**** **Department of Statistical and Actuarial Sciences**

 **DS 3000B – Introduction to Machine Learning Course Outline**

**Cross listed with DS 9000B**

**1. Course Information**

**Course Information**

Academic Term: Winter 2024/25

Lectures:

Tuesday 4:30pm to 6:30pm, MC 110.

Thursday 4:30pm to 6:30pm, UCC146.

**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.

Unless you have either the requisites for this course or written special permission from your Science Academic Advisors 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.

**List of Antirequisites**

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

**2. Instructor Information**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Instructors** | **Email** | **Office** | **Phone** | **Office Hours** |
| Dr. Cristián Bravo(Course Coordinator) | cbravoro@uwo.ca | WSC 280 | (519) 661-2098 x87565 | Tuesdays 3pm – 4pm |
| TA Team TBD |  |  |  |  |

Students must use their Western (@uwo.ca) email addresses when contacting their instructors. The course will have an active Discord channel for fast communication. Students must use their real name as nicknames, or they will not be accepted/removed from the Discord.

Office hours will be in person unless agreed-to by email.

**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 Regularized 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
* Classification, evaluation, logistic regression regularization, multi-class problems.
* Test error, Cross-validation, Model Selection, Bias-Variance tradeoff, uncertainty and bootstrapping.
* Feature Selection and Regularization (L1 and L2)
* Trees, XGBoost, Random Forest
* Dimensionality reduction, PCA, NMF, tSNE, UMAP.
* Unsupervised learning: Clustering, K-means, hierarchical clustering, autoencoders.
* Fairness, Accountability, Explainability and Transparency.
* Neural networks and an introduction to deep learning.

**Table of Contents and Schedule**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Week** | **Week Starting** | **Lecture** | **Lab** | **Homework** |
| 1 | Jan 06 | Supervised Learning and Model Fitting   | Pandas and Numpy  | Linear regression, Loss function  Squared error   Mean absolute deviation |
| 2 | Jan 13 | Probability and Maximum Likelihood   | Polars | Regression  |
| 3 | Jan 20 | Introduce test set/out-of-sample idea, Test error.  Classification, Evaluation, Logistic regression Regularization, Multi-class problems   | Running a logistic model | Logistic regression, Classification, Evaluation |
| 4 | Jan 27 | Model Selection, Cross-validation, Bias-Variance tradeoff , Bootstrap | Uncertainty | Uncertainty and model selection |
|  5 | Feb 03 | Feature Selection and Regularization (L1 and L2)   | Regularization   | Regularization and nested cross-validation   |
| 6 | Feb 10 | Trees, XGB, Random Forest    | Tree Lab   | Tree homework (**deadline Mar 03 after midterm**) |
| 7 | Feb 17 | Reading Week | - | - |
| 8 | Feb 24 | **Midterm prep**  | **Midterm**   | - |
| 9 | Mar 03 | Dimensionality reduction, PCA, NMF, t-SNE, UMAP | Dimensionality reduction  | Dimensionality reduction  |
| 10 | Mar 10 | Clustering, K-means, hierarchical clustering, autoencoders |   Clustering  |  Clustering  |
| 11 | Mar 17 | Fairness  | Fairness | Fairness |
| 12 | Mar 24 | Neural Networks, Gradients, learning   | Dense Networks | Dense Networks |
| 13 | Mar 31 | Deep Learning  | Introduction to Deep Learning | No homework. **Term project deadline Friday 5pm.** |

Classes begin: September 5, 2024; January 6, 2025

Fall Reading Week: October 12 – 20

Spring Reading Week: February 15 – 23, 2025

Classes end: December 6, 2024; April 4, 2025

Exam period: December 9 – 22, 2024; April 7 – 30, 2025

**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/**](https://web.stanford.edu/~hastie/ElemStatLearn/) or with a Western account at <https://link.springer.com/book/10.1007/978-0-387-84858-7>]

Recommended complementary book:

Machine Learning: A Probabilistic Perspective by P. Kevin Murphy [**Free** [online](https://www.cs.ubc.ca/~murphyk/MLbook/)]

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](https://www.google.ca/chrome/?brand=CHBD&gclid=CjwKCAjwxLH3BRApEiwAqX9arfg8JaH6fWGASk9bHTkfW_dyBir93A1-TliP-7u1Kguf-WZsoGAPhBoC9NYQAvD_BwE&gclsrc=aw.ds) or [Mozilla Firefox](https://www.mozilla.org/en-CA/) are the preferred browsers to optimally use OWL; update your browsers frequently. Students interested in evaluating their internet speed, please click [here.](https://www.google.com/search?q=internet+speed+test+google&rlz=1C5CHFA_enCA702CA702&oq=internet+speed+test+google&aqs=chrome..69i57j0l7.3608j0j4&sourceid=chrome&ie=UTF-8)

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**

**Grading Scheme and Assessment Dates**

The overall course grade will be calculated as listed below:

Homework 10%

Midterm Test 25%

Term Project 30%

Final Exam 35%

* Every week, a pass / fail **homework** will be provided. They will have a deadline on Mondays 5pm. No late submissions will be accepted. The assignments will come with a selection of multiple-choice questions (50%) and a code-based problem (50%) dealing with the class topics of the week. **These are the study guides of the course**. The mark for this activity will be calculated by the best 9 out of the 10 homework assignments of the semester, using the formula (# passed homework) / (# total homework) \* 100.
* The midterm is **a closed-book test** with multiple choice and/or written problems, following the structure of the analysis and multiple-choice questions in the homework. The midterm is the designated examination that cannot be skipped using self-reported undocumented absences. It will include up to week 7. It will occur in the second slot of the class, after reading week. Grad students have an extra question they must complete in the given time. This question is optional for UG students, providing up to 10% extra marks (TBC, max. 100%). Note no extra time will be provided for PG students.
* The exam is **a closed-book test** that covers all topics of the semester following the same structure as the midterm. It will be scheduled by the registrar. Grad students have an extra question they must complete in the given time. This question is optional for UG students, providing up to 10% extra marks (TBC, max. 100%). Note no extra time will be provided for PG students.
* Western has signed an agreement with the company Environics, which has provided unique, real data that we will use for a term project. The project, **to be done in groups of up to 4 students**, will be released the week after the midterm with a deadline on **Friday April 4th, 5pm. Starting at 5:05pm** that day, a 10-point late penalty will be incurred, increasing by 10 points every 24 hours (i.e. if you submit at 5.05pm on Friday the 4th, and your mark is 70%, you will get 60%. If you submit at 5.05pm on Monday, you will get 70% - 3\* 10% = 40%). We thank Environics for their generous support for our learning experience.

**GENERAL INFORMATION ABOUT MISSED COURSEWORK**

**Graduate students:**

You must provide valid documentation for medical-related absences or adequate supporting documentation for non-medical absences to the instructor as soon as possible and no later than 48 hours after the missed assessment.

Documentation for medical illness must include the completion of a Western Student Medical Certificate (SMC) or, where that is not possible, equivalent documentation, by a health care practitioner.

For other types of absences, proper documentation must be provided.

**Undergraduate students:**

Students must familiarize themselves with the *University Policy on Academic Consideration – Undergraduate Students in First Entry Programs* posted on the Academic Calendar:

<https://www.uwo.ca/univsec/pdf/academic_policies/appeals/academic_consideration_Sep24.pdf>,

This policy does not apply to requests for Academic Consideration submitted for **attempted or completed work**, whether online or in person.

The policy also does not apply to students experiencing longer-term impacts on their academic responsibilities. These students should consult [Accessible Education](http://academicsupport.uwo.ca/accessible_education/).

For procedures on how to submit Academic Consideration requests, please see the information posted on the Office of the Registrar’s webpage:

<https://registrar.uwo.ca/academics/academic_considerations/>

All requests for Academic Consideration must be made within 48 hours after the assessment date or submission deadline.

All Academic Consideration requests normally must include supporting documentation; however, recognizing that formal documentation may not be available in some extenuating circumstances, the policy allows students to make one Academic Consideration request **without supporting documentation** in this course. However, the following assessments are excluded from this, and therefore always require formal supporting documentation:

* Examinations scheduled during official examination periods (Defined by policy)
* Midterm (designated by the instructor as the one assessment that always requires documentation when requesting Academic Consideration).

When a student *mistakenly* submits their one allowed Academic Consideration request **without supporting documentation** for the assessments listed above or those in the **Coursework with Assessment Flexibility** section below, the request cannot be recalled and reapplied. This privilege is forfeited.

**EVALUATION SCHEME FOR MISSED ASSESSMENTS (GRADUATE AND UNDERGRADUATE STUDENTS)**

**Exam and/or Midterm**: When a student misses the Final Exam and their Academic Consideration has been granted, they will be allowed to write the Special Examination (the name given by the University to a makeup Final Exam). See the Academic Calendar for details (under [Special Examinations](https://www.westerncalendar.uwo.ca/PolicyPages.cfm?Command=showCategory&PolicyCategoryID=5&SelectedCalendar=Live&ArchiveID=#SubHeading_70)), especially for those who miss multiple final exams within one examination period. This Special Examination **will test** **all topics of the semester, independently if the exam or midterm was missed**. The mark in this examination will replace the mark of the missed assignment(s).

**Term Coursework:** If the deadline for the coursework is missed using Academic Considerations, a one-week extension will be provided. No further extensions will be allowed.

**Homework:** This course has 10 homework assignments, and we will drop up to 1 failed or missed homework, using the rest to calculate your homework grade. Should extenuating circumstances arise, students do not need to request Academic Consideration or provide documentation for the first missed homework (both UG and PG). Academic consideration requests will be denied for this missed submission. You can use Academic Considerations (UG) or proper documentation (PG) for any extra homework, and in that case, we will not count that one for your homework grade.

**TO PASS THE COURSE, A MARK ABOVE 50% (60% PG) MUST BE OBTAINED IN THE WRITTEN EXAMINATION SECTION OF THE COURSE (5/12 \* MIDTERM + 7/12 \* EXAM).**

**6. Additional Statements**

**Religious Accommodation**

When conflicts with a religious holiday that requires an absence from the University or prohibits certain activities, students should request an accommodation for their absence in writing to the course instructor and/or the Academic Advising office of their Faculty of Registration. This notice should be made as early as possible but not later than two weeks prior to the writing or the examination (or one week prior to the writing of the test).

Please visit the Diversity Calendars posted on our university’s EDID website for the recognized religious holidays:

[https://www.edi.uwo.ca](https://www.edi.uwo.ca/).

**Accommodation Policies**

Students with disabilities are encouraged to contact Accessible Education, 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](https://www.uwo.ca/univsec/pdf/academic_policies/appeals/Academic%20Accommodation_disabilities.pdf).

**Academic Policies**

The website for Registrar Services is <https://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 their official university address is attended to in a timely manner.

**No electronic devices will be allowed in the exam, midterm, and special examination. If a student is found using an electronic device, it will be considered a Scholastic Offence (see below).**

**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:

<https://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf>.

All required papers (homework, exams, term project, midterm, etc.) 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](http://www.turnitin.com/)).

Computer-marked multiple-choice tests and exams may be subject to submission for similarity review by software that will check for unusual coincidences in answer patterns that may indicate cheating.

**Only in case of health lockdowns**, 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](https://remoteproctoring.uwo.ca/).

**All lectures will be recorded and made available at the course’s OWL site. Please note your voice may be recorded when attending lectures.**

**Support Services**

Please visit the Science & Basic Medical Sciences Academic Advising 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/>.

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

Western is committed to reducing incidents of gender-based and sexual violence and providing compassionate support to anyone who has gone through these traumatic events. If you have experienced sexual or gender-based violence (either recently or in the past), you will find information about support services for survivors, including emergency contacts at

<https://www.uwo.ca/health/student_support/survivor_support/get-help.html>.

To connect with a case manager or set up an appointment, please contact support@uwo.ca.

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

<http://academicsupport.uwo.ca/accessible_education/index.html>

if you have any questions regarding accommodations.

Learning-skills counsellors at Learning Development and Success ([https://learning.uwo.ca](https://learning.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.

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/>.

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