Machine Learning
Fall 2017

Course Team

- Lecturers
  - Yevgeny Seldin (course organizer)
  - Christian Igel
- Master TA
  - Lennard Oliver Hildendorf
- TAs
  - Artem Chupryna
  - Chloé Rouyer
  - Malte Bødkergaard Nielsen
  - Mathias Højgaard Jensen
  - Mostafa Mehdipour
  - Sandu Ursu
  - Steffen Czolbe
  - Steven Cheung
  - Xuwen Zhang
  - Yiran Zhang

Weekly Plan

	Mon	Tue	Wed	Thu	Fri
9:15 – 10:00	TA (x1)	TA (x1)			TA (x1)
10:15 – 11:00	TA (x1)	TA (x1)	Lecture 1		TA (x1)
11:15 – 12:00	TA (x1)	TA (x1)	Lecture 1		TA (x1)
12:00 – 13:15					
13:15 – 14:00	TA (x3)	TA (x1)	Lecture 2	TA (x1)	TA (x2)
14:15 – 15:00	TA (x3)	TA (x1)	Lecture 2	TA (x1)	TA (x2)
15:15 – 16:00	TA (x3)	TA (x1)		TA (x1)	TA (x2)
16:15 – 17:00					
23:59		**DEADLINE**			

- The first TA classes take place tomorrow
- You can attend any and as many TA classes as you like
- There will be some changes in the schedule in the first two weeks of January

Tentative Lecture Plan

	Morning	Afternoon
Yevgeny | Christian | Lennard | Y + C		
Mon, 19, Nov		Introduction to ML; Introduction to Supervised Learning; K-Nearest Neighbors; Validation; Cross-validation
Wed, 21, Nov	Regression	Markov's and Chebyshev's Inequalities; Q&A
Wed, 28, Nov	Linear classification; Perceptron; Logistic regression	Hoeffding's inequality; validation
Wed, 5, Dec	Kernels and basic kernel methods	Occam's razor, decision trees
Wed, 12, Dec	SVMs	VC analysis
Wed, 19, Dec	Random forests	VC analysis of SVMs
Wed, 26, Dec	Christmas	Christmas
Wed, 2, Jan		Clustering
Mon, 7, Jan		Neural Networks 1
Wed, 9, Jan	Neural Networks 2	PCA
Wed, 16, Jan	Summary Lecture	Course Evaluation

* *The exact order of the shaded lectures is still under consideration*

Home Assignments

- Weekly home assignments
- Every student must submit his/her own report
  - It is allowed to discuss the questions in small groups
- You have to score at least 50% to be admitted to the final exam
  - You can score 50% on all assignments or 100% on half of the assignments or anything in between – we are taking the average
  - We determine eligibility by taking a lower confidence bound on your score
  - Your submissions demonstrate your work throughout the course; if you end up close to the borderline 50% we will take the number of submissions into account

Exam Eligibility & Late Submissions

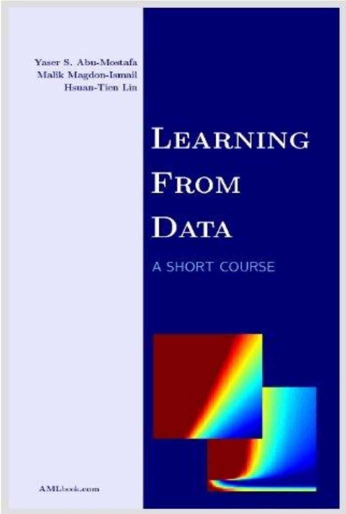
- You have to score at least 50% in the home assignments to be admitted to the final exam
  - Exam eligibility is determined in early January (the exact date is decided by the exam office) and only assignments graded by that date count
  - For re-exam eligibility all the assignments count
- Late submissions will not be graded
  - Irrespective of the reason
  - But we will count positively if you submit all assignments
- Do NOT notify us about late submissions
  - Irrespective of the reason
  - I.e., even if you were sick, do NOT notify us
- No resubmissions
- No grade complaints under 25 points mistake
  - We are happy to help you with the material, but we do not want to waste time on point counting: you only have to score above 50%

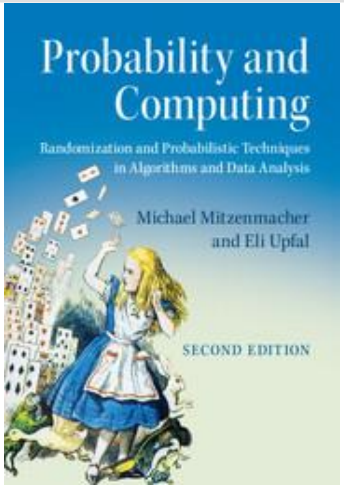
Final Exam

- Final exam: 5-days take-home exam in the exam week (deadline on Friday, 25 January)
- **Final exam must be solved individually**
  - You are NOT allowed to work in groups on the final exam
  - We will be very strict about cheating; if proven guilty you may be expelled from the university
- Final grade = final exam grade

Course Material

- **Primary material:** our lecture notes, slides, and blackboard
- **Supplementary material [MANDATORY]:** “Learning from Data” textbook
  - An excellent complimentary reading
  - It has additional chapters online at http://amlbook.com
- **Supplementary material [OPTIONAL]:** “Probability and Computing”
  - An excellent reading on Probability Theory
  - Highly recommended if you do not have a good background in probability theory
  - The first four chapters are relevant for the course





What is Machine Learning?

Examples of Learning Systems

- Biological
  - Animals
  - Humans
  - Plants (?)
- Machine learning

What is Learning?

- Ability to use past experience to take better actions in new situations

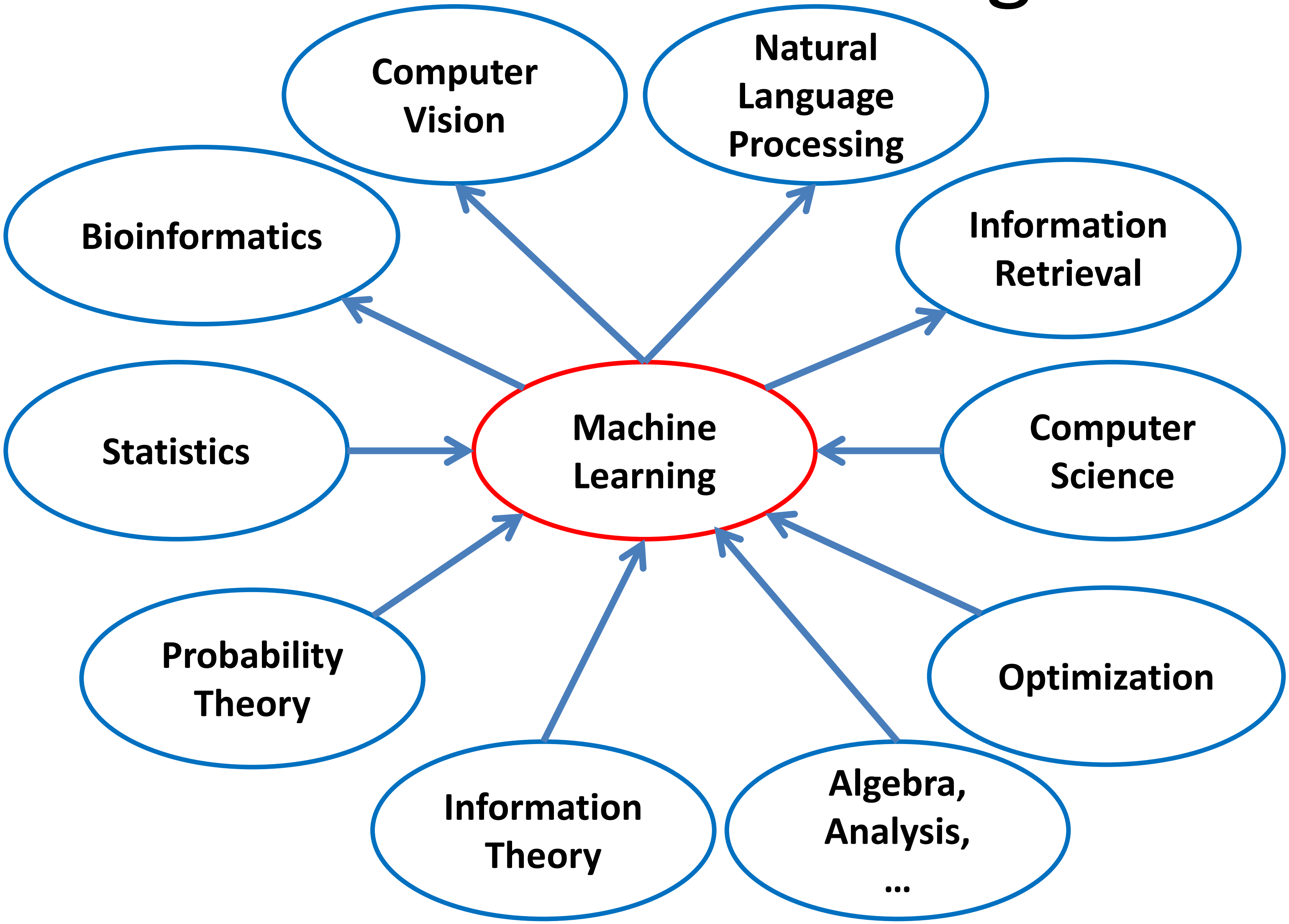
Success stories of Machine Learning

- Speech recognition
- Handwriting recognition
- Machine translation (e.g., Google translate)
- Car driving
- Human genome analysis

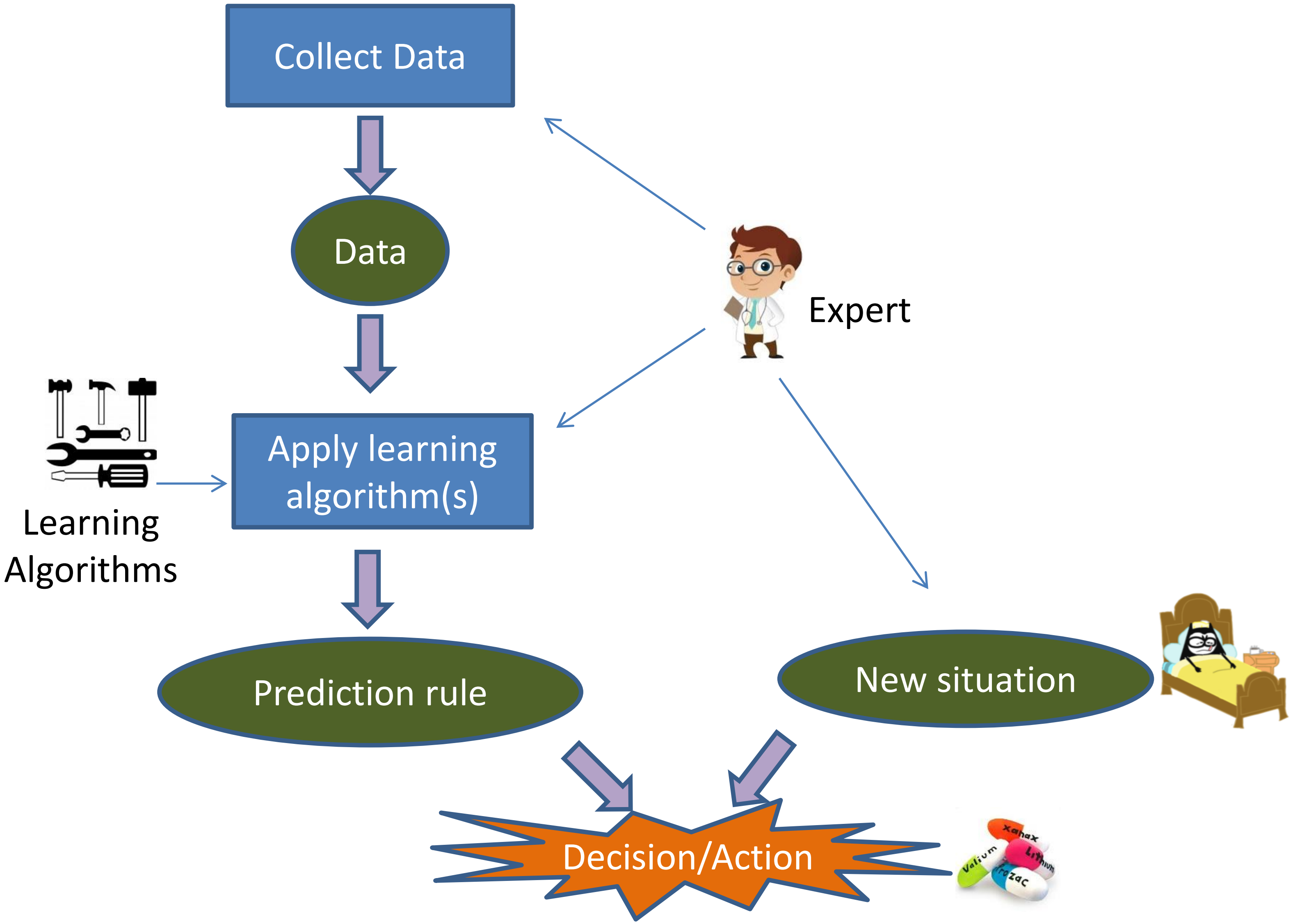
When do we need Machine Learning?

- Tasks that are too complex to program
  - Tasks performed by humans
    - Speech recognition, image understanding, etc.
  - Tasks beyond human capabilities
    - Analysis of large datasets: genetic data, medical data, Internet data, etc.
- Adaptivity
  - Adaptation in handwriting recognition; spam filtering; advertising; etc.

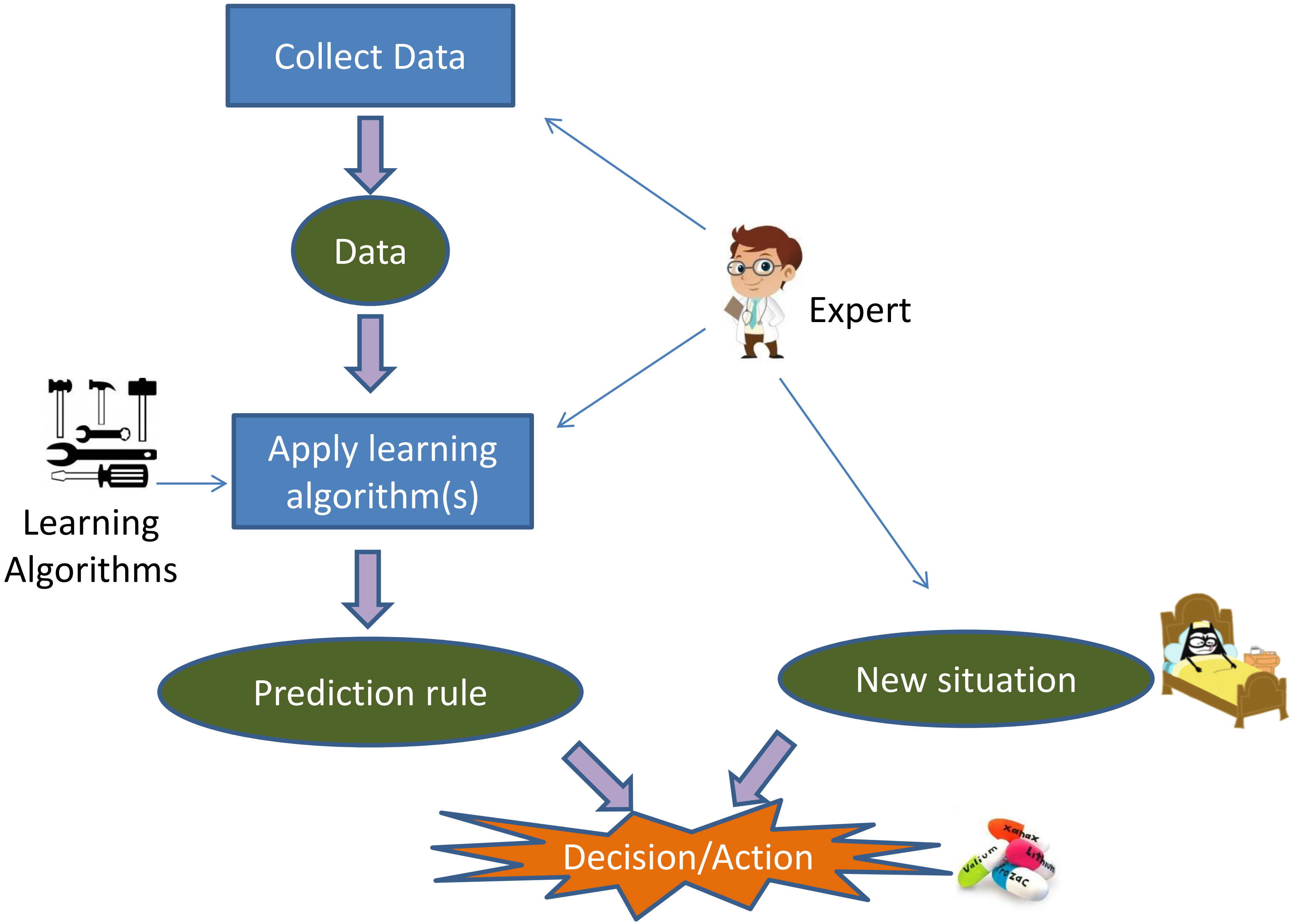
Machine Learning



“Classical” Learning Process



Where are you in this picture?



Be careful with Machine Learning!

- If you have learned to feed your data into ML algorithms and get something out, it does not yet mean that you know how to apply machine learning!

Common Pitfalls in Machine Learning

- Overfitting
  - May be internal and external to an algorithm
- Sampling Bias
  - External to an algorithm

Overfitting Example: Part 1

- A judge flips a coin to make a “fair” decision
- You ask the judge to give you the coin, you flip it 10 times and get all “heads”
- What is the probability of this happening if the coin is fair?
    - $1 / 2^{10} = 1 / 1024$
- Would you believe that the coin is fair?

Overfitting Example: Part 2

- A judge flips a coin to make a “fair” decision
- Everyone out of 10,000 spectators takes the coin home (in presence of a witness), flips it 10 times, and makes a video recording of the experiment
- If the outcome is unsuspicious, the spectator deletes the video and forgets about the experiment
- If a spectator gets “all heads” he/she uploads the video on youtube and loudly claims the decision was unfair
- How many times out of 10,000 experiments would you expect to observe “all heads” if the coin is fair?
  - $10{,}000 * (1 / 1024) \approx 10$
- What is the probability that in 10,000 experiments you never observe “all heads” if the coin is fair?
  - $(1 - 1/1024)^{10{,}000} \approx 1/20{,}000$
- Would you believe such a youtube video proof that the coin is unfair?

What does it have to do with ML?

- If we repeat an experiment sufficiently many times, even unlikely events may be observed with high probability
- Most applications of machine learning involve some form of randomness
- First and foremost, the data is a random sample from a population
- Machine learning tools are widely available and many people are applying them
  - = multiple experiments
- … but only “outstanding” results are reported
- The problem may be both internal and external to an algorithm

Sampling Bias Example

- US elections 1948
- Dewey predicted strong win by polls done over landline phones
- … but only upper class Americans had landline phones at the time…

What can we learn from this example?

- If we feed an algorithm with data that is not representative for the problem of interest we cannot expect to get reliable predictions
  - "Garbage in, garbage out" principle
- The problem is external to the algorithm
  - Algorithm's assumptions are violated

Recap

- Machine Learning algorithms cannot be applied as a black box
- We have to understand the effects of:
  - Overfitting
    - May be internal and external to an algorithm
  - Sampling Bias
    - External to an algorithm

Responsible Machine Learning

- Great power should be used with great responsibility
- Tremendous influence on all aspects of our life
  - Internet search
  - Social media
  - Medical research
  - …
- Machine Learning and Privacy
  - Inevitable conflict

Course goals

- Teach how to work with uncertainty
- Teach some machine learning algorithms
- Teach the assumptions behind learning algorithms
- Teach tools for analyzing the algorithms