

# Course Outline

## Land Acknowledgement

Land Acknowledgement I acknowledge that I live and work on Treaty 6 Territory and the homeland of the Métis. I pay my respect to the First Nations and Métis ancestors of this place past and present and reaffirm our relationship with one another. I respect the treaties that were made on these Territories, I acknowledge the harms and mistakes of the past, I recognize the ongoing present-day colonial violence that is faced by Indigenous peoples within healthcare, education, justice, child welfare and government systems and I dedicate myself to moving forward in partnership towards decolonization in the spirit of reconciliation and collaboration.

## Course Details

**Course Name:** Statistical Machine Learning for Data Science

**Course Number:** Stat 847

**Course Code (CRN):** to be announced

**Year & Term:** 2023-2024 Term 1

**Required Text:** Lantz, B. (2019). Machine Learning with R. Packt Publishing Ltd.

### Readings/Reference:

[1] Wickham, H., & Grolemund, G. (2016). R for data science: import, tidy, transform, visualize, and model data. O'Reilly Media, Inc.

[2] Grolemund, G. (2015) Hands-On Programming with R: Write Your Own Functions and Simulations. O'Reilly Media, Inc.

[3] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning. New York: Springer.

**Course Website:** on Canvas

**Prerequisites:** STAT 344 or STAT 345 or CMPT 317 or CMPT 318.

(For graduate students, students should have basic statistical theoretical knowledge, a good understanding of linear regression, and sufficient R coding skills.)

## Instructor Details

**Dr. Li Xing**

Office: Room 215 McLean Hall

Email: For the course related matters, please contact us at [dr.xing.course@usask.ca](mailto:dr.xing.course@usask.ca)

## Schedule

**Course Delivery:**

**Lecture Section:** Wednesday 6:00pm-9:00pm (online via zoom)

**Lab Section:** Thursday 3:30pm-4:50pm (online via zoom for outside USask students, or onsite in a lab room to be announced).

**Office Hours:** Friday 5:00pm-6:00pm, and by appointment (online via zoom, or onsite by appointment)

## Catalogue Description

Based on a mathematical and statistical theory foundation, the course introduces statistical methods for supervised and unsupervised learning, focusing on hands-on skills with statistical software, R, and applications to real data. The course covers resampling methods, regression and classification, tree-based methods, dimension reduction and clustering. It embeds R training throughout the entire class.

## Learning Objectives

By the completion of this course, students is expected to

1. use statistical software R on data management and visualization.
2. identify and apply the right tools from a critical statistical learning toolkit provided in the course to extract useful information from real data.
3. understand the theoretical basis of the methods.
4. given a real data problem, specify an appropriate research hypothesis and then manage a proper data analysis process using R software.
5. demonstrate and explain these skills in writing and through an oral presentation.
6. design and assemble machine learning tools to combine base learners and build an advanced learning tool.

## Content Overview

1. Introduction to machine learning.
2. Managing and understanding data with R.
3. Reproducibility and visualization.
4. Supervised and unsupervised learnings.

## Tentative Schedule:

Week	Date	Topic/Section	Assignments, Term Tests, and Holidays
1	Sept 6	Introduction to the Course and Introduction to Statistical Learning	Sept 6 the first day of school
2	Sept 11	Introduction to R for Data Science	
3	Sept 18	Reproducible Research and Version Control (RMarkdown + Github)	
4	Sept 25	Visualization (base plot, ggplot)	Assignment 1 due
5	Oct 2	Linear regression and Logistic regression	Quiz 1

6	Oct 9	Cross-validation and Bootstrap	Quiz 2 Assignment 2 due
7	Oct 16	Classification and Confusion matrix	Quiz 3
8	Oct 23	KNN and Navie Bayes Method	Quiz 4 Assignment 3 due
9	Oct 30	Subset Selection and Penalized Regression	Quiz 5
10	Nov 6	No School	Reading Break
11	Nov 13	Decision Tree	Quiz 6 Assignment 4 due
12	Nov 20	Unsupervised Learning	Quiz 7
13	Nov 27	Student Presentation	Presentation Due
14	Dec 4	Student Presentation	Dec 8 end of the term

### Midterm and Final Examination Scheduling

There are no midterm and final exams for this course.

### Grading Scheme

10 In-class Assignments	10% (1% for each assignment)
7 In-class Quizzes	15% (3% for each quiz. Choose 5 maximum scores out of 7)
4 Take-home Assignments	32% (8% for each assignment)
1 Course Project including Proposal + Report + Presentation	43% (5% for proposal, 20% for final presentation and 18% for final report)
Total	100%

### Evaluation Components

#### In-class Assignments 1-10

**Description:** Problem based assignments.

**Value:** 10% of final grade

**Submission:** The in-class assignment is due mid-night of the class delivery date (based on CMT-6 saskatoon time zone).

#### In-class Quizzes 1-7

**Description:** Problem based assignments.

**Value:** 15% of final grade

**Submission:** The quiz submission is due within 15-minutes time window during its designated lecture session.

#### **Lecture Assignments 1-4**

**Description:** Problem based assignments.

**Value:** 32% of the final grade

**Due Date:** (A1) Oct 1, 2023, (A2) Oct 15, 2023, (A3) Oct 29, 2023, and (A4) Nov 19, 2023

**Submission:** Assignment submission is online. Detailed instruction will be provided during the course.

#### **Course Project**

**Value:** 45% of the final grade

**Date:** Proposal Due Nov 5, 2023

Presentation Due Nov 26, 2023

Final Report Due Dec 18, 2023

**Type:** Online submission of proposal; in-class presentation, and take-home project.

**Submission:** Project related coursework should be directly submitted to the course instructor via email [dr.xing.course@usask.ca](mailto:dr.xing.course@usask.ca)

#### **Late Coursework**

We will not accept any late in-class assignment. We will accept lecture assignments only for seven (7) days beyond the due date. The penalty for your delay is 10 percent per day of lateness from the value of the assignment, including weekend days. Extensions may be granted only in exceptional circumstances (such as significant illness or emergency).

#### **Criteria That Must Be Met to Pass**

Students must complete at least six in-class assignments, at least two lecture assignments, submit the project proposal, conduct the presentation, and submit the report to be eligible to pass the course.

#### **Recommended Technology for Remote Learning**

Students can access course materials via the course platform on Canvas. Zoom will be used for office hours, online discussions, and personal meetings.

Students are reminded of the importance of having the appropriate technology for remote learning. The list of recommendations can be found at <https://students.usask.ca/remote-learning/tech-requirements.php>.

#### **Recording of the Course**

##### **Use of video and recording of the course:**

Please note that the lectures will be recorded and videos are available on canvas, which belong to the instructor and the University and are protected by copyright. Do not download, copy, or share recordings without the explicit permission of the instructor.

For questions about recording and use of sessions in which you have participated, including any concerns related to your privacy, please contact your instructor. More information on class recordings can be found in the Academic Courses Policy <https://policies.usask.ca/policies/academic-affairs/academic-courses.php#5ClassRecordings>.

### **Copyright**

Course materials are provided to you based on your registration in a class, and anything created by your professors and instructors is their intellectual property, unless materials are designated as open education resources. This includes exams, PowerPoint/PDF slides and other course notes. Additionally, other copyright-protected materials created by textbook publishers and authors may be provided to you based on license terms and educational exceptions in the Canadian Copyright Act (see <http://laws-lois.justice.gc.ca/eng/acts/C-42/index.html>).

**Before you copy or distribute others' copyright-protected materials, please ensure that your use of the materials is covered under the University's Fair Dealing Copyright Guidelines available at <https://library.usask.ca/copyright/general-information/fair-dealing-guidelines.php>.** For example, posting others' copyright-protected materials on the open web is not covered under the University's Fair Dealing Copyright Guidelines, and doing so requires permission from the copyright holder.

For more information about copyright, please visit <https://library.usask.ca/copyright/index.php> where there is information for students available at <https://library.usask.ca/copyright/students/rights.php>, or contact the University's Copyright Coordinator at <mailto:copyright.coordinator@usask.ca> or 306-966-8817.

### **Integrity in a Remote Learning Context**

Although the face of teaching and learning has changed due to covid-19, the rules and principles governing academic integrity remain the same. If you ever have questions about what may or may not be permitted, ask your instructor. Students have found it especially important to clarify rules related to exams administered remotely and to follow these carefully and completely.

The University of Saskatchewan is committed to the highest standards of academic integrity and honesty. Students are expected to be familiar with these standards regarding academic honesty and to uphold the policies of the University in this respect. Students are particularly urged to familiarize themselves with the provisions of the Student Conduct & Appeals section of the University Secretary Website and avoid any behavior that could potentially result in suspicions of cheating, plagiarism, misrepresentation of facts and/or participation in an offence. Academic dishonesty is a serious offence and can result in suspension or expulsion from the University.

All students should read and be familiar with the Regulations on Academic Student Misconduct (<https://secretariat.usask.ca/student-conduct-appeals/academic-misconduct.php>) as well as

the Standard of Student Conduct in Non-Academic Matters and Procedures for Resolution of Complaints and Appeals (<https://secretariat.usask.ca/student-conduct-appeals/academic-misconduct.php#IXXIIAPPEALS>)

For more information on what academic integrity means for students see the Academic Integrity section of the University Library Website at: <https://library.usask.ca/academic-integrity#AboutAcademicIntegrity>

You are encouraged to complete the Academic Integrity Tutorial to understand the fundamental values of academic integrity and how to be a responsible scholar and member of the USask community - <https://library.usask.ca/academic-integrity.php#AcademicIntegrityTutorial>

### **Access and Equity Services (AES) for Students**

Students who have disabilities (learning, medical, physical, or mental health) are strongly encouraged to register with Access and Equity Services (AES) if they have not already done so. Students who suspect they may have disabilities should contact AES for advice and referrals at any time. Those students who are registered with AES with mental health disabilities and who anticipate that they may have responses to certain course materials or topics, should discuss course content with their instructors prior to course add / drop dates. In order to access AES programs and supports, students must follow AES policy and procedures. For more information or advice, visit <https://students.usask.ca/health/centres/access-equity-services.php>, or contact AES at 306-966-7273 or [aes@usask.ca](mailto:aes@usask.ca).

Students registered with AES may request alternative arrangements for mid-term and final examinations. Students must arrange such accommodations through AES by the stated deadlines. Instructors shall provide the examinations for students who are being accommodated by the deadlines established by AES.

For information on AES services and remote learning please visit <https://updates.usask.ca/info/current/accessibility.php#AccessandEquityServices>

### **Student Supports**

#### **Academic Help for Students**

The University Library offers a range of learning and academic support to assist USask undergrad and graduate students. For information on specific services, please see the Learning page on the Library web site <https://library.usask.ca/support/learning.php>.

- Remote learning support information <https://students.usask.ca/study/remote-learning.php>
- Remote learning tutorial [https://libguides.usask.ca/remote\\_learning](https://libguides.usask.ca/remote_learning)
- Study skills materials for online learning <https://libguides.usask.ca/studyskills>
- A guide on netiquette, principles to guide respectful online learning interactions <https://teaching.usask.ca/remote-teaching/netiquette.php>

### **Teaching, Learning and Student Experience**

Teaching, Learning and Student Experience (TLSE) provides developmental and support services and programs to students and the university community. For more information, see the students' web site <http://students.usask.ca>.

### **Financial Support**

Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is urged to contact Student Central (<https://students.usask.ca/student-central.php>).

### **Aboriginal Students' Centre**

The Aboriginal Students' Centre (ASC) is dedicated to supporting Aboriginal student academic and personal success. The centre offers personal, social, cultural and some academic supports to Métis, First Nations, and Inuit students. The centre is also dedicated to intercultural education, bringing Aboriginal and non-Aboriginal students together to learn from, with and about one another in a respectful, inclusive and safe environment. Students are encouraged to visit the ASC's Facebook page (<https://www.facebook.com/aboriginalstudentscentre/>) to learn more.

### **International Student and Study Abroad Centre**

The International Student and Study Abroad Centre (ISSAC) supports student success and facilitates international education experiences at USask and abroad. ISSAC is here to assist all international undergraduate, graduate, exchange and English as a Second Language students in their transition to the University of Saskatchewan and to life in Canada. ISSAC offers advising and support on matters that affect international students and their families and on matters related to studying abroad as University of Saskatchewan students. Please visit [students.usask.ca](http://students.usask.ca) or [updates.usask.ca](http://updates.usask.ca) for more information.

### **Acknowledgements**

The instructor would like to thank the Pacific Institute for the Mathematical Sciences (PIMS) for generous support for facilitating remote teaching and enhancing invited sessions. She appreciates that PIMS has selected the course as a PIMS network course, which provide opportunities for a more general student audience among our PIMS network universities.

The instructor also would like to thank the Gwenna Moss Centre for Teaching and Learning, Mathematics and Statistics Department, and University Service for their support on course preparation and delivery.