Warehouse and Databases
- Mining frequent patterns, associations and correlations
- Types of patterns
- Data mining primitives

2) *Data preprocessing:*

- Data summarization
- Data cleaning
- Data integration and transformation
- Data reduction
- Data discretization and concept hierarchy generation

1) *Mining frequent patterns, associations and correlations:*

- Scalable frequent itemset mining methods
- Mining association rules
- Correlation analysis
- Constraint-based association mining

2) *Classification and prediction:*

- Classification by decision tree induction
- Bayesian classification
- Rule-based classification
- Classification by backpropagation
- Support Vector Machines
- Prediction

3) *Cluster Analysis:*

- Partitioning methods
- Density-based methods
- Grid-based methods
- Model-based clustering methods
- Clustering high-dimensional data
- Constraint-based cluster analysis

4) *Mining stream, time-series, and sequence data:*

- Mining time-series data
- Mining sequence patterns in databases and biological data

Assignment Schedule and Grading:

Project 1	25%
Project 2	25%
Project 3	25%
Project 4	25%