# Machine Learning

Jump to Today

# CMPT 726 Machine Learning

Machine learning is the study of computer algorithms that improve automatically through experience, which play an increasingly important role in artificial intelligence, computer science and beyond. The goal of this course is to introduce students to machine learning, starting from the foundations and gradually building up to modern techniques. Students in the course will learn about the theoretical underpinnings, modern applications and software tools for applying deep learning. This course is intended to be an introductory course for students interested in conducting research in machine learning or applying machine learning, and should prepare students for more advanced courses, such as CMPT 727 and CMPT 728. No previous knowledge of machine learning is assumed, but students are expected to have solid background in calculus, linear algebra, probability and programming using Python.

## Instruction Format, Times and Location

### Lecture Times and Format

Tuesdays 16:30-18:20
Thursdays 16:30-17:20

**Location:** YouTube (Synchronous with live Q&A)

**Lectures and Recordings:**

- Full playlist (https://www.youtube.com/playlist?list=PLUBop1d3Zm2vcC90zpexLkD4U0sI07kmF)
- Lecture notes are posted under "Files (https://canvas.sfu.ca/courses/71925/files) "

1. Introduction and Linear Algebra Basics [ Lecture (https://youtu.be/mxjmSxt7R90) ]
2. Linear Algebra, SVD, Eigendecomposition, Norms [ Lecture (https://youtu.be/14qrG2VeMI4) | Missing p-norms slide (https://youtu.be/hXpsPyXqo9k) ]
3. Taylor Expansion, Quadratic Forms [ Lecture (https://youtu.be/2KK0q0wXC80) ]
4. Quadratic Forms, Convexity, Optimality Conditions, Linear Regression Intro [ Lecture (https://youtu.be/yB85fBSQ4WI) ]

5. Linear Regression [ **Lecture** (https://youtu.be/1CSUDlg3tWU) ]
6. Linear Regression Continued [ **Lecture** (https://youtu.be/65tEJsZ9rrY) ]
7. Linear Regression, Ridge Regression[ **Lecture** (https://youtu.be/S7iknheQ9ZM) ]
8. Probability Review [ **Lecture** (https://youtu.be/LL5bUCqXQEs) ]
9. Gaussian distributions [ **Lecture** (https://youtu.be/kDllf_RnwqM) ]
10. Probabilistic Interpretation of Linear Regression [ **Lecture** (https://youtu.be/-GBJo2qdTXI) ]
11. Bias Variance Decomposition, Optimization Intro [ **Lecture** (https://youtu.be/UR0kbymL6wE) ]
12. Optimization Algorithms [ **Lecture** (https://youtu.be/FPw47xV9BUA) ]
13. Stochastic Optimization, Neural Networks [ **Lecture** (https://youtu.be/1VCDRPZ2hLw) ]
14. Neural Networks and Backpropagation [ **Lecture** (https://youtu.be/DiXtOF22nuI) ]
15. Backpropagation in MLPs [ **Lecture** (https://youtu.be/49zHsI-hLG4) ]
16. Backpropagation and Gradient Issues [ **Lecture** (https://youtu.be/-DG74iqVJVM) ]
17. Residual Connections [ **Lecture** (https://youtu.be/0_iXnf6Zn4w) ]
18. Autoencoders, CNNs, Classification, Support Vector Machines [ **Lecture** (https://youtu.be/AYH7Rfgs9XY) ]
19. The SVM Dual Problem [ **Lecture** (https://youtu.be/ckZNdSC1EnA) ]
20. SVM Strong Duality and Solution, Kernels, Soft Margin SVM [ **Lecture** (https://youtu.be/mBII53noLIE) ]
21. Soft-Margin SVM, Other Classification [ **Lecture** (https://youtu.be/ESz4g9CiwPA) ]
22. Logistic and Softmax Classification, Review [ **Lecture** (https://youtu.be/g5AM1XPP1fI) ]

This course will be offered in a partially online format. Lectures will be broadcast on YouTube during the lecture time, and I will be also watching the lecture at the same time to answer any questions in the chat in real time. This format has been effective for large classes since it is difficult to have meaningful interactions during the large lecture.

**Tips:**

- Feel free to briefly pause the video to think about a slide or ask questions before proceeding.
- Pay attention to the chat in case there are important questions and answers
- Review lecture videos after the lecture

# Final Exam

**When:** Dec. 7, 15:30-18:30

**Location:** Images Theatre

# Instructional Team

**Instructor:**

- Mo Chen
- Office Hours: on **Zoom** (https://sfu.zoom.us/j/64279881680?pwd=bndDcFdUV2JuWC8yM2N3b0krRnppUT09)
  - Wednesdays and Thursdays 17:30-18:30

**Teaching Assistants**:

- Mohammad Soltanshah
  - Office Hours: Mondays 10:00 - 11:00, ASB 9808
- Wanyi Su
  - Office Hours: Tuesdays 9:00 - 10:00, ASB 9700
- Payam Jome Yazdian
  - Office hours: Wednesdays 15:30 - 16:30, ASB 9808
  - Sep 21&28, October 19&26 will be held online (Zoom **Link** (https://sfu.zoom.us/j/61108750715?pwd=c3hqWXBVRWRBUktQcGJ5cHdXekxlQT09) ),
- Michael Lu
  - Office Hours: Thursdays:  13:00 - 14:00, ASB 9808
  - Octotber 20 will be head on **Zoom** (https://sfu.zoom.us/j/62869835117?pwd=citVOWdTa09kWTJxTnJvVzZQTjZqUT09)

**Discussion Board:** **CourSys forum** (https://coursys.sfu.ca/2022fa-cmpt-726-x1/forum/)

# Grading

- Quizzes: 15%, participation
- Assignment 0: bonus 1%
- Assignment 1: 15%
- Assignment 2: 15%

- Assignment 3: 15%
- Final Exam: 40%

Late policy:

- Quizzes: No late submissions
- Assignments: 2% deduction per hour for up to 48 hours

## Quizzes

We will have weekly quizzes, the participation of which together will make up 15% of your grade. Each quiz is released on Thursdays and covers the material taught in lectures during the week, and is due the following Monday at 11:59pm Pacific time. Your lowest quiz score will be dropped.

The quizzes are designed to test your understanding of the material; if you do not do well on a quiz, you should review the content covered in the previous week and make sure you understand it thoroughly. Success in this course requires keeping pace with the lectures; materials in later weeks build upon materials covered in earlier weeks, so if you miss one thing, you will miss everything that follows. Do not expect to be able to cram before the exam and do well in this course.

## Assignments

Assignments are designed to assess a deeper level of understanding than quizzes and may include extensions of materials covered in class. Each assignment is worth 15% of your grade (except for Assignment 0). You may collaborate with other students in the course on assignments, under the following conditions:

- You must declare who you collaborated with in your submission, and
- You must write up the solutions on your own.

You may ask questions about the assignment and engage in discussions with other students on the discussion board. It is your responsibility to check the discussion board – clarifications may be posted in response to questions, and updates/corrections may be posted on the discussion board.

The following activities are prohibited:

- Posting assignment problems or your solutions on the web

- Sharing your solutions with other students
- Looking up solutions from previous semesters, other courses or other students
- Discussing assignment problems with others not in the course

These are considered academic offences and will result in harsh penalties (see Academic Integrity section below).

## How to Get Help

If you have questions about the course material, you should come to office hours or post them on the discussion board. Use discussion boards for quick clarifications about a specific concept and office hours for help with understanding a large chunk of the material. The latter tends to take a while to explain and is not easily distilled in written form. Do not send email for content-related questions - other students may have the same questions as you do and would benefit from a public response.

If you have questions about the assignments, you should make sure not to post your solutions or reveal the details of the approach you took. Instead, pose your questions about specific parts you would like clarification on and abstract away details about other parts. Try to break down your questions and ask each part separately. For questions that take more than five minutes to explain, you should ask them during office hours; otherwise, you should post on the discussion board. When posting in the discussion board, be sure to start a thread in the post dedicated to the problem in the assignment, rather than making a new post. This will ensure that your question is attended to in a timely fashion.

If you cannot ask a question without revealing your solutions, you should send us an email with your question. Note however that we will prioritize answering questions on the discussion board, so be sure to email us with your question **at least 24 hours** before the deadline. You should not use this to ask us to check your solutions before submission and effectively pre-grade them; questions should be narrow in scope and pertain to a specific part of a question.

## Academic Integrity

We take academic integrity very seriously and have a zero-tolerance policy for cheating. Any offence would result in a grade of zero for the assignment. Serious cases would result in failing the course and a referral to the University Board on Student Discipline, which can result in a permanent record on your transcript and/or suspension/expulsion from the university.

Please note that online communication mediums outside of the Canvas discussion board should not be used to discuss assignment content. You are not permitted to upload any of the assignment problems, solutions, or your own solutions to assignment problems to any

website that is accessible by others. The only exceptions to this policy are online communication mediums between you and other students in the course who are explicitly listed on your assignment. Looking at online solutions from previous semesters, other courses or other students is forbidden, as is sharing of your solutions with others. Furthermore, all students have an *affirmative duty* to report possible cases of cheating or unauthorized communication to the course staff immediately. Awareness of and failure to report cheating constitutes academic misconduct on the part of the bystander.

We patrol the web for instances of cheating and run rigorous checks on all assignment and exam submissions at the end of the semester. Penalties may be imposed retroactively.

Instances of cheating may be reported via email to Mo. Identities of reporting students will be kept confidential.

## Topics may include

- Linear Regression
- Hyperparameters, Cross-Validation, Regularization
- Maximum Likelihood, Maximum A Posteriori
- Bias/Variance Tradeoff
- Kernel Ridge Regression
- Nonlinear Least Squares
- Optimization Methods
- Neural Networks
- Logistic Regression, Multinomial Logistic Regression
- Support Vector Machines
- k-Nearest Neighbours
- Autoencoders

## Resources

### Practice

- Redo quizzes, assignments, and in-class reviews
- Berkeley CS189/289A Material from Jonathan Shewchuk: **https://people.eecs.berkeley.edu/~jrs/189/ (https://people.eecs.berkeley.edu/~jrs/189/)**
- Stanford CS229 Material: **https://github.com/maxim5/cs229-2018-autumn (https://github.com/maxim5/cs229-2018-autumn)**

- Some Exercises in "An Introduction to Statistical Learning (https://www.statlearning.com/) ", Ch. 3, 4, 6, 9, 10, are relevant

### Math Background

- Math for ML: http://gwthomas.github.io/docs/math4ml.pdf (http://gwthomas.github.io/docs/math4ml.pdf)
- Matrix Cookbook: http://www2.imm.dtu.dk/pubdb/edoc/imm3274.pdf (http://www2.imm.dtu.dk/pubdb/edoc/imm3274.pdf)
- Linear Algebra Review and Reference: https://cs229.stanford.edu/section/cs229-linalg.pdf (https://cs229.stanford.edu/section/cs229-linalg.pdf)
- Review of Probability Theory: https://cs229.stanford.edu/section/cs229-prob.pdf (https://cs229.stanford.edu/section/cs229-prob.pdf)

### Programming Background

- NumPy Tutorial: https://www.datacamp.com/community/tutorials/python-numpy-tutorial (https://www.datacamp.com/community/tutorials/python-numpy-tutorial)
- PyTorch Tutorial: https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html (https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html)
- Step-by-Step PyTorch Example: https://towardsdatascience.com/understanding-pytorch-with-an-example-a-step-by-step-tutorial-81fc5f8c4e8e (https://towardsdatascience.com/understanding-pytorch-with-an-example-a-step-by-step-tutorial-81fc5f8c4e8e)

### Visualization

- Neural Network Playground: https://playground.tensorflow.org (https://playground.tensorflow.org)
- Decision boundary visualizer: https://ml-visualizer.herokuapp.com/ (https://ml-visualizer.herokuapp.com/)

### Textbook for a different perspective

- An Introduction to Statistical Learning: https://www.statlearning.com/ (https://www.statlearning.com/)

## Credit

This course is modelled after CS189 at UC Berkeley taught by Prof. Anant Sahai (https://www2.eecs.berkeley.edu/Faculty/Homepages/sahai.html) .

The lecture slides are kindly provided by Prof. Ke Li (https://www.sfu.ca/~keli/) .

# Course Summary:

| Date | Details | Due |
|---|---|---|
| Fri Sep 16, 2022 | Assignment 0: Math Background (https://canvas.sfu.ca/courses/71925/assignments/806780) | due by 11:59pm |
| Mon Sep 19, 2022 | Quiz 1: Linear Algebra (https://canvas.sfu.ca/courses/71925/assignments/807574) | due by 11:59pm |
| | Quiz 1: Linear Algebra (https://canvas.sfu.ca/courses/71925/assignments/807574) (1 student) | due by 11:59pm |
| Mon Sep 26, 2022 | Quiz 2: Linear Algebra, Convexity, Optimality Conditions, Linear Regression (https://canvas.sfu.ca/courses/71925/assignments/809813) | due by 11:59pm |
| Mon Oct 3, 2022 | Quiz 3: Linear Regression (https://canvas.sfu.ca/courses/71925/assignments/811105) | due by 11:59pm |
| Mon Oct 10, 2022 | Quiz 4: Probability Review (https://canvas.sfu.ca/courses/71925/assignments/812455) | due by 11:59pm |
| Mon Oct 17, 2022 | Quiz 5: Probabilistic Interpretation of Linear Regression (https://canvas.sfu.ca/courses/71925/assignments/814408) | due by 11:59pm |

| Date | Details | Due |
|---|---|---|
| Mon Oct 24, 2022 | Quiz 6: Optimization (https://canvas.sfu.ca/courses/71925/assignments/815588) | due by 11:59pm |
| Mon Oct 31, 2022 | Quiz 7: Neural Networks (https://canvas.sfu.ca/courses/71925/assignments/817730) | due by 11:59pm |
| Fri Nov 4, 2022 | Assignment 2 Code (https://canvas.sfu.ca/courses/71925/assignments/818198) | due by 11:59pm |
| Tue Nov 8, 2022 | Assignment 2 Code (https://canvas.sfu.ca/courses/71925/assignments/818198) (1 student) | due by 11:59pm |
| Thu Nov 10, 2022 | Quiz 8: Neural Networks (https://canvas.sfu.ca/courses/71925/assignments/818882) | due by 11:59pm |
| Wed Nov 16, 2022 | Quiz 9: Support Vector Machine (https://canvas.sfu.ca/courses/71925/assignments/821124) | due by 11:59pm |
| Mon Nov 28, 2022 | Quiz 10: Classification (https://canvas.sfu.ca/courses/71925/assignments/823986) | due by 11:59pm |
| Tue Nov 29, 2022 | Assignment 3 Code (https://canvas.sfu.ca/courses/71925/assignments/820595) | due by 11:59pm |
| | Final Exam (For Special Accommodations) (https://canvas.sfu.ca/courses/71925/assignments/827787) | |