What is Machine Learning?

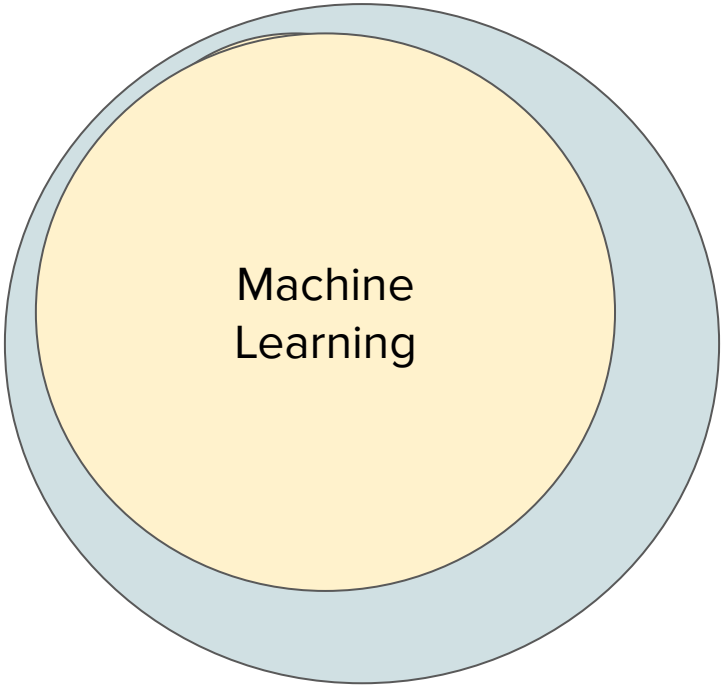
ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

Outline

- Why machine learning?
- What is machine learning?
  - How does it work?
  - How can we train it?
  - What can it do?
- Going to the real world
- Challenges
- Wrap-up

ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Is machine learning the same as AI?



Not to scale

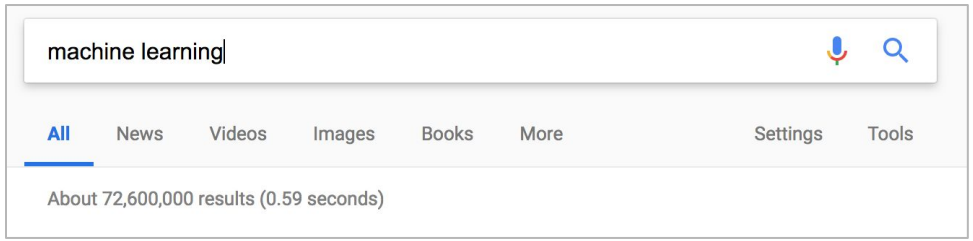
Why is machine learning so exciting?

Find patterns in data that are too complicated for a human to detect.

Find solutions to problems that are difficult to solve with traditional programming.

AI2 ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

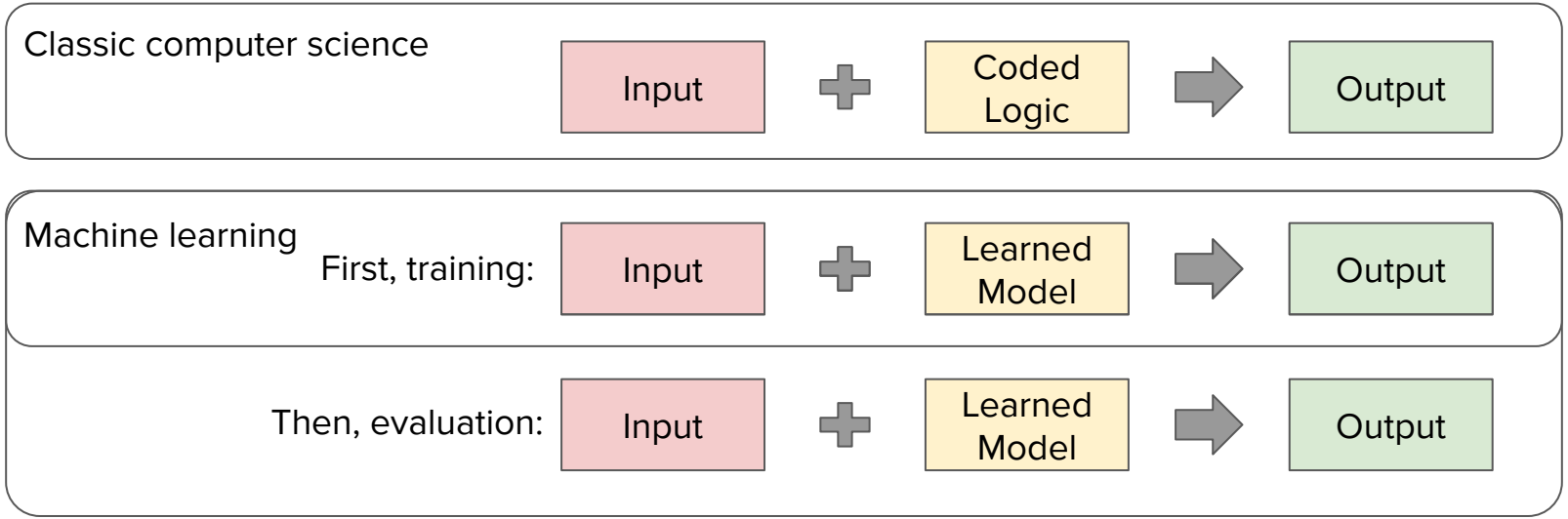
Why is machine learning gaining popularity?



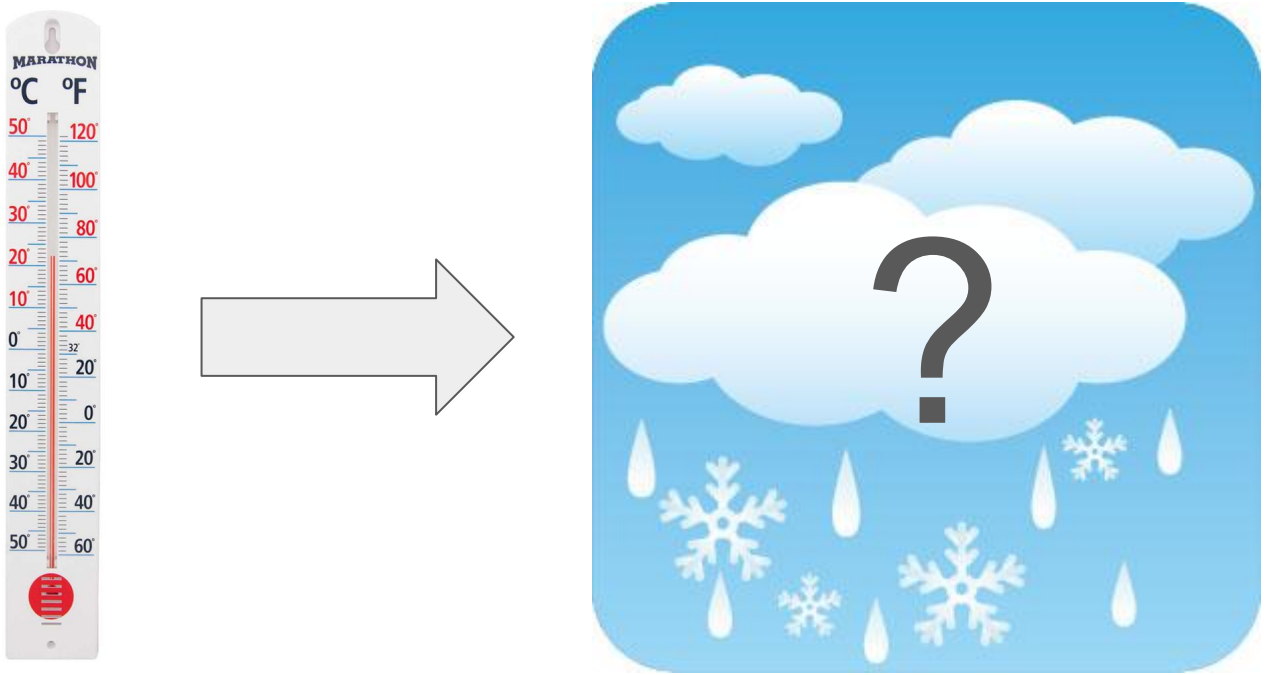
- Solves hard problems
- More accessible all the time
  - More data
  - More compute resources
  - Better tools

What is machine learning?

Computers learn to solve problems without being explicitly programmed.

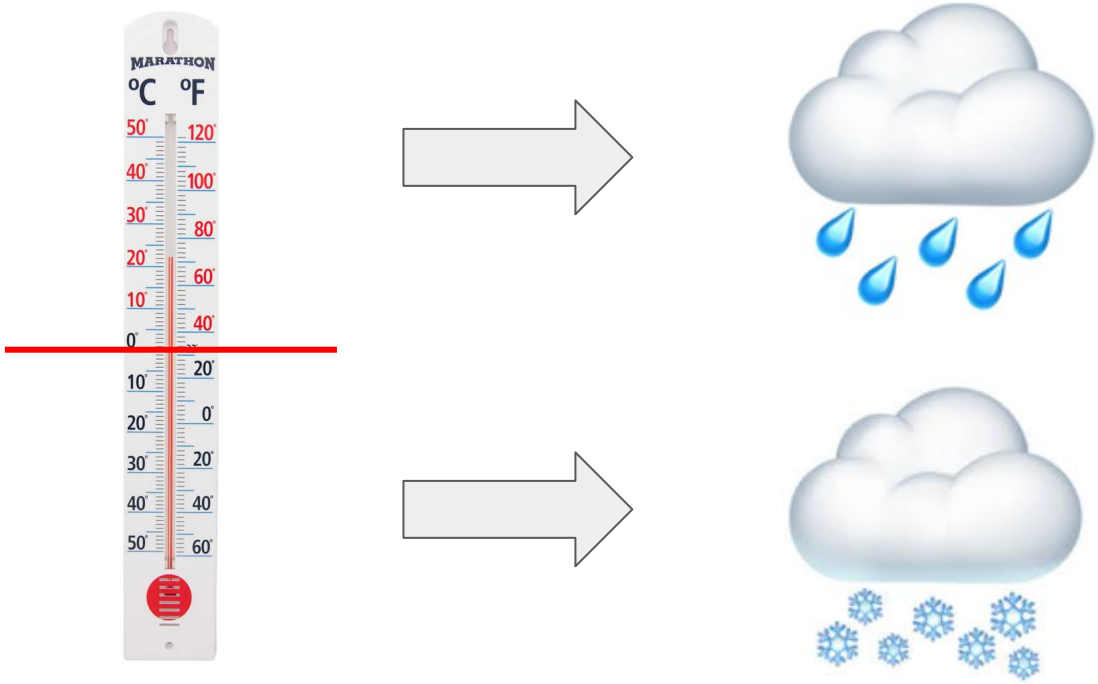


Given temperature, can we predict rain or snow?



ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

How would you solve the problem?



Computer needs to figure it out from data.

Recorded past weather data.

Temperature in °F	Precipitation
27	Snow
52	Rain
65	Rain
10	Snow
71	Rain
...	...

Training the model

1. Find min and max temps.
2. For each value between min and max temps, calculate how accurately it separates “rain” temps from “snow” temps.
3. Return best predictor temp and its accuracy.

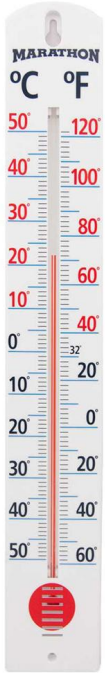
[Watch training happen.]

27	Snow
52	Rain
65	Rain
10	Snow
71	Rain
17	Snow
...	...

Training data

variable	value
min temp	
max temp	
current temp	
current error	
best predictor temp so far	
lowest error so far	

Training state



Trained model

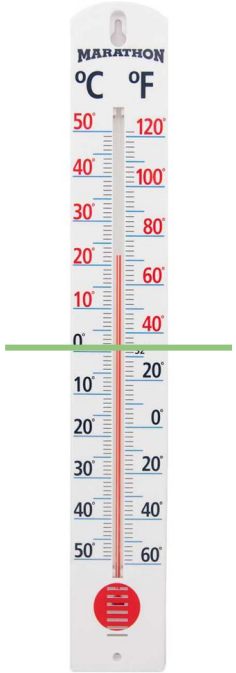
1. Find min and max temps.
2. For each value between min and max temps, calculate how accurately it separates “rain” temps from “snow” temps.
3. Return best predictor temp and its accuracy.

27	Snow
52	Rain
65	Rain
10	Snow
71	Rain
17	Snow
...	...

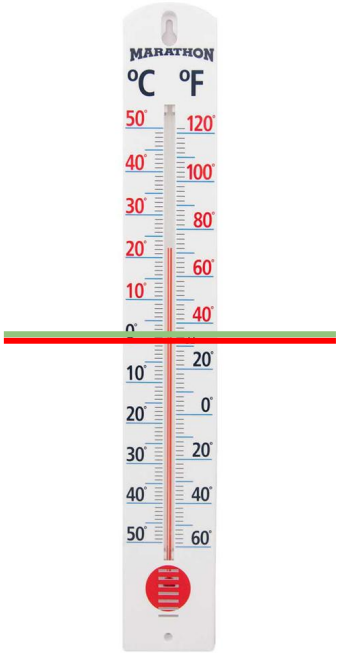
Training data

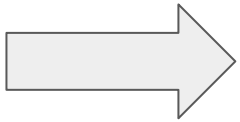
variable	value
min temp	-1
max temp	91
current temp	91
current error	0.66
best predictor temp so far	34
lowest error so far	0.03

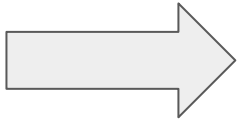
Training state



What did the model learn?

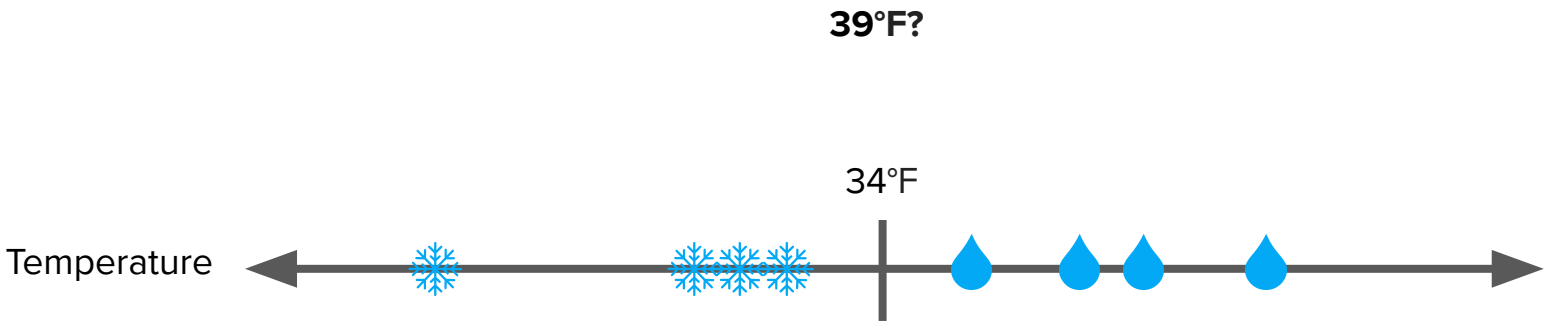






ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

What does the model do with new data?



The model predicts snow or rain based on temperature!

ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

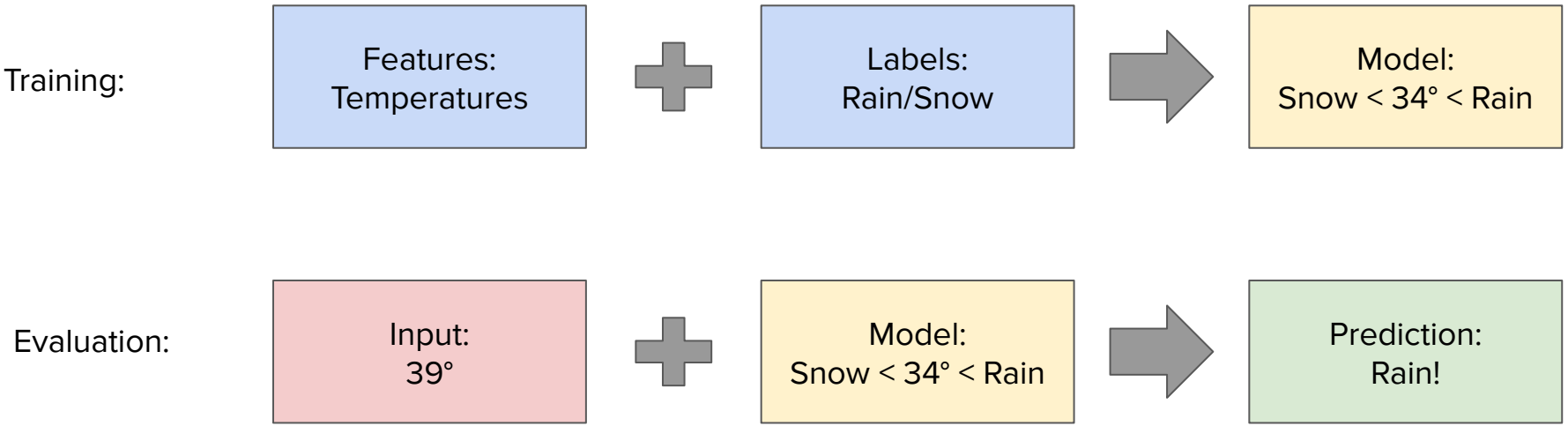
Some terminology

A **machine learning** algorithm uses the **features** and **labels** in **training data** to **train** a **model**.

The model can then be used to **evaluate** new data to **predict** labels.

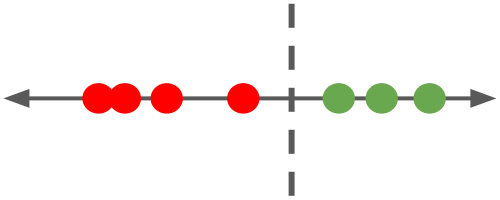
Term	**Example**
Features	Temperature
Labels	Rain/Snow
Model	Snow < 34° < Rain
Training data	Table of temperatures and precipitation type
Training	Try every temperature to see if it produces the most accurate results
Evaluation	Compare new input to learned temperature, predict rain or snow

Features, labels, model, prediction

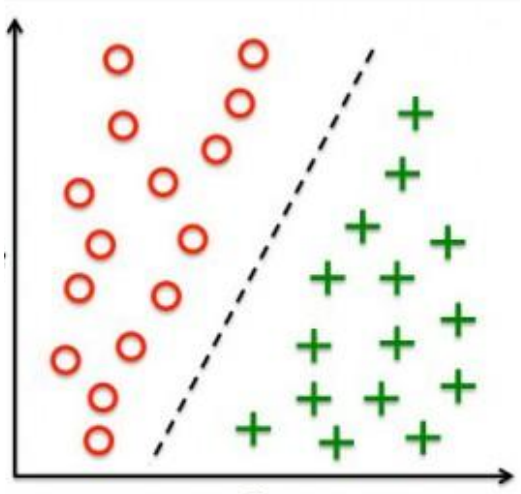


ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

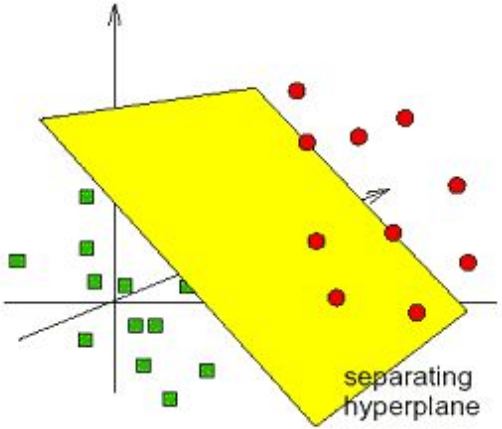
More features? More dimensions.



- Temperature



- Temperature
- Humidity



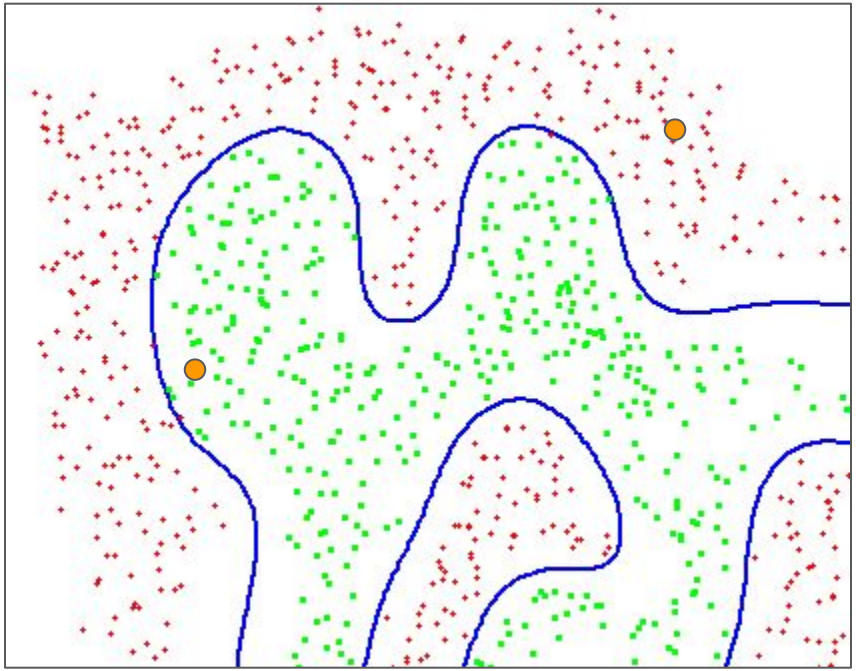
- Temperature
- Humidity
- Altitude

Types of training

- Supervised
- Unsupervised
- Semi-supervised
- Reinforcement

ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

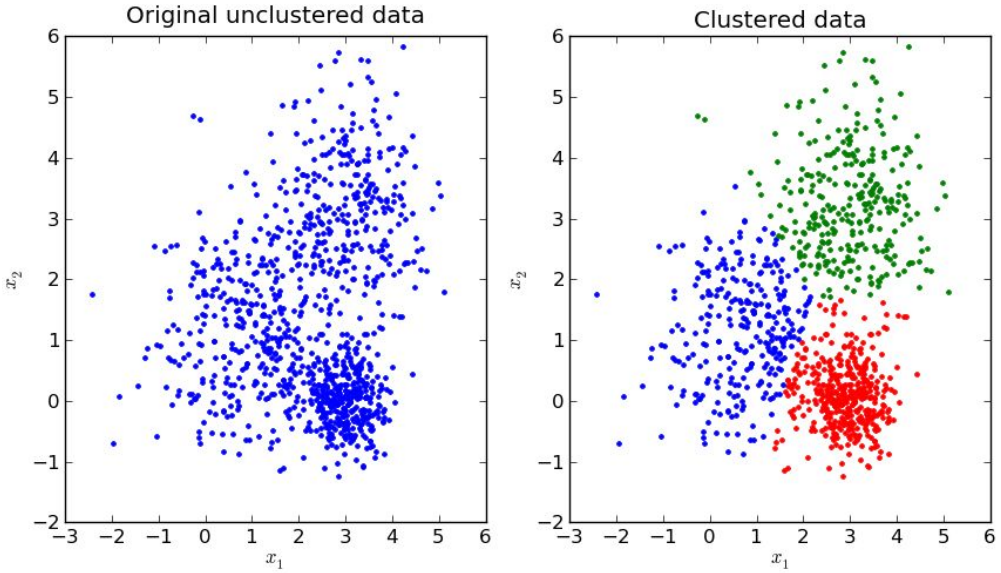
Supervised Learning



Model learns based on knowing the correct output for given input.

- Weather data
- Mail labeled as spam
- Labeled photos of animals

Unsupervised Learning

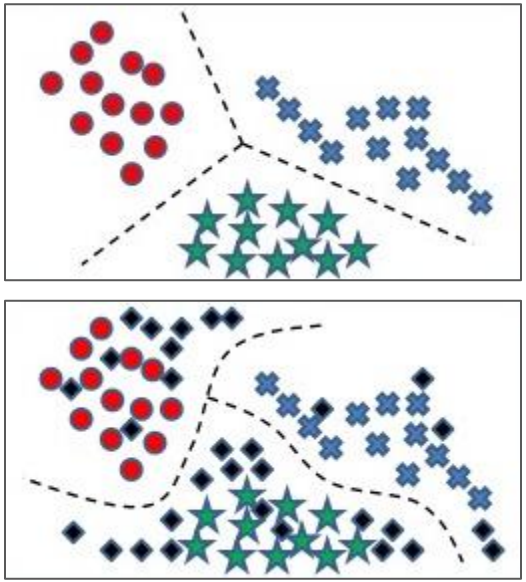


Model learns based only on the input.

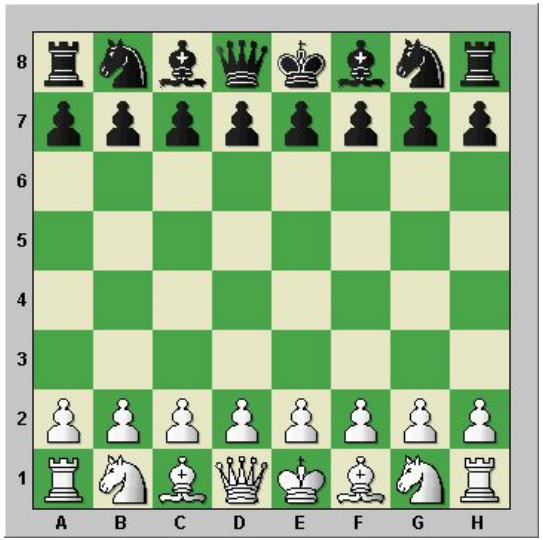
- Purchase data
- Movie preference
- Crime reports

ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Other training techniques



Semi-Supervised Learning



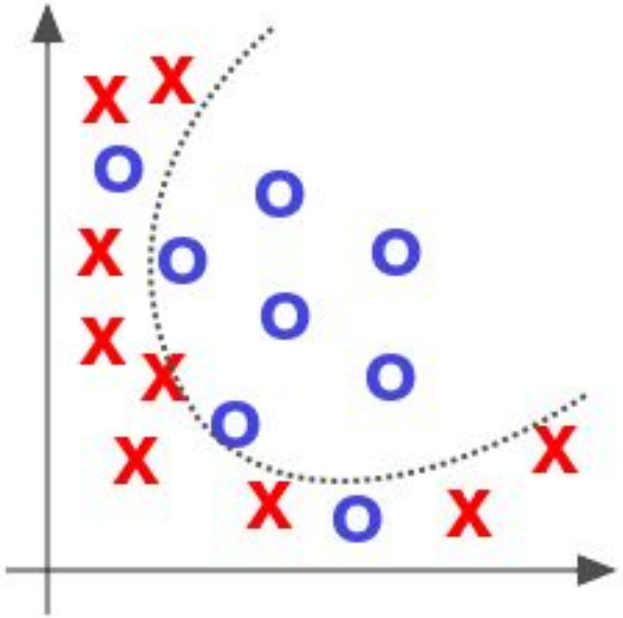
Reinforcement Learning

What can these models do?

- **Classification**
- **Regression**
- **Clustering**

ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

Classification



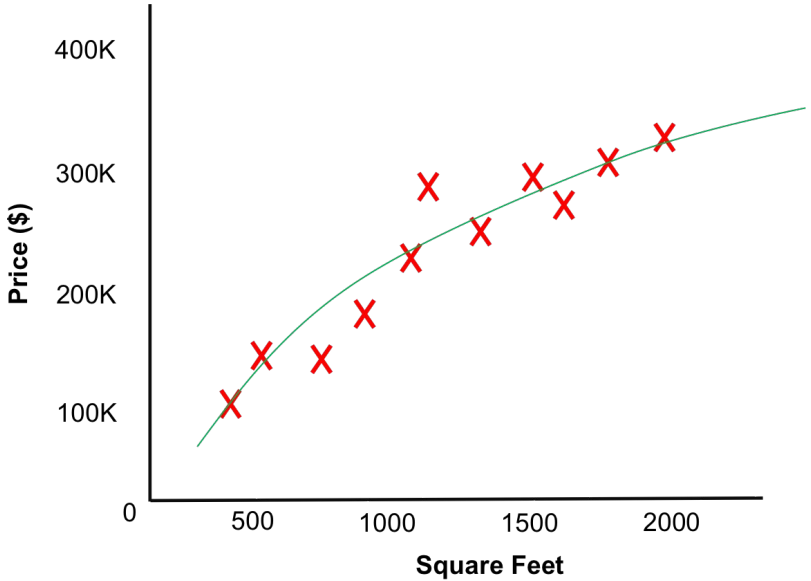
Predict which category input belongs in.

Example uses:

- Rain or snow prediction
- Spam detection
- Object detection in images

ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Regression



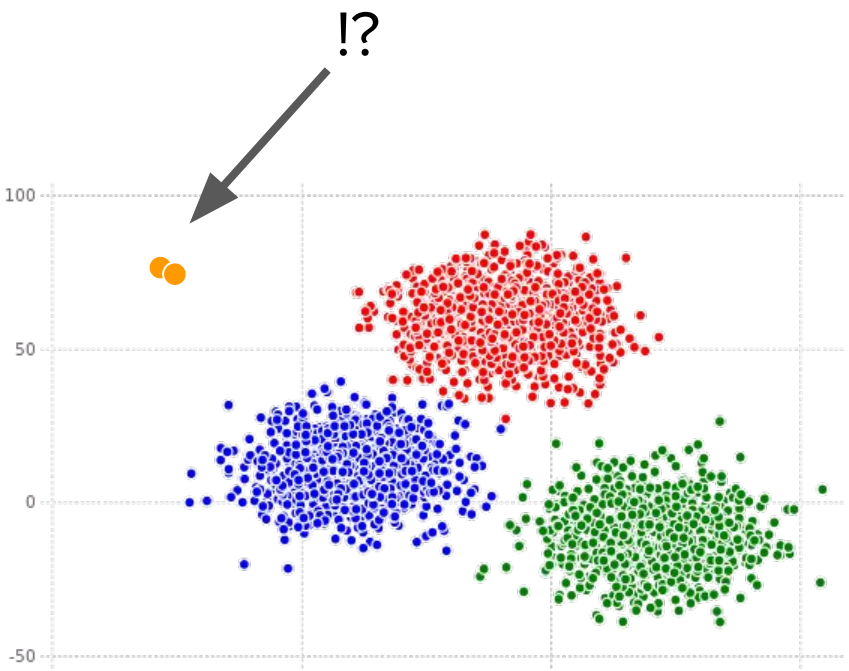
Predict continuous, numeric output based on input.

Example uses:

- Predict property value based on sales of similar homes
- Project future spending based on past spending

ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Clustering

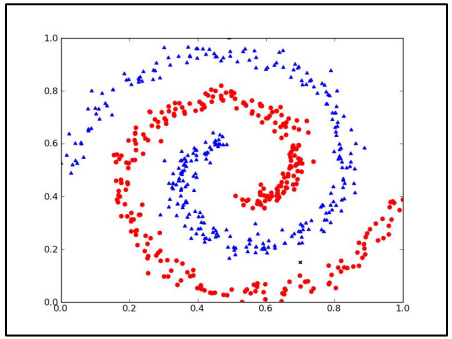


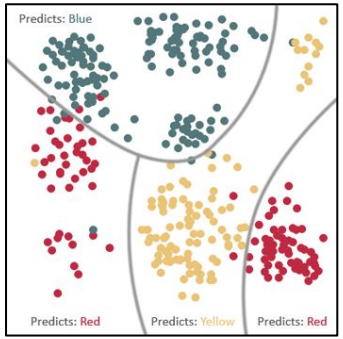
Locate naturally occurring groups within the input data.

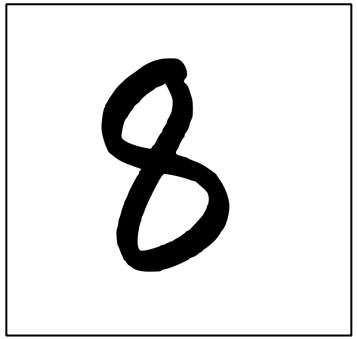
Example uses

- Security anomaly detection
- Movie recommendations
- Clique analysis in social networks

How do we go from dots to the real world?

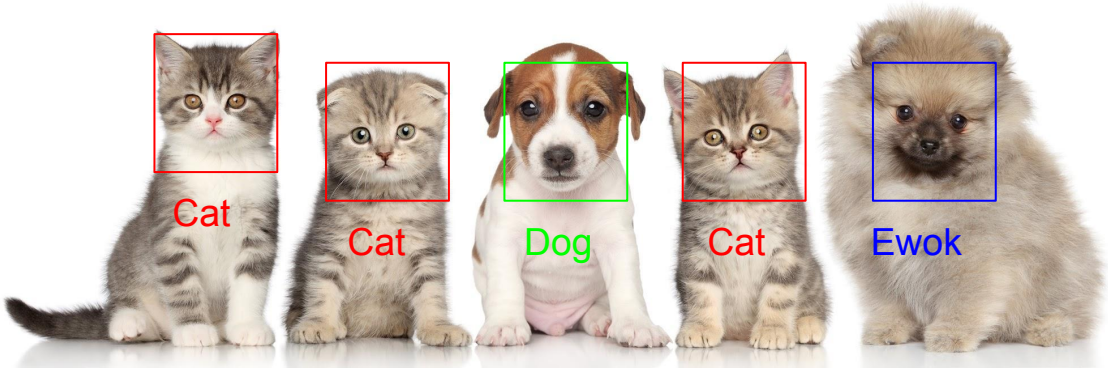






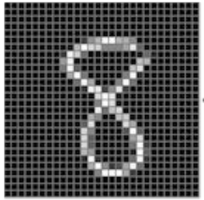
How stellar is the weather today, am I right?!

Is this message Happy or Sad?



ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Images as numeric input

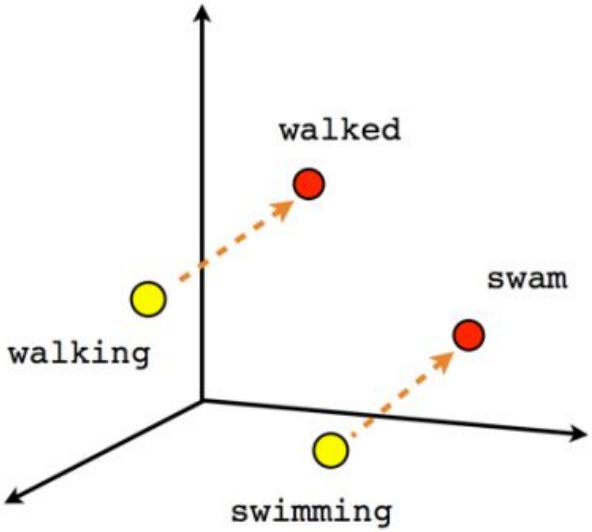


28 x 28
784 pixels

→

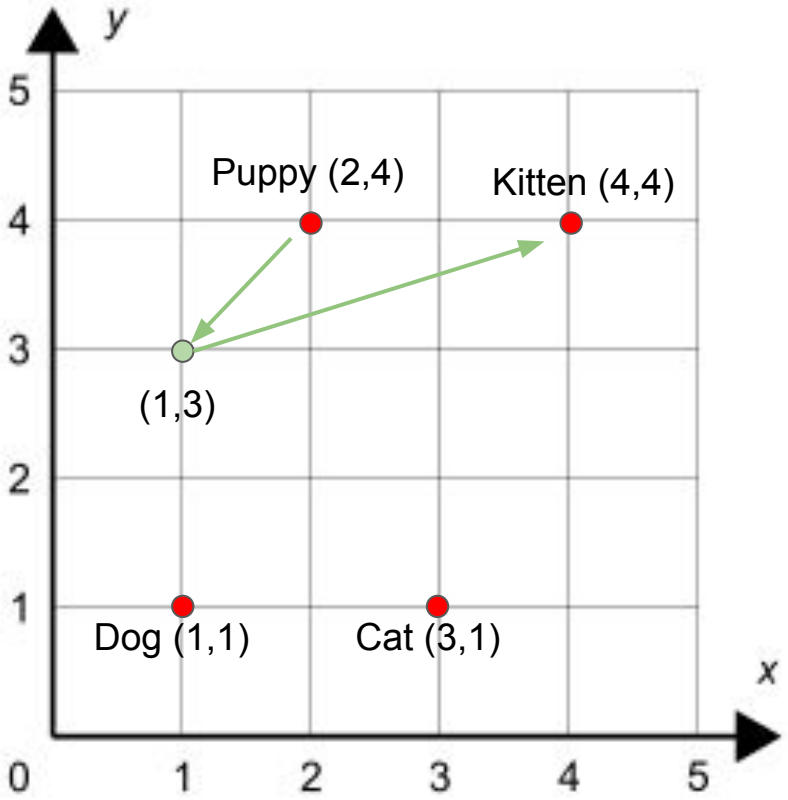
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	4	62	146	182	254	254	181	176	139	15	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	34	186	253	217	208	136	136	136	166	232	99	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	61	242	208	111	3	0	0	0	0	0	18	32	107	43	0	0	0	0	0	0
0	0	0	0	0	0	0	0	156	242	23	0	0	0	0	0	0	0	13	191	181	6	0	0	0	0	0	0
0	0	0	0	0	0	0	0	121	255	98	3	0	0	0	0	0	8	194	225	12	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	169	253	120	3	0	0	0	0	128	247	51	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	3	111	244	169	19	0	14	131	249	117	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	59	241	235	72	142	229	66	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	25	218	254	231	36	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	133	253	221	33	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	19	237	111	196	217	19	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	174	138	0	23	193	204	18	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	96	224	0	0	0	25	218	169	3	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	138	0	0	0	0	86	253	99	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	97	0	0	0	0	3	162	214	11	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	97	0	0	0	0	0	118	253	68	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	185	157	0	0	0	0	0	40	254	98	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	50	244	61	0	0	0	0	112	244	58	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	174	251	142	59	83	167	244	111	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	6	133	253	253	253	169	61	3	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Words as numeric input



ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

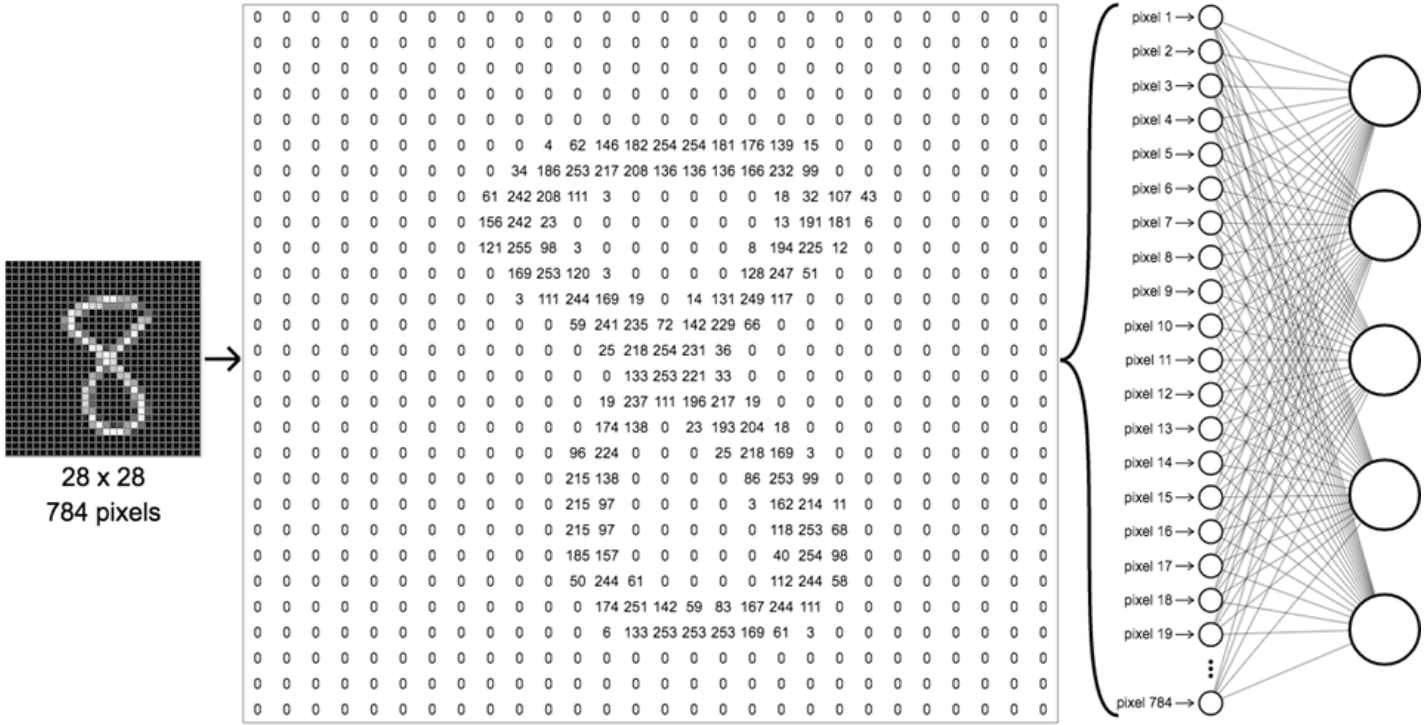
Words as numeric input



Puppy - Dog + Cat = Kitten

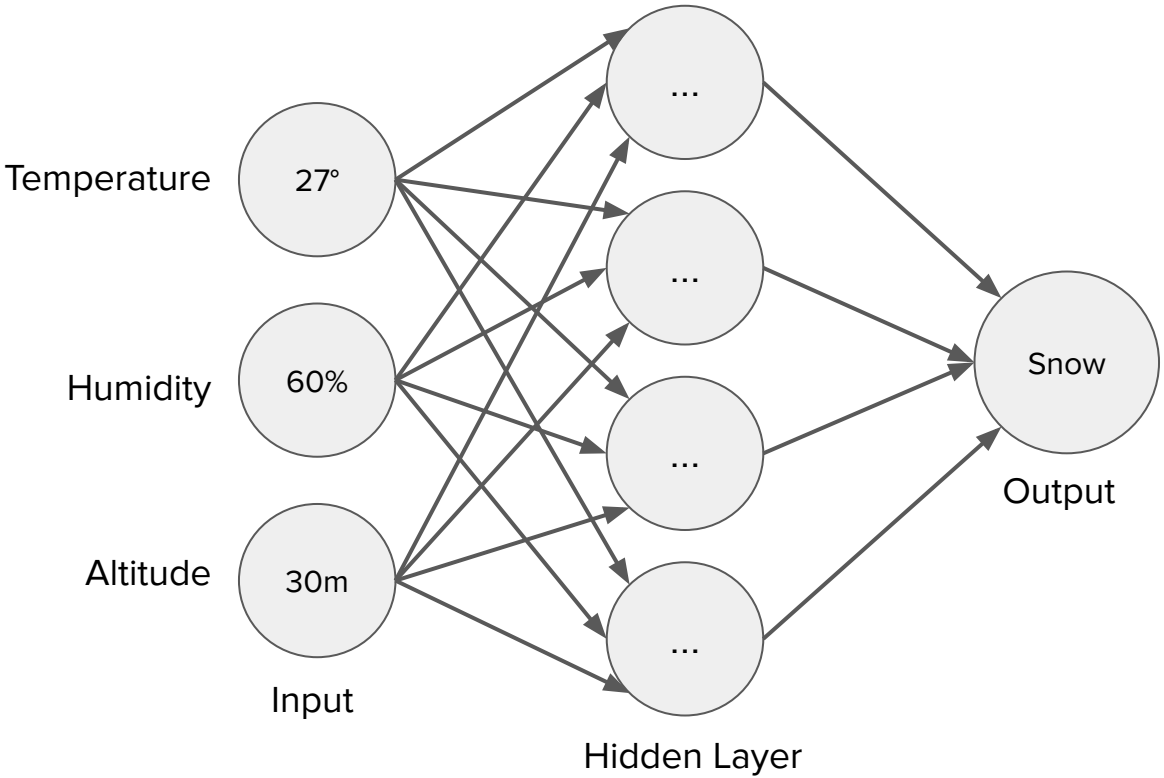
ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Neural networks



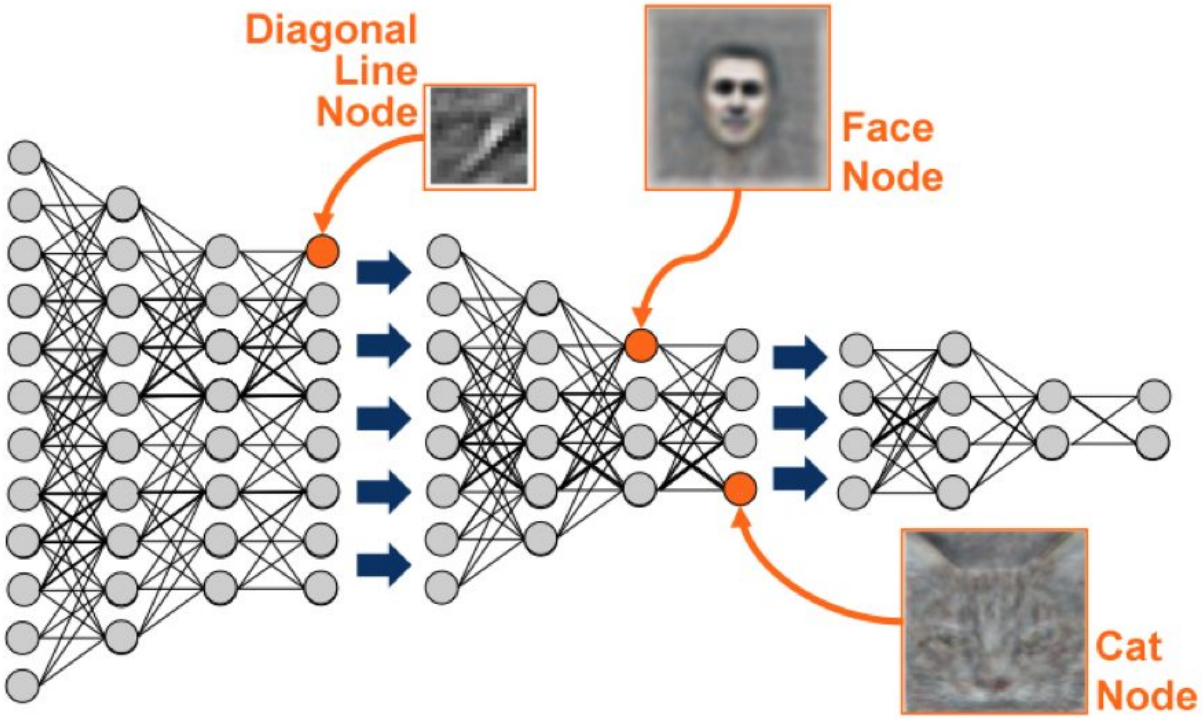
ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Hidden layers combine inputs

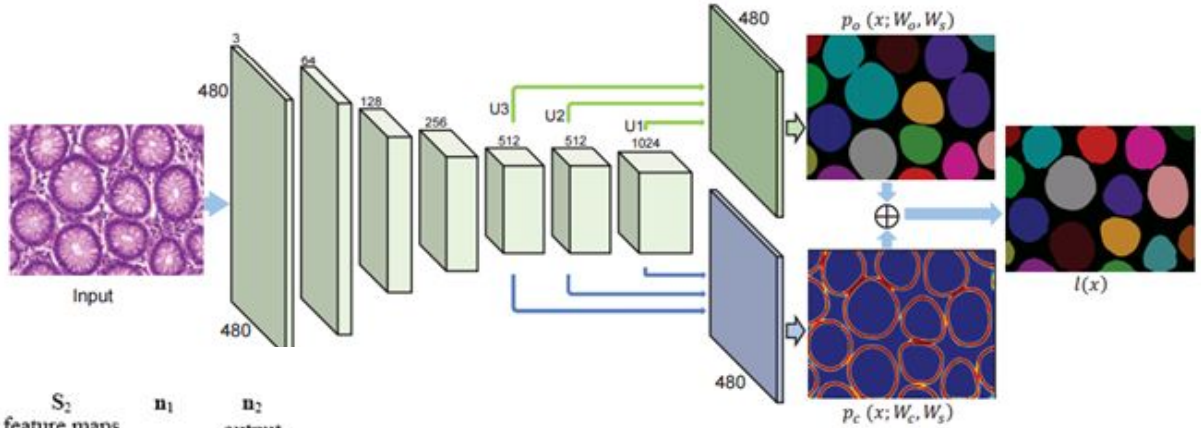


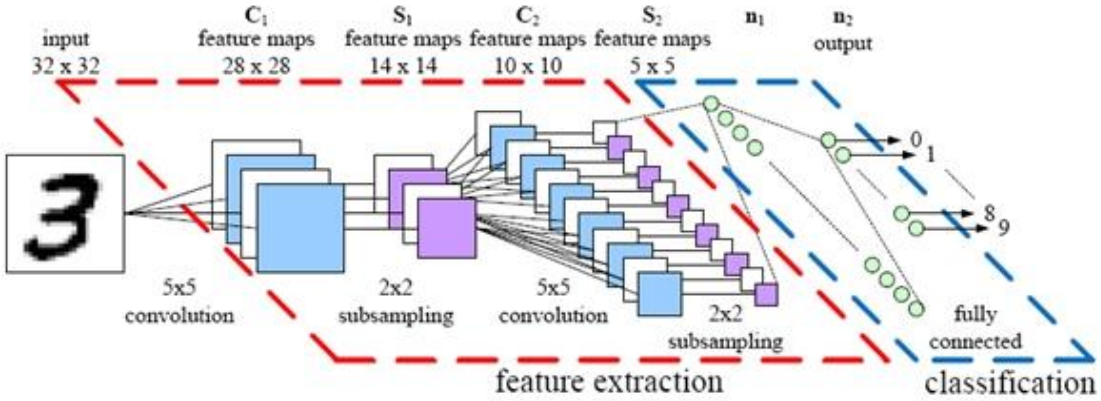
ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Each layer solves part of the problem



Deep Learning networks have lots of layers





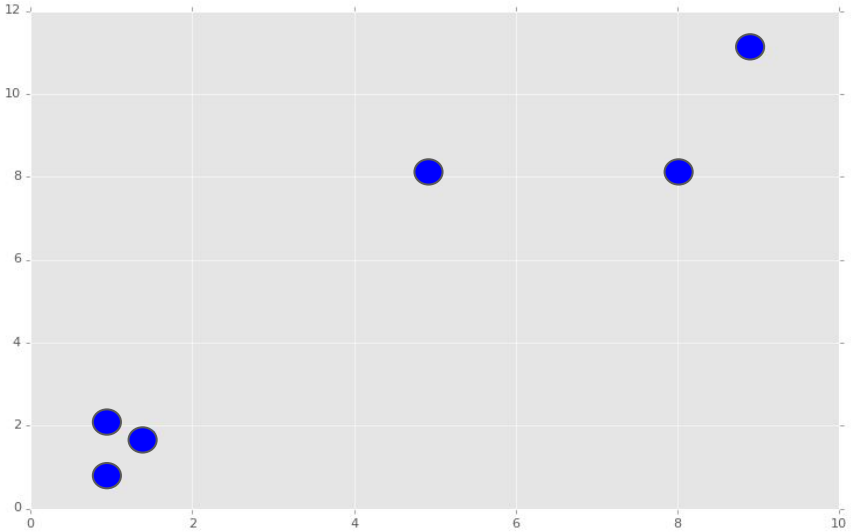
ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

Challenges

- Sparse data
- Overfitting
- Biased data
- Unintended proxy variables

What can go wrong: Sparse data

Insufficient training data for model to learn.

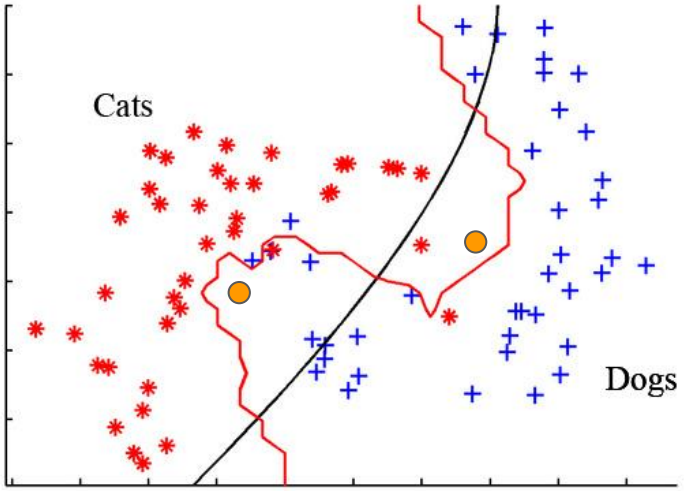


°F	Precipitation
47	Rain
52	Rain
65	Rain
38	Rain
…	…

ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

What can go wrong: Overfitting

Model matches the noise instead of the signal in the training data.

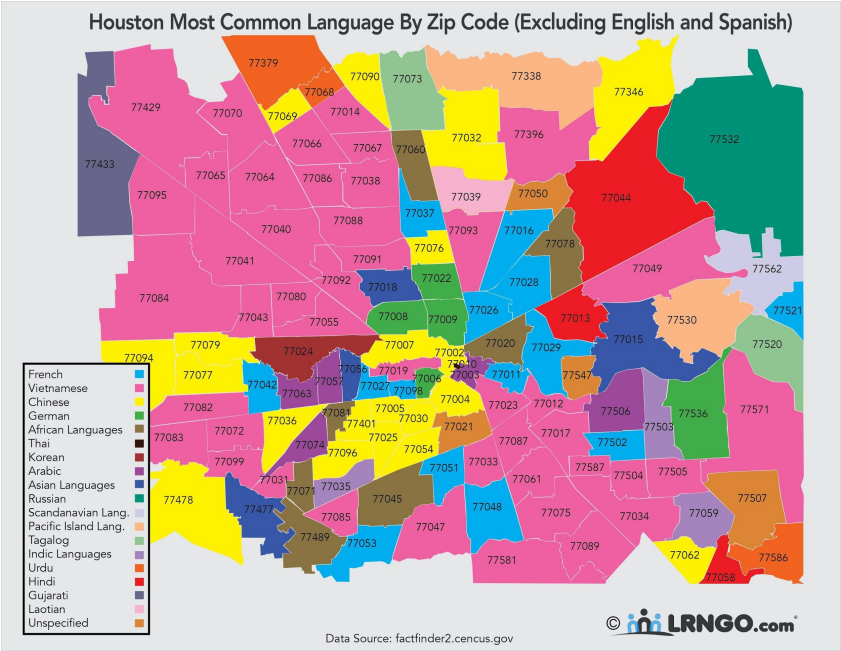


What can go wrong: Biased data

Incomplete or non-representative training data results in skewed predictions.

What can go wrong: Unintended proxy variables

Keying off an unintended signal via correlated data.



ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Current challenges

Labeled data is limited.

ML models are highly specialized to domain of training data.

Opaque ML models make it hard to determine if predictions are biased.

Experimental results are often not reproducible.

Wrap-up

- Industry roles
- Machine learning vs Artificial intelligence?

Industry roles

Data scientist

Data engineer

Software engineer

QA engineer for models

Data annotator

Domain expert

Computational linguist

AI/ML researcher

Ethicist

Lawyer

Marketer

Social media strategist

Public relations specialist

Product manager

Designer

HCI/UX expert

...

More to read and play with

http://playground.tensorflow.org

http://cs.stanford.edu/people/karpathy/convnetjs/demo/classify2d.html

http://colah.github.io/

http://allenai.org

milesc@allenai.org