

Course Project

- Each student will be required to choose a unique research project and write a term paper.
- 70% of your project grade will be on your final paper,
- 15% will be based on your proposal, and
- 15% will be based on your in-class presentation.

Your performance will be graded based on the following criteria:

- Accuracy of results.
- Toughness of investigation.
- Validity of conclusions.
- Novelty of approach.
- Quality of written/oral presentations.

- The topic of your project may be related to your current research, but should not be research that has been submitted for publication anywhere (accepted or not). It can be an extension of your previous work.

Possible topics:

1. Your own idea.
2. Comparison of different classifiers.
 - How do you evaluate performance?
 - Which classifiers are best for what types of data/application?
 - What is the sensitivity of the results on the parameters?
 - What are the limits of learnability?
What are the error bounds?

3. Combining classifiers:

- Bagging, boosting, your new method
- Combining classifiers for unsupervised learning
- Combining supervised and unsupervised learning

4. Unsupervised Learning

- How do you initialize clustering?
- How do you find the number of clusters?
- How do you evaluate performance?
- Comparison of different unsupervised learners.

5. Feature Selection and Dimensionality Reduction

6. Learning from mixed data (e.g., discrete and continuous data)

7. Learning when features are missing

8. Finding the best model or hypothesis

9. Speeding up algorithms

10. On-line versions of algorithms

11. Active Learning

11. Multi-Label problem

12. Multi-Task problem

13. Class imbalance problem

14. anomaly detection

15. Dealing with concept drift

16. Visualization tool for algorithms

17. A machine learning application.

Data Sources:

<http://www.ics.uci.edu/~mllearn/MLRepository.html>

<http://kdd.ics.uci.edu/>

Conferences:

International Conference on Machine Learning (ICML)

Int'l Conference on Neural Information Processing Systems (NIPS)

ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD)

Journals:

IEEE Transactions on Pattern Analysis and Machine Intelligence

Journal of Machine Learning Research

Machine Learning

Guidelines for writing your project proposal

The project will comprise a substantial portion of your grade. Your proposal will provide an opportunity for you to receive feedback on your research idea, and will count for 15% of your project grade. Your proposal should be one to two pages long (not including the bibliography). Proposals should be in 12 point font.

In your proposal, you should answer the following questions (*briefly*).

1. What research issue(s) interest you most? Why?
2. Who else has worked in this vein? What did they accomplish? What can't they do?
3. What kind of progress would you like to see? Why?
4. Do you have an idea for making some such progress? Explain.
5. What do you expect to discover from your investigation?
6. How will your expected result(s) affect the research community?

A large part of your proposal should address Question 4. In addition, your proposal should contain a bibliography of related articles. This does not need to be the full bibliography, just an indication that you have done a preliminary search for related work.

- If you are having trouble deciding on a research project, then set up an appointment with me **before** the deadline.

Basic Concepts

- Bayesian Decision Theory
- ML, MAP, Bayesian

Supervised Learning Methods:

- Naïve Bayes
- Bayesian Networks
- Hidden Markov Models
- Linear Discriminant Analysis
- K-nn classifiers
- Decision Trees
- Support Vector Machines
- Neural Networks
- Bagging
- Boosting

Unsupervised Learning Methods:

- K-means
- Gaussian Mixture Models
- Hierarchical Methods
- PCA
- ICA
- Spectral Clustering