



STAT*6801 Statistical Learning

Fall 2020

Section(s): C01

Department of Mathematics & Statistics

Credit Weight: 0.50

Version 1.00 - August 24, 2020

1 Course Details

1.1 Calendar Description

Topics include: nonparametric and semiparametric regression; kernel methods; regression splines; local polynomial models; generalized additive models; classification and regression trees; neural networks. This course deals with both the methodology and its application with appropriate software. Areas of application include biology, economics, engineering and medicine.

1.2 Course Description

Please note that the ongoing COVID-19 pandemic may necessitate a revision of the format of course offerings and academic schedules. Any such changes will be announced via CourseLink and/or class email. All University-wide decisions will be posted on the COVID-19 website (<https://news.uoguelph.ca/2019-novel-coronavirus-information/>) and circulated by email.

In this course we will discuss modern regression methods otherwise referred to as statistical learning. We will begin by reviewing classical (unsupervised) learning in the form of the linear model. We will then examine other learning methods that extends the linear model. These will encompass nonparametric and semiparametric regression; kernel methods; regression splines; local polynomial models; and, generalized additive models. We will then introduce penalization procedures turning regression into a supervised learning problem. Different penalizations lead to different shrinkage methods with the primary ones being ridge regression and LASSO regression, in various forms. This is the basis of what may be called modern machine learning methods which is heavily computational. We will end with classification and regression trees, as well as neural networks and what is called deep learning.

1.3 Timetable

Synchronous online delivery (via ZOOM) Monday 1430 - 1720.

1.4 Final Exam

Dec 17, 2020 8:30-10:30.

2 Instructional Support

2.1 Instructional Support Team

Instructor:	Peter Kim
Email:	pkim@uoguelph.ca
Telephone:	519-824-4120 x58165
Office:	MACN 515
Office Hours:	Office hours will be held via Zoom at the conclusion of each class on Mondays at 5:20 pm, or by appointment.

3 Learning Resources

3.1 Required Resources

Lectures (Notes)

Lectures will be provided interactively via ZOOM. Information concerning assignments and readings will be included there as well as on the courselink website.

Modern Applied Statistics with S, Fourth Edition, W.N. Venables and B.D. Ripley, Springer (Textbook)

Available online through the library. Most of the software codes are available in this text except LASSO.

R package (Software)

<https://cran.r-project.org/package=glmnet>

Documentation is available at:

<https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>

Glmnet documentation (Article)

<https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>

Documentation of glmnet. This package contains the codes for LASSO.

4 Learning Outcomes

This is a course on advanced data analysis using modern statistical learning methods.

4.1 Course Learning Outcomes

By the end of this course, you should be able to:

1. Following this course students will have a firm grasp of:
 - supervised and unsupervised learning
 - Kernel methods
 - Generalized Additive Models
 - Linear models (supervised and unsupervised)
 - LASSO and Ridge Regression
 - Local Polynomial and Splines
 - CART
 - Neural Networks

5 Teaching and Learning Activities

Lectures will be conducted virtually. Students will have to have access to high speed internet as the presentations will be done online. The presentation slides will be uploaded to courselink prior to each lecture. We will have discussions during each lecture period where students can be engaged.

5.1 Lecture

Week 1

Topics: Linear models

References: Venables and Ripley, Chapter 6, Linear Statistical Models

Week 2

Topics: Linear models cont'd

References: Venables and Ripley, Chapter 6, Linear Statistical Models

Week 3

Topics: Extending linear models to nonparametric and semi-parametric specifications; kernel method and generalized

additive models (GAM)

References: Venables and Ripley, Chapter 8, Non-Linear and Smooth Regression

Week 4

Topics: Extending linear models to nonparametric and semi-parametric specifications; local polynomials and regression splines

References: Venables and Ripley, Chapter 8, Non-Linear and Smooth Regression

Week 5

Topics: Review of material and Exam I. The exam will be conducted in a virtual class. Students will be given the exam through course link. They will write with the camera on and upload their solutions via Drop-Box.

Week 6

Topics: Thanksgiving break. Reading assignments Singular Value Decomposition.

References: Files will be posted on Courselink.

Week 7

Topics: Supervised learning. This involves what had been covered until now with constraints put on the parameters. We will start with the L_2 constraint which is called ridge regression.

References: GLMNET documentations: <https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>

Week 8

Topics: Supervised learning continued. We will then look at the L_1 constraint which is called least absolute shrinkage and selection operator, or, the LASSO.

References: GLMNET documentations: <https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>

Week 9

Topics: Supervised learning computations. We will go over how the computation is performed for ridge and lasso regression. This involves selecting tuning parameters via cross-validation.

Week 10

Topics: Supervised learning computations and calculations continued. Exam II.

References: GLMNET documentations: <https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>

Week 11

Topics: Classification and regression trees, and possibly support vector machines.

References: Venables and Ripley, Chapter 9, Tree-Based Methods

Week 12

Topics: Neural networks and deep learning.

References: Venables and Ripley, Chapter 8, Section 4.8, Neural Networks

6 Assessments

6.1 Marking Schemes & Distributions

Assignments: There will be regular assignments that will not be collected for grading but will be discussed in class.

Midterms: There will be two in class 90 minute (approximate) midterms.

Midterm 1: 25%

Midterm 2: 25%

Final Exam: 50%

6.2 Assessment Details

Midterm 1 (25%)

Date: Mon, Oct 5

Approximately 90 minutes in-class.

Midterm 2 (25%)

Date: Mon, Nov 9

Approximately 90 minutes in-class.

Final Examination (50%)

Date: Week 13

Details will be announced as we approach the date.

7 University Statements

7.1 Email Communication

As per university regulations, all students are required to check their e-mail account regularly: e-mail is the official route of communication between the University and its students.

7.2 When You Cannot Meet a Course Requirement

When you find yourself unable to meet an in-course requirement because of illness or compassionate reasons please advise the course instructor (or designated person, such as a teaching assistant) in writing, with your name, id#, and e-mail contact. The grounds for Academic Consideration are detailed in the Undergraduate and Graduate Calendars.

Undergraduate Calendar - Academic Consideration and Appeals

<https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml>

Graduate Calendar - Grounds for Academic Consideration

<https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/index.shtml>

Associate Diploma Calendar - Academic Consideration, Appeals and Petitions

<https://www.uoguelph.ca/registrar/calendars/diploma/current/index.shtml>

7.3 Drop Date

Students will have until the last day of classes to drop courses without academic penalty. The

deadline to drop two-semester courses will be the last day of classes in the second semester. This applies to all students (undergraduate, graduate and diploma) except for Doctor of Veterinary Medicine and Associate Diploma in Veterinary Technology (conventional and alternative delivery) students. The regulations and procedures for course registration are available in their respective Academic Calendars.

Undergraduate Calendar - Dropping Courses

<https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-drop.shtml>

Graduate Calendar - Registration Changes

<https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/genreg-reg-regchg.shtml>

Associate Diploma Calendar - Dropping Courses

<https://www.uoguelph.ca/registrar/calendars/diploma/current/c08/c08-drop.shtml>

7.4 Copies of Out-of-class Assignments

Keep paper and/or other reliable back-up copies of all out-of-class assignments: you may be asked to resubmit work at any time.

7.5 Accessibility

The University promotes the full participation of students who experience disabilities in their academic programs. To that end, the provision of academic accommodation is a shared responsibility between the University and the student.

When accommodations are needed, the student is required to first register with Student Accessibility Services (SAS). Documentation to substantiate the existence of a disability is required; however, interim accommodations may be possible while that process is underway.

Accommodations are available for both permanent and temporary disabilities. It should be noted that common illnesses such as a cold or the flu do not constitute a disability.

Use of the SAS Exam Centre requires students to book their exams at least 7 days in advance and not later than the 40th Class Day.

For Guelph students, information can be found on the SAS website

<https://www.uoguelph.ca/sas>

For Ridgetown students, information can be found on the Ridgetown SAS website

<https://www.ridgetownc.com/services/accessibilityservices.cfm>

7.6 Academic Integrity

The University of Guelph is committed to upholding the highest standards of academic integrity, and it is the responsibility of all members of the University community-faculty, staff, and students-to be aware of what constitutes academic misconduct and to do as much as possible to prevent academic offences from occurring. University of Guelph students have

the responsibility of abiding by the University's policy on academic misconduct regardless of their location of study; faculty, staff, and students have the responsibility of supporting an environment that encourages academic integrity. Students need to remain aware that instructors have access to and the right to use electronic and other means of detection.

Please note: Whether or not a student intended to commit academic misconduct is not relevant for a finding of guilt. Hurried or careless submission of assignments does not excuse students from responsibility for verifying the academic integrity of their work before submitting it. Students who are in any doubt as to whether an action on their part could be construed as an academic offence should consult with a faculty member or faculty advisor.

Undergraduate Calendar - Academic Misconduct

<https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml>

Graduate Calendar - Academic Misconduct

<https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/index.shtml>

7.7 Recording of Materials

Presentations that are made in relation to course work - including lectures - cannot be recorded or copied without the permission of the presenter, whether the instructor, a student, or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted.

7.8 Resources

The Academic Calendars are the source of information about the University of Guelph's procedures, policies, and regulations that apply to undergraduate, graduate, and diploma programs.

Academic Calendars

<https://www.uoguelph.ca/academics/calendars>

7.9 Disclaimer

Please note that the ongoing COVID-19 pandemic may necessitate a revision of the format of course offerings and academic schedules. Any such changes will be announced via CourseLink and/or class email. All University-wide decisions will be posted on the COVID-19 website (<https://news.uoguelph.ca/2019-novel-coronavirus-information/>) and circulated by email.

7.10 Illness

The University will not require verification of illness (doctor's notes) for the fall 2020 or winter 2021 semesters.
