

Department of Population Medicine
POPM*6950:
Health Sciences Analytics
Winter 2020
First class: Tuesday, Jan. 7, 10:00, room TBD

Course Coordinator and co-instructor

Olaf Berke
oberke@uoguelph.ca

Co-instructor:

Zvonimir Poljak
zpoljak@uoguelph.ca

Office Hours:

After class or by appointment

Teaching Assistants:

TBD

Class schedule:

Tuesdays and Thursdays, 10:00 – 11:20 am in room TBD

Description:

Machine Learning and Data Science are presented, and are compared to traditional statistical/epidemiological data analysis. The topics will focus on application in clinical as well as public health fields. Hands on examples from textbooks and the instructors own research projects will provide a rich experience and proper skill training.

Objectives:

Students who complete this course should be able to:

- Distinguish between statistics and data science
- Manage data in R/RStudio
- Understand the concepts of various machine learning algorithms for predictive analysis
- Conduct some basic data analysis using case study data

Course approach:

This course will consist of two 80 minute sessions per week. Generally, classes start with discussions around methods and applications as per assigned class readings. This is followed by demonstrations of data analysis using the open software R/RStudio. Students will further their practical skills through homework exercises.

Pre-requisites:

POPM*6210 Epidemiology II is a co-requisite. Students not taking Epidemiology II concurrently and who have not previously taken Epidemiology I will require permission of the course coordinator. Students seeking permission to take this course without the Epidemiology I course require at least one graduate-level course in epidemiology.

Pre-class preparation:

To enhance learning during classroom sessions, pre-class readings (or other preparation) will be assigned. Details will be provided for each session. You are expected to complete the readings prior to the class.

Course Topics

"This is a tentative topics list. The exact list and timing of topics will be provided in class."

1. Data Science and Statistics - Prediction versus Explanation
2. Supervised and Unsupervised Learning
3. Study Design, Variable Pre-processing, General Modelling Strategies
4. (Logistic) Regression and Friends: Lasso and Ridge Regression
5. Evaluation of Predictive Models: Cross-validation strategy and Accuracy Measures
6. Cluster Analysis and Classification and Regression Trees (CART)
7. Machine Learning Algorithms (Random Forrest, Neural Networks, Support Vector Machines and more)
8. Time Series Analysis and Forecasting with Neural Networks
9. Spatial Data Analysis using Ensemble Learning
10. Network Analysis
11. Ethical issues of Machine Learning in the Health Sciences

Overview of evaluation methods:

- Midterm exam = 40% (in-class, *written exam*, e.g. quiz plus open questions)
- Final exam = 60% (this will likely be a *term project*)

Due dates

Assignments and project reports are due on due date and time as indicated. Late submissions will be penalized. Complete details will be provided on CourseLink.

Reading material (available on-line from the UofG library):

Baumer, Kaplan and Horton (2017) *Modern Data Science with R*. CRC Press.

James, Witten, Hastie and Tibshirani (2013) *An Introduction to Statistical Learning*.
Springer

Steyerberg, E. (2019) *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* (Second Edn.). Springer.

Additional research paper as assigned

Software

The statistics section of the course will make use of the open and free “R” software package in combination with “RStudio”. You will have to install R and RStudio onto your own computer (a

respective handout on CourseLink explains the process). **You will have to use your own computer in the final exam!**

CourseLink

Additional course materials and information regarding all assignments will be posted on the CourseLink website. The website can only be accessed by persons who are registered for the class, therefore it is imperative that students ensure that they are registered for the class and have access to CourseLink. If you have any problems accessing CourseLink, contact CCS at x58888 or 58888help@uoguelph.ca

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.

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.

Graduate Calendar - Grounds for Academic Consideration:

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

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. Graduate Calendar - Registration Changes:

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

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.

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. More information can be found on the SAS website: <https://www.uoguelph.ca/sas>

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.

Graduate Calendar - Academic Misconduct

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

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.

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>