



DS3000B/9000 – Introduction to Machine Learning

1. Course Information

Course Information

| | |
|----------------------------|------------------------|
| Academic term: | Winter 2023/24 |
| Undergraduate course code: | DATASCI 3000B |
| Graduate course code: | DATASCI 9000 |
| Lectures: | Tuesdays, 4:30-6:30PM |
| Tutorials: | Thursdays, 4:30-6:30PM |
| Classroom: | HSB 236 |

This course is part of the University Accreditation Program meeting certain requirements by the Canadian Institute of Actuaries. Please consult the [CIA website](#) for full details on CIA accreditation.

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. For detailed prerequisites (and alternative courses), see the [academic calendar](#).

Anti-requisites

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, you may be removed from this course, and it will be deleted from your records. This decision may not be appealed. You will receive no adjustment to your fees in the event you are dropped for failing to have the necessary prerequisites.

2. Instructors' Information

| Instructor | Email | Office Hours |
|-----------------|-----------------|--|
| Fazeli, Alireza | afazeli2@uwo.ca | Mondays 9:00 – 11:00AM Location: WSC-272 |

| Teaching Assistants | Office Hours |
|---------------------|--------------|
| TBD | TBD |
| TBD | |
| TBD | |

The Forums section of the course's site on OWL is the preferred channel for all types of inquiries. In case of a matter which needs to be addressed privately, students may email the instructor directly using their "@uwo.ca" email accounts. Emails sent from miscellaneous accounts will be ignored.

3. Course Syllabus, Schedule, Delivery Mode

Introduces machine learning and statistical methods for data analysis through applied examples. The goal of this course is 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, t-SNE
- Clustering, K-means, Hierarchical Clustering
- Model Limitations, Causality

Session format:

Both lecture and lab sessions are in-person.

Schedule:

| Week | Topic |
|--------------------|--|
| 1 Jan. 8 – 14 | Introduction and preliminaries |
| 2 Jan. 15 – 21 | supervised learning and model fitting |
| 3 Jan. 22 – 28 | Probability and maximum likelihood |
| 4 Jan. 29 – Feb. 4 | Test set/out-of-sample idea, classification, evaluation, logistic regression |
| 5 Feb. 5 – 11 | Quantifying uncertainty |
| 6 Feb. 12 – 18 | Test error, cross-validation, model selection, bias-variance trade-off |
| 7 Feb. 19 – 25 | Spring reading week |
| 8 Feb. 26 – Mar. 3 | Tuesday: summary session, Thursday: midterm Exam |
| 9 Mar. 4 – 10 | Feature selection and regularization |
| 10 Mar. 11 – 17 | Tree-based models |
| 11 Mar. 18 – 24 | Dimensionality reduction |
| 12 Mar. 25 – 31 | Clustering |
| 13 Apr. 1 – 7 | Fairness and transparency and/or artificial neural networks (ANN) |

Classes begin: January 8, 2024

Last day of classes: April 8, 2024

Examination period: April 11 – 30, 2024

The university will announce the exact date of the final exam.

Contingency plan

Although the intent is for this course to be delivered in person, should any university-declared emergency require some or all the course to be delivered online, either synchronously or asynchronously, the course will adapt accordingly. The grading scheme will not change. Any assessments affected will be conducted online as determined by the course instructor.

4. Course Materials**Books:**

The Elements of Statistical Learning by Hastie, Tibshirani and Friedman. [[online](#)]

Machine Learning: A Probabilistic Perspective by P. Kevin Murphy [[online](#)]

Students are responsible for checking the course [OWL](#) site regularly 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](#).

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 on 519-661-3800 or ext. 83800.

[Google Chrome](#) or [Mozilla Firefox](#) are the preferred browsers to optimally use OWL; please keep your browsers updated. Students interested in evaluating their internet speed, please click [here](#).

While self-installation of the software on your own computer is recommended, there is also the possibility of using online platforms. For example:

- [Google Colab](#)

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 is recommended. If this were not available, we recommend using an online environment.

5. Methods of Evaluation

The final course grade will be calculated as listed below:

| | |
|----------------|-----|
| 10 Assignments | 30% |
| Midterm Exam | 30% |
| Final Exam | 40% |

For undergrads (DS3000) to pass the course they must earn **a minimum mark of 50% on the final exam** and the same cumulatively (*i.e.*, assignments + midterm + final). Otherwise, a final course grade of 48% or less will be reported.

Example: An undergrad student obtains 100% on the assignments, 100% on the midterm, and 45% on the final exam. The final course grade is calculated as $(100\% \times 30\%) + (100\% \times 30\%) + (45\% \times 40\%) = 78\%$. However, since they did not obtain a minimum mark of 50% on the final exam, the calculated final course grade is superseded by 48%. Therefore, the student's transcript will reflect a final course grade of 48%.

For grads (DS9000) to pass the course they must earn **a minimum mark of 60% on the final exam** and the same cumulatively (*i.e.*, assignments + midterm + final). Otherwise, a final course grade of 58% or less will be reported.

Example: A grad student obtains 100% on the assignments, 100% on the midterm, and 55% on the final exam. The final course grade is calculated as $(100\% \times 30\%) + (100\% \times 30\%) + (55\% \times 40\%) = 82\%$. However, since they did not obtain a minimum mark of 60% on the final exam, the calculated final course grade is superseded by 58%. Therefore, the student's transcript will reflect a final course grade of 58%.

Weekly Assignment:

Assignments will be released each week, with due dates in the following week. There will be no make-up for missed weekly assignments.

Midterm:

The midterm will be an in-person 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, like the structure of the weekly assignments. Each student will need a **personal** laptop to complete the midterm. The exam is “open book” meaning students can access any offline and online contents. However, any sort of communication with people inside or outside the class is prohibited. Using generative AI tools such as ChatGPT is NOT allowed and will be counted as cheating. The midterm will run during the regular lecture/lab hours. For more details, please see the schedule in Section 3 of this document.

Final Exam:

The final exam date will be scheduled by the Office of the Registrar. It will be in-person and cover concepts from the entire course. It will be in both “close-book” and “open-book” structures. Any sort of offline/online communication with people inside or outside the class is prohibited. Using generative AI tools such as ChatGPT is NOT allowed and will be counted as cheating. Each student will need a **personal** laptop to complete the open-book portion of the exam. The open-book portion will be like the midterm in structure.

6. Student Absences

If you are unable to meet a course requirement due to illness or other serious circumstances, please follow the procedures below.

Assessments worth less than 10% of the overall course grade:

- Late assignment submission will be subject to a late penalty discount of 10% a day (this means if your coursework gets an 80%, and you submit n days late, your final mark will be 80% minus $n \times 10\%$. The day late starts one minute after the deadline of the original assessment has passed. There is NO EXCEPTION to this policy unless the student informs the instructor of their circumstances at least 36 hours prior to deadline and the instructor deems the request justified and grants an exception.
- An assignment cannot be submitted after it has been returned to the class. In case of a missed assignment with justified cause determined by the instructor or Academic Counselling Office, the weight will be transferred to other assignments (or to the midterm/final exam).
- Note that in all cases, the instructor reserves the right to require supporting documentation (medical or otherwise) to be submitted by the student to their Dean’s Office Academic Counselling unit before finalizing a decision.

Assessments worth 10% or more of the overall course grade:

For work totalling 10% or more of the final course grade, you must provide valid medical or supporting documentation to the Academic Counselling Office of your Faculty of Registration as soon as possible.

For further information, please consult the University's medical illness [policy](#). The Student Medical Certificate is available [here](#).

Students must note that instructors are not permitted to receive documentation directly from a student, whether in support of an application for consideration on medical grounds or else.

Absence from Midterm Examination

If you miss the Midterm Exam, please contact the Academic Counselling Office of your Faculty of Registration as soon as you can do so, and shifting the weight to the final exam may be granted.

Absence from Final Examination

If you miss the Final Exam, please contact the Academic Counselling office of your Faculty of Registration as soon as you can 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).

Note: Missed work can *only* be excused through one of the mechanisms above. Being asked not to attend an in-person course requirement due to potential COVID-19 symptoms is **not** sufficient on its own.

7. Accommodation and Accessibility

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) [here](#).

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 [here](#).

8. Academic Policies

The website for Registrarial Services is [this](#).

In accordance with the [policy](#), 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 final exams.

Scholastic offences are taken seriously, and students are directed to read the appropriate policy, specifically, the definition of what constitutes a Scholastic Offence, at this [website](#).

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](#).

In the event of health lock-down, 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](#).

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/>.

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 the Student Development Centre (<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.