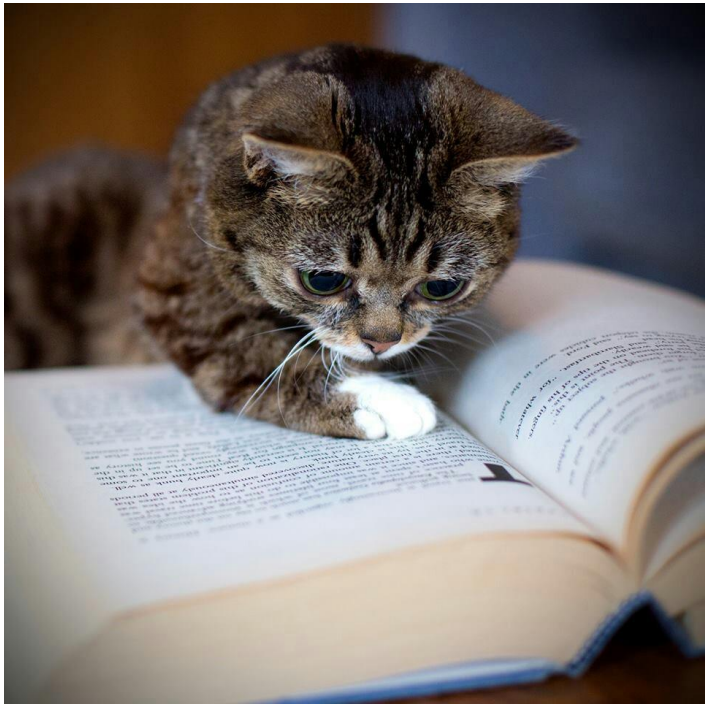
BONUS: MACHINE LEARNING

BME 121 2016

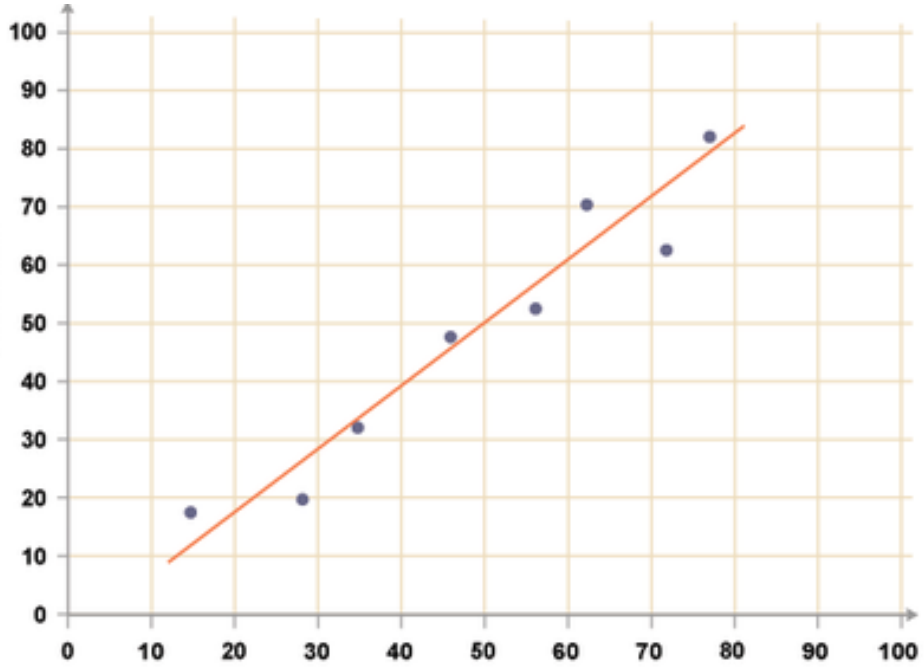
Jeff Luo



Line of Best Fit

- “Models of best fit”: true purpose is to find a mathematical model that best approximates real world data
- Line: Find m and b for

  $y = mx + b$
- A model is always imperfect (has error), best model wins

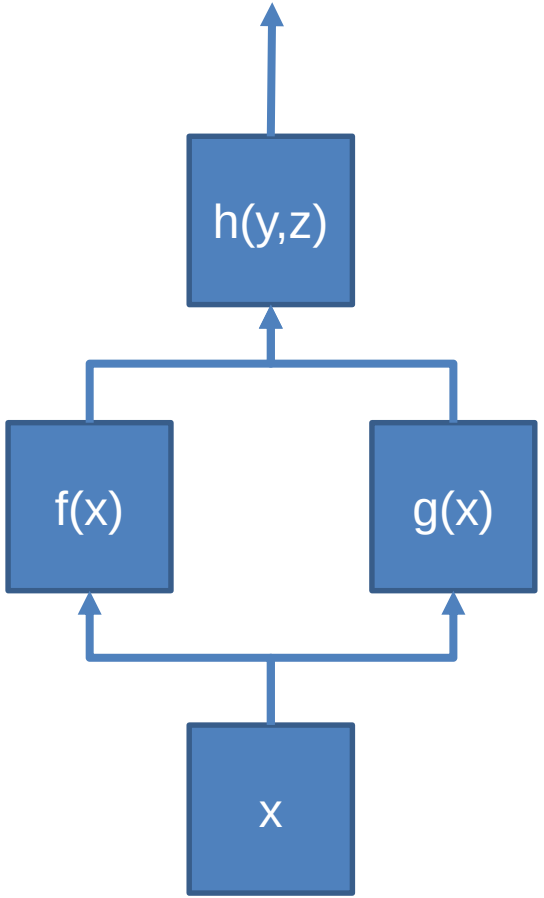


Curve of Best Fit

- This can actually be broken down into multiple lines of best fit:

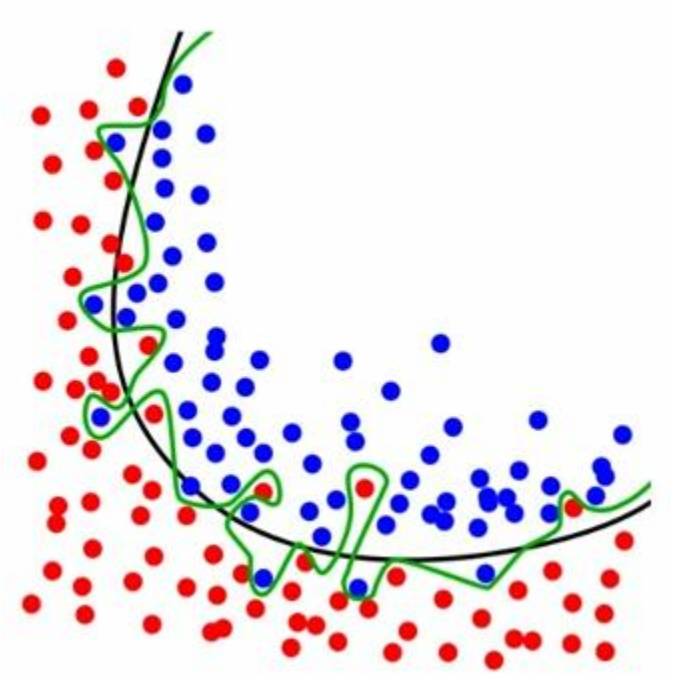
- $y = f(x) = mx + b$
- $z = g(x) = nx + c$
- $a = h(y, z) = yz + d = (mx + b)(nx + c) + d$

- And so on…

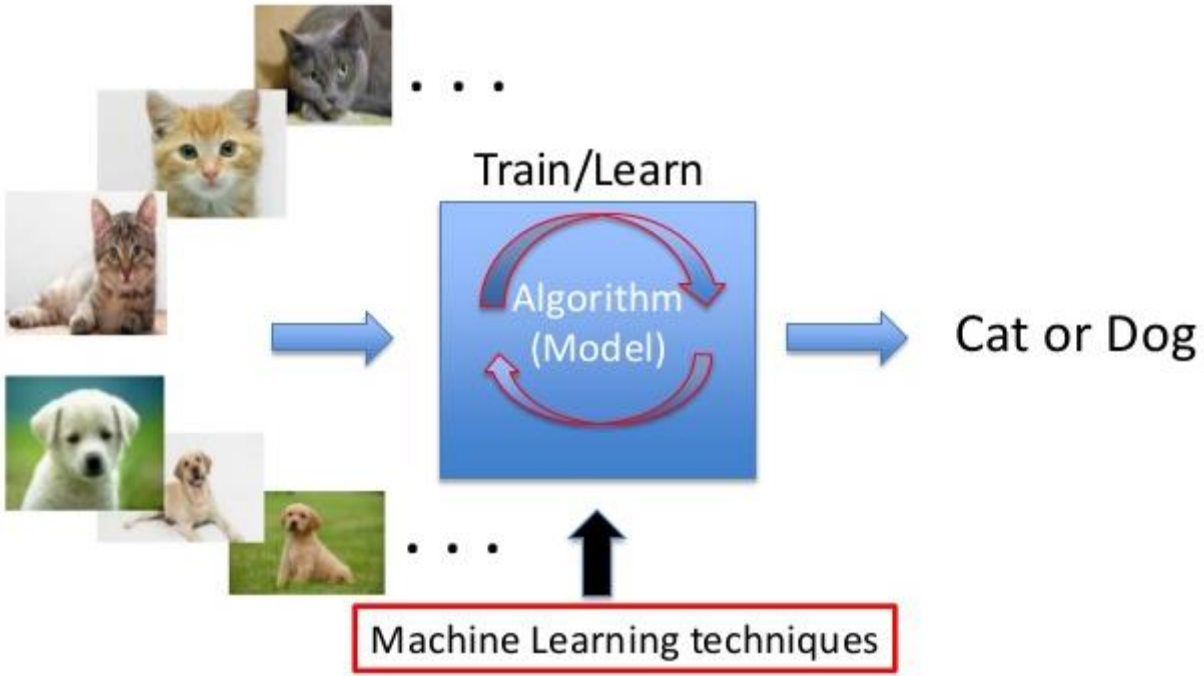


Machine Learning

- Basic idea:
- Gather a lot of data (millions of data points)
- Learn (find) a model that best fits the data:
  - Able to discriminate between two kinds of things
- But don’t overfit the data
  - Green curve – very accurate, but can’t generalize to new data
  - Black curve – some error, but more general
- Statistical idea: the data gathered for finding the model is a sample of all possible data, and there’s always a bit of randomness to real data



Machine Learning

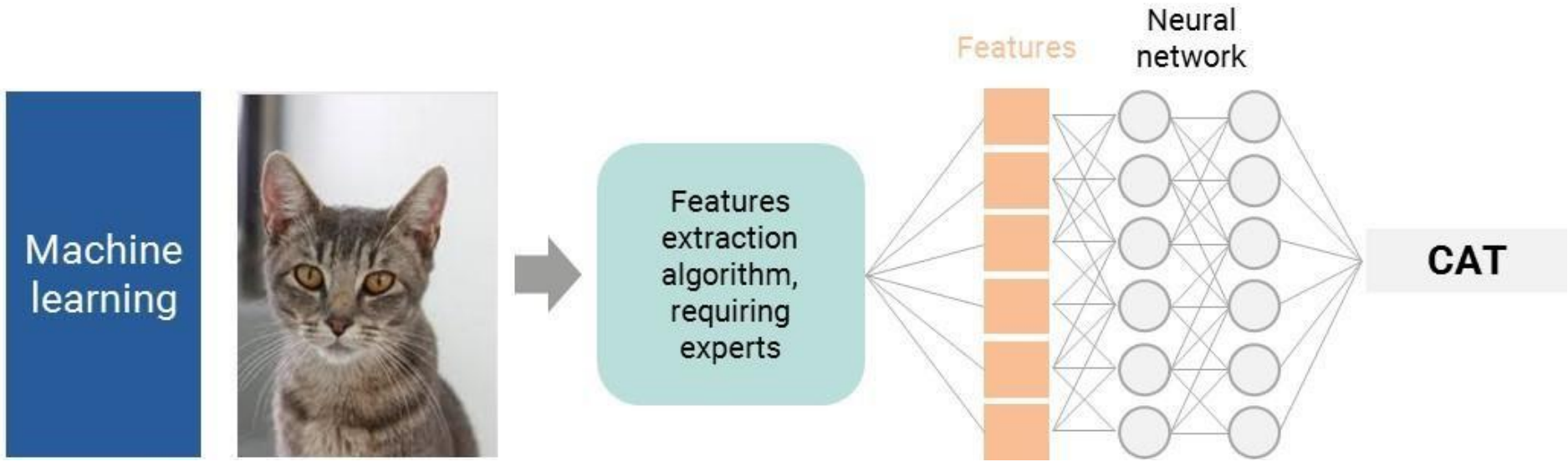


Working with Complex Data

- Booleans:
  - y = f(x) where f(x) could be some Boolean operation, eg !x, x || true, x && false, …
  - Usually Boolean building blocks are functions of two input Booleans:
    - x && y, x || y, x == y, etc
- Numbers:
  - mx + b where m, x, and b are numbers
- Strings, music, images, videos (anything normally stored as an array):
  - xm + b where m, x and b are matrices:

$$\begin{bmatrix} x_{00} & \cdots & x_{0j} \\ \vdots & \ddots & \vdots \\ x_{i0} & \cdots & x_{ij} \end{bmatrix} \begin{bmatrix} m_{00} & \cdots & m_{0j} \\ \vdots & \ddots & \vdots \\ m_{i0} & \cdots & m_{ij} \end{bmatrix} + \begin{bmatrix} b_{00} & \cdots & b_{0j} \\ \vdots & \ddots & \vdots \\ b_{i0} & \cdots & b_{ij} \end{bmatrix}$$

Neural Network



A neural network is a complex model built up using simple neurons (nodes)

Eg each neuron computes a line of best fit

Features are things that strongly help categorize and discriminate input data

Eg for cats, shape (via edge detection), and colour (cat's aren't purple…) are some good features.

Machine Learning – Google Search Demo

- Open up an incognito / private web browser tab
- Go to google.com
- Type in University of Waterloo
- Clear the search, type in the letter P and see what are the auto complete suggestions
- Repeat (with a new incognito / private tab) using another university's name
  - Try a university in a different country
- Repeat by first entering a musician's name, movie name, …
- What's happening? Google learns related search queries from everyone, including yours
  - Incognito / private tabs are required for experiments so that your own past search history doesn't influence the suggestions

Machine Learning – Google Drawing Demo

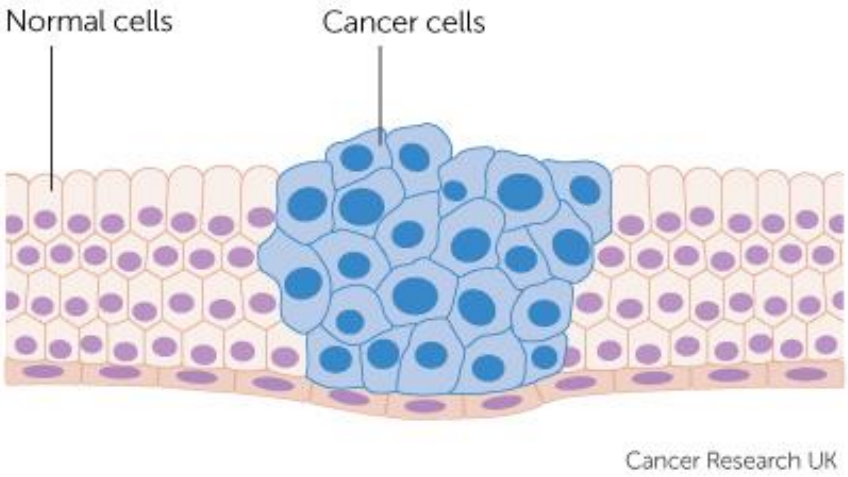
- Set your speakers to moderate volume
- Go to https://quickdraw.withgoogle.com in a normal browser tab

- Task: sketch 6 pictures, each under 20 seconds
- Google will predict what you're drawing along the way

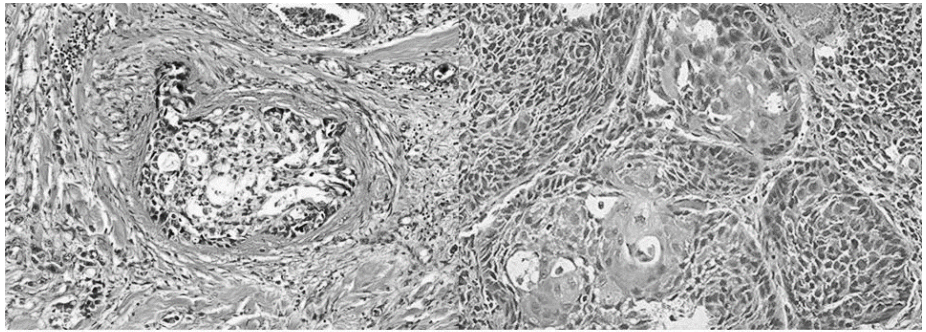
- At the end, click on each of your drawings to see what kind of drawings other users have given to Google to train a model of how something is drawn
- Image search is this model used backwards: given words, find the pictures with the strongest correlation

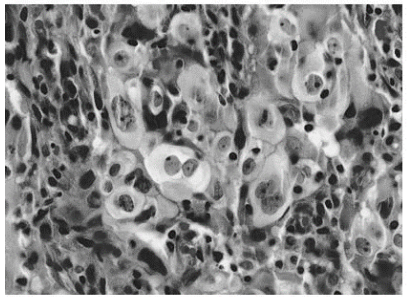
Machine Learning in Medical Use

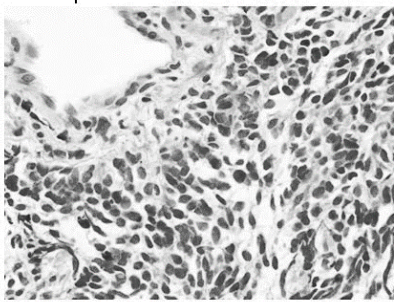
- Cancer shown in text books:



- Scans of cancer cells in reality:


Adenocarcinoma

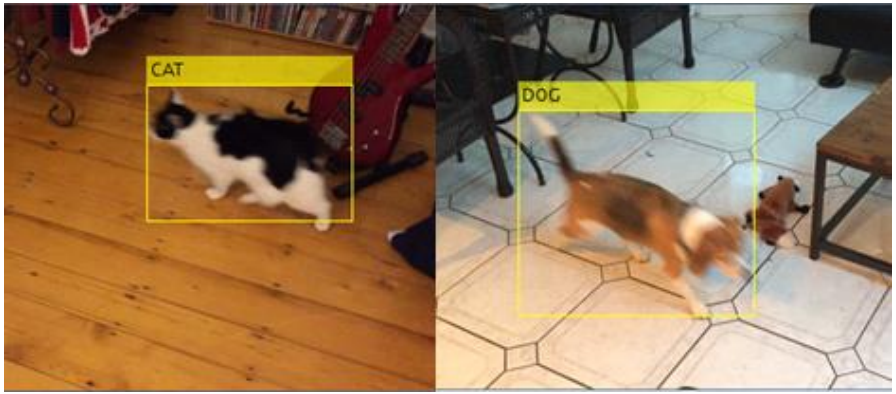

Large Cell Tumor


Small Cell Lung Cancer

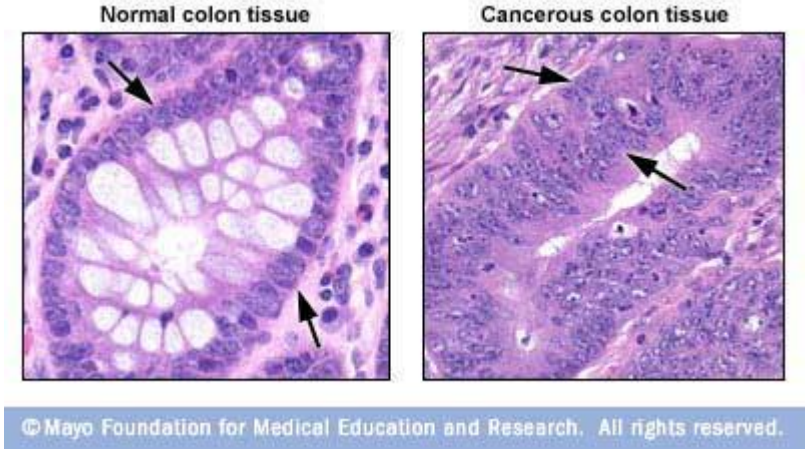
Squamous cell Carcinoma

Machine Learning in Medical Use

- Can we train an image model that predicts whether a picture is a cat or dog?



- Can we train an image model that predicts whether a cell is normal or cancer?



Machine Learning in Medical Use

- Not all medical images are equal
- CAT / MRI scans reveal structure
  - CAT subjects drink contrast substance
  - MRI subjects can't have metal implants
- PET scan reveals functionality
  - Radioactive tracers injected in subject body, which detects differences in metabolic and chemical activity in the body
- CAT + PET most accurate for cancer detection (currently)
- Can we learn a model that predicts cell functionality based on cell structure?
- Yes (2016 results): given CAT/MRI scan, predict whether cells are:
  - normal / cancerous
  - alive / dead
  - Etc…
- Future idea: surgeons can use live video feed of a CAT/MRI scan to guide cancer surgery
  - + Augmented Reality

Machine Learning in Medical Use

- Recent news (Nov 29, 2016):

- “Google researchers trained an algorithm to recognize a common form of eye disease as well as many experts can.”

- https://www.technologyreview.com/s/602958/an-ai-ophthalmologist-shows-how-machine-learning-may-transform-medicine/