The Algorithms Aren't Alright
Why Machine Learning Still Need Us

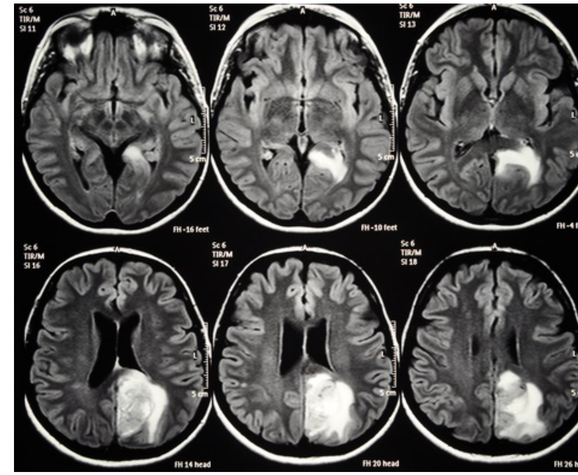
Professor Brett H. Meyer
Electrical and Computer Engineering
McGill University

September 17, 2019

Machine Learning: The 4th Revolution?

- Machine learning *is everywhere*
- ML **has a lot to offer**
- Medicine!
  - Automatic diagnosis using *computer vision*
  - Outperform human docs
- Transportation!
  - Self-driving vehicles will be *safer* and *more efficient*

rsipvision.com



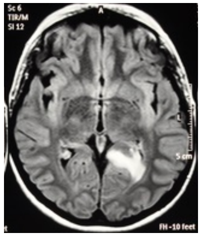
Machine Learning: The 4th Revolution?

- Productivity!
  - *Natural language processing* enables voice assistants, chat bots, and automatic translation
  - Helps us connect with each other and institutions

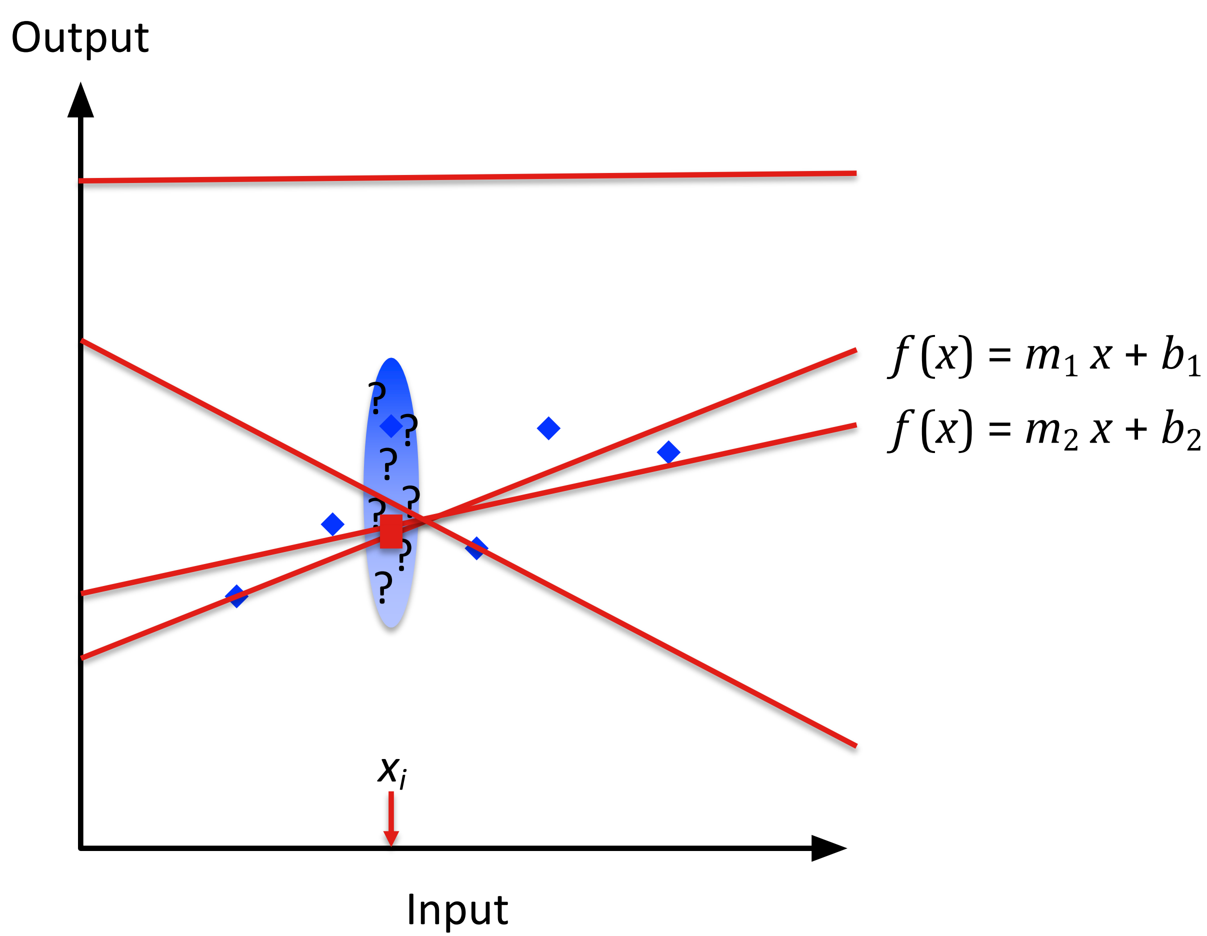
- Farming!
  - *Time-series forecasting* makes it possible to predict crop yield
  - Reduces costs for farmers and consumers alike

ML is All About the Data

- Basic idea: given some data describing some *system*, can SW build a model?
- Computer vision
  - Given lots of MRIs, learn which have tumors, and which do not
  - Given lots of pictures of road signs, learn which are stop signs, and which are not
- Forecasting
  - Given past environmental conditions and tomato yields, learn to predict future yield




Example: Linear Regression



Machine learning is an automatic ***approach to finding "***good***" values for*** m ***and*** b

What Could *Possibly* Go Wrong?

- f can be **fragile**
  - Small (imperceptible, even) changes in input can result in dramatic changes in output
- f can be **obtuse**
  - Your doctor says you have a tumor
  - ML says you do not …
  - … but it isn't clear why, from the math
- f can be **biased**
  - ML learns relationships between data
  - But correlation is not causation, and ML cannot tell the difference

The Algorithms Aren't Alright

- A little about me
- Brief overview of *machine learning*
- Introduction to *deep learning*
- Challenges in deep learning
  - *Robustness:* can learning algorithms be defeated?
  - *Explainability:* can we justify why deep learning makes any given choice?
  - *Bias:* can we make learning algorithms fair?

First, a Little About Me

- Computer scientist and engineer by training
  - U of Wisconsin BS'03
  - Carnegie Mellon U, MS'05, PhD'09
  - University of Virginia, Postdoc, 2009-2011
- Professor of ECE at McGill since 2011
- Research on computer system design
  - Making computers work when they're broken
  - Making it hard to hack airplanes
  - Making machine learning software (hardware) faster

RSSL

RELIABLE SILICON SYSTEMS LAB

Brief History of Machine Learning

- Artificial intelligence is as old as computing
  - Computing is older than you think
  - The first programmers debated it!
- *Deep learning* dates back to the 1940s
  - Has fallen into and out of fashion several times, and
  - Has not been practical until recently

Ada Lovelace, 1815-1852

Why is Machine Learning Hot Today?

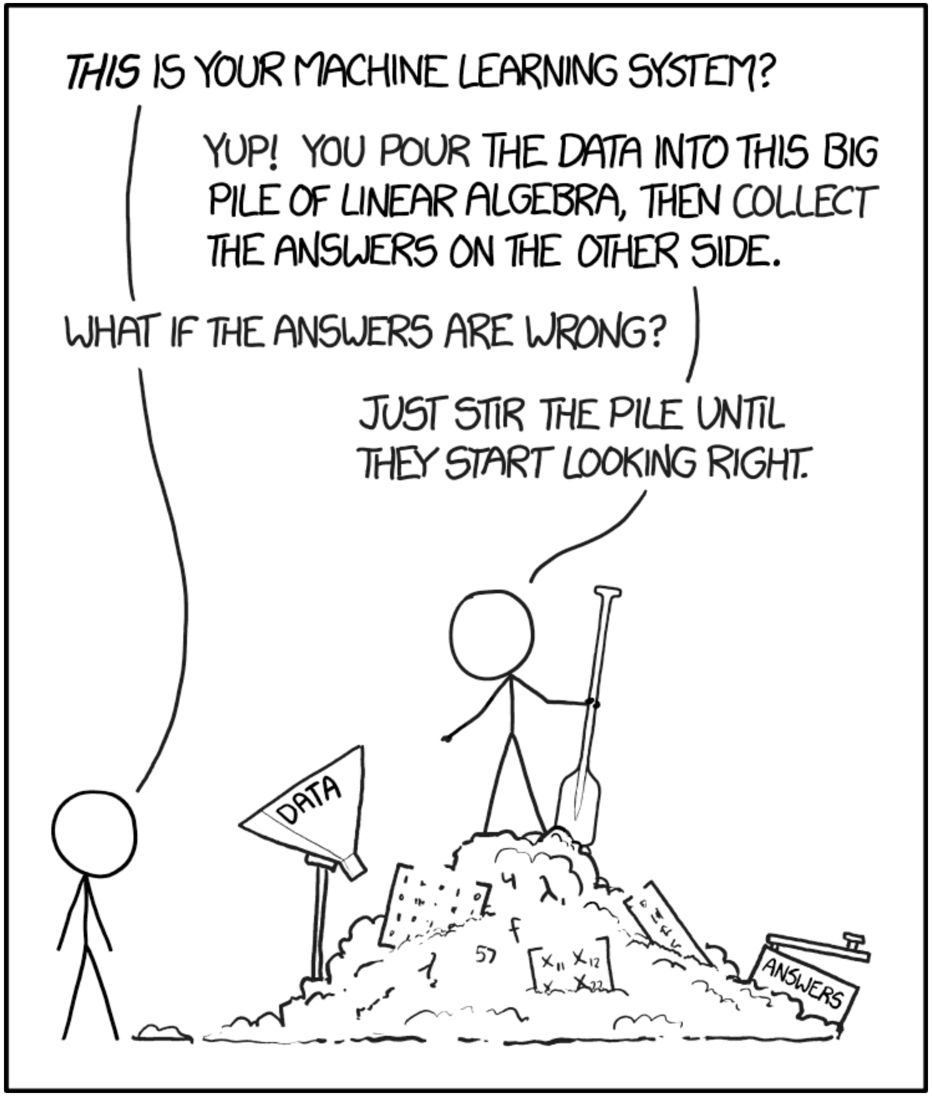
- Unprecedented computing power


- Unprecedented data

amazon

Typical Machine Learning Flow

- Collect, prepare data
- Configure ML model
  - Model structure, etc
- Train and evaluate
- Deploy ML model!



Source: xkcd.com

Data Collection and Preparation

- **Data is destiny**
  - What you collect determines what you can learn
- *Input features*
  - Columns describing the characteristics of a data point
- *Output features*
  - Columns representing what should be learned, given inputs

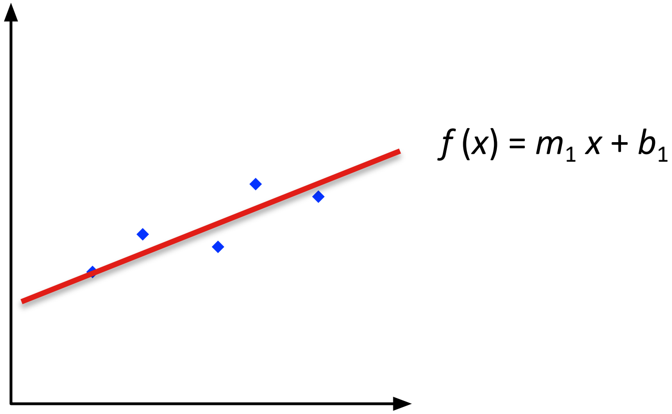
Sunlight (Hours)	CO_2 (ppm)	Tomatoes (Kg)
8	500	10
10	650	12

Data Collection and Preparation

- **Data is imperfect**
- Noisy measurement ***500? Or 550?***
- Missing columns ***Temperature?***
- Correlated columns ***Temp1 and Temp2?***
- Insufficient rows ***Enough days?***
- Unrepresentative rows ***Enough variation in days?***
- Unbalanced classes ***Enough variation in yield?***

Training and Evaluation

- ML algorithms start as *blank slates*
- *Training* adjusts internal variables to reduce error
  - Try a data point
  - Make adjustments
  - Repeat!
- Deploy!
  - … *and hope you used the right data*

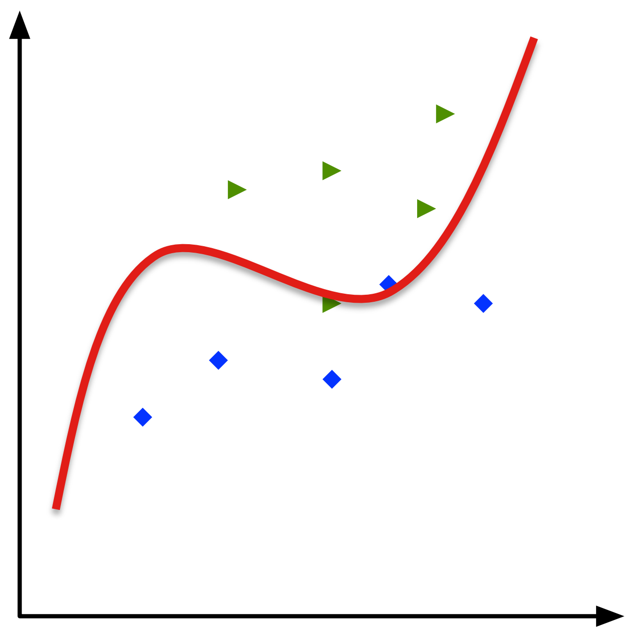


Types of Machine Learning

- Regression and kernel methods
- Decision trees
  - Random forests
- Neural networks
  - Deep learning

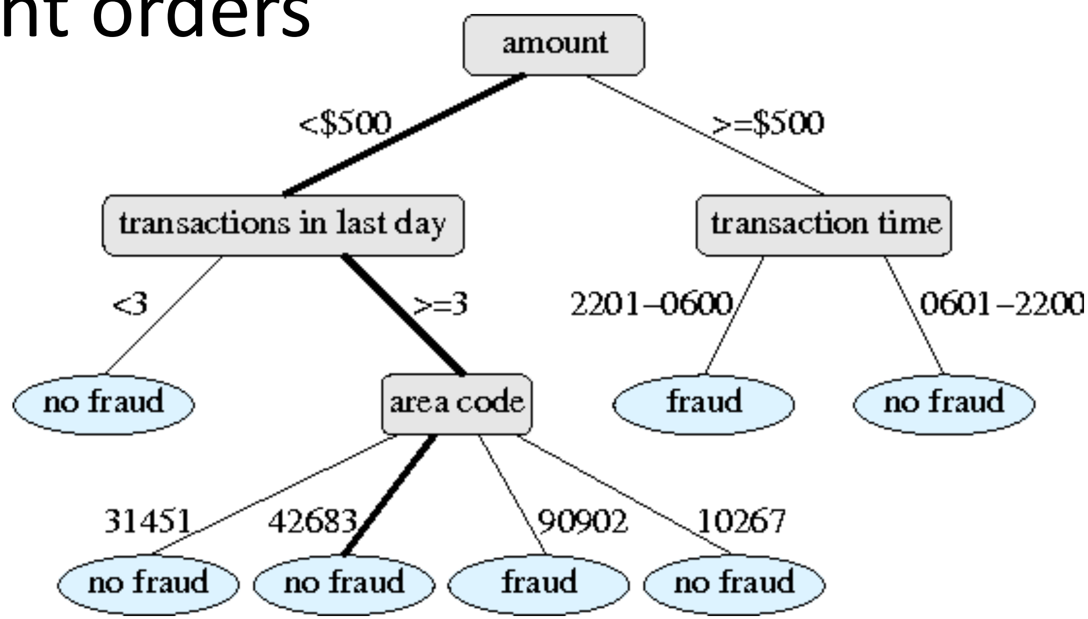
Regression and Kernel Methods

- *Regression*: fit an equation (e.g., a line or polynomial) to data
- *Kernel methods*: fit an equation so it divides data (e.g., above the line, cats, below, dogs)



Decision Trees

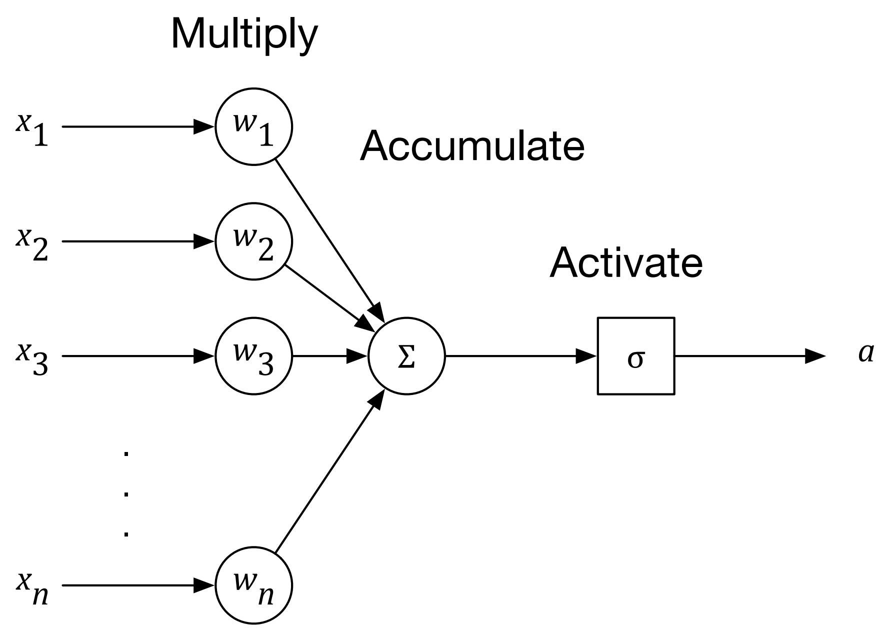
- Subdivide data based on input values
  - E.g., if *sunlight* > 8 hrs, and CO_2 < 500 ppm, then 10 Kg of tomatoes
- Random forests combine many trees with decisions in different orders



Kalyankrishnan, et al., CIKM 2014

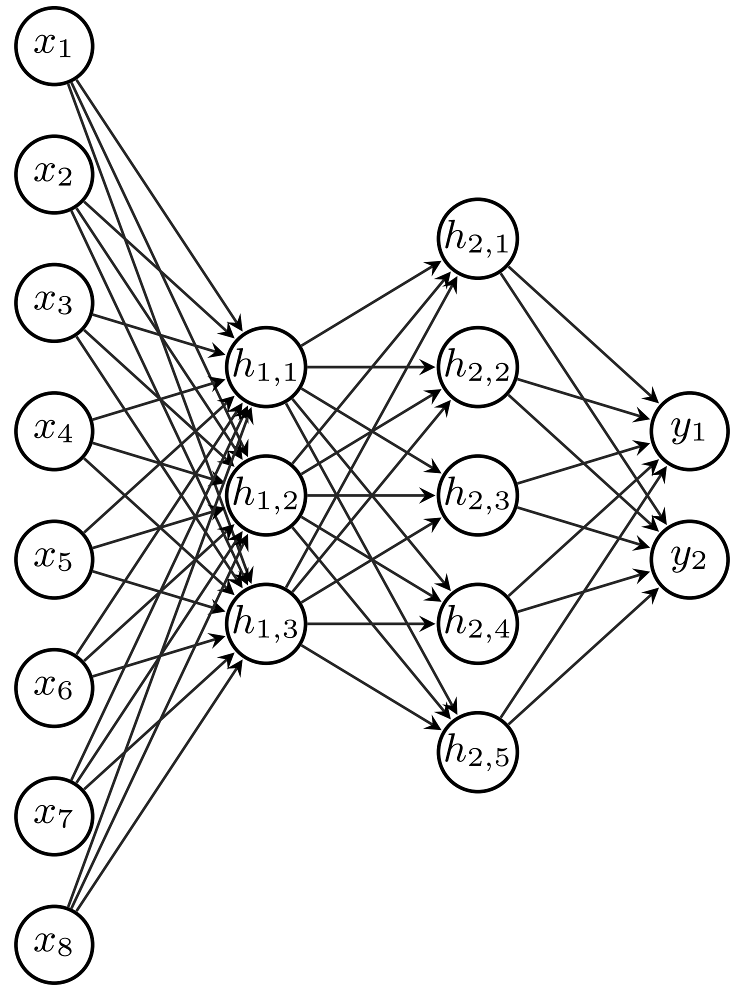
Neural Networks

- From a single neuron (*perceptron*), to 10s of layers of 100s of neurons (*deep learning*)
- Input features are carefully selected
- Weights *w* are selected through training

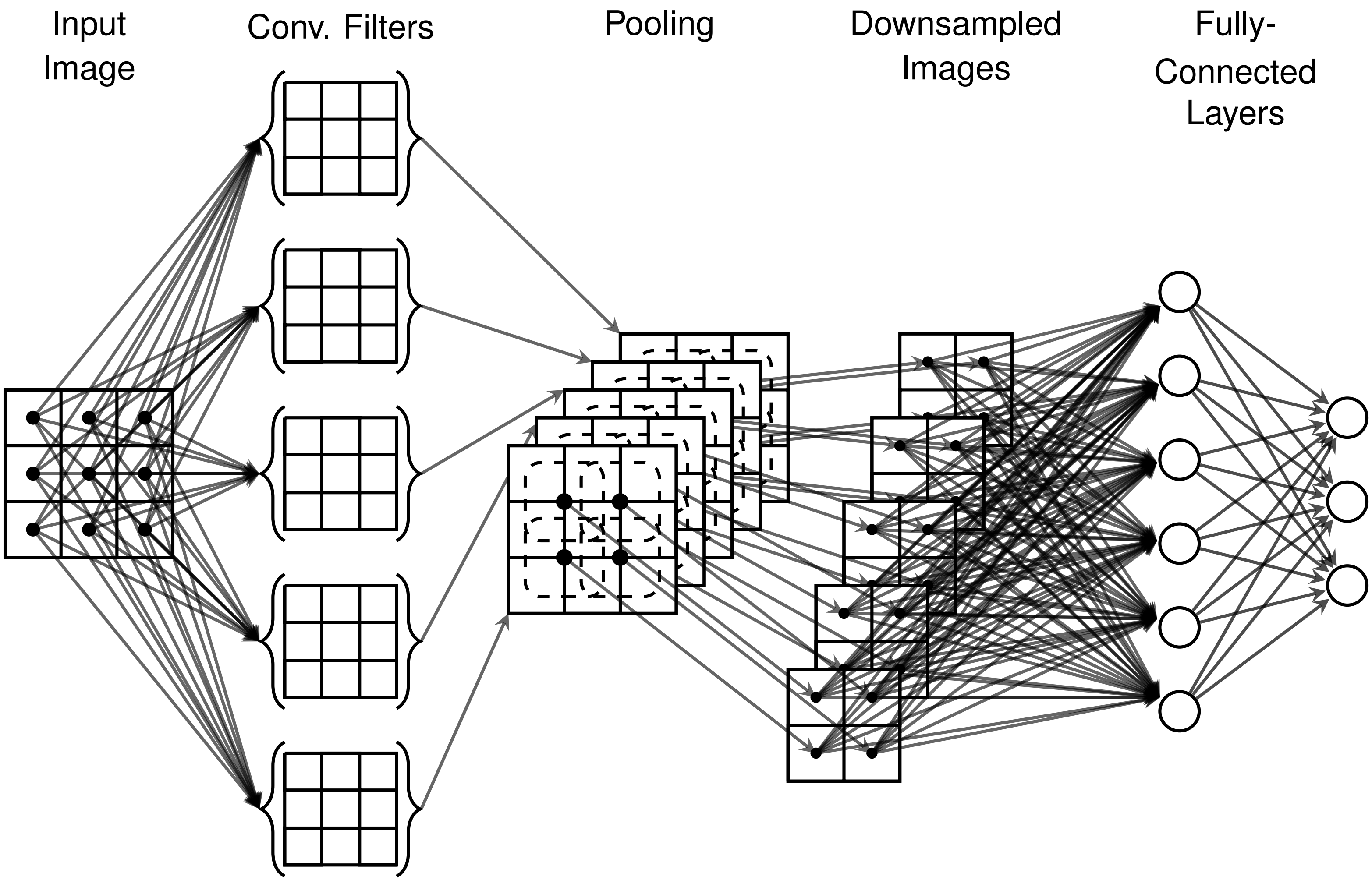


Multilayer Perceptrons

- Many neurons!
  - Learn more complex relationships
- Requires more data
- Takes longer to compute
  - Pick *hyperparameters* to balance **accuracy** and **latency**

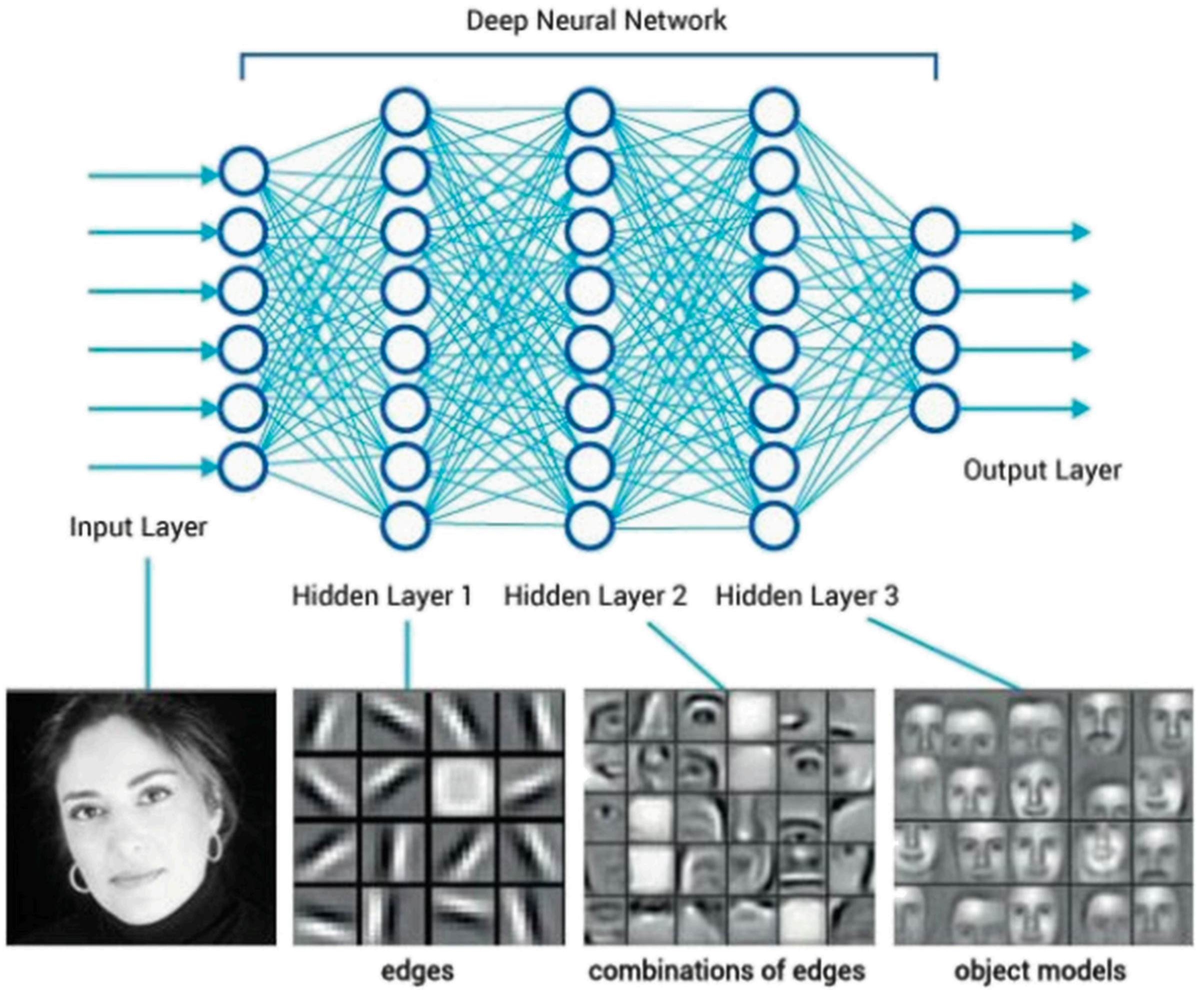


Convolutional Neural Networks



Deep Learning

- Deep networks extract features automatically



A Deeper Understanding of Deep Learning

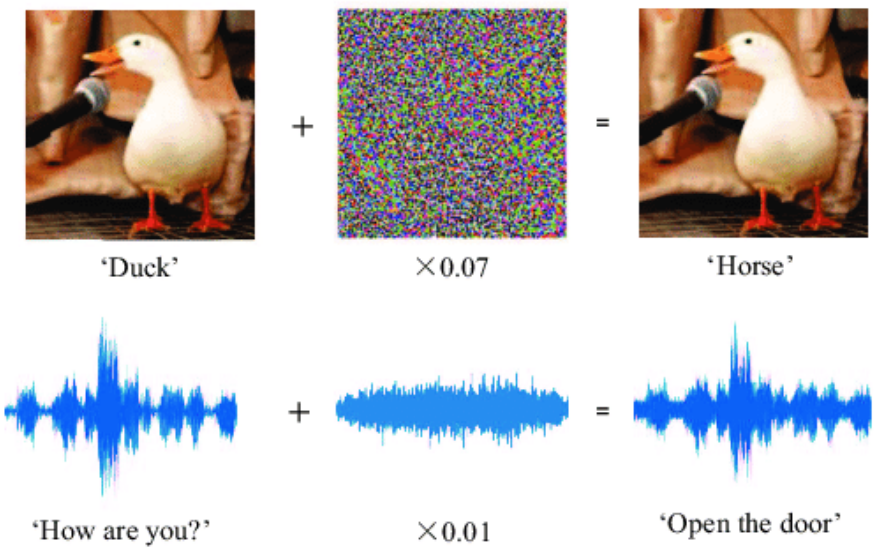
Deep Learning is Hard

- Deep learning requires **a lot** of *data*
  - For computer vision, 10s of gigabytes
- Deep learning requires **a lot** of *computation*
  - 100s of millions of *computations* per *data point*
  - 100s of millions of *data points*
  - 100s of *training runs* to get the weights right
- Correctly structuring the algorithm requires the right **tools** and **expertise**
  - 100s of trillions of different graphs are possible

And that's not all that can go wrong

Robustness

- *f* can be **fragile**
  ... because algorithms don't learn the way we do
- If we change the right thing in an input, we can control (disrupt) the output of the algorithm



Gong and Poellabauer, IoTSec 2018

Robustness

Eykholt, et al., CVPR 2018

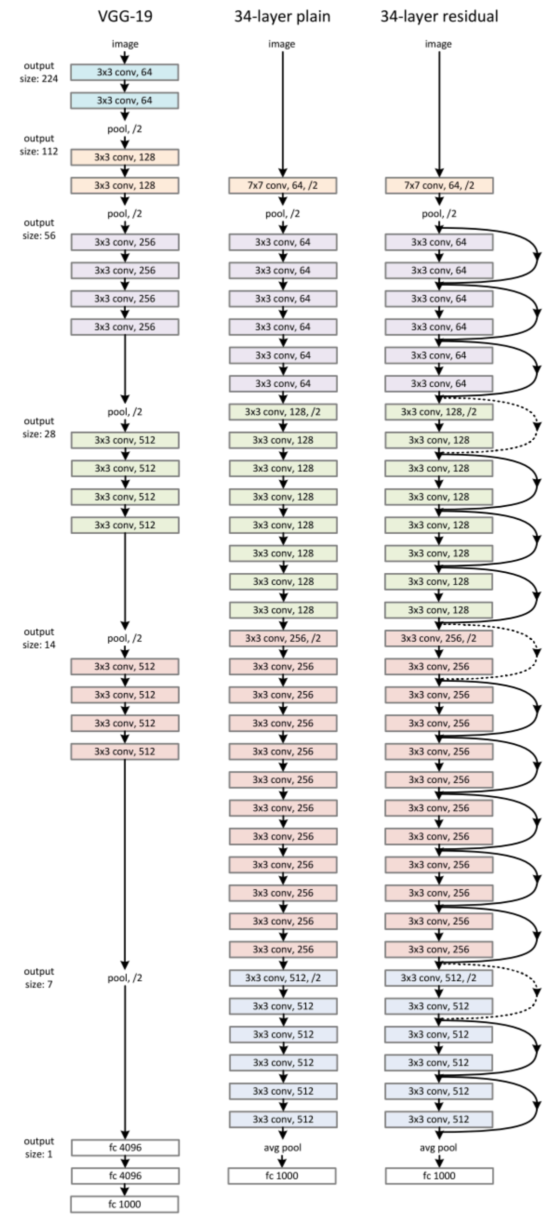
Athalye, et al., ICLR 2018

Caveats

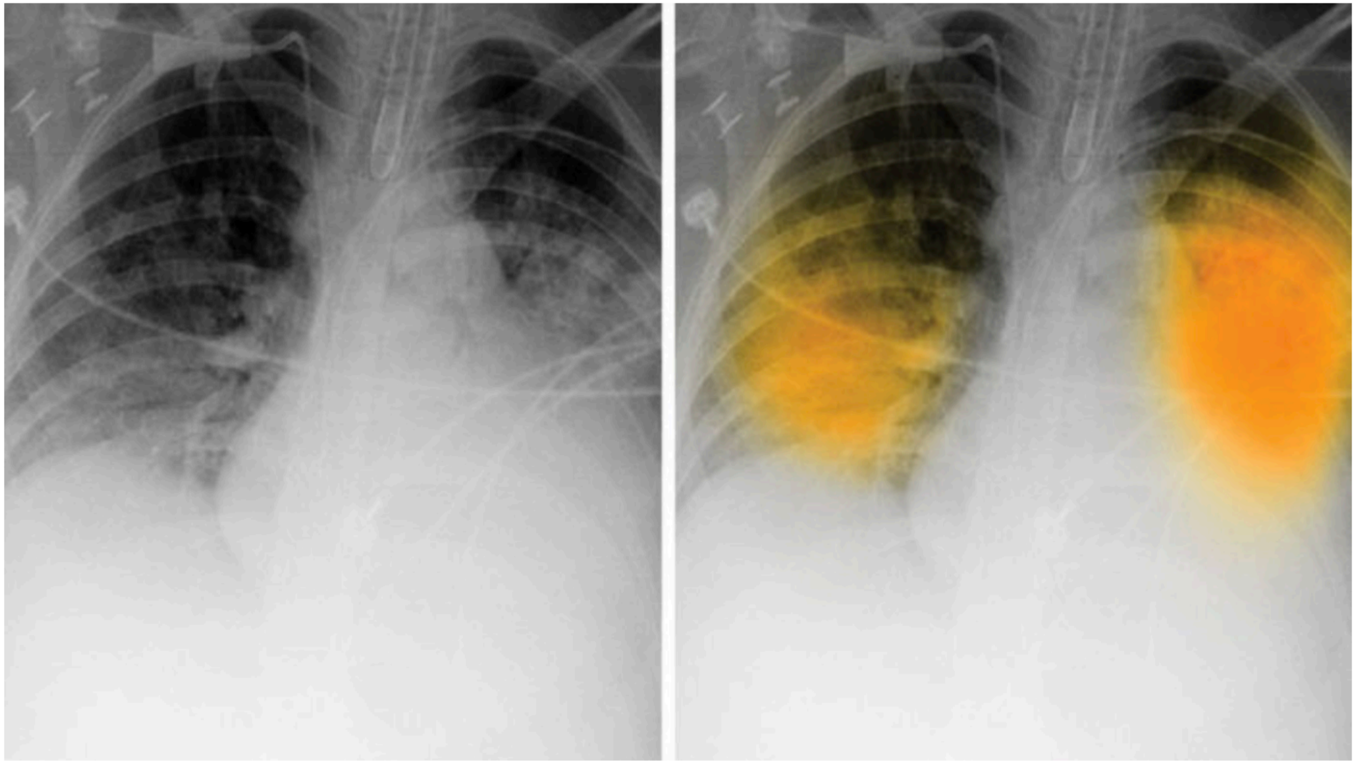
- *Adversarial examples* require knowledge of the algorithm being attacked
- Self-driving cars face greater challenges than strategically defaced stop signs
- *Take away*: be careful you entrust to ML, because it can be attacked

Explainability

- *f* can be **obtuse**
  … because, you know, *100s of millions of calculations*
- If there is a human in the loop, they must be able to trust the ML



Where are we on trust?



Scientists are developing a multitude of artificial intelligence algorithms to help radiologists, like this one that lights up likely pneumonia in the lungs. ALBERT HSIAO AND BRIAN HURT/UC SAN DIEGO AIDA LABORATORY

Artificial intelligence could revolutionize medical care. But don't trust it to read your x-ray just yet

By **Jennifer Couzin-Frankel** | Jun. 17, 2019 , 12:45 PM

[Science]

Caveats

- New approaches to algorithm design are needed to increase transparency
- New protocols are needed for collaboration with ML algorithms
- *Take away*: be careful what you entrust to ML because it may not be able to explain itself

Bias

- *f* can be **biased**
  … because humans are biased
- If algorithms make decisions that affect people, extra care is needed to ensure fairness

Algorithmic Decision-making FTW!

- Machine learning can improve consistency in decision making
- Consider: asylum judges and loan officers
  - Timing of the decision is correlated with decision
  - Past decisions are anticorrelated with future decisions

Bias in AI

- **Do we know we have the right data?**
- College admissions
  – What makes a successful student?
- Insurance
  – What makes someone a risk?
- Mortgages
  – Why do people default?
- Sentencing
  – Why causes recidivism?

Crosscutting Issue: Accountability

- When something *does* go wrong, who is at fault?
- **All stakeholders!**
  - Data providers
  - Algorithm designers
  - Algorithm integrators
- European Union is a world leader in ethical AI

Conclusions

- ML is *really* here, and has **a lot** to offer!
  – Medicine, transportation, productivity, agriculture, …
- Data is destiny
  – If you haven't measured it, ML can't learn it
- ML must be made **robust**
- ML would benefit from being **explainable**
- ML cannot be allowed to be **biased**
- Practitioners are responsible for appropriate data collection, training, evaluation, and deployment!

Thank you!

http://rssl.ece.mcgill.ca