STA414H1 S

Statistical Methods for Machine Learning II

Winter 2026 Syllabus

Course Meetings

STA414H1 S

Section	Day & Time	Delivery Mode & Location
LEC5101	Tuesday, 6:00 PM - 9:00 PM	In Person: MC 102

Refer to ACORN for the most up-to-date information about the location of the course meetings.

Course Contacts

Instructor: Thibault Randrianarisoa
Email: t.randrianarisoa@utoronto.ca
Office Hours and Location: Tuesday 9:30-11:30 @ UY 9179
Additional Notes: Questions about course content can be asked in lectures and tutorials, or directed to the course discussion board. We will be using Piazza as our primary discussion board for the course https://piazza.com/utoronto.ca/winter2026/sta414sta2014 (access code is r7anu46950q if required). Instructors and TAs will check in on a regular basis and participate in discussions. Please do not email questions about course content to instructors or TAs. Questions regarding personal matters should be directly addressed to the instructor by email. Be-

Course Overview

Probabilistic foundations of supervised and unsupervised learning methods such as naive Bayes, mixture models, and logistic regression. Gradient-based fitting of composite models including neural nets. Exact inference, stochastic variational inference, and Marko chain Monte Carlo. Variational autoencoders and generative adversarial networks.

The language of probability allows us to coherently and automatically account for uncertainty. This course will teach you how to build, fit, and do inference in probabilistic models. These models let us generate novel images and text, find meaningful latent representations of data, take advantage of large unlabeled datasets, and even let us do analogical reasoning automatically. It will offer a broad view of model-building and optimization techniques that are based on probabilistic building blocks which will serve as a foundation for more advanced machine learning courses.

Course Learning Outcomes

What you will learn:

- **Standard statistical learning algorithms:** When to use them, and their limitations.
- **The main elements of probabilistic models:** Distributions, expectations, latent variables, neural networks, and how to combine them.
- **Standard computational tools:** Monte Carlo, Stochastic optimization, Variational Inference, and automatic differentiation.

Skill Outcomes

Research Stream:

- **Derive and Extend:** Students will be able to mathematically derive custom model architectures and extend existing frameworks to handle different data structures.
- **Algorithmic Reasoning:** Identify when exact inference is feasible versus when approximate methods are required based on the characteristics of the model.

Applied Stream:

- **Feasibility Assessment:** Evaluate the "cost of entry" for a project by estimating how much data and compute are required to reach target performance.
- **Iterative Refinement:** Apply the "Linear-to-Complex" modeling workflow, starting with simple baselines to justify the use of more resource-intensive nonlinear models.

Prerequisites: STA314H1/ CSC311H1/ CSC311H5/ (STA314H5, STA315H5)/ CSCC11H3; STA302H1/ STAC67H3/ STA302H5; CSC108H1/ CSC110Y1/ CSC120H1/ CSC148H1/ CSCA08H3/ CSCA48H3/ CSCA20H3/ CSC108H5/ CSC148H5; MAT235Y1/ MAT237Y1/ MAT257Y1/ (MATB41H3, MATB42H3)/ (MAT232H5, MAT236H5)/ (MAT233H5, MAT236H5); MAT223H1/ MAT224H1/ MAT240H1/ MATA22H3/ MATA23H3/ MAT223H5/ MAT240H5/ MATB24H3/ MAT224H5
Corequisites: None
Exclusions: CSC412H1, STAD68H3
Recommended Preparation: STA303H1
Credit Value: 0.5

Students should possess a strong command of **linear algebra** (vector/matrix manipulation and geometric intuition), **calculus** (partial derivatives and gradients), **probability** (common distributions and Bayes' Rule), and **foundational statistics** (expectation, variance, median and maximum likelihood).

Course Materials

No required textbooks.

However, some suggested readings are:

- (PRML) Christopher M. Bishop (2006) Pattern Recognition and Machine Learning
- (ITIL) David MacKay (2003) Information Theory, Inference, and Learning Algorithms
- (PML1) Kevin P. Murphy (2022), Probabilistic Machine Learning: An Introduction

- (PML2) Kevin P. Murphy (2023), [Probabilistic Machine Learning: Advanced topics]

Other useful references are:

- (DL) Ian Goodfellow, Yoshua Bengio and Aaron Courville (2016), [Deep Learning]
- (MLPP) Kevin P. Murphy (2013), [Machine Learning: A Probabilistic Perspective]
- (ESL) Trevor Hastie, Robert Tibshirani, Jerome Friedman (2009) [The Elements of Statistical Learning]
- (ISL) Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani (2017) [Introduction to Statistical Learning]

Marking Scheme

Assessment	Percent	Details	Due Date
Midterm	30%	Covers topics from first 5 weeks, 2-hour test (pen & paper derivations)	2026-02-24
Quizzes	15%	10-minute quizzes at the end of each lecture (except Midterm week), covers material from preceding week. 10 quizzes in total, keep 8 best for the grade.	2026-01-13, 2026-01-20, 2026-01-27, 2026-02-03, 2026-02-10, 2026-03-03, 2026-03-10, 2026-03-17, 2026-03-24, 2026-03-31
Homework Assignments	10%	Pen & paper derivations + Coding (Python/Numpy) / Four Homework Assignments (2.5% each). The assignments will be released on the course webpage. Submission instructions will be provided with each assignment. Assignments must be your own individual work. After attempting the problems on an individual basis, you may discuss and work together on the homework assignments with up to two classmates. However, you must write your own code and write up your own solutions individually and explicitly name	2026-02-01, 2026-02-15, 2026-03-15, 2026-03-29

Assessment	Percent	Details	Due Date
		any collaborators at the top of the homework.	
In-Person Final Exam	45%		Final Exam Period

Late Assessment Submissions Policy

Assignments should be handed in by deadline; a late penalty of 10% per day will be assessed thereafter (up to 3 days, then submission is blocked).

Course Schedule

Week	Description
Week 1 5–11 January	Introduction and Probabilistic Models
Week 2 12–18 January	Directed Acyclic Graphs (DAGs) & Decision Theory
Week 3 19–25 January	Exact Inference & Message passing
Week 4 26 January – 1 February	Hidden Markov Models & Monte-Carlo Methods
Week 5 2–8 February	MCMC
Week 6 9–15 February	Variational Inference
Reading Week	Reading Week
Midterm 16–22 February	Midterm
Week 7 23 February – 1 March	Neural networks
Week 8	Kernel Methods and Gaussian Processes

2–8 March	
Week 9 9–15 March	Embeddings/Attention/Transformers
Week 10 16–22 March	Variational AutoEncoders
Week 11 23–29 March	Diffusion Models

Policies & Statements

Late/Missed Assignments

Ten percent of the value will be deducted for each late day (up to 3 days, then submission is blocked). No credit will be given for assignments submitted after 3 days.

Missed test.

- If the midterm is missed for a valid reason, you must submit documentation to the course instructor.
- If a test is missed for a valid medical reason, you must submit the absence declaration form and let your instructor know immediately.
- The form will only be accepted as valid if the form is filled out according to the instructions on the form.
- If the midterm test is missed for a valid reason, then the final test will be worth 75% of your final grade. Other reasons for missing a test will require prior approval by your instructor. If prior approval is not received for non-medical reasons then you will receive a term test grade of zero.

Online Communication

Questions about course content can be asked in lectures and tutorials, or directed to the course discussion board. We will be using Piazza as our primary discussion board for the course https://piazza.com/utoronto.ca/winter2026/sta414sta2014/home (access code is r7anu46950q if required).
Instructors and TAs will check in on a regular basis and selectively participate in discussions. Please
do not email questions about course content to instructors or TAs.
Questions regarding personal matters should be directly addressed to the instructor by email and from your UofT mailing account. Before sending an email, make sure that you are not

asking for information that is already available in
the course outline/website/announcements. All student emails to instructors/TAs should include:

a. the course code and term in the subject (e.g., [STA414] or [STA2104]), and
b. your student number in the body or signature.

If either is missing, your email will not be considered. Allow at least 48hr for a response before following up.

Use of Generative Artificial Intelligence Tools

Students may use artificial intelligence tools, including generative AI, in this course as learning aids or to help produce assignments. However, students are ultimately accountable for the work they submit.

Students may not use artificial intelligence tools for taking tests or completing major course assignments. However, these tools may be useful when gathering information from across sources and assimilating it for understanding.

The knowing use of generative artificial intelligence tools, including ChatGPT and other AI writing and coding assistants, for the completion of, or to support the completion of, an examination, term test or any other form of academic assessment, may be considered an academic offense in this course.

Academic Integrity

All suspected cases of academic dishonesty will be investigated following procedures outlined in the Code of Behaviour on Academic Matters (https://governingcouncil.utoronto.ca/secretariat/policies/code-behaviour-academic-matters-july-1-2019). If you have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, please reach out to me. Note that you are expected to seek out additional information on academic integrity from me or from other institutional resources. For example, to learn more about how to cite and use source material appropriately and for other writing support, see the U of T writing support website at http://www.writing.utoronto.ca. Consult the Code of Behaviour on Academic Matters for a complete outline of the University's policy and expectations. For more information, please see A&S Student Academic Integrity (https://www.artsci.utoronto.ca/current/academic-advising-and-support/student-academic-integrity) and the University of Toronto Website on Academic Integrity (https://www.academicintegrity.utoronto.ca).

Students with Disabilities or Accommodation Requirements

Students with diverse learning styles and needs are welcome in this course. If you have an acute or ongoing disability issue or accommodation need, you should register with Accessibility Services (AS) at the beginning of the academic year by visiting https://studentlife.utoronto.ca/department/accessibility-services/. Without registration, you will not be able to verify your situation with your instructors, and instructors will not be advised about your accommodation needs. AS will assess your situation, develop an accommodation plan with you, and support you in requesting accommodation for your course work. Remember that the process of accommodation is private: AS will not share details of your needs or condition with any instructor, and your instructors will not reveal that you are registered with AS.

Accommodations

Students with diverse learning styles and needs are welcome in this course. In particular, if you have a disability/health consideration that may require accommodations, please feel free to approach me and/or the AccessAbility Services Office as soon as possible.

AccessAbility Services staff (located in Rm IA5105, Sam Ibrahim Building) are available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations 416-287-7560 or email ability.utsc@utoronto.ca. The sooner you let us know your needs the quicker we can assist you in achieving your learning goals in this course.

Absence declaration.

Students who are absent from academic participation for any reason (e.g., COVID, cold, flu and other illness or injury, family situation) and who require consideration for missed academic work have been asked to record their absence through the ACORN online absence declaration. The absence declaration is considered sufficient documentation to indicate an absence and no additional information or documentation should be required when seeking consideration from an instructor. Students should also advise their instructor of their absence. Instructors will not be automatically alerted when a student declares an absence. **It is student's responsibility to let instructors know that they have used the Absence Declaration so that you can discuss any needed consideration, where appropriate.**

Additional Content

The marking scheme for graduate students following STA2104 is the following:

Assessment	Percent	Details	Due Date
Homework Assignments	12%	• Three Homework Assignments (4% each) • Pen & paper derivations + Coding (Python/Numpy)	20260201, 20260215, 20260315,
Quizzes	8%	• Every week (starting next week, except Midterm week), before the tutorial • About 10 minutes each • 10 quizzes in total, keep top 8	
Project	35%	See below	• Proposal due on Feb 24 • Report due on

			April 24
Final Exam	45%	• About 3 hours • Conceptual/theoretical, minimal coding.	TBA

Project details:

- Students can work on projects individually, or in groups of up to four.
- The grade will depend on the ideas, how well you present them in the report, how clearly you position your work relative to existing literature, how illuminating your experiments are, and well-supported your conclusions are. Full marks will require a novel contribution.
- More detailed expectations will be published on the course webpage.
- Each group of students will write a short (around 2 pages) research project proposal, which ideally will be structured similarly to a standard paper. It should include a description of a minimum viable project, some nice-to-haves if time allows, and a short review of related work. You don’t have to do exactly what your project proposal says. The point of the proposal is mainly to have *a* plan and to make it easy for me to give you feedback.
- At the end of the class you’ll hand in a project report (around 4 to 8 pages), ideally in the format of a machine learning conference paper.