

Ensemble Based Systems in Decision Making

CS-795 (INFS-795; IT 803)

Dr. Carlotta Domeniconi
carlotta@cs.gmu.edu

Spring 2009

General Info

- Instructor: Carlotta Domeniconi
 - Office: S&T2, Rm 449
 - Email: carlotta@cs.gmu.edu
 - Phone: (703) 993-1697
- Office hours: Mon. 5-7pm, or by appointment, or stop by
 - <http://www.cs.gmu.edu/~carlotta/teaching//CS-795-s09/info.html>
- Visit the class webpage often!

Why a course on Ensemble Methods

"The task of improving classification accuracy by learning ensemble of classifiers is considered as one of the most four important directions in machine learning research." (T. G. Dietterich, Machine Learning Research: four current directions, AI Magazine, 18(4), pages 97-136, 1997).

Ensemble Methods: the basic idea

- Given a task, we train a collection of learners
- Given a test point, we observe the prediction of each component
- Combine the predictions (according to a given scheme) to derive a "consensus" prediction

Ensemble Methods: we use them all the time!

- We gather different doctors' opinions before undergoing a major surgery
- We read users' reviews before purchasing a product
- We request references before hiring someone

Ensemble Methods... or

- Multiple learner systems
- Committee of learners
- Mixture of experts
- Ensemble based systems
- ...

Objective of this course

- In depth study and critical analysis of current methodologies to construct effective ensembles of learners
- Technical tools from... linear algebra, probability, statistics, multivariate calculus, optimization

Topics (tentative)

- Fundamental issues with learning, model selection, over-fitting, decision theory, curse-of-dimensionality
- Overview on Classification
- Overview on Clustering
- Semi-supervised learning
- Introduction to ensemble methods
- Requirements for effective ensembles: accuracy and diversity

More topics

- Methods to construct diverse learners
- Methods to combine the collection of learners
- Ensembles of classifiers: Bagging and Boosting and ...
- Clustering ensembles
- Semi-supervised ensembles

Course Format

- Lectures by the instructor
- Presentations of papers by students
- 1 midterm/homeworks (?)
- Questions on papers presented in class
- Project: proposal, presentation, paper

Some useful books and reading material

- **On Pattern Recognition and Machine Learning:**

- Christopher M Bishop,
“Pattern Recognition and Machine Learning”,
Springer, 2006.

Companion website:

[http://research.microsoft.com/~cmbishop/PRML/
index.htm](http://research.microsoft.com/~cmbishop/PRML/index.htm)

- **On Pattern Classification:**

- R. O. Duda, P. E. Hart, D. G. Stork,
“[Pattern Classification](#)”,
Second Edition, Wiley, 2001.

- **On Ensemble Methods:**

- Ludmila I. Kuncheva,
[Combining Pattern Classifiers – Methods and Algorithms](#),
Wiley, 2004.

- **International Workshop Series:**

- *Multiple Classifier Systems (MCS) Workshops*

- First one was held in 2000.

- Current developments can be found in its proceedings, and in proceedings of other ML/Data Mining/Pattern Recognition Conferences (KDD, ICDM, SDM, ICPR, ICML)