

---

## CSE 446 Machine Learning

---

### Credits

3.0 (3 hrs lecture)

### Lead Instructor

Dan Weld

### Textbook

- *Machine Learning*, Murphy

### Course Description

Methods for designing systems that learn from data and improve with experience. Supervised learning and predictive modeling: decision trees, rule induction, nearest neighbors, Bayesian methods, neural networks, support vector machines, and model ensembles. Unsupervised learning and clustering.

### Prerequisites

either CSE 326 or CSE 332; either STAT 390, STAT 391, or CSE 312.

### CE Major Status

Selected Elective

### Course Objectives

#### ABET Outcomes

- (a) an ability to apply knowledge of mathematics, science, and engineering
- (b) an ability to design and conduct experiments, as well as to analyze and interpret data
- (c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability
- (d) an ability to function on multi-disciplinary teams
- (e) an ability to identify, formulate, and solve engineering problems
- (f) an understanding of professional and ethical responsibility
- (g) an ability to communicate effectively
- (h) the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context
- (i) a recognition of the need for, and an ability to engage in life-long learning
- (j) knowledge of contemporary issues
- (k) an ability to use the techniques, skills, and modern engineering tools necessary for

engineering practice

### **Course Topics**

- Architecture of a Relation Extractor
- Supervised Learning & Logistic Regression
- Instaread and Features for ML  
Project Discussion and Crawling the Web
- IR Models & Index Construction and Link Analysis & Pagerank
- SE Query Processing: Alta Vista
- NYU's 2011 KBP System
- Computational Advertising
- Crowdsourcing
- Cryptography & Practical Internet Security
- Mining unstructured healthcare data