CIS4020Fall 2019

Data Science

contact

Dr. Daniel Gillis

Pronouns: he/him/his

dgillis@uoguelph.ca

http://danielgillis.ca

twitter:@drdanielgillis

community partner

Cooperators

land acknowledgment

The Dish With One Spoon Covenant speaks to our collective responsibility to steward and sustain the land and environment in which we live and work, so that all peoples, present and future, may benefit from the sustenance it provides. As we continue to strive to strengthen our relationships with and continue to learn from our Indigenous neighbours, we recognize the partnerships and knowledge that have guided the learning and research conducted in and for this class. We acknowledge that the University of Guelph resides in the ancestral and treaty lands of several Indigenous peoples, including the Attawandaron people and the Mississaugas of the Credit, and we recognize and honour our Anishinaabe, Haudenosaunee, and Métis neighbours. We acknowledge that the work we do here occurs on their traditional lands so that we might work to build lasting partnerships that respect, honour, and value the culture, traditions, and wisdom of those who have lived here since time immemorial.

course description

Data Science focuses on extracting the important relations in data. The course is intended as a survey of the discipline and focuses on applied computational methods for data analysis. Topics include algorithms, computational and machine learning methods, software tools, and modeling, as they apply to the analysis of and discovery in big data.

The Academic Calendars are the source of information about the University of Guelph's procedures, policies, and regulations which apply to undergraduate, graduate, and diploma programs. You can find more information here.

learning outcomes

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

1 **Explain the primary concepts and tools used in data analysis.**

2 **Create systems which analyze, capture, and format large data sets.**

3 **Apply the appropriate algorithms for analyzing different types of data.**

4 **Describe different data types and their characteristics.**

5 **Describe the different statistical and learning algorithms for data analysis.**

6 **Create software which implements analysis algorithms.**

7 **Integrate existing software tools and libraries into an analysis system.**

CIS4020 Fall 2019

Data Science

contact
Dr. Daniel Gillis
Pronouns: he/him/his

dgillis@uoguelph.ca
http://danielgillis.ca
twitter:@drdanielgillis

community partner
Cooperators

lecture, lab & final exam times

You are expected to attend every class and every lab. You are also expected to come to class having reviewed certain topics so that classroom time can be spent applying what you've learned to your course project.

Lecture	**Tues/Thurs 8:30 am - 9:50 am**	MINS 106
Lab	**Mon 8:30 am - 9:20 am**	MINS 103

prerequisites & restrictions

Students enrolled in CIS4020 are expected to have successfully completed the following courses:

CIS2750	**Software Systems Development and Integration**
STAT2040	**Statistics I**
MATH1160	**Linear Algebra I**

course website

Course material, news, announcements, and grades will be regularly posted to the CIS4020 website which can be found here. You are responsible for checking the website regularly.

Community Engagement Materials
You will be spending the entire semester working on a project with a community partner. As such, materials describing community-engaged projects and partners will be provided. It is your responsibility to review these materials before interacting with the client.

Challenge Materials
Articles and other supporting data about the client and their overarching challenge will be provided on the website. It is your responsibility to review these materials before interacting with the client.

Lecture Materials
Selected notes will be posted on the course website as the instructor and TAs have time to make them available. You are expected to take your own notes during class and lab times. Please contact me if you need any accommodations to support your education.

Course Deliverables
Descriptions and rubrics will be made available on the course website for each of the Course Deliverables.

CIS4020Fall 2019

Data Science

contact

Dr. Daniel Gillis
Pronouns: he/him/his

dgillis@uoguelph.ca
http://danielgillis.ca
twitter:@drdanielgillis

community partner

Cooperators

texts

Required	**Doing Data Science** Cathie O'Neil & Rachel Schutt	2014
Suggested	**Data Visualization: A Practical Introduction** Kieran Healy	2018
Suggested	**Statistics for Data Science** James D. Miller	2017
Suggested	**Statistics for Machine Learning** Pratap Dangeti	2017

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 discourages misconduct. All students who take a SOCS course must pass the Academic Integrity Self Test.

For educational purposes, instructors impose conditions on assignments that may limit students' permission to collaborate with others or to utilize external sources (including, but not limited to, software, data, images, text, etc.). Any permitted utilization must be done with proper references. Aiding and abetting is a punishable offence; students must be careful not to help others commit offences by giving out solutions or providing to access computer accounts. Instructors may use automated tools to detect possible cases of academic misconduct.

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.

- The Academic Misconduct Policy is detailed in the Undergraduate Calendar here.
- The SOCS Academic Integrity Unit is here. Login with your central login credentials. The key to use is "imhonest".

CIS4020Fall 2019

Data Science

contact

Dr. Daniel Gillis
Pronouns: he/him/his

dgillis@uoguelph.ca
http://danielgillis.ca
twitter:@drdanielgillis

community partner

Cooperators

grading scheme

The following grading scheme will be used to assess your progress in the course. Your assignments should be submitted in PDF format unless otherwise stated by the instructor. If you are having difficulty with an assignment, please discuss this with the instructor in advance of the due date. Do not wait until the last day to do so.

Deliverable 1 **Review Assignment: due any time between September 27 at 5:00 pm & September 30th at 9:00 am** 10%
Topics: Data science methods, statistics, linear algebra, philosophy of data science

Deliverable 2 **Midterm Presentation & Critical Review: due in class October 8 or October 10** 20%
Topics: Problem identification, data description, proposed method(s)

Deliverable 3 **Method Chapter, Presentation & Critical Review: due in class November 5, 7, 12, 14, 19, or 21** 25%
Topics: Literature review, mathematical/algorithmic review of method, code, questions, and application of method

Deliverable 4 **Poster Presentation & Critical Review: due in class November 28** 20%
Topics: Poster and oral presentation of your project, highlighting research question, methods used, results, interpretations, future work

Deliverable 5 **Final Report & Reflections: due any time between November 29 at 5:00 pm & December 2nd at 9:00 am** 25%
Topics: Final Project Report

Missed Labs: if you miss a lab, you are affecting your team's ability to complete the final project. If you are going to miss a lab, please discuss this with your team first. You must then present to the instructor a plan for how you and your team will deal with your absence.

Missed Assessments: if you are unable to meet an in-course requirement due to medical, psychological, or compassionate reasons, please make an appointment to meet your course instructor. Please see here for specific details and consult the undergraduate calendar for information on regulations and procedures for Academic Consideration.

Accommodation of Religious Obligations: if you are unable to meet an in-course requirement due to religious obligations, please email the instructor ***within two weeks of the start of the semester*** to make alternate arrangements. See here for information on regulations and procedures for Academic Accommodation of Religious Obligations.

Late Assigments/Lab Demos: the teaching team will begin grading course deliverables shortly after the last moment they are due. If you have not submitted a course deliverable on time, it will be considered late. You may submit any course deliverable late, however, your final grade will be reduced depending on how late it is. For example, if you achieved a grade of 95% on assignment one, your recorded grade would be:

95%	**if the assignment was submitted between September 27 at 5:00 pm and September 30th at 9:00 am (i.e. not late)**
85%	**if the assignment was submitted between September 30th at 9:01 am and October 1 at 9:00 am (i.e. late by up to one day)**
75%	**if the assignment was submitted between October 1 at 9:01 am and October 2 at 9:00 am (i.e. late by up to two days)**
65%	**if the assignment was submitted between October 2 at 9:01 am and October 3 at 9:00 am (i.e. late by up to three days)**
0%	**if the assignment was submitted after October 3 at 9:00 am (i.e. late more than three days)**

Please note that you will receive a 0 on any course deliverable that is submitted more than three days late. If there are extenuating circumstances that could explain why you need more than three days extension, you must speak with the instructor before the last due date of the deliverable. For example, you must contact the instructor by September 30th at 9:00 am to explain why you'd need more than three days extension to complete assignment 1. As part of this conversation, we will work to identify a new due date, and the expected penalty for the late submission.

Regrades: if you feel your assignment has been graded incorrectly, you must, in writing, present your case to the instructor. Be specific about what you believe was graded incorrectly. All regraded material will be *completely* regraded. This could result in your grade being reduced.

roles and responsibilities

- **Communication & Email Policy** Major announcements will be posted to the course website. It is your responsibility to check the course website regularly. As per university regulations, all students are required to check their mail.uoguelph.ca e-mail account regularly: e-mail is the official route of communication between the University and its students.
- **Recording of Materials** Presentations which are made in relation to course work, including lectures, cannot be recorded or copied without the permission of the presenter, whether the instructor, classmate, or guest lecturer. Material recorded with permission is restricted to use for that course and may not be posted on any public space unless further permission is granted.
- **Instructor's Role and Responsibility to Students** The instructor's role is to develop and deliver course material in ways that facilitate learning for a variety of students. Selected notes will be made available to students on the course website but are not intended to be stand-alone. During lectures, the instructor will expand and explain the content of notes and provide example problems that supplement posted notes. Scheduled classes will be the principal venue to provide information and feedback for exams and assignments.
- **Students' Learning Responsibilities** Students are expected to take advantage of the learning opportunities provided during lectures, labs, and help sessions. Students, especially those having difficulty with the course content, should also make use of other resources recommended by the instructor. Students who fall behind due to illness, work, or extra-curricular activities are advised to keep the instructor informed as early as possible. This will allow the instructor to recommend extra resources in a timely manner and/or provide consideration if appropriate.

CIS4020Fall 2019

Data Science

contact

Dr. Daniel Gillis
Pronouns: he/him/his

dgillis@uoguelph.ca
http://danielgillis.ca
twitter:@drdanielgillis

community partner

Cooperators

code of conduct

Because this class needs to be a participatory community if students are to fulfill their potential for learning, people who disrupt the community by their words or actions disrupt that community. Disrespectful speech and disruptive behaviour have a negative impact on everyone's learning. Aggressive behaviour, bullying, abusive language, sexism, agism, ablism, racism, homophobia, transphobia, islamophobia, etc. will not be tolerated. We are all entitled to ask questions and hold opinions, but these must be presented in a respectful manner. All discussions must be respectful, regardless of how much you disagree with another opinion. We will refer to and address our peers using their names and pronouns. We will work towards a space where everyone has a voice, can be heard, and is accepted and welcomed.

Our learning environment must be a friendly, safe, and welcoming environment for all, regardless of ethnicity, gender, sexual orientation, ability, socioeconomic status, and/or religion (or lack thereof). As we wish to facilitate and encourage the fullest participation from everyone, this code of conduct outlines the expectations for all participants [1].

- **Expected Behaviour:** Members of our community are expected to 1) participate in an authentic and active way, and in doing so, contribute to the health and value of our community, 2) exercise consideration and respect in their speech and actions, 3) attempt collaboration before conflict, 4) refrain from demeaning, discriminatory, or harassing behaviour and speech, and 5) be mindful of your surroundings and your fellow participants, and alert community leaders (e.g. your instructor) if you notice a dangerous situation, someone in distress, or violation of this Code of Conduct, even if they seem inconsequential.
- **Citizenship & Participation:** Communities mirror the societies in which they exist and positive action is essential to counteract the many forms of inequality and abuses of power that exist in society. If you see someone who is making an extra effort to ensure our community is welcoming, friendly, and encourages all participants to contribute to the fullest extent, we want to know.
- **Unacceptable Behaviour:** Unacceptable behaviours include: intimidating, harassing, abusive, discriminatory, derogatory or demeaning speech or actions by any participant in our community, either in person, online, at any related events, or in one-on-one communications carried out in the context of community business. Harassment includes: harmful or prejudicial verbal or written comments related to race, religion, disability, gender, sexual orientation; inappropriate use of nudity and/or sexual images in public spaces (including computer labs and presentation slides); deliberate intimidation, stalking or following; harassing photography or recording; sustained disruption of talks or other events; inappropriate physical contact, and unwelcome sexual attention.
- **Consequences of Unacceptable Behaviour:** Unacceptable behaviour from any community member, including the course instructor and those members with decision-making authority, will not be tolerated. Anyone asked to stop unacceptable behaviour is expected to comply immediately. If a community member engages in unacceptable behaviour, action will be taken to ensure that such behaviour ends, beginning with action on the part of the course instructor, and escalating if necessary. Additional information on University policy regarding harassment, conduct, and human rights is available here.

[1] Based on citizen code of conduct, distributed under a Creative Commons Attribution-ShareAlike license

CIS4020Fall 2019

Data Science

contact

Dr. Daniel Gillis
Pronouns: he/him/his

dgillis@uoguelph.ca
http://danielgillis.ca
twitter:@drdanielgillis

community partner

Cooperators

- **If You Witness or Are Subject to Unacceptable Behaviour or Have Any Other Concerns:** please notify the course instructor as soon as possible. If you feel that the course instructor cannot or will not provide remedy for the situation, please contact any of these alternate resources: 1) Associate Director, Undergraduate (click here to email), 2) Director of the School (click here to email), 3) Associate Dean, Academic (click here to email), 4) Office of Diversity and Human Rights (click here to email, or ext 53000), or 5) Campus Community Police (ext 52245).

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 for a short-term disability should contact Student Accessibility Services (SAS) as soon as possible.

For more information, contact SAS at 1.519.824.4120 ext 56208 or email csd@uoguelph.ca or see the website here.

children in the class

Currently, the University of Guelph does not have a formal policy on children in the classroom. The policy described here is just a reflection of my own beliefs and commitments to student, staff, and faculty parents.

- All breastfeeding babies are welcome in class as often as is necessary.
- For older children and babies, I understand that unforeseen disruptions in childcare often put parents in the position of having to miss class to stay home with a child. While this is not meant to be a long-term childcare solution, occasionally bringing a child to class in order to cover gaps in care is perfectly acceptable.
- I ask that all students work with me to create a welcoming environment that is respectful of all forms of diversity, including diversity in parenting status.
- In all cases where babies and children come to class, I ask that you sit close to the door so that if your little one needs special attention and is disrupting learning for other students, you may step outside until their need has been met.
- Finally, I understand that often the largest barrier to completing your coursework once you become a parent is the tiredness many parents feel in the evening once children have finally gone to sleep. While I maintain the high expectations for all students in my classes regardless of parenting status, I am happy to problem-solve with you in a way that makes you feel supported as you strive for school-parenting balance.

CIS4020Fall 2019

Data Science

contact

Dr. Daniel Gillis
Pronouns: he/him/his

dgillis@uoguelph.ca
http://danielgillis.ca
twitter:@drdanielgillis

community partner

Cooperators

health & wellness

All members of campus play a role in fostering and promoting a safe, supportive environment, as well as good physical, emotional, spiritual, cultural, and mental health and wellness.

> ***If you are sick, heartbroken, or exhausted, go home. Work is not more important than your health.*** -Dr. Max Liboiron
>
> ***If you are experiencing any challenges, please do not hesitate to contact the instructor, and know that there are resources on campus set up to help you out.***
>
> ***Medical concerns?*** Student Health Services at x52131
>
> ***Threats of violence, personal safety?*** Campus police at x2000
>
> ***Psychological or emotional concerns?*** Counselling services at x53244
>
> ***Accessibility concerns?*** SAS at x56208
>
> ***Sexual assault?*** Campus police at x2000, or counselling services at x53244
>
> ***Mental Health concerns?*** Please see the Mental Health Resources page here.

Other sources of help can be found at the following links:

- Student Health Services, Monday to Friday, 8:30am-4:30pm, x52131, J.T. Powell Building
- Couselling Services, Monday to Friday, 8:15am-4:15pm, x53244, Level 3, University Centre
- Wellness Education Centre, Monday to Friday, 8:30am-4:30pm, x53327, J.T. Powell Building
- Student Support Network, Monday to Friday, 12:00pm-10:00pm, Raithby House
- Campus Community Police, 24/7, x2000, Trent Building
- Good2Talk, 1.866.925.5454
- Here 24/7, 1.844.437.3427

CIS4020Fall 2019

Data Science

contact

Dr. Daniel Gillis
Pronouns: he/him/his

dgillis@uoguelph.ca
http://danielgillis.ca
twitter:@drdanielgillis

community partner

Cooperators

topics by date

WEEK	TOPICS COVERED†
Intro	**Introductions, Project & CES**
Week 1-4	**What is Data Science?** **Meet the Client** **Statistics Review** **Linear Algebra Review** **Critical Review**
Week 5	**Midterm Presentations**
Week 6-8	**Project Development**
Week 9-11	**Presentation of Data Science Methods**
Week 12	**Final Poster Presentation**