OR&IE 674 – Statistical Learning
Theory For Data Mining
Spring 2004
Professor D. Ruppert

1 Home Page
All course materials can be found at:
http://www.orie.cornell.edu/~davidr/or674/

This is a link to ORIE 674 webpage at CIT’s “Blackboard.” Be sure to sign up as a 674 student with Blackboard so that you receive email announcements.

2 Class Times
MWF, 1:25–2:15, Upson 111

3 Prerequisites
Knowledge of basis statistics including multiple linear regression, hypothesis testing, and estimation. OR&IE 670 is sufficient background.

4 Instructor
David Ruppert
255-9136
dr24@cornell.edu
225 Rhodes Hall
http://www.orie.cornell.edu/~davidr/

5 Syllabus
The lectures will include the following topics:
- Linear regression
- Local regression
- Penalized least-squares
- Basis expansion and spline bases
- Classification
  - Logistic regression
  - Linear discriminant analysis
  - Nonparametric classification
- Model assessment
- Model averaging
- CART
- Additive models
- Boosting
- Support vector machines
- Cluster analysis

6 Textbook
Hastie, T., Tibshirani, R., and Friedman, J. (2001), The Elements of Statistical Learning, Springer.

7 Reserve Books
In addition to the textbook, the following will be placed on reserve in the Engineering Library.

8 Grades
Grades will be based upon class participation, homework, and a course project that will involve reading the literature on a topic and presenting oral and written reports. There will be no exams.

If you have questions, please contact me.