**Department of Statistical and Actuarial Sciences**

**SS 4850G/SS 9850G (Advanced Data Analysis)**

**Winter 2024 Course Syllabus**

**1. Course Information**

**Course Information**

|  |  |  |  |
| --- | --- | --- | --- |
| **Instructor** | **Day/Time** | **Location** | **Contact** |
| Dr. Camila de Souza | Monday, Wednesday and Friday12:30 pm to 1:30 pm | **NCB-295** | use **OWL messages** (contact “Instructor Role”)  |

**Prerequisite(s):**

* For SS4850G: A minimum mark of 60% in both Statistical Sciences 3843A/B and Statistical Sciences 3859A/B.
* Knowledge of R/RStudio, including the use of markdown to produce beautiful PDF technical reports, is assumed or that you are willing to learn.

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** |
| Dr. Camila de Souza | use OWL messages (contact “Instructor Role”) | WSC 225 | 519-661-2111 x83618 | TBD |

|  |
| --- |
|  Warning |
| Students must use OWL messages to contact Dr. Camila de Souza. Messages sent to the instructor’s UWO email will **NOT** be replied. |

You can expect a response to a message to the instructor within **approximately 48 hours** during the work week (during busy times, it may take a little longer). ***Note that messages will not be answered within the 24-hour period before exams or project deadlines***, nor can I guarantee responses over weekends/holidays.

**3. Course Syllabus, Schedule, Delivery Mode**

**Course description**

Modern methods of data analysis including ridge regression, lasso, linear discriminant analysis, nonparametric regression, bootstrap, EM algorithm, Bayesian inference, classification trees, neural networks, dimensionality reduction techniques, and clustering.

**Tentative Course Schedule**

|  |  |  |
| --- | --- | --- |
| **Week** | **Topics** | **Reminders** |
| 1 | Jan 8-12  | * Statistical learning
* Supervised vs unsupervised learning
 |  |
| 2 | Jan 15-19 | **Supervised Learning*** Ridge regression and LASSO
 |  |
| 3 | Jan 22-26 | * Classification: Discriminant Analysis, Logistic regression, Naïve Bayes, kNN
* ROC
 |  |
| 4 | Jan 29-Feb 2 | * Mitigating class imbalance
 | Feb 2 – **Assignment 1** |
| 5 | Feb 5-9  | * Nonparametric regression
 |  |
| 6 | Feb 12-16 | * Model assessment
 |  |
| 7 | Feb 19-23  | **Reading week** |  |
| 8 | Feb 26-Mar 1 | - Overview of tree-based methods and neural networks | **Midterm exam** – to be scheduled |
| 9 | Mar 4-8 | **Inference**- Bootstrap versus maximum likelihood- Expectation-Maximization algorithm | March 8 – **Assignment 2** |
| 10 | Mar 11-15 | - Introduction to Bayesian Inference  |  |
| 11 | Mar 18-22 | - More on Bayesian Inference | **First part of final project** is due Friday Mar 22 |
| 12 | Mar 25-29 | **Unsupervised Learning**- Clustering: K-means, hierarchical clustering |  |
| 13 | Apr 1-5 | - Dimensionality reduction techniques: PCA NMF, MDS  |  |
| 14 | April 8-12 | April 8 is the last day of classes | **Assignment 3** – April 8**Final project** due April 12 |

**Key Sessional Dates**

Classes begin: January 8, 2024

Winter Reading Week: February 17 – 25, 2024

Classes end: April 8, 2024

Exam period: April 11 – 30, 2024

**Contingency plan**

Although the intent is for this course to be delivered in person, should any university-declared emergency require some or all of 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**

Chapters and sections from the following books (**all available online at UWO library**):

* The Elements of Statistical Learning – also available at author’s webpage: <https://web.stanford.edu/~hastie/ElemStatLearn//>
* An Introduction to Statistical Learning with Applications in R – also available at author’s webpage: <https://trevorhastie.github.io/ISLR/>
* Kuhn, Max, and Kjell Johnson. *Applied predictive modeling*. New York: Springer, 2013
* Hoff, Peter D. *A first course in Bayesian statistical methods*. New York: Springer, 2009

**R statistical software package:** This course is heavily based on R; therefore, all assignments will require **coding in R and Rmarkdown**. Please ensure the latest R version is installed on our computer and R studio (<https://posit.co/download/rstudio-desktop/>).

**OWL**

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. 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 by phone at 519-661-3800 or ext. 83800.

**5. Methods of Evaluation**

|  |  |  |
| --- | --- | --- |
| **Component** | **Weight** | **Deadlines/Due dates** |
| Assignment 1 | 5% | Friday February 2 |
| Assignment 2 | 7.5% | Friday March 8 |
| Assignment 3 | 7.5% | Friday April 8 |
| Midterm exam | 35% | To be scheduled by the Registrar’s officeTentative on the week of Feb 26 – March 1 |
| Final Exam: data analysis project report (individual or in pairs) | 45%(10% for first part and 35% for complete final report) | * First project part (dataset description and methodology plan) is due Friday March 22
* Complete final project report is due Friday April 12
 |

**Assignments**

* **Assignments will be available on the course OWL site. However, you will not submit your solutions to OWL. Instead, assignments must be submitted through Gradescope**,an online collaborative grading system – there will be a link to Gradescope on our OWL course site. You are responsible for ensuring that your assignment is successfully uploaded and legible. Submissions that cannot be read by the grader will receive a grade of zero. Assignment questions must be properly assigned to each page, or the submission will not be graded.
* After receiving the grades from an assignment, **students will have seven days to submit any regrade requests on that assignment**, **with the exception of the final assignment**. After this seven-day period, regrade requests will NOT be accepted. Regrade requests must be made using the Gradescope tool “Regrade Request”.
* **Assignment submissions are due 11:55 pm (Eastern Time) on the due date**. Assignments that are up to 24 hours late will receive a deduction of 15% on their mark unless late coupons are used. Late assignments up to 48 hours will receive a deduction of 30% on their mark unless late coupons are used. **No extensions will be given beyond 48 hours.**
* Each student will have **2 late coupons** worth 24 hours each that they can use at their own discretion for whatever reason (Note: I do not need to know the reason) towards their assignments. You can use them together on one assignment for a 48-hour extension or on two separate assignments for a 24-hour extension each. **No extensions will be given for any reason beyond the use of the late coupons; therefore, I suggest saving them until you absolutely need them.** You do not need to tell me when you will use the coupons; they will automatically come off the first two late days.
* Solutions to assignments **will not** be posted; however, TAs will provide comments on incorrect answers using Gradescope, which will allow students to find out the correct solutions. In addition, students can ask the instructor and TAs for more details on solutions via the Regrade Request tool on Gradescope and during office hours.

**Midterm exam**

* There will be a **2-hour in-person closed-book midterm exam**. The exact day and time are yet to be confirmed by the Registrar’s Office.

**Final exam**

* The final exam consists of a project report, which can be conducted individually or in pairs (two students). The project involves a comprehensive, advanced statistical analysis of a dataset. Each student (or pair of students) will work on a different dataset. More details will be given in the first weeks of classes. Submissions of the first part and complete report are **due 11:55pm (Eastern Time) on the due dates and must be conducted via Gradescope**. No credit will be given for submissions beyond due dates unless a valid academic accommodation is obtained (see Section 6 for details on accommodation). Those students obtaining appropriate accommodation must contact the instructor immediately to set a new deadline.

**Rounding of marks**

Across the Sciences Undergraduate Education programs, we strive to maintain high standards that reflect the effort that both students and faculty put into the teaching and learning experience during this course. All students will be treated equally and evaluated based only on their actual achievement. ***Final grades*** on this course, irrespective of the number of decimal places used in marking individual assignments and tests, will be calculated to one decimal place and rounded to the nearest integer, e.g., 74.4 becomes 74, and 74.5 becomes 75. Marks WILL NOT be bumped to the next grade or GPA, e.g. a 79 will NOT be bumped up to an 80, an 84 WILL NOT be bumped up to an 85, etc. The mark attained is the mark you achieved and the mark assigned; requests for mark “bumping” will be denied.

**6. Student Absences**

If you are unable to meet a course requirement due to illness or other serious circumstances, please follow the procedures below, which depend on whether you are an undergraduate or graduate student.

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

* **For both undergraduate and graduate students:**

The assignments have late coupons to account for any extensions needed for any reason. No extensions will be given beyond this.

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

* **For undergraduate students:**

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

The Student Medical Certificate is available at <https://www.uwo.ca/univsec/pdf/academic_policies/appeals/medicalform.pdf>.

If the midterm exam is missed and proper accommodation is obtained, the student will be eligible to write the makeup exam at another time selected by the instructor.

**Absences from Final Examinations**

If you miss the deadline for the final project (partial or final report), please contact the Academic Counselling office of your Faculty of Registration as soon as you can. They will assess your eligibility for a deadline extension. Those students obtaining appropriate accommodation must then contact the instructor immediately to set a new deadline.

* **For graduate students:**

For medical absences, get a Student Medical Certificate (SMC) signed by a licensed medical or mental health. Get appropriate documentation for non-medical absences. Note that in all cases, students are required to contact their instructors within 24 hours of the end of the period covered to get approved unless otherwise instructed in the course outline.  If the midterm exam is missed and proper documentation is obtained, the student will be eligible to write the makeup exam at another time selected by the instructor. If the deadline for the final project (partial or final report) is missed and proper documentation is obtained, the student must then contact the instructor immediately to set a new deadline.

|  |
| --- |
|  Danger |
| 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) at <https://multiculturalcalendar.com/ecal/index.php?s=c-univwo>.

**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.

**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 their official university address is attended to in a timely manner.

Participants in this course are not permitted to record lectures and/or labs, except where recording is an approved accommodation and the participant has the prior written permission of the instructor.

Online services such as Chegg are actively monitored. Any questions that are coming out

from assignments and are posted to an online service will be searched. Such an activity will be considered as a scholastic offence and will result in an academic penalty.

A scientific calculator will be allowed on the midterm and final exams. A graphics calculator will not be allowed.

**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>.

Clickers, specifically Mentimeter and iClicker, will be used in this class for instant feedback and non-graded assessments. Clickers will not be used for any marks towards the course.

**Professionalism & Privacy**

|  |  |
| --- | --- |
| Lock | Western students are expected to follow the Student Code of Conduct (<https://www.uwo.ca/univsec/pdf/board/code.pdf>). Additionally, the following expectations and professional conduct apply to this course:All course materials created by the instructor(s) are copyrighted and cannot be sold/sharedRecordings are not permitted (audio or video) without explicit permission Permitted recordings are not to be distributed |

**Copyright Statement**

Please be aware that all course materials created by the instructor(s) are copyrighted and **cannot be sold/shared**. Those include materials used in assignments, tests/quizzes, midterms, and finals. Any posting/sharing of such materials in part or whole without the owner’s consent is considered a violation of the Copyright Act and will be considered as a scholastic offence.

**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.

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/>.