Machine learning to forecast and improve clinical outcomes and healthy aging using data collected during daily living

Alvin Rajkomar, MD

Google Health

Disclosures

Outline

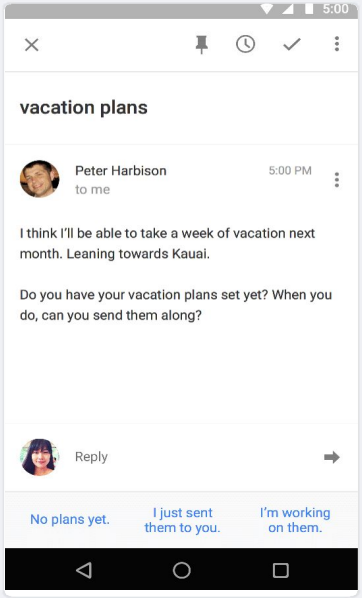
What is machine learning?	
Elements of ML	Input Data with sensors
	Output Data
Putting it together	Cohorts
	Getting the right labels
	Creating interventions
	Ensuring fairness

What is machine learning?

Machine Learning is common outside healthcare

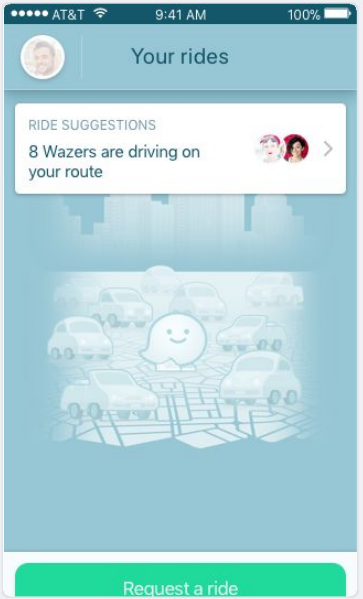
Translate

Inbox



Photos

Waze



The NEW ENGLAND JOURNAL of MEDICINE

REVIEW ARTICLE

FRONTIERS IN MEDICINE

Machine Learning in Medicine

Alvin Rajkomar, M.D., Jeffrey Dean, Ph.D., and Isaac Kohane, M.D., Ph.D.

Clinical Review & Education

JAMA | Users' Guides to the Medical Literature

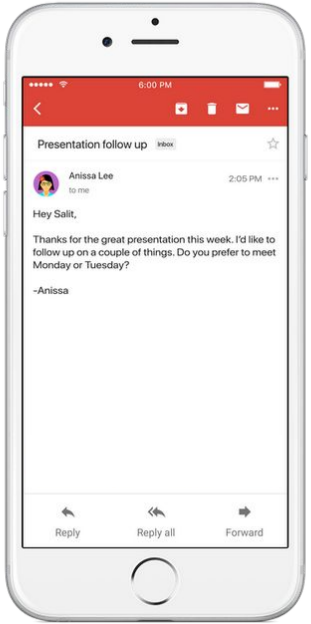
How to Read Articles That Use Machine Learning
Users' Guides to the Medical Literature

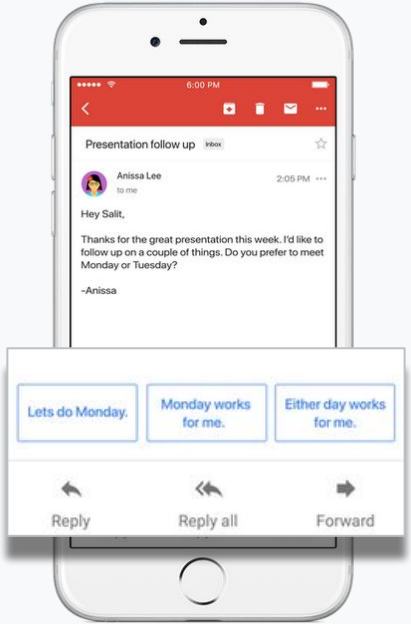
Yun Liu, PhD; Po-Hsuan Cameron Chen, PhD; Jonathan Krause, PhD; Lily Peng, MD, PhD

Input Output

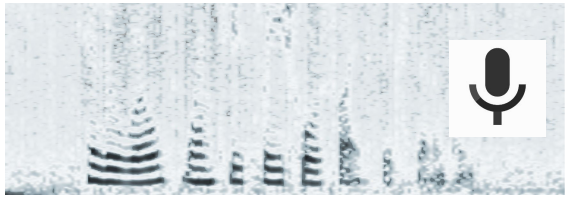
Input

Output

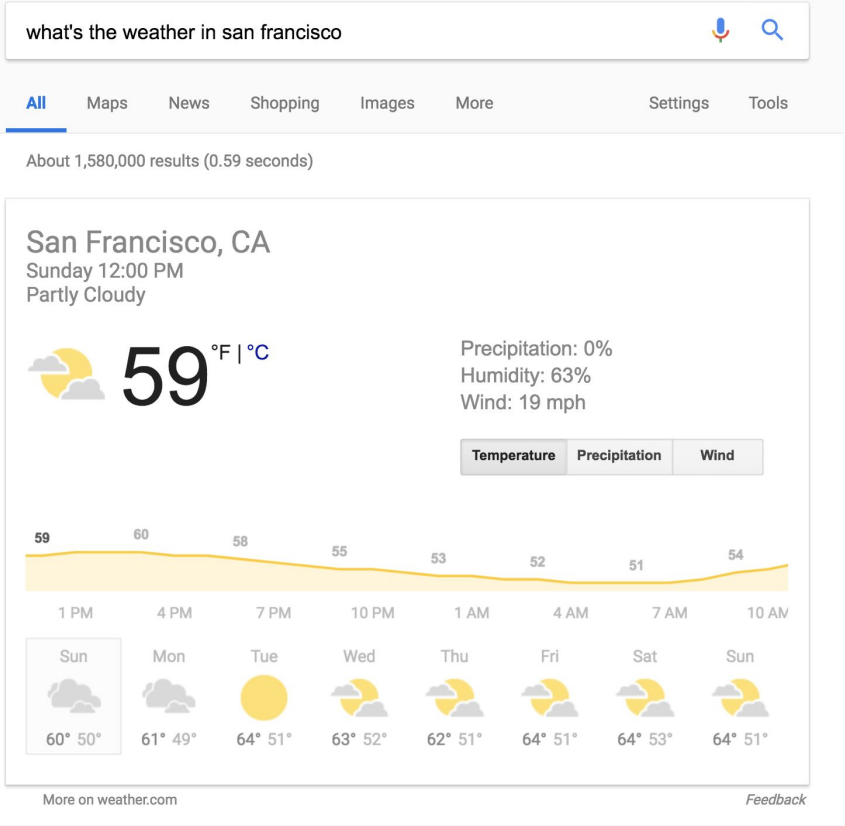




Input



Output



Input

Output

JPG × describe image here

All Images Maps Shopping More Settings Tools

About 6,090,000 results (1.13 seconds)

Image size:
729 × 425

Find other sizes of this image:
All sizes - Small - Medium

Best guess for this image: ***tiger in the rainforest***

Rules and contingencies

Spam the old way:
Write a computer program with ***explicit*** rules to follow

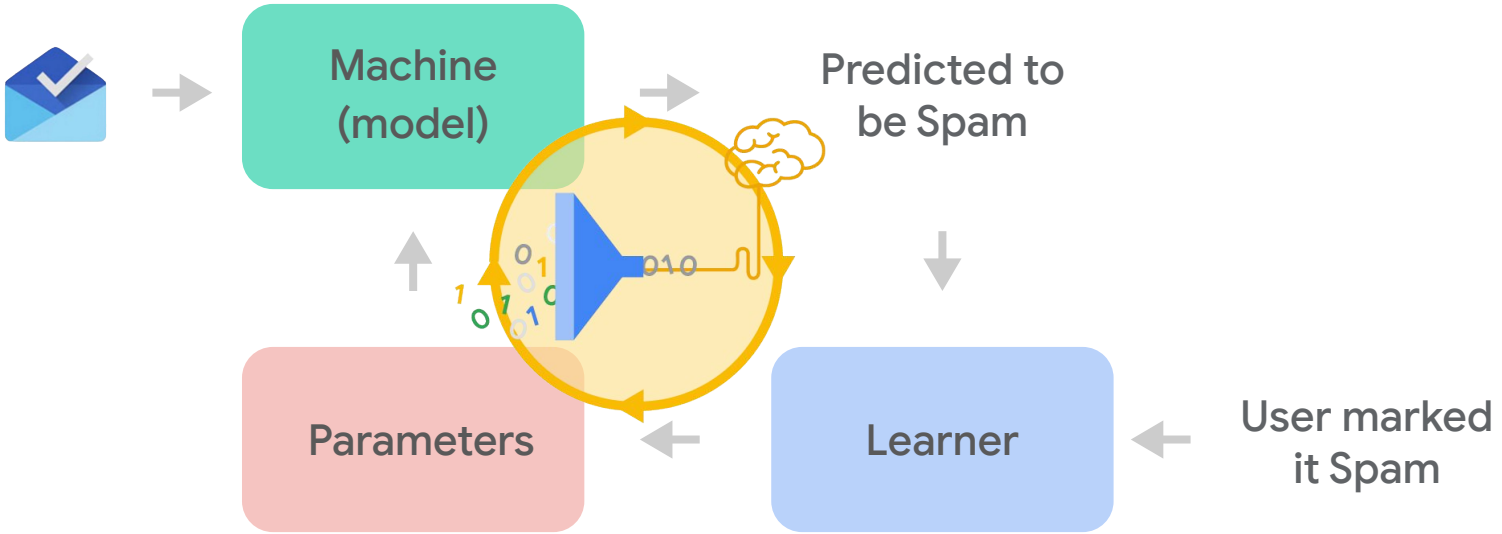
```
if email contains V!agrå
   then mark 'is spam';
if email contains ...
if email contains ...
```

Learning from data

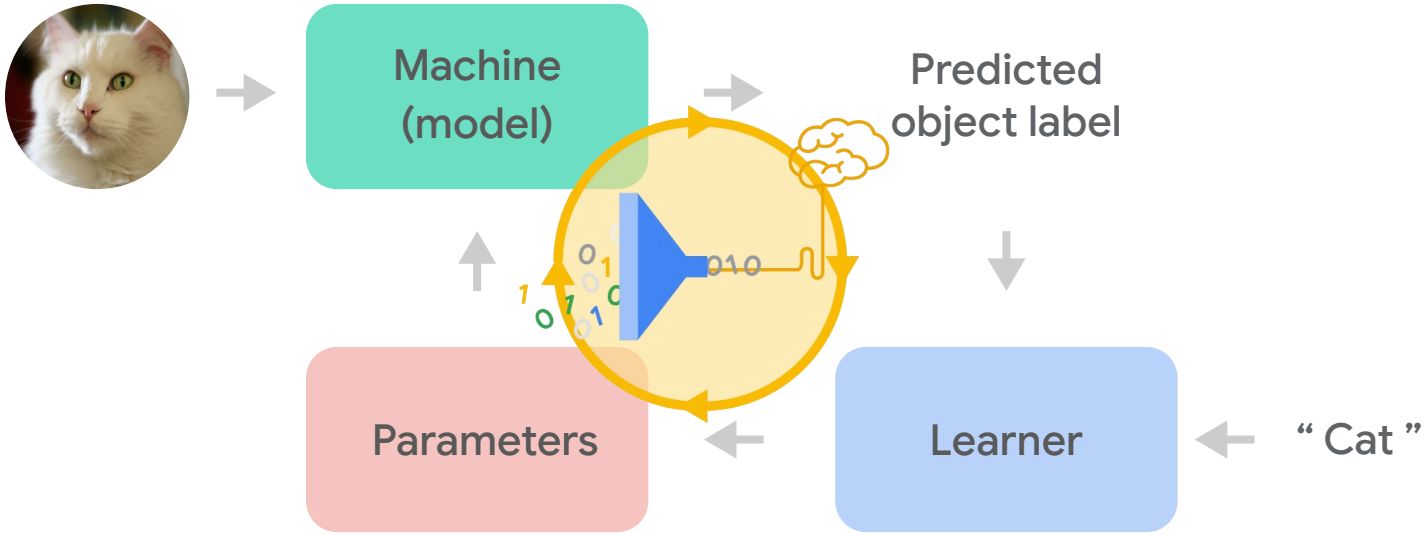
Spam the new way:
Write a computer program to ***learn*** from examples

```
try to classify some emails ;
change self to reduce errors ;
repeat ;
```

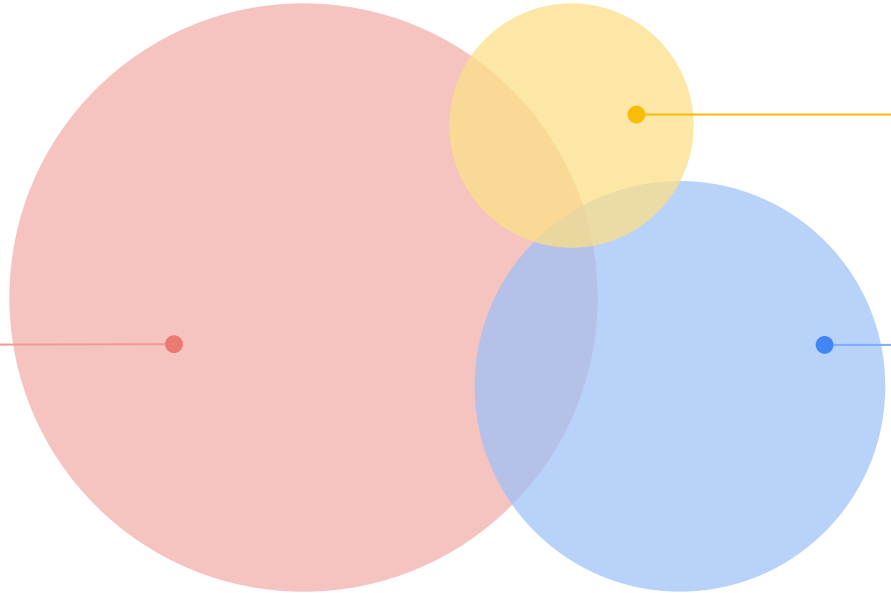
Machines Learning by Example



Machines Learning by Example



A tale of two disciplines



Artificial Intelligence

Making machines intelligent

Robotics

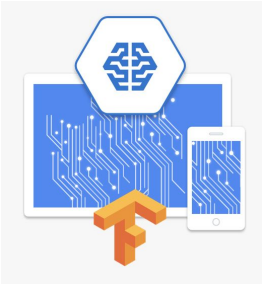
Machine Learning

Making machines that learn

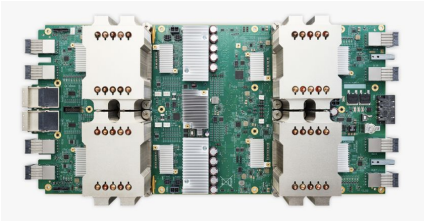
- Traditional AI Systems are ***programmed*** to be clever.
- Modern ML-based AI systems to ***learn*** to be clever.

Why widespread use of ML is occurring now

Training examples



Algorithms + tools



Computational resources

Creativity + ingenuity

What do you need to build a model?

Input Data - Terminology

Wearables
- Movement (activity tracking)
- Electrical activity (EKG)

mHealth (phone)

Ambient sensors
- Cameras
- Microphones
- Keyboards on phones

Examples of Sensors

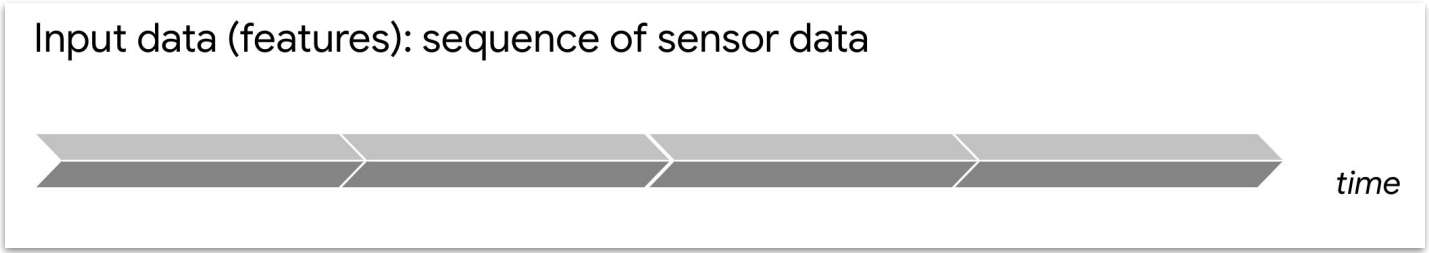
	Category of Measurement	Examples of specific sensors	Examples of Measurements	Derived Measurements
Data measured from wearable sensors				
	Inertia	Accelerometer, Gyroscope, Magnetometer	Linear and angular motion	Types of activity (e.g. walking), step length, falls
	Light transmittance through skin	Photoplethysmographic (PPG)	Oxygen saturation, heart rate, heart rate variability	Measurements of cardiovascular health
	Electrical activity	Electrodes	Electrocardiograms (EKG),	Heart rhythms, sleep states,
			electroencephalograms (EEG), Galvanic skin responses	emotional state
	Mechanical movements	Piezoelectric sensors	Pulsations on skin from heart beats	
	Chemical analytes on skin	Potentiometric and amperometric devices	Glucose, lactate, sodium measurements in sweat	
	Temperature	Thermistor	Body temperature	Elevated risk of infection [41]
	Location	Global position satellite measurements	Movement	Location entropy to indicate depression

Measurements from ambient sources				
	Video	Cameras	Pixels	Activity classification in the home, vital signs[42], gait
	Audio	Microphones	Waveforms	Respiratory status from breath sounds, emotion from voice
	Interactions with computing devices	Smartphones, Tablets, Keyboards	Patterns of typing and scrolling	Fine motor control that tracks development of Alzheimer disease.[38] Digital phenotypes for psychiatric diseases[31]
	Smart devices	Smart pill caps	How often medication bottles are opened	Medication adherence

Input Data - Terminology

Clinical Studies
- Predictors
- Covariates

Machine Learning
- Features



Input Data - Terminology

Passive

- No additional interaction necessary
- ***Easier to collect, less standard***

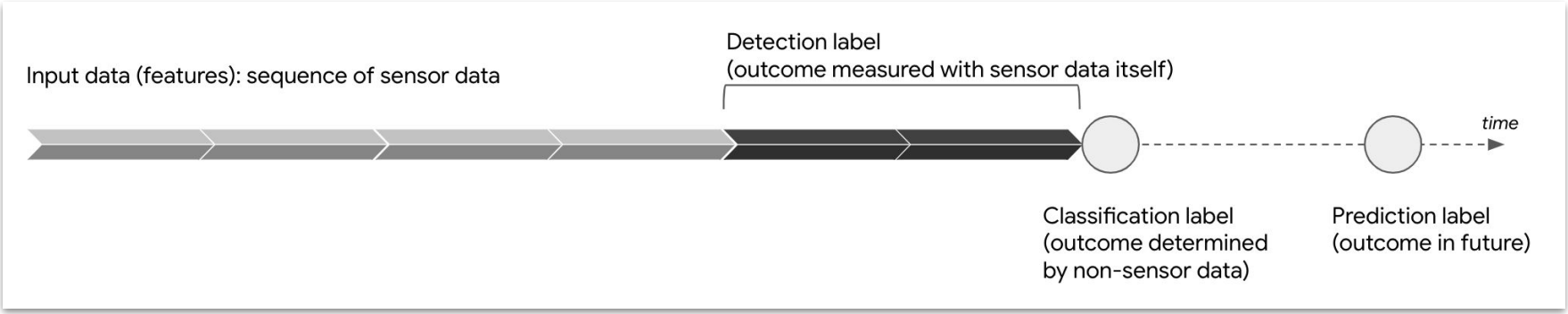
Active

- Interaction with device
  - E.g. Start a six minute walk test
  - E.gl Answer patient reported outcome
- ***Harder to collect, more standard***

Output Data - Terminology

Clinical Trials - Outcomes
- Primary Outcomes (usually major clinical event)
- Secondary Outcomes

Machine Learning - Labels, Output
- Detection
- Classification
- Prediction



Isn't this all very straightforward?

In-person discussion

No further slides