UNIVERSITY
*of* GUELPH

# Course Outline STAT*6801: Fall 2021

# Provisional outline awaiting approval
## 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. This includes on-campus scheduling during the semester, midterms and final examination schedules. All University-wide decisions will be posted on the COVID-19 website and circulated by email.
For information on current safety protocols, follow these links:
https://news.uoguelph.ca/return-to-campuses/how-u-of-g-is-preparing-for-your-safe-return/
https://news.uoguelph.ca/return-to-campuses/spaces/#ClassroomSpaces
Please note, these guidelines may be updated as required in response to evolving University, Public Health or government directives.

**Accommodation due to Illness:**

Medical notes will not normally be required for singular instances of academic consideration, although students may be required to provide supporting documentation for multiple missed assessments or when involving a large part of a course (e.g. final exam or major assignment).

**Mental Health Services:**

**One out of every five students in Canada experiences some sort of mental health issue at some point in their academic career. If you find yourself facing a mental health crisis, or just need to talk to someone, please consider taking advantage of one of the following resources available to University of Guelph students:**

*Counselling Services:* **Visit the Counselling Services website (https://wellness.uoguelph.ca/counselling) to get information on resources available to you, both online and in-person. You can also visit them at Health Services (J.T. Powell Building, ext 53244) where they offer individual and group counselling sessions by**

appointment or walk-in.

*Student Support Network:* is located in the Wellness & Education Promotion Centre in the J.T. Powell Building and offers confidential, peer-based, drop-in support.

*Good2Talk:* (**1-866-925-5454**) is a free, 24/7 student hotline that provides professional counselling and referrals for mental health, addictions and well-being.

*Here 24/7:* (**1-844-437-3247**) specializes in assessment, referral and appointment booking and is available 24/7 for crisis support.

**You are not alone and you will not be judged for asking for help.**

## General Information

**Course Title: Statistical Learning**

**Course Description (from Graduate Calendar):** 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.

**Credit Weight: 0.5**

**Academic Department (or campus): Math and Stats**

**Campus: Guelph**

**Semester Offering: Fall 2021**

**Class Schedule: Remote Asynchronous + 1 Synchronous Meeting on Friday 1:30 – 2:20 via Zoom**

## Instructor Information

**Instructor Name: Tony Desmond**
**Instructor Email: tdesmond@uoguelph.ca**

# Course Content

**This course will deal with a variety of topics in statistical learning and their implementation in R. In lectures I will briefly review recent research in generalized linear models. One focus of the course will be nonparametric and semiparametric versions of these models. An important example is generalized additive models (GAMs), which will be treated in some depth. In addition modern nonparametric regression via kernels, splines, etc. will be studied. Other topics, which will be covered, include: classification and regression trees, random forests, boosting and neural networks. Time permitting, topics such as wavelets and MARS (Multivariate Adaptive Regression Splines) may also be treated. In the project component of the course the student is encouraged to work in areas (both applied and theoretical), of his or her own interest, with the instructor's permission. Much of the material in the required and recommended texts relates to research published in the last two decades or so. Areas of application include medicine, finance, agriculture, economics, pharmacokinetics, bioassay, engineering reliability, to name only a few. Familiarity with R will be assumed. The best way to acquire familiarity is via the manuals (available on line). Also simply working through the required texts is of great value.**

**Learning Outcomes**:

1. Understand basic statistical learning concepts such as: generalization, predictive accuracy, overfitting, training, test and validation sets, parsimony, cross-validation.
2. Explore and understand how standard parametric models such as linear and generalized models can be viewed from a statistical learning perspective.
3. Explore and understand non-parametric approaches to statistical learning, which extend the flexibility and enhance the predictive accuracy of parametric supervised learning.
4. Explore and understand algorithmic approaches such as neural nets, classification and regression trees.
5. Implement the approaches in 2, 3, and 4 using the software package R on real data from various subject matter areas.

**Lecture Content:**

1. Statistical Learning: Prediction vs Inference; The 2 cultures; Algorithmic vs Data models; The bias-variance tradeoff; The prediction accuracy/ model-interpretability tradeoff; generalizability and validation; supervised and unsupervised learning; regression vs classification.

2. Linear and Generalized Linear Models from a statistical learning perspective. Difficulties with high-dimensional data. Ridge Regression and the LASSO; The glmnet package.

3. Moving beyond Linearity: Regression splines; smoothing splines; local regression; generalized additive models.

4. Tree-based Methods: Regression and Classification Trees; Trees vs Linear and Generalized Linear Models; Random Forests, Bagging and Boosting.

5. Neural Networks

6. Other topics: Wavelets, MARS (Multivariate Adaptive Regression Splines); Support Vector Machines.


**Course Assignments and Tests:**

**4 Assignments, each worth 12.50%. The post dates and due dates are as follows:**

**A1: Post date, Thurs Sept 16, 5pm. Due: Wed Sept 29 11:59pm**

**A2: Post date, Thurs Sept 30, 5pm. Due: Wed Oct 13 11:59pm**

**A3: Post date, Thurs Oct 14, 5pm. Due: Wed Oct 27 11:59pm**

**A4: Post date, Thurs Oct 28, 5pm. Due: Wed Nov 10 11:59pm**

**Final Term Project, worth 50%. Due date Friday Dec 10 5pm.**


## **Course Resources**


**Required Texts:**

**Extending the Linear Model with R, by Julian Faraway, 2nd Ed. Chapman and Hall 2016.**

**The Elements of Statistical Learning:  Data Mining, Inference and Prediction, by Hastie, Tibshirani, and Friedman. Springer 2009 2nd Edition.**

**Recommended Texts:**

**An Introduction to Statistical Learning with Applications in R, 2$^{nd}$**

**Edition, by James, G et al., Springer 2021.**

**Statistical Learning with Sparsity: The LASSO and its Generalizations, by Hastie et al, Chapman and Hall 2016**

**Modern Applied Statistics with S, 4$^{th}$ Edition, by W.N. Venables and B.D. Ripley. Springer 2004.**

**Statistical Learning from a Regression Perspective, by Berk, R. 3$^{rd}$ Edition, Springer 2021.**

**Semiparametric Regression, by Ruppert, Wand and Carroll, Cambridge University Press 2003.**

**Generalized Additive Models, by Hastie and Tibshirani, Chapman and Hall, 1990.**

**Statistical Learning for Biomedical Data, by Mallet et al, CUP, 2011.**

**NB: Copies of each of these texts have been placed on reserve in the library. With the exception of the last 2 these are electronic copies.**

## Course Policies

## Late Assignments will not be accepted except under very exceptional circumstances.

**Course Policy on Group Work:**

**Assignment solutions should be your own work, be clear, legible and well organized. You may discuss assignments with other classmates, but the work handed in should be your own.**

**Course Policy regarding use of electronic devices and recording of lectures**

*Electronic recording of classes is expressly forbidden without consent of the instructor. When recordings are permitted they are solely for the use of the authorized student and may*

## University Policies

### Academic Consideration

When you find yourself unable to meet an in-course requirement because of illness or compassionate reasons, please advise the course instructor in writing, with your name, id#, and e-mail contact. See the academic calendar for information on regulations and procedures for Academic Consideration:
http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml

### Academic Misconduct

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 discourages misconduct. 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.

The Academic Misconduct Policy is detailed in the Undergraduate Calendar:
http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml

### Accessibility

The University of Guelph is committed to creating a barrier-free environment. Providing services for students is a shared responsibility among students, faculty and administrators. This relationship is based on respect of individual rights, the dignity of the individual and the University community's shared commitment to an open and supportive learning environment. Students requiring service or accommodation, whether due to an identified, ongoing disability or a short-term disability should contact the Centre for Students with Disabilities as soon as possible.

For more information, contact SAS at 519-824-4120 ext. 56208 or email accessibility@uoguelph.ca or see the website: http://www.uoguelph.ca/sas


**Course Evaluation Information**

Please see http://www.mathstat.uoguelph.ca/files/TeachevaluationformF10.pdf

**Drop date**

The last date to drop one-semester courses, without academic penalty is *Friday,* **December 3 2021**. For regulations and procedures for Dropping Courses, see the Academic Calendar: http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-drop.shtml

## Additional Course Information

*Additional Course Information will be provided in class.*