
Term Information

Effective Term Spring 2023

General Information

Course Bulletin Listing/Subject Area Mathematics
Fiscal Unit/Academic Org Mathematics - D0671
College/Academic Group Arts and Sciences
Level/Career Graduate, Undergraduate
Course Number/Catalog 5637
Course Title Topics in Predictive Modeling
Transcript Abbreviation Topics Pred Mod
Course Description An introduction to some basic concepts and methods in statistical learning with emphasis on the mathematics behind these concepts and methods.
Semester Credit Hours/Units Fixed: 3

Offering Information

Length Of Course 14 Week, 7 Week, 6 Week
Flexibly Scheduled Course Never
Does any section of this course have a distance education component? No
Grading Basis Letter Grade
Repeatable No
Course Components Lecture
Grade Roster Component Lecture
Credit Available by Exam No
Admission Condition Course No
Off Campus Never
Campus of Offering Columbus

Prerequisites and Exclusions

Prerequisites/Corequisites A grade of C- or better in 2568 or 5520H and a grade of C- or better in Stat 4202; or permission of department.
Exclusions
Electronically Enforced Yes

Cross-Listings

Cross-Listings

Subject/CIP Code

Subject/CIP Code 27.0101
Subsidy Level Doctoral Course
Intended Rank Junior, Senior, Masters, Doctoral

Requirement/Elective Designation

The course is an elective (for this or other units) or is a service course for other units

Course Details

Course goals or learning objectives/outcomes

- Understanding supervised and unsupervised learning.
- Understanding the lasso, sparse regression, classification and regression trees.
- Understanding boosting and support vector machines.

Content Topic List

- Supervised/unsupervised learning, loss function, regression and classification.
- Model Accuracy: training and test error, bias-variance trade-off, KNN.
- Linear Regression: simple and multiple linear regression, hypothesis testing, variable selection, model fit
- Introduction to GLM: general linear regression and link function.
- Classification: logistic regression, linear and quadratic discriminant analysis.
- Resampling: cross-validation, bootstrap.
- Model Selection: variable selection, ridge regression and Lasso, dimension reduction, principal component analysis, partial least squares.
- Nonlinear Regression: polynomial regression, regression splines, smooth splines, generalized additive model.
- Tree Based Methods: decision trees, pruning, classification and regression trees, error rate, bagging, random forests, boosting.
- Neural Networks.
- Support Vector Machines: maximal margin classifier, support vector classifier, kernel functions, support vector machines.
- Introduction to Unsupervised Learning: PCA in unsupervised learning, K-means and hierarchical clusterings.

Sought Concurrence

Yes

Attachments

- Curriculum_map_actsci_09102021.docx: Curriculum map - Act Sci
(Other Supporting Documentation. Owner: Husen, William J)
- Curriculum_map_math_09102021.docx: Curriculum map - Math
(Other Supporting Documentation. Owner: Husen, William J)
- 5637.pdf: Syllabus - updated
(Syllabus. Owner: Husen, William J)
- 5637_concurrence.pdf: Concurrence - Stats
(Concurrence. Owner: Husen, William J)

Comments

- Syllabus updated and concurrence from Statistics added. *(by Husen, William J on 09/29/2022 01:07 PM)*
- See email feedback to dept. 10-6-21 RLS *(by Steele, Rachel Lea on 10/06/2021 05:06 PM)*

COURSE REQUEST
5637 - Status: PENDING

Last Updated: Vankeerbergen, Bernadette
Chantal
10/15/2022

Workflow Information

Status	User(s)	Date/Time	Step
Submitted	Husen, William J	09/17/2021 11:06 AM	Submitted for Approval
Approved	Husen, William J	09/17/2021 11:29 AM	Unit Approval
Approved	Vankeerbergen, Bernadette Chantal	09/20/2021 12:01 PM	College Approval
Revision Requested	Steele, Rachel Lea	10/06/2021 05:06 PM	ASCCAO Approval
Submitted	Husen, William J	09/29/2022 01:07 PM	Submitted for Approval
Approved	Husen, William J	09/29/2022 01:07 PM	Unit Approval
Approved	Vankeerbergen, Bernadette Chantal	10/15/2022 12:37 PM	College Approval
Pending Approval	Cody, Emily Kathryn Jenkins, Mary Ellen Bigler Hanlin, Deborah Kay Hilty, Michael Vankeerbergen, Bernadette Chantal Steele, Rachel Lea	10/15/2022 12:37 PM	ASCCAO Approval

Math 5637: Topics in Risk Modeling

- DESCRIPTION

Risk Modeling is an important area of Actuarial Science, which is based on statistical or machine learning and the underlying mathematics. The professional societies administer exams for the actuarial profession. This course covers most topics of Exam SRM administered by the Society of Actuaries and exams MAS I and MAS II administered by the Casualty Actuarial Society.

Many mathematical concepts and results are used in statistical or machine learning, and many concepts and methods in statistical or machine learning are based on mathematical thinking. However, most books and courses on statistical or machine learning do not go deep in the mathematics which statistical or machine learning is based on, and they usually just quote the mathematical results as formulas. This course will present the fundamental mathematics used in statistical learning, and introduce basic concepts and methods in statistical learning as applications of these mathematics.

This course emphasizes mathematics and the theory side of statistical learning. For a more practical treatment with serious hands-on experience, a statistics or computer science course on the subject is recommended, for example Stat 4620 or Stat 6500.

- COURSE OBJECTIVES/LEARNING OUTCOMES

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

- Understand the mathematical concepts and results on which statistical or machine learning is based.
- Understand basic concepts and principles of statistical learning, and considerations when applying statistical learning.
- Understand and apply the main methods in statistical learning.
- Evaluate and select learning models or methods.

- CLASS FORMAT

Lecture – 3 hours per week

- PREREQUISITE

Linear algebra (Math 2568 or equivalent), probability (Math 4530 or Stat 4201 or equivalent), and statistics (Stat 4202 or equivalent); or by department permission.

- TEXTS

Class notes will be distributed. The following are recommended textbooks.

1. *Mathematics for Machine Learning*
by Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong

2. *An Introduction to Statistical Learning with Applications in R*
by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani
OSU students can access this book freely at

<https://link-springer-com.proxy.lib.ohio-state.edu/book/10.1007/978-1-4614-7138-7>

3. *The Elements of Statistical Learning – Data Mining, Inference, and Prediction*
by Trevor Hastie, Robert Tibshirani, Jerome Friedman
<https://web.stanford.edu/~hastie/ElemStatLearn/>

- **HOMEWORK AND EXAMS**

There will be

- Weekly homework assignment
- Two midterm exams
- Final exam

- **EXPECTED WORKLOAD**

students will be expected to be working on homework for an approximate total of 6 hours per week.

- **GRADE**

The course grade will be based on

- Homework, 20%
- Two midterm exams, 50%
- Final exam, 30%

Course grade will be determined by the total percentage obtained, roughly as 90–100 for an A, 80–89 for a B, 65–79 for a C, and 50–64 for a D.

- **SCHEDULE**

A tentative weekly schedule is attached. This schedule and material covered may be changed without notice. It is the students responsibility to keep track of these changes. Changes may be announced in class verbally, through Carmen, or through email.

- **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 3335-5-48.7). For additional information, see the Code of Student Conduct at <http://studentaffairs.osu.edu/csc/> .

- Disability Services Statement
Students with disabilities that have been certified by Student Life Disabilities Services (SLDS) will be appropriately accommodated and should inform the instructor as soon as possible of their needs. SLDS contact information: sllds@osu.edu; 614-292-3307; 098 Baker Hall, 113 W. 12th Avenue.

Week 1	Introduction to Statistical Learning: supervised/unsupervised learning, regression and classification, loss function, conditional mean and median, model accuracy
Week 2	Vector Calculus: multivariable vector functions, gradient, functions defined by matrix operations Application: simple linear regression
Week 3	Matrix and Vector Geometry: matrix and its row and column spaces, orthogonal projection, minimum distance Application: multiple regression, extensions of linear model
Week 4	Classification Problems: KNN, logistic regression
Week 5	Multivariate Normal Distribution Application: linear and quadratic discriminant analysis
Week 6	Resampling: cross-validation, bootstrap
Week 7 & 8	Convex Optimization: optimization and constrained optimization, duality, convex optimization Application: ridge regression and lasso
Week 9	Symmetric Matrix and Quadratic Forms: symmetric matrix and its eigenvalues, eigenspaces, diagonalization, quadratic forms Application: principal component analysis, dimension reduction
Week 10	Nonlinear Regression and Splines
Week 11	Vector Geometry: hyperplane and its normal form, distance to a hyperplane Applications: maximal margin classifier, support vector classifier, support vector machine
Week 12	Tree Based Methods: regression and classification trees, bagging, random forest, boosting
Week 13	Neural Networks
Week 14	Unsupervised Learning: PCA, K-Means and Hierarchical Clusterings

Re: Math 5637

Craigmile, Peter <pfc@stat.osu.edu>

Fri 9/16/2022 9:20 AM

To: Ban, Chunsheng <cban@math.ohio-state.edu>

Cc: MacEachern, Steven <snm@stat.osu.edu>; Kaizar, Elly <kaizar.1@osu.edu>; Xu, Xinyi <xinyi@stat.osu.edu>; Hans, Christopher <hans@stat.osu.edu>

Dear Chunsheng,

Thank you again for meeting earlier this week. I enjoyed our conversation, as always. I went back to the curriculum committee. With your assurance that the major focus on the course is the math (especially linear algebra) and that you will recommend students take either Stat 4620 or Stat 6500 if they want to learn more about statistical learning, the Department of Statistics gives concurrence for Math 5637.

We look forward to collaborating more in the future. Currently, Elly Kaizar, our vice chair of admin can assist with scheduling and general administrative issues for Stat courses that actuarial science students are interested in taking, Xinyi Xu, our grad chair is happy to correspond about fostering links with our PhD students, and Chris Hans manages our undergraduate programs. Feel free to reach out to us, or to the chair Steve MacEachern, if there are more substantive connections that you see we can make in the future.

I will also remind you that our R course, Stat 5730 (<https://stat.osu.edu/courses/stat-5730>), will be revised in the future to be more modular. Stat 5730 is already a great resource for any student who wants to learn the R programming language, but we hope that providing choices of different modules will allow students to customize the skills they want to learn.

My best,
Peter

Peter Craigmile, Ph.D.,
Professor, Department of Statistics, The Ohio State University.
Chair, Curriculum Committee, Department of Statistics

Math - BS/BA Curriculum Map					
Goal 1	Learn conceptual frameworks needed to study higher mathematics, including an introduction to mathematical reasoning and an understanding of how to read and write proofs.				
Goal 2	Aquire basic mastery of core areas of mathematics including calculus, analysis and algebra.				
Goal 3	Develop powerful mathematical problem solving skills.				
Goal 4	Learn to communicate mathematical understanding effectively.				
Goal 5	Become proficient in chosen tracks within the major.				
Course	Goal 1	Goal 2	Goal 3	Goal 4	Goal 5
AcctMIS 2000			Beginning		Intermediate
Biochem 4511					Advanced
Biology 1113			Beginning		Intermediate
Biology 1114			Beginning		Intermediate
Biology 3401					Intermediate
BusFin 3120			Intermediate	Intermediate	Advanced
BusFin 3220			Intermediate	Intermediate	Advanced
Chem 1210			Beginning		Intermediate
Chem 1220			Beginning		Intermediate
Chem 2210					Advanced
Chem 2510					Advanced
Chem 4300					Advanced
Chem 4310					Advanced
CSE 1222			Beginning		Intermediate
CSE 1223			Beginning		Intermediate
CSE 2221			Beginning	Beginning	
CSE 2111			Beginning		Intermediate
Econ 2001.01			Beginning		Intermediate
Econ 2002.01			Beginning		Intermediate
EEOB 3310					Advanced
EEOB 3420					Advanced
EEOB 4520					Advanced
Math 1151	Beginning	Beginning	Beginning		
Math 1152	Beginning	Beginning	Beginning		
Math 1181H	Intermediate	Intermediate	Beginning		
Math 1295				Intermediate	Beginning
Math 2153	Intermediate	Intermediate	Beginning		
Math 2182H	Intermediate	Intermediate	Beginning		
Math 2255	Beginning	Intermediate	Intermediate	Beginning	

Math 2568	Beginning	Beginning	Beginning		Beginning
Math 2568H	Intermediate	Beginning	Intermediate	Beginning	Beginning
Math 3345	Advanced	Advanced	Intermediate	Intermediate	Intermediate
Math 3345H	Advanced	Advanced	Intermediate	Intermediate	Intermediate
Math 3350				Intermediate	Beginning
Math 3589			Intermediate	Intermediate	Advanced
Math 3607			Intermediate	Intermediate	Advanced
Math 3618			Intermediate	Advanced	Advanced
Math 4181H	Advanced	Advanced	Advanced	Advanced	Advanced
Math 4182H	Advanced	Advanced	Advanced	Advanced	Advanced
Math 4350			Intermediate	Advanced	Advanced
Math 4504	Advanced	Intermediate	Intermediate	Advanced	Advanced
Math 4507	Advanced	Intermediate	Intermediate	Advanced	Advanced
Math 4512	Intermediate		Intermediate	Intermediate	Intermediate
Math 4530	Intermediate	Beginning	Intermediate	Intermediate	Intermediate
Math 4547	Advanced	Advanced	Intermediate	Advanced	Beginning
Math 4548	Advanced	Advanced	Intermediate	Advanced	Beginning
Math 4551	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate
Math 4552	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate
Math 4556			Intermediate	Advanced	Advanced
Math 4557	Intermediate		Intermediate	Intermediate	Intermediate
Math 4570	Intermediate	Intermediate	Advanced	Intermediate	Intermediate
Math 4573	Advanced	Intermediate	Intermediate	Intermediate	Intermediate
Math 4575	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate
Math 4578	Intermediate	Intermediate	Intermediate	Intermediate	Advanced
Math 4580	Advanced	Advanced	Intermediate	Advanced	Beginning
Math 4581	Advanced	Advanced	Intermediate	Advanced	Beginning
Math 5101	Beginning	Advanced	Intermediate		Intermediate
Math 5102	Beginning	Advanced	Intermediate		Intermediate
Math 5421	Beginning	Beginning	Intermediate	Beginning	Advanced
Math 5451	Beginning	Beginning	Intermediate	Beginning	Advanced
Math 5520H	Advanced	Advanced	Advanced	Advanced	Intermediate
Math 5522H	Advanced	Advanced	Advanced	Advanced	Intermediate
Math 5529H	Advanced	Advanced	Advanced	Advanced	Intermediate
Math 5530H	Advanced	Advanced	Advanced	Advanced	Intermediate
Math 5540H	Advanced	Advanced	Advanced	Advanced	Advanced
Math 5540H	Advanced	Advanced	Advanced	Intermediate	Beginning
Math 5576H	Advanced	Advanced	Advanced	Advanced	Advanced
Math 5590H	Advanced	Advanced	Advanced	Advanced	Advanced
Math 5591H	Advanced	Advanced	Advanced	Advanced	Advanced
Math 5632			Intermediate	Advanced	Advanced

Math 5635			Intermediate	Advanced	Advanced
Math 5636			Intermediate	Advanced	Advanced
Math 5637			Intermediate	Advanced	Advanced
Math 5660					Intermediate
Math 5756			Beginning	Intermediate	Intermediate
Math 5757			Beginning	Intermediate	Intermediate
MolGen 4500					Advanced
MolGen 5601					Advanced
Physics 1250			Beginning		Intermediate
Physics 1251			Beginning		Intermediate
Physics 2300					Advanced
Physics 2301					Advanced
Stat 4201	Intermediate	Beginning	Intermediate	Intermediate	Intermediate
Stat 4202	Intermediate		Intermediate		Intermediate

Actuarial Science BS/BA Curriculum Map					
Goal 1	To supply a strong general background in mathematics, statistics, and relevant concepts from the insurance industry				
Goal 2	To prepare students to take some of the national actuarial examinations administered by the Society of Actuaries and the Casualty Actuarial Society				
Course	Goal 1	Goal 2			
Math 1151	Beginning	Beginning			
Math 1152	Beginning	Beginning			
ACCTMIS 2000	Beginning				
Econ 2001.01	Beginning				
Econ 2002.01	Beginning				
CSE 1222	Beginning	Intermediate			
CSE 1223	Beginning	Intermediate			
CSE 2111	Beginning	Intermediate			
Comm 2110	Beginning				
Comm 2131	Beginning				
Comm 2367	Beginning				
BusFin 3120	Intermediate	Beginning			
English 3304	Beginning				
Math 2153	Intermediate	Beginning			
Math 2568	Intermediate	Beginning			
Math 3588	Intermediate	Advanced			
Math 3618	Intermediate	Advanced			
Math 4530	Advanced	Advanced			
Stat 4201	Advanced	Advanced			
Math 5632	Advanced	Advanced			
Stat 4202	Advanced	Advanced			
Math 5630	Advanced	Advanced			
Math 5631	Advanced	Advanced			
Math 5633	Advanced	Advanced			
Math 5634	Advanced	Advanced			
Math 5635	Advanced	Advanced			
Math 5636	Advanced	Advanced			
Math 5637	Advanced	Advanced			