Statistics 4620
Introduction to Statistical Learning
2-semester-hour course

Prerequisite: Stat 3302 (Statistical Modeling for Discovery II)
Exclusions:
Class distribution: Two 55-minute lectures per week

Course Description and Learning Outcomes
The course provides an introduction to the principles of statistical learning and standard learning techniques for regression, classification, clustering, dimensionality reduction, and feature extraction.

Upon successful completion of the course, students will be able to

1. Recognize the types of learning problems and understand their statistical formulations.
2. Understand the foundational principles of statistical learning including statistical modeling, computation and evaluation.
3. Comprehend the rationale and algorithms behind statistical learning techniques and know their relative merits and limitations.
4. Evaluate and compare different learning techniques numerically in terms of generalization error.
5. Use statistical learning methods for data analysis and interpret the results in the context of the data problem.

Required Text and Other Course Materials
The required textbook for the course is (books currently under review). The book is available for purchase at the official University bookstore (ohiostate.bkstore.com) and elsewhere online. The book is available on reserve in the 18th Avenue Library.

Students will be required to use the R^1 software environment for statistical computing and graphics. R can be downloaded for free at [http://www.r-project.org](http://www.r-project.org). Instructions for using the software will be given in class. Many students prefer to use RStudio, an IDE designed for use with R. RStudio is available for free at [http://www.rstudio.com](http://www.rstudio.com).

---

^1 For information on the use of R in data analytics, see:
Assignments

Homework will be assigned (approximately) bi-weekly, will be due on the dates announced in class and will be graded. Assignments will consist of a mix of technical questions to assess students’ understanding of the statistical models, and questions asking students to perform analyses of data sets. The grade for the analysis portion of each assignment will be based on both the accurateness and appropriateness of the analysis, as well as the clarity of the description of the analysis and results.

Project: Each student will be responsible for completing an individual project. Proposals for project ideas will be due mid-way through the semester, and the project will be due near the end of the semester. The project will consist of finding a data set, formulating questions that can be answered with the data, and performing an appropriate analysis to answer the questions.

Exams

There will be two in-class midterms that cover material from lecture, the assigned readings and homework.

A cumulative final examination will be given during the university’s examination period.

Grading Information

The final course grade will be based on homework assignments, two projects, two midterms and a comprehensive final examination. The weights for each component of the grade are:

<table>
<thead>
<tr>
<th>Component</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>15%</td>
</tr>
<tr>
<td>Midterm 1</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm 2</td>
<td>20%</td>
</tr>
<tr>
<td>Project</td>
<td>15%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>30%</td>
</tr>
</tbody>
</table>

Outline of topics

1. Overview of predictive modeling and model evaluation
2. Penalized regression and nonparametric regression
3. Nearest neighbor methods
4. Classification and regression trees
5. Model selection and validation
6. High dimensional data and variable selection
**Statement on Academic Misconduct**

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term “academic misconduct” includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct [http://studentlife.osu.edu/csc/](http://studentlife.osu.edu/csc/).

**Special Accommodations**

Students with disabilities that have been certified by the Office for Disability Services will be appropriately accommodated and should inform the instructor as soon as possible of their needs. The Office for Disability Services is located in 150 Pomerene Hall, 1760 Neil Avenue; telephone 292-3307, TDD 292-0901; [http://www.ods.ohio-state.edu/](http://www.ods.ohio-state.edu/).