Western Science

Department of Statistical and Actuarial Sciences

DS 3000B – Introduction to Machine Learning
Crosslisted with DS 9000B

1. Course Information

Course Information
Academic Term: Winter 2021/22

Lectures:
Tuesday 4:30pm to 6:30pm
Thursday 4:30pm to 6:30pm
Place: NS-7 / Online Asynchronous with a Synchronous component via Zoom until the Faculty of Science says otherwise

List of Prerequisites
(Data Science 1200A/B or Computer Science 1026A/B or Computer Science 1027A/B or Computer Science 2120A/B or Digital Humanities 2220A/B or Engineering Science 1036A/B or Data Science 2000A/B or Integrated Science 2002B or Statistical Sciences 2864A/B); (Data Science 2000A/B or Integrated Science 2002B or Statistical Sciences 2857A/B or 0.5 course from the Introductory Statistics Course List); (Mathematics 1600A/B or Numerical and Mathematical Methods 1411A/B or the former Applied Mathematics 1411A/B or Data Science 2100A); (Calculus 1000A/B or Calculus 1500A/B or Numerical and Mathematical Methods 1412A/B or the former Applied Mathematics 1412A/B or Data Science 2100A). Note that Data Science 2000A/B, Integrated Science 2002B and Data Science 2100A can be used to fulfill multiple prerequisites.

List of Antirerequisites
The former Computer Science 4414A/B, the former Statistical Sciences 3850F/G, the former Software Engineering 4460A/B

Unless you have either the requisites for this course or written special permission from your Dean to enroll in it, you may be removed from this course and it will be deleted from your record. This decision may not be appealed. You will receive no adjustment to your fees in the event that you are dropped from a course for failing to have the necessary prerequisites.

2. Instructor Information

Instructors	Email	Office	Phone	Office Hours
Cristián Bravo Román	cbravoro@uwo.ca	WSC 280	x87665	Wed 11am – 12pm

Sherly Alfonso-Sánchez	salfonso@uwo.ca	WSC 208		See OWL
Duo Xu	dxu258@uwo.ca			See OWL
Sahab Zandi	szandi@uwo.ca	WSC 208		See OWL
Yuhao (Jet) Zhou	yzho82@uwo.ca	WSC 208		See OWL

Students must use their Western (@uwo.ca) email addresses when contacting their instructors.

3. Course Syllabus, Schedule, Delivery Mode

Introduces machine learning and statistical methods for data analysis through applied examples. The goal of this course to expose students to topics related to statistical learning such as, Linear Regression, Logistic Regression, Discriminant Analysis, Model Selection and Regularization, Cross Validation, Tree Based Methods and Clustering. The course emphasizes the ability to apply techniques to real data sets and critically evaluate their performance.

Topics include:

- Supervised Learning and Model Fitting
- Statistics, Prediction, and Maximum Likelihood
- Introduce test set/out-of-sample idea.
- Classification, Evaluation, Logistic regression Regularization, Multi-class problems
- Estimating Performance, Quantifying Uncertainty on parameter estimates and on model predictions
- Test error, Cross-validation, Model Selection, Bias-Variance tradeoff
- Feature Selection and Regularization (L1 and L2)
- Trees, Random Forest
- Neural Networks, Gradients, learning
- Autoencoders, Dimensionality reduction, PCA, NMF, tSNE
- Clustering, K-means, hierarchical clustering
- Model limitations, Causality.

Table of Contents and Schedule

Week	Dates	Lecture	Lab	Assignment
1	Jan 10 – 14	Supervised Learning and Model Fitting	Pandas and Numpy, Optimization	Linear regression, Loss function Squared error Mean absolute deviation
2	Jan 17 – 21	Probability and Maximum Likelihood	Pandas	Regression
3	Jan 24 – 28	Introduce test set/out-of-sample idea. Classification, Evaluation, Logistic regression Regularization,	Implementation of topics	Logistic regression, Classification, Evaluation

		Multi-class problems		
4	Jan 31 – Feb 4	Estimating Performance, Quantifying Uncertainty	Bootstrap	Bootstrap and confidence intervals
5	Feb 7 – 11	Test error, Cross-validation, Model Selection, Bias-Variance tradeoff	Cross-validation	Cross-validation and model selection
6	Feb 14 – 18	Feature Selection and Regularization (L1 and L2)	Regularization	Regularization and nested cross-validation
7	Feb 21 – 25	Reading Week	-	-
8	Feb 28 – Mar 4	**Midterm**	**Midterm**	-
9	Mar 7 – 11	Trees, Random Forest	Tree Lab	Tree homework
10	Mar 14 – 18	Autoencoders, Dimensionality reduction, PCA, NMF, t-SNE	Dimensionality reduction	Dimensionality reduction
11	Mar 21 – 25	Clustering, K-means, hierarchical clustering	Clustering	Clustering
12	Mar 28 – Apr 1	Model limitations, Causality.	Deploying models	Interpretability
13	Apr 4 – Apr 8	Neural Networks, Gradients, learning	Simple 1-hidden layer network	Study guide for the exam including ANN

Classes begin January 10th, 2021
Reading Week: February 19–27, 2022
Classes end April 7th, 2022

Contingency plan for an in-person class pivoting to 100% online learning

Although the intent is for this course to be delivered in-person to the extent possible, the changing COVID-19 landscape may necessitate some or all the course to be delivered online, with both a synchronous (i.e., at the times indicated in the timetable) and an asynchronous (e.g., posted on OWL for students to view at their convenience) component. The grading scheme will not change. Any assessments affected will be conducted online as determined by the course instructor.

When deemed necessary, tests and examinations in this course will be conducted using a remote proctoring service. By taking this course, you are consenting to the use of this software and acknowledging that you will be required to provide personal information (including some biometric data) and that the session will be recorded. Completion of this course will require you to have a reliable

internet connection and a device that meets the technical requirements for this service. More information about this remote proctoring service, including technical requirements, is available on Western's Remote Proctoring website at: https://remoteproctoring.uwo.ca.

4. Course Materials

Core book:

The Elements of Statistical Learning by Hastie, Tibshirani and Friedman. [**Available for Free Online at https://web.stanford.edu/~hastie/ElemStatLearn/**]

Recommended complementary book:
Machine Learning: A Probabilistic Perspective by P. Kevin Murphy [**Free** online]

Students are responsible for checking the course OWL site (http://owl.uwo.ca) on a regular basis for news and updates. This is the primary method by which information will be disseminated to all students in the class.

All course material will be posted to OWL: http://owl.uwo.ca.

If students need assistance with the course OWL site, they can seek support on the OWL Help page. Alternatively, they can contact the Western Technology Services Helpdesk. They can be contacted by phone at 519-661-3800 or ext. 83800.

Google Chrome or Mozilla Firefox are the preferred browsers to optimally use OWL; update your browsers frequently. Students interested in evaluating their internet speed, please click here.

While self-installation of the software in your own computer is possible, there is also the possibility of using online platforms. Two are available:

- Google Colab: https://colab.research.google.com
- Kaggle Kernels: https://www.kaggle.com/code

Technical Requirements

This is a mostly code-based course so a laptop with internet connection is required. If making your own local installation, a computer with a sufficiently powerful processor (at least two cores @2.2 GHz) with at least 8GB of RAM are recommended. If this were not available, we recommend using an online environment.

5. Methods of Evaluation

The overall course grade will be calculated as listed below:

Weekly Assignments (9)	36%
Midterm Exam	31%
Final Exam	33%

Weekly Assignment:

Assignments will be released each week, with due dates of the following week. There will be no make-up for missed weekly assignments. If the student submits a self-reported absence before the due date of the assignment, an extension of the deadline will be granted in line with the University policies.

Midterm:
The midterm will be a practical examination in the form of a timed assignment. Students will be given a data set and a set of practical data analytic problems to solve, similar to the structure of the weekly assignments. The exam is "open book & open web", meaning that students can access any notes or any documents on the web. Electronic communication with other people inside or outside class is prohibited.

Final Exam:
The final exam will be scheduled by the Registrar. The exam is "open book & open web", meaning that students can access any notes or any documents on the web. Electronic communication with other people inside or outside class is prohibited. It will cover concepts from the entire course and is in structure similar to the midterm exam. The exam will be a practical examination; each student will need a laptop to complete the midterm.

In order to pass the course, a mark above 50% must be obtained in the written examination section of the course (midterm + exam).

Accommodated Evaluations

- Late assessments without illness self-reports will be subject to a late penalty discount of 10%/day (this means if your coursework gets an 80%, and you submit one day late, your final mark will be 80% - 10% = 70%). The day late starts one minute after the deadline of the original assessment has passed. There are **NO EXCEPTIONS** to this policy.
- Late assessments with illness self-reports should be submitted within 24 hours of submission of the last illness self-report.
- An assessment cannot be submitted after it has been returned to the class. In case of a missed assignment with justified cause, the weight will be transferred to the other assignments in the same category (assessments or midterm/exam).
- If permission to waive the requirement that students receive evaluation on work totaling 15% of their final grade at least three days prior to the deadline for withdrawal without academic penalty has been obtained from the Dean's Office, a statement to this effect must be made.
- If the final examination was missed, there will be a special examination during normal special examination period.

6. Student Absences

Academic Consideration for Student Absences
Students who experience an extenuating circumstance (illness, injury or other extenuating circumstance) sufficiently significant to temporarily render them unable to meet academic requirements may submit a request for academic consideration through the following routes:

(i) Submitting a Self-Reported Absence (SRA) form provided that the conditions for submission are met. To be eligible for a Self-Reported Absence:

- an absence must be no more than 48 hours
- the assessments must be worth no more than 30% of the student's final grade (so only assessment can be SRAd in this course).
- no more than two SRAs may be submitted during the Fall/Winter term

(ii) For medical absences, submitting a Student Medical Certificate (SMC) signed by a licensed medical or mental health practitioner to the Academic Counselling office of their Faculty of Registration.

(iii) Submitting appropriate documentation for non-medical absences to the Academic Counselling office in their Faculty of Registration.

Note that in all cases, students are required to contact their instructors within 24 hours of the end of the period covered. **If an SRA is not notified, it will be considered invalid, and the work will receive a mark of 0.**

Students should also note that individual instructors are not permitted to receive documentation directly from a student, whether in support of an application for consideration on medical grounds, or for other reasons. **All documentation required for absences that are not covered by the Self-Reported Absence Policy must be submitted to the Academic Counselling office of a student's Home Faculty.**

For the policy on Academic Consideration for Student Absences – Undergraduate Students in First Entry Programs, see:

https://www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_illness.pdf

and for the Student Medical Certificate (SMC), see:

http://www.uwo.ca/univsec/pdf/academic_policies/appeals/medicalform.pdf.

Religious Accommodation

When a course requirement conflicts with a religious holiday that requires an absence from the University or prohibits certain activities, students should request accommodation for their absence in writing at least two weeks prior to the holiday to the course instructor and/or the Academic Counselling office of their Faculty of Registration. Please consult University's list of recognized religious holidays (updated annually) at

https://multiculturalcalendar.com/ecal/index.php?s=c-univwo.

Absences from Final Examinations

If you miss the Final Exam, please contact the Academic Counselling office of your Faculty of Registration as soon as you are able to do so. They will assess your eligibility to write the Special Examination (the name given by the University to a makeup Final Exam).

You may also be eligible to write the Special Exam if you are in a "Multiple Exam Situation" (e.g., more than 2 exams in 23-hour period, more than 3 exams in a 47-hour period).

If a student fails to write a scheduled Special Examination, the date of the next Special Examination (if granted) normally will be the scheduled date for the final exam the next time this course is offered. The maximum course load for that term will be reduced by the credit of the course(s) for which the final examination has been deferred. See the Academic Calendar for details (under Special Examinations).

7. Accommodation and Accessibility

Accommodation Policies
Students with disabilities work with Accessible Education (formerly SSD), which provides recommendations for accommodation based on medical documentation or psychological and cognitive testing. The policy on Academic Accommodation for Students with Disabilities can be found at:

https://www.uwo.ca/univsec/pdf/academic_policies/appeals/Academic Accommodation_disabilities.pdf,

8. Academic Policies

The website for Registrarial Services is http://www.registrar.uwo.ca.

In accordance with policy,

https://www.uwo.ca/univsec/pdf/policies_procedures/section1/mapp113.pdf,

the centrally administered e-mail account provided to students will be considered the individual's official university e-mail address. It is the responsibility of the account holder to ensure that e-mail received from the University at his/her official university address is attended to in a timely manner.

A laptop computer with internet connection is required for both the midterm and the exam.

Scholastic offences are taken seriously and students are directed to read the appropriate policy, specifically, the definition of what constitutes a Scholastic Offence, at the following Web site:

http://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf.

All required papers may be subject to submission for textual similarity review to the commercial plagiarism detection software under license to the University for the detection of plagiarism. All papers submitted for such checking will be included as source documents in the reference database for the purpose of detecting plagiarism of papers subsequently submitted to the system. Use of the service is subject to the licensing agreement, currently between The University of Western Ontario and Turnitin.com (http://www.turnitin.com).

Tests and examinations in this course will be conducted using a remote proctoring service. By taking this course, you are consenting to the use of this software and acknowledge that you will be required to provide **personal information** (including some biometric data) and the session will be **recorded**. Completion of this course will require you to have a reliable internet connection and a device that meets the technical requirements for this service. More information about this remote proctoring service, including technical requirements, is available on Western's Remote Proctoring website at:

https://remoteproctoring.uwo.ca.

9. Support Services

Please visit the Science & Basic Medical Sciences Academic Counselling webpage for information on adding/dropping courses, academic considerations for absences, appeals, exam conflicts, and many other academic related matters: https://www.uwo.ca/sci/counselling/.

Please contact the course instructor if you require lecture or printed material in an alternate format or if any other arrangements can make this course more accessible to you. You may also wish to contact Accessible Education at (519) 661-2147 if you have any questions regarding accommodations.

Western University is committed to a thriving campus as we deliver our courses in the mixed model of both virtual and face-to-face formats. We encourage you to check out the Digital Student Experience website to manage your academics and well-being: https://www.uwo.ca/se/digital/.

Learning-skills counsellors at the Student Development Centre (http://www.sdc.uwo.ca) are ready to help you improve your learning skills. They offer presentations on strategies for improving time management, multiple-choice exam preparation/writing, textbook reading, and more. Individual support is offered throughout the Fall/Winter terms in the drop-in Learning Help Centre, and year-round through individual counselling.

Students who are in emotional/mental distress should refer to Mental Health@Western (http://www.health.uwo.ca/mentalhealth) for a complete list of options about how to obtain help.

Additional student-run support services are offered by the USC, http://westernusc.ca/services.

10. Accreditation

This course is accredited under the Canadian Institute of Actuaries (CIA) University Accreditation Program (UAP) for the 2021-22 academic year. Achievement of the established exemption grade in this course may qualify a student from exemptions from writing certain preliminary exams.

Please see the following link for full details:

http://www.cia-ica.ca/membership/university-accreditation-program---home

In addition to the university's internal policies on conduct, including academic misconduct, candidates pursuing credits for writing professional examinations shall also be subject to the **Code of Conduct and Ethics for Candidates in the CIA Education System** and the associated **Policy on Conduct and Ethics for Candidates in the CIA Education System**.

This course with a minimum mark of 80%, along with Statistics 3859A (minimum mark of 80%), can give you an exemption for the SRM exam.