Machine Learning

CMPT 410/726

Mo Chen
SFU School of Computing Science
2025-09-03

About This Course

Instructor: Mo Chen, mochen@cs.sfu.ca

Website (Canvas page): https://canvas.sfu.ca/courses/92500

Discussion forum: https://coursys.sfu.ca/2025fa-cmpt-726-x1/forum/

Office Hours: See Canvas

Lectures: See Canvas

Teaching Assistants: See Canvas

Grading

- Assignments: 45% over 4 assignments
- Quizzes: 10%. Completion and correctness
- Peer grading: 5%. Timeliness, completion and correctness

- Final Exam: 40%
  - Time and Location TBA

Quizzes

- Weekly quizzes are assigned on Thursdays (starting next week) and due the following Tuesday at 23:59. (Will be available in Canvas under the "Quizzes" section.)
- Each quiz covers concepts taught in lectures in the same week.
- Designed to test your understanding of the basic concepts covered each week.
- **Important:** Materials build upon materials covered earlier. If you miss one thing, you'll miss everything that follows.
- You will be given participation marks for getting all questions correct (unlimited tries). You can omit up to two quizzes.

Assignments

Worth 45% total

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.
- While you may discuss ideas for solving a problem, you must write up 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.

Assignments

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 offenses and will result in harsh penalties.

Academic Integrity

We take this very seriously and have zero tolerance for cheating or assistance with cheating.

Consequences include but not limited to:

- Zero on the assignment
- Failing the course
- Referral to University Board on Student Discipline
- Permanent record on your transcript
- Suspension or expulsion from the university

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.

Double Blind Peer Grading

Everyone grades 3 random assignments from peers

- Follow detailed solutions and rubrics
- If marks are taken off, write a few words (~1 sentence) to explain why
- Due 1 week after solutions are released
- 5% participation marks for timeliness, completeness, correctness of grading

Final grade will be determined by teaching assistants

Disputes to be handled directly by teaching staff

Double Blind Peer Grading

Some benefits:

- Frees up teaching staff's time for more help with assignments and course material
- Ensures assignments are thoroughly reviewed
- Ensures thorough review of assignment solutions

Assignment Summary

	Release	Due / Solutions available	Peer reviews due	Portion of final grade
A0	Sept. 3	Sept. 8	Sept. 15	1%
A1	Sept. 15	Oct. 6	Oct. 13	16%
A2	Oct. 13	Nov. 3	Nov. 10	16%
A3	Nov. 10	Dec. 1	Dec. 8	12%

Required Background

Calculus: $E = mc^2 \Rightarrow \frac{\partial E}{\partial c} = 2mc$

Linear Algebra: $Au_i = \lambda_i u_i$; $\frac{\partial}{\partial x}(x^\top a) = a$

Probability: $p(X) = \sum_Y p(X, Y)\,; p(x) = \int p(x, y) dy; \mathbb{E}_x[f] = \int p(x) f(x) dx$

Python Programming: Review NumPy and PyTorch tutorials linked from the website

Good Resources: (See Canvas syllabus)

Warning: While it will be possible to refresh, if you’ve never seen these before, this course will be **very** difficult.

Administrivia - Resources

- No required textbook
- Closest to the course content is the course notes (will be available on website)
- Exam will cover content in the lectures, assignments, and quizzes.
- Reference books (tend to cover more content at a faster pace):
  - Machine Learning: A Probabilistic Perspective, Kevin P. Murphy, MIT Press, 2012, 9780262018029
  - The Elements of Statistical Learning, Trevor Hastie, Robert Tibshirani, and Jerome Friedman, Springer-Verlag, 2009, 9780387848570
  - All of Statistics, Larry Wasserman, Springer, 2010, 9781441923226
  - Pattern Recognition and Machine Learning, Christopher M. Bishop, Springer, 2006, 9780387310732
  - Machine Learning, Tom Mitchell, McGraw Hill, 1997, 9780070428072

Administrivia - How to Get Help

Regarding Course Material:

- TA Office Hours: Help with a large chunk of the material.
- Discussion Forum: Quick clarifications about a specific concept.
- Do not send email for content-related questions – others may have the same question and would benefit from a public response.

Administrivia - How to Get Help

Regarding Assignments:

- Office Hours: Any questions that are broad in scope or not concretely formulated, or take more than five minutes to lay out or explain.
- Discussion Board: Any specific and concretely formulated questions.
- **Do NOT post your solutions or details of how you approach the problem.**
- If you cannot ask your question without revealing your solution, send us an email.
- We prioritize discussion board questions, so be sure to email **at least 24 hours** before the deadline.

How to Succeed in This Course

This is a challenging course – it is important to both understand the theoretical principles and practice applying machine learning.

Very important: Goal is to understand, not memorize.

- Pay attention to lectures, including questions and answers. Take notes as appropriate.
- If you can't follow the derivations, review them after the lecture and make sure you understand each step.
- If you do not understand a step in the derivation, ask on the discussion board or come to office hours.
- After the lecture, reflect on the material and relate it to concepts covered earlier. Think about the similarities and differences.
- If needed, take extra time to become familiar with new terminology in English
- Everything builds upon material covered earlier, **so do not fall behind!**
- Do not expect to be able to cram before the exam and do well in this course.

Possible Learning Pipeline

Before each lecture:

- Preview / skim through lecture slides
- Catch up on the relevant math background
    - Essence of Linear Algebra: https://www.youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab
    - Essence of Calculus: https://www.youtube.com/playlist?list=PLZHQObOWTQDMsr9K-rj53DwVRMYO3t5Yr

After each lecture:

- Catch up on anything you missed.
- Fall 2022 recording:

  https://www.youtube.com/playlist?list=PLUBop1d3Zm2vcC90zpexLkD4U0sI07kmF

Questions?

Why Machine Learning?

What is Artificial Intelligence (AI)?

- The study of how to engineer *intelligent* systems/machines.
- What is *intelligence*? Anything that humans can do that machines can't do easily.
  - The ability to see and interpret visual input
  - The ability to read and understand language
  - The ability to move and interact with the world
  - The ability to reason and perform logical deduction

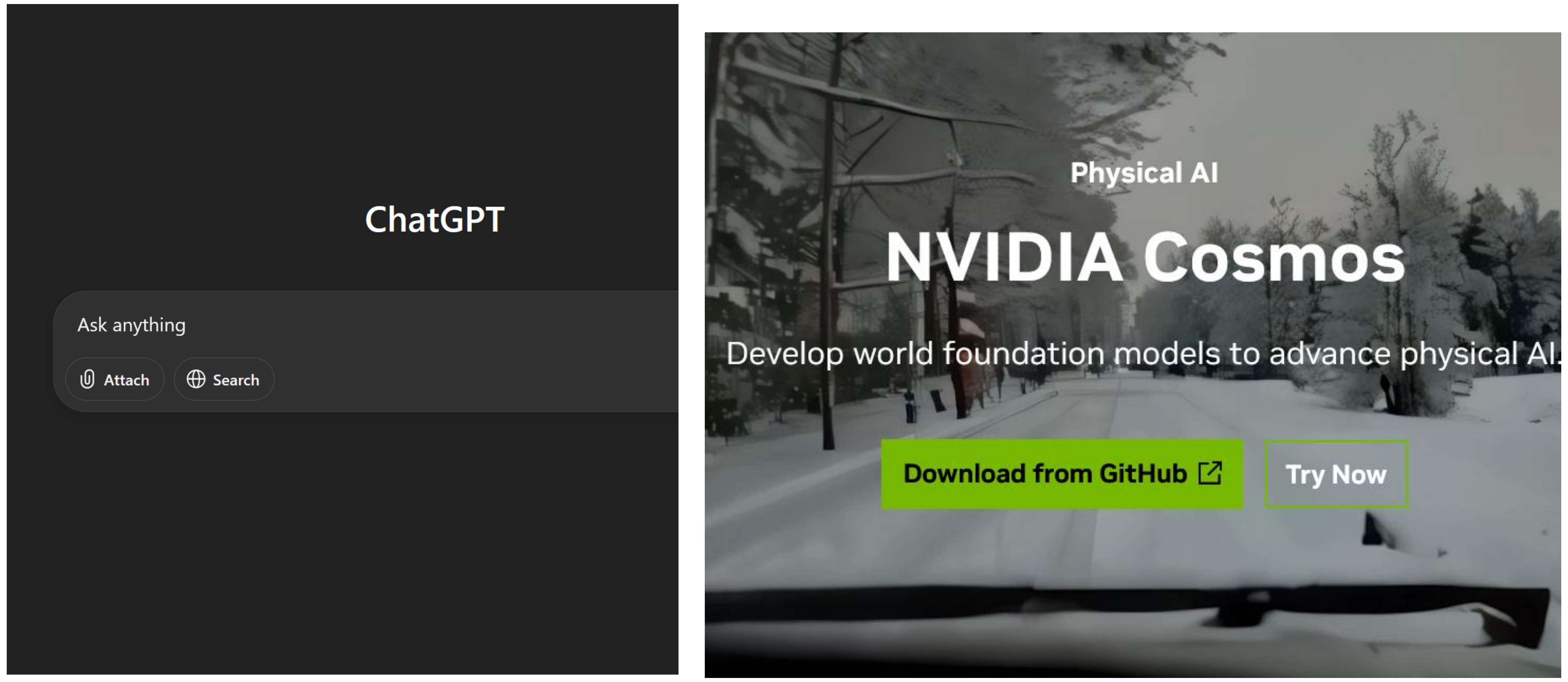
What is Artificial Intelligence (AI)?

- The study of how to engineer *intelligent* systems/machines.
- What is *intelligence*? Anything that humans can do that machines can't do easily.
  - The ability to see and interpret visual input Computer Vision
  - The ability to read and understand language Natural Language Processing (NLP)
  - The ability to move and interact with the world Robotics
  - The ability to reason and perform logical deduction Traditional AI

Successes of AI

Computer Vision

Natural Language Processing (NLP)



Robotics

Traditional AI

Why is AI hard?

- Suppose we want to write a program that recognizes chairs.

- First attempt:
  - A chair is something that has a seat, a back and four legs.

Images taken from ShapeNet

Why is AI hard?

- Suppose we want to write a program that recognizes chairs.


- Second attempt:
    - A chair is something that has a seat, a back and multiple legs.

Images taken from ShapeNet

Why is AI hard?

- Suppose we want to write a program that recognizes chairs.

- Third attempt:
  - A chair is something that has a seat, a back and a frame.

Images taken from ShapeNet

Why is AI hard?

- Suppose we want to write a program that recognizes chairs.

- Fourth attempt:
    - A chair is something that has a seat and a back.

Images taken from ShapeNet

Why is AI hard?

- Suppose we want to write a program that recognizes chairs.

- Fifth attempt:
  - A chair is something that has a seat.

Images taken from ShapeNet

Why is AI hard?

- Suppose we want to write a program that recognizes chairs.

- Why is this not a chair?

- There are exceptions to every rule, and exceptions to every exception.

Images taken from ShapeNet

Why is AI hard?

- Problem: The inner workings of our brain are not well understood.

- We don't know *how* our brain converts input to output, so we can't write a program to do so.

- This problem perplexed early computer scientists:

"The Analytical Engine[1] has no pretensions to *originate* anything. It can do *whatever we know how to order it* to perform. "

Ada Lovelace

[1]The Analytical Engine was the first conception of a general-purpose (i.e. Turing complete) computer.

Learning Machines

- Alan Turing proposed the concept of a *learning machine* in 1950 (in the same paper that proposed the Turing test).
- Idea: Divide the problem into two parts:
  - A machine that simulates a child's brain (analogous to a blank notebook: should function by simple mechanisms and have lots of blank sheets)
  - A way of teaching the child machine (should be simple since we know how to teach a human child)
- Teacher rewards good behaviour and penalizes bad behaviour.

Learning Machines

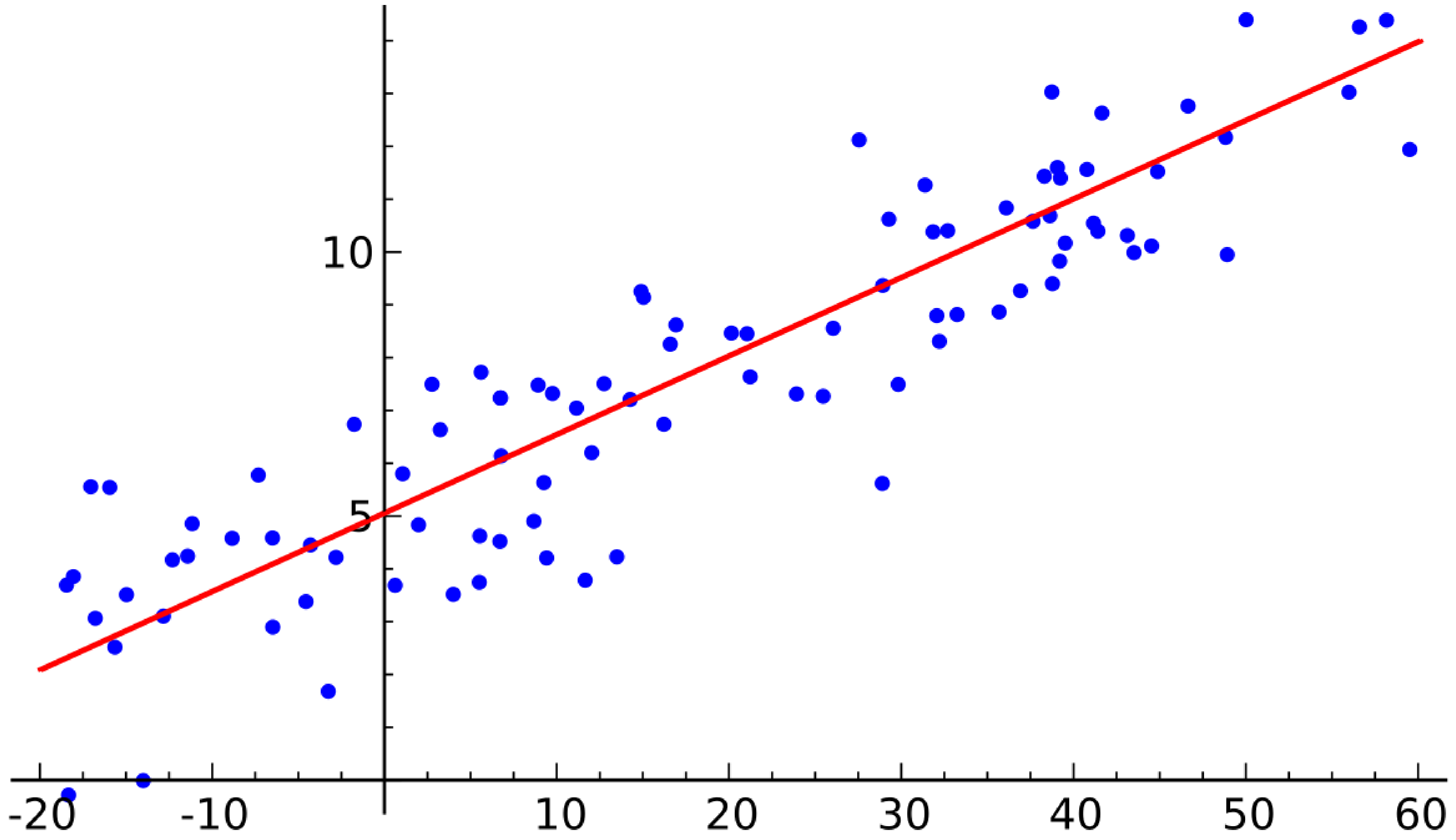
"An important feature of a learning machine is that its teacher will often be very largely ignorant of quite what is going on inside"

Alan Turing

- While we don't know *how* our brain converts input to output, we know what the output should be for every input.
- We can use this knowledge to teach the machine.

Machine Learning

- In modern terms:
  - Child machine: *Model*
  - Blank sheets: *Model parameters*
  - Teacher: *Loss function*



Input

Predicted Output

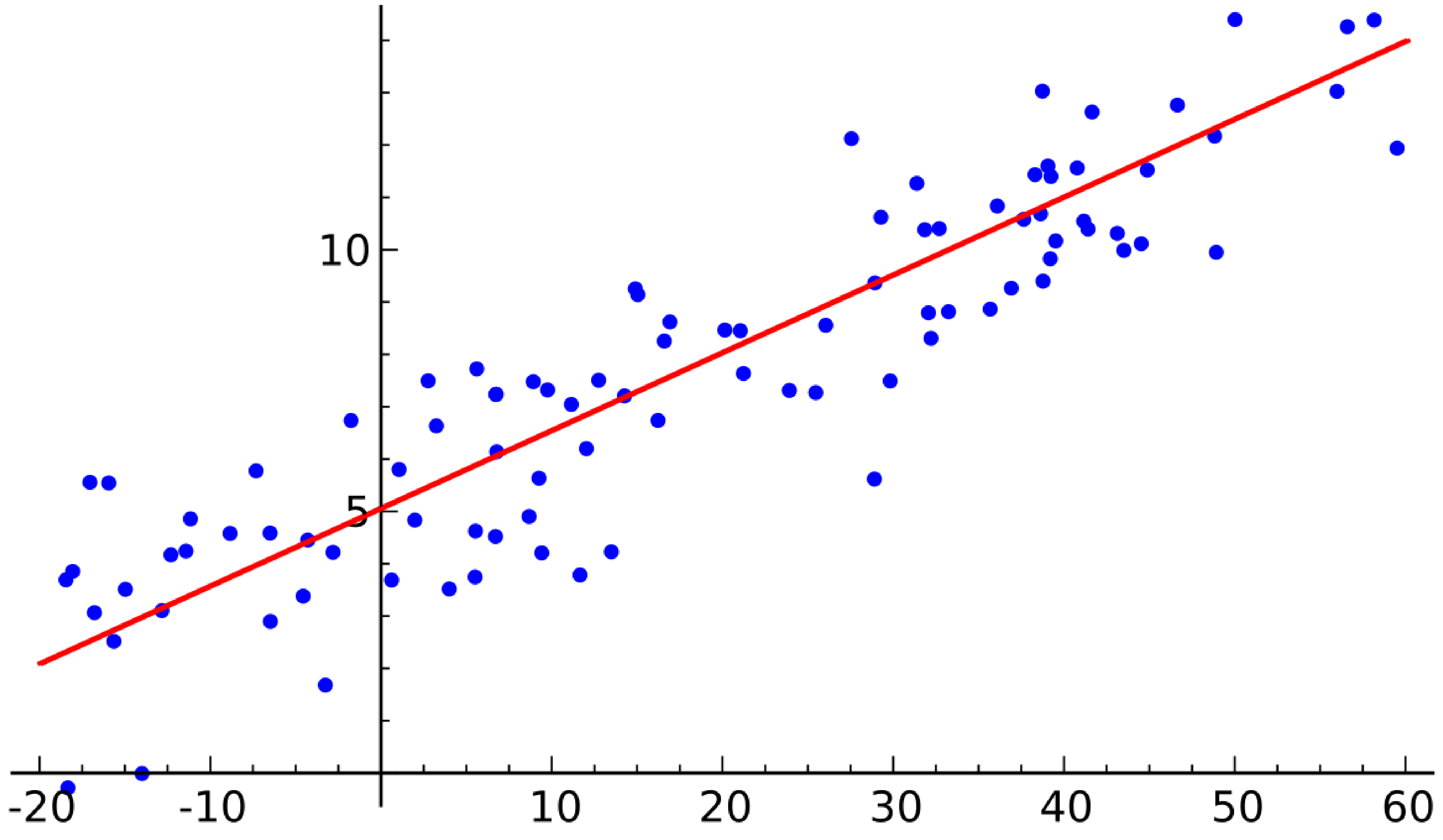
$$\hat{y} = wx + b$$

$$L = (y - \hat{y})^2$$

Desired Output

Machine Learning

- In modern terms:
  - Child machine: *Model*
  - Blank sheets: *Model parameters*
  - Teacher: *Loss function*



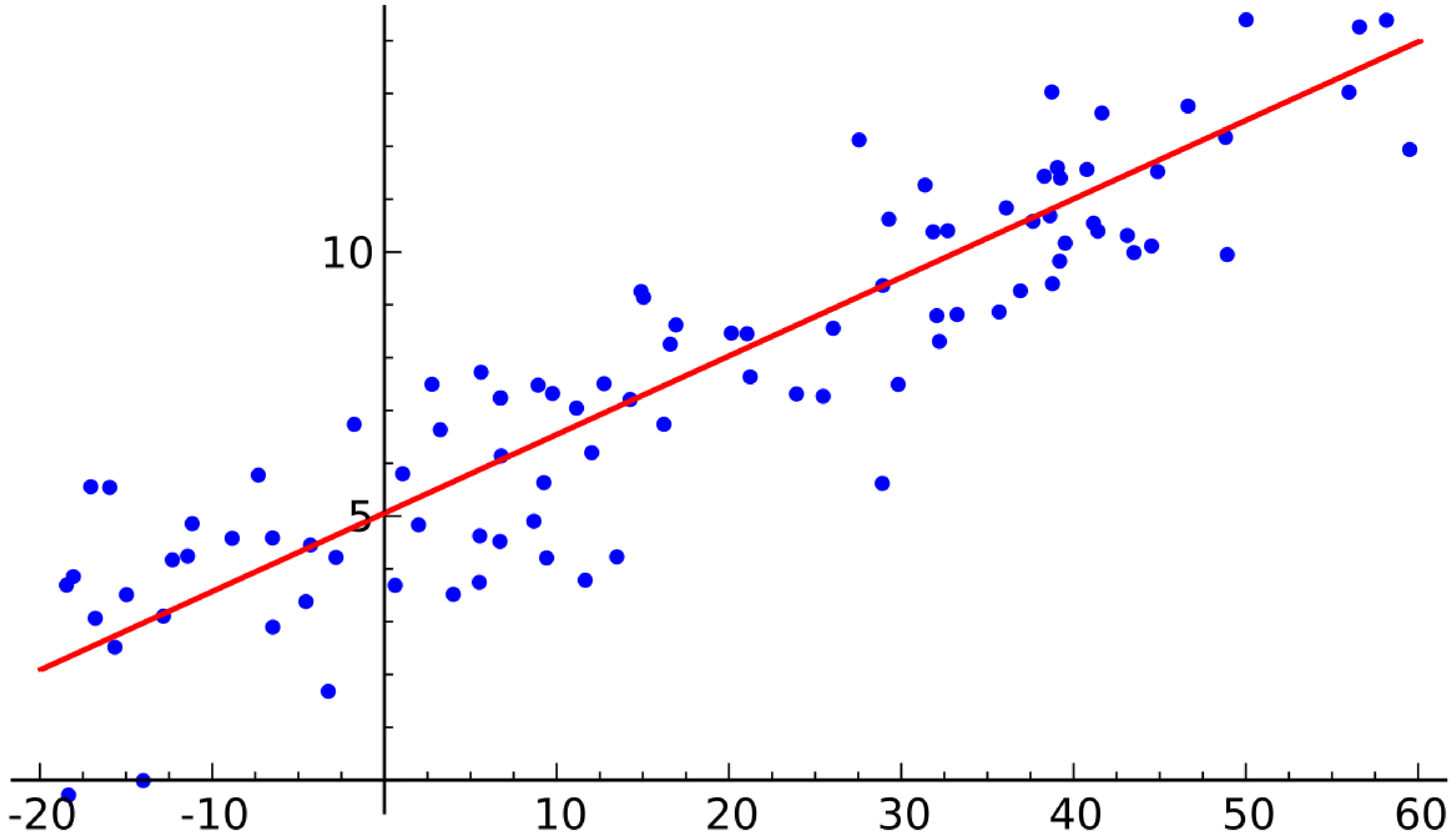
Parameters

Model → $\boxed{\hat{y} = wx + b}$

Loss Function → $L = (y - \hat{y})^2$

Machine Learning

- In modern terms:
  - Child machine: *Model*
  - Blank sheets: *Model parameters*
  - Teacher: *Loss function*



We want to find the parameter values that minimize the loss:

$$w^*, b^* = \arg\min_{w,b} L$$

Acknowledgement

- This course is modelled after CS189 at UC Berkeley taught by Prof. Anant Sahai.
- Slides for this course are kindly provided by Prof. Ke Li

How to Engage in Research as an Undergrad

- Take “Special Topics” courses

How to Engage in Research as an Undergrad

- Take “Special Topics” courses
- Volunteer in a research lab

How to Engage in Research as an Undergrad

- Take “Special Topics” courses
- Volunteer in a research lab
- Apply for an NSERC or VPR USRA

How to Engage in Research as an Undergrad

- Take "Special Topics" courses
- Volunteer in a research lab
- Apply for an NSERC or VPR USRA
  - Undergraduate Student Research Award
  - Stipend for 16 weeks of research
  - Usually apply by early January

How to Engage in Research as an Undergrad

- Take “Special Topics” courses
- Volunteer in a research lab
- Apply for an NSERC or VPR USRA
  - Undergraduate Student Research Award
  - Stipend for 16 weeks of research
  - Usually apply by early January
- Special research projects (CMPT 415, 416)

Multi-Agent Robotic Systems (MARS) Lab

Human-robot interactions

Human intent inference
Human-robot modelling

Control

Safety verification
Computational challenges
Multi-agent control
System identification

Learning

Data efficiency
Generalization
Multi-agent learning
Representation learning

MARS LAB

SFU SCHOOL OF COMPUTING SCIENCE

Real-time Formal Verification of Autonomous Systems with An FPGA

Minh Bui, Michael Lu, Rezah Hojabr,
Mo Chen, Arrvindh Shriraman