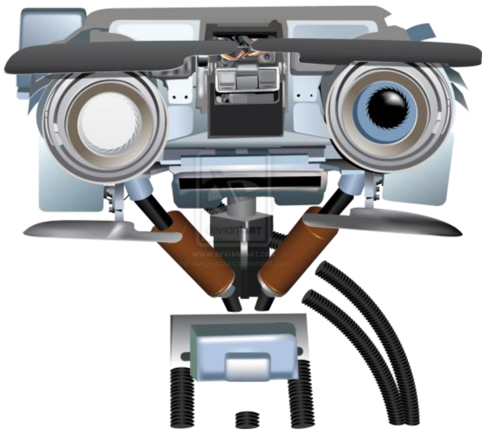
Today

- Discussing Pre-lecture Material
- Learning from Examples

Reminders: If your timezone is *not* EST, please let me know if you have not already done so

If you are *not* added to the course Piazza, please email me ASAP

Announcement: I will post classroom in-person meeting status by Monday



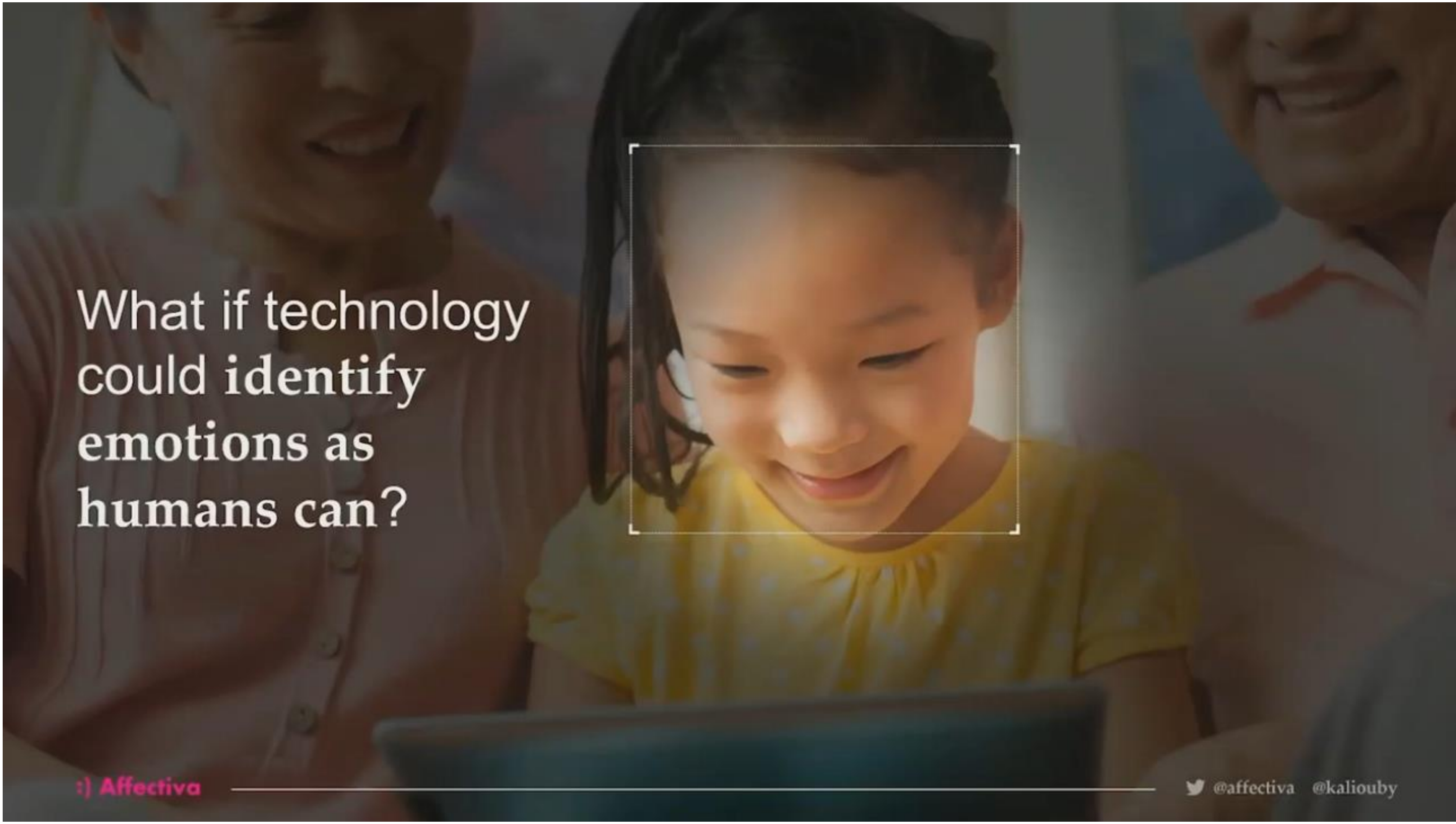
Pre-lecture Material

Humanizing Technology

Affectiva

Rana el Kaliouby

Co-founder and CEO


What if technology could **identify emotions as humans can?**
:) Affectiva
@affectiva @kaliouby

Emotional Intelligence

- What is the main problem definition?

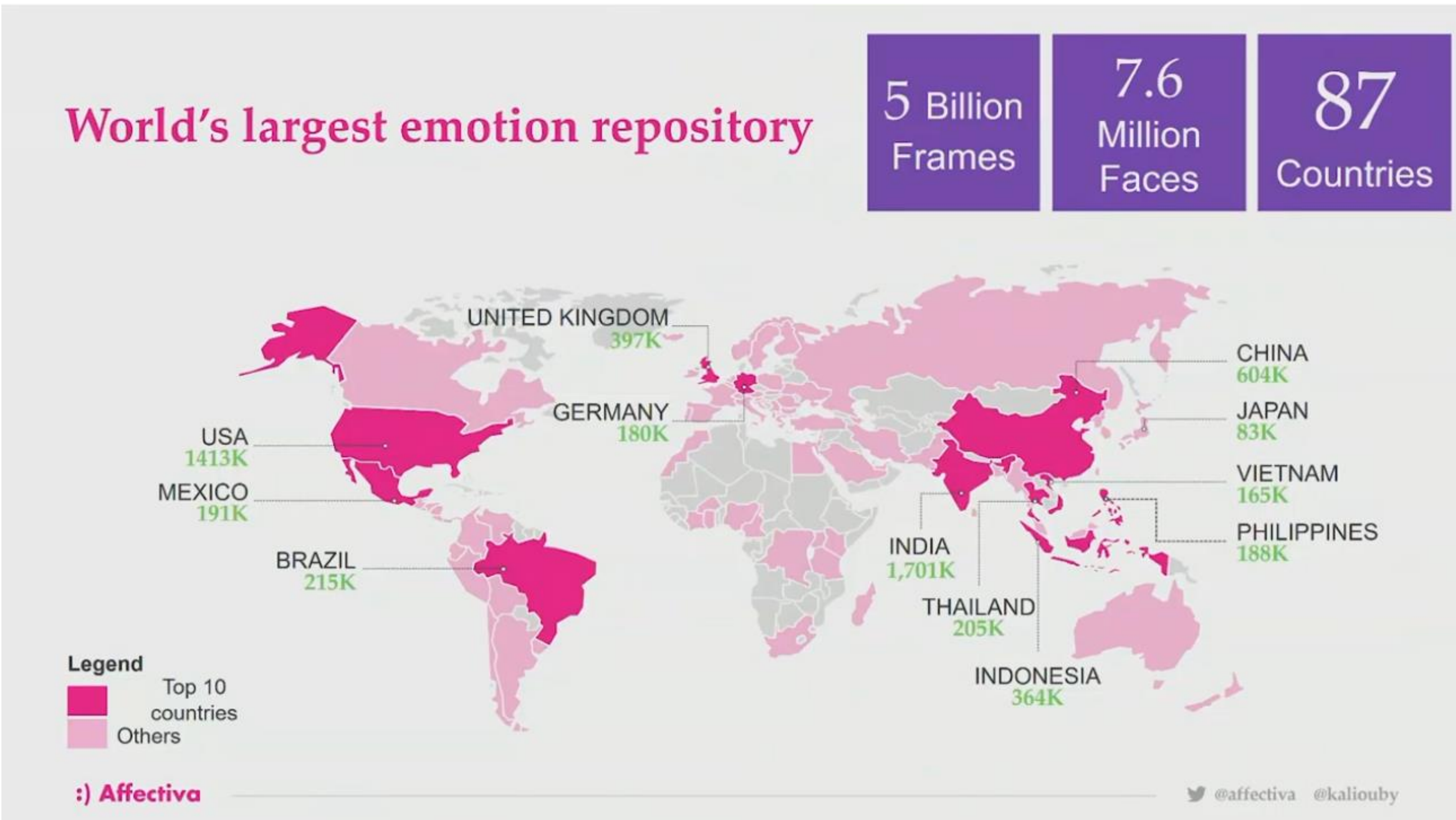
Emotional Intelligence

- What is the main problem definition?
  - Machine Learning Problem

Why?

Data Collection

- Consent, diversity, cultures



Emotional Intelligence

- What is the main problem definition?
  - Machine Learning Problem
  - Supervised vs. unsupervised

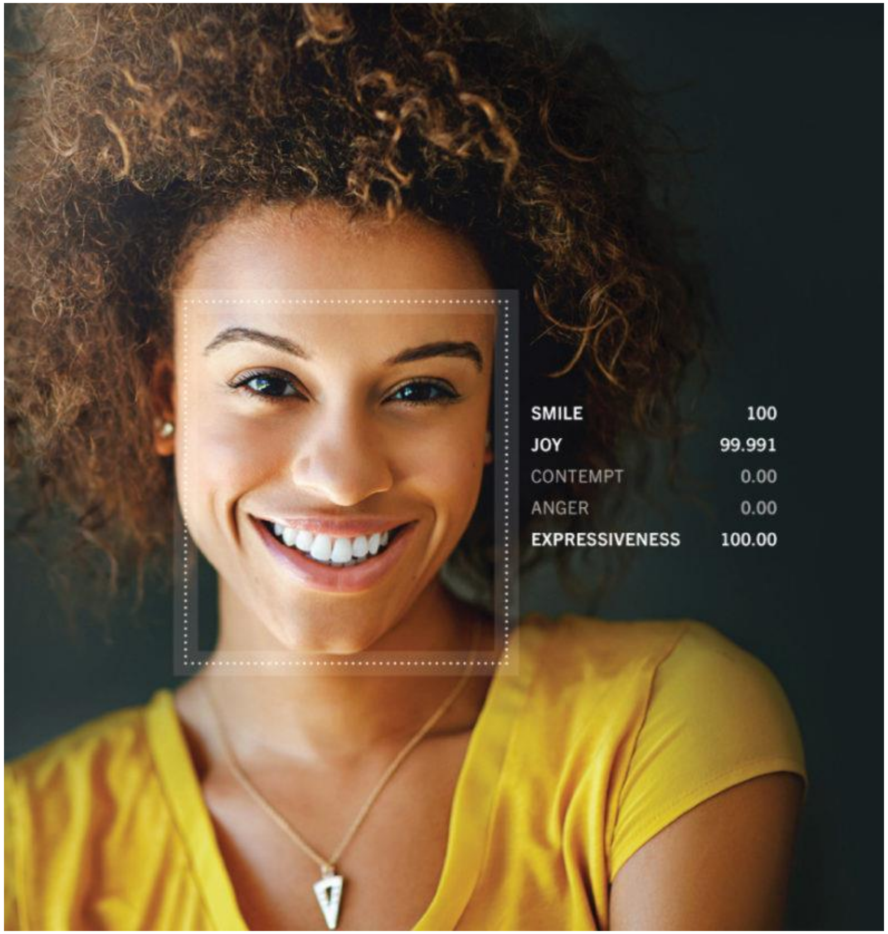
Labels *vs.* No labels

Emotional Intelligence

- What is the main problem definition?
  - Machine Learning Problem
  - Supervised vs. unsupervised
  - Classification: Happy, Sad, Angry, Surprised, Fear, …

Classification: Scale for Each Class

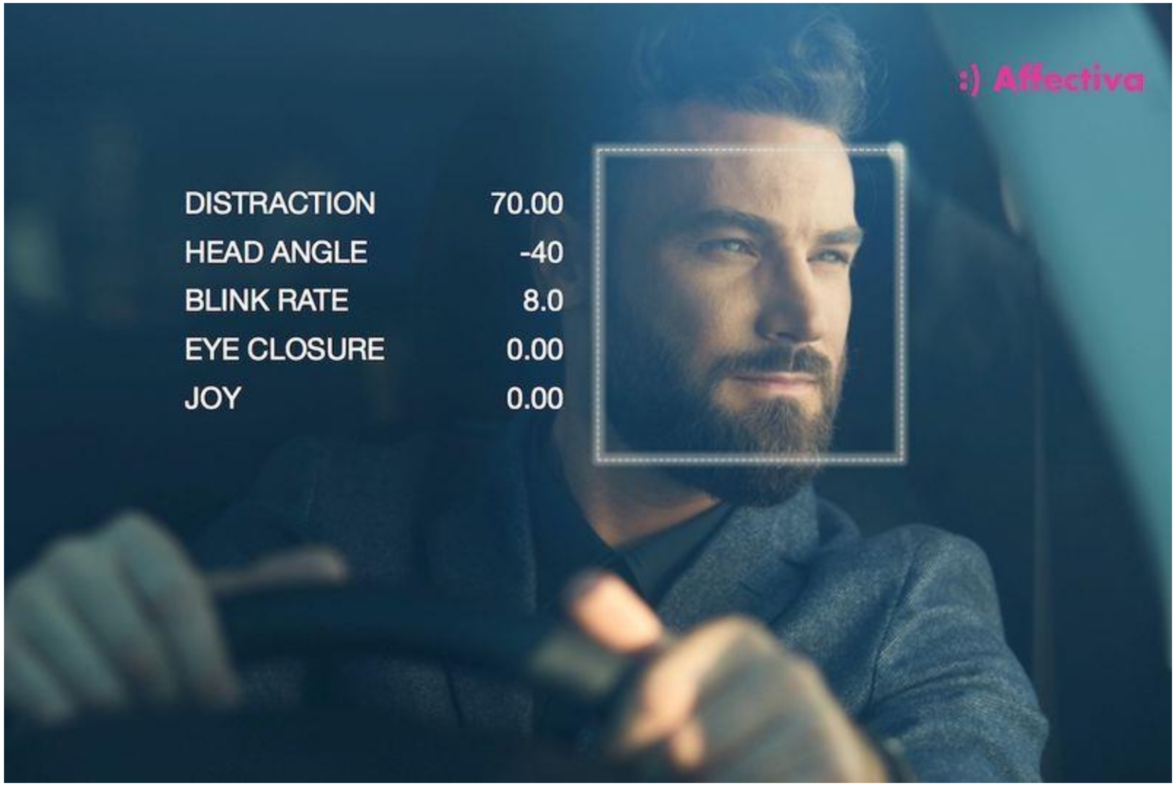
- Multi-class *vs.* Smile Classifier



:) Affectiva

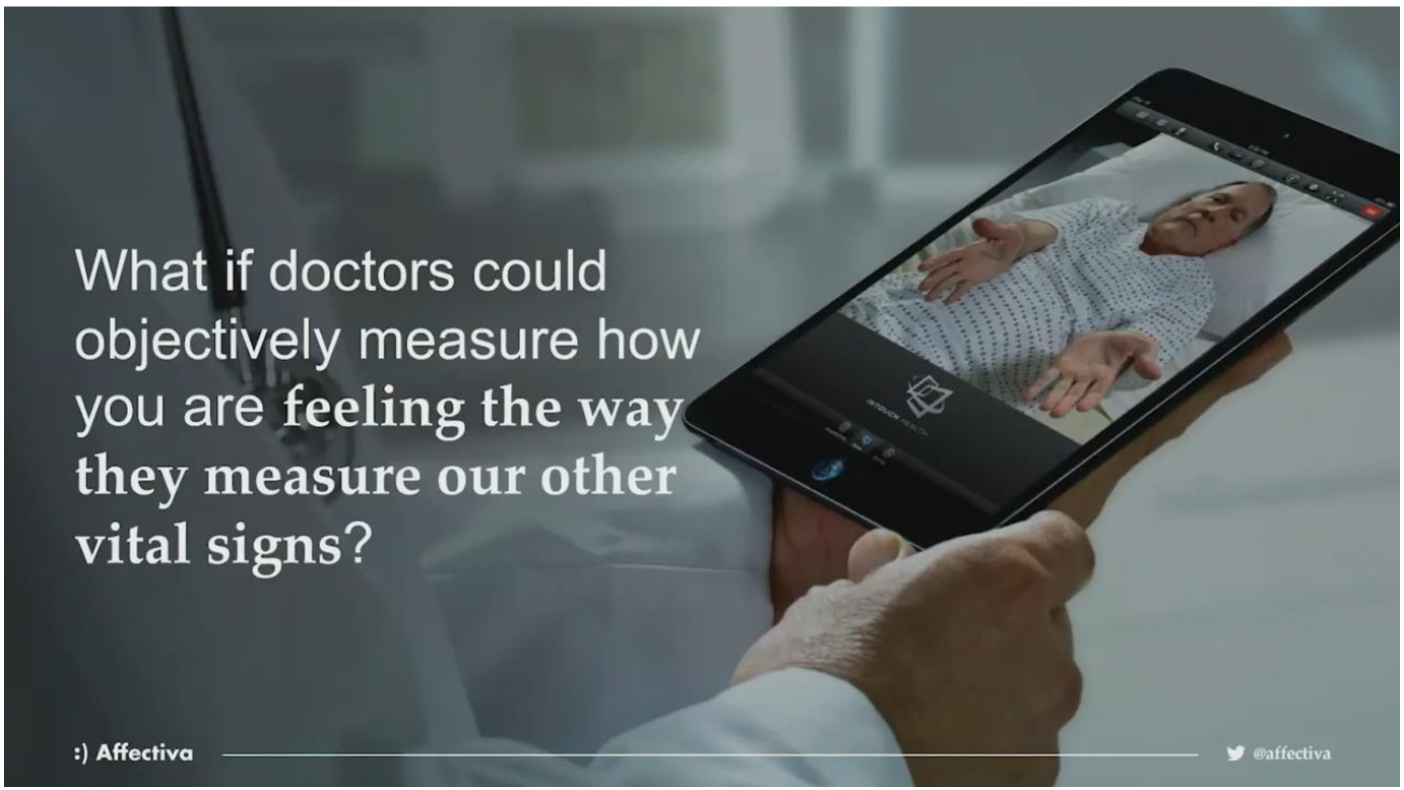
Applications

- Applications that benefit society:
  – Automotive Safety



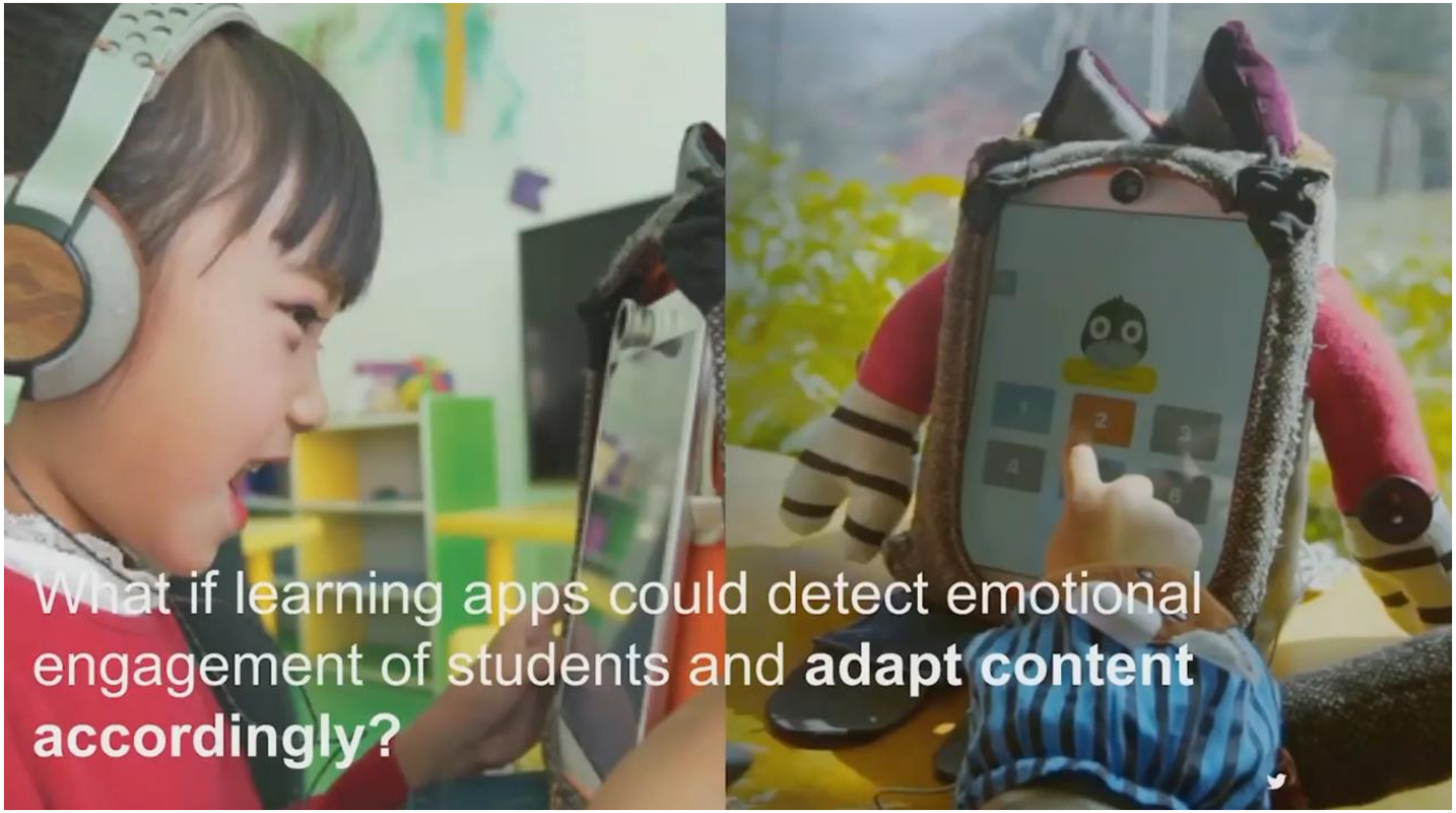
Applications

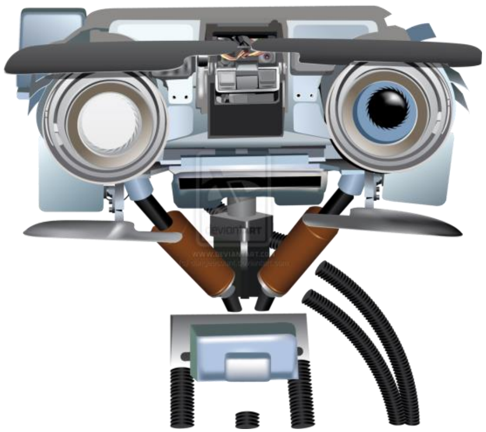
- Applications that benefit society:
  - Mental Health



Applications

- Applications that benefit society:
  - Education



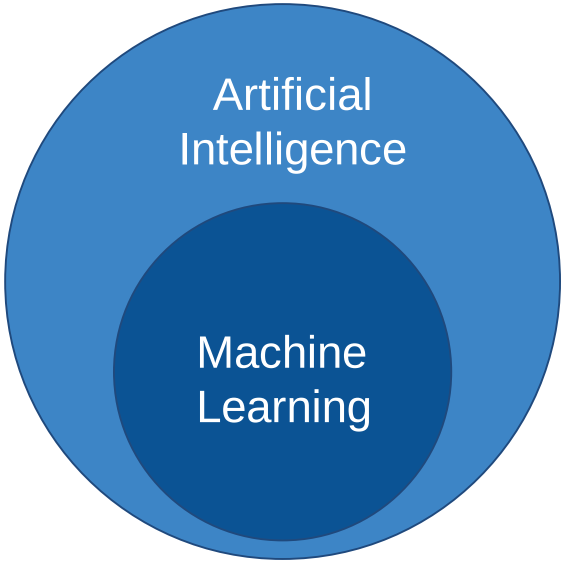


AI *and* ML

Recap

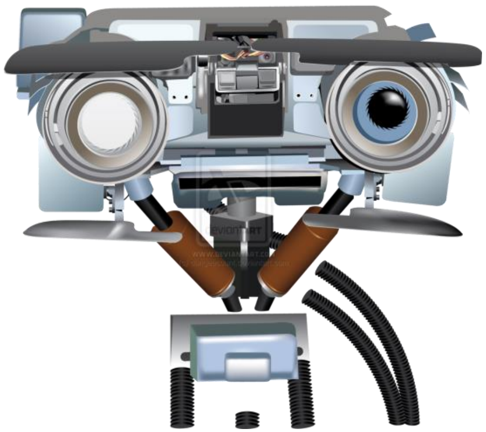
How is Machine Learning (ML) different from AI?

- Machine learning is a type of artificial intelligence
- Machine learning makes decisions based on data it has seen
- Not all AI algorithms need to do this
- Many of the latest AI systems all make use of ML
- For this reason, many people use the terms AI and ML interchangeably



Types of Machine Learning

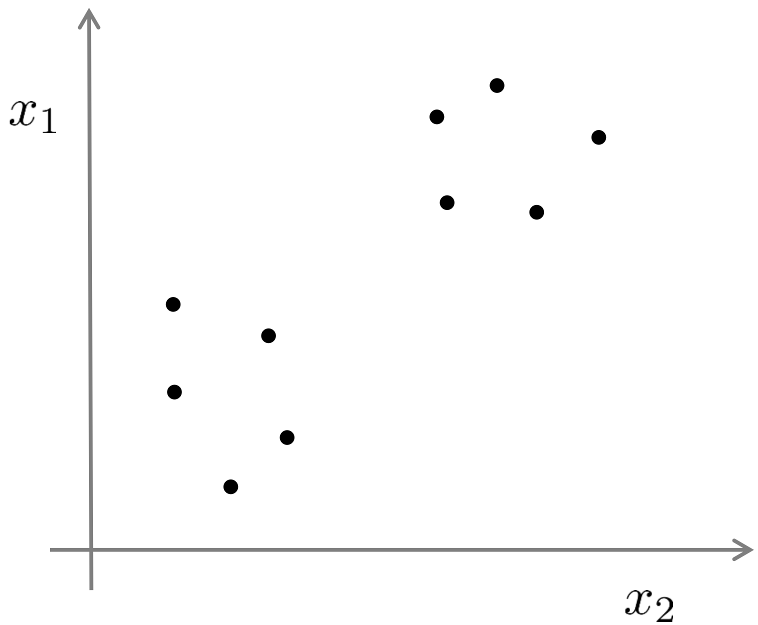
- Unsupervised Learning
- Supervised Learning
- Semi-supervised Learning
- Reinforcement Learning



Unsupervised Learning

Example:
Clustering using K-means Algorithm for

Unsupervised learning

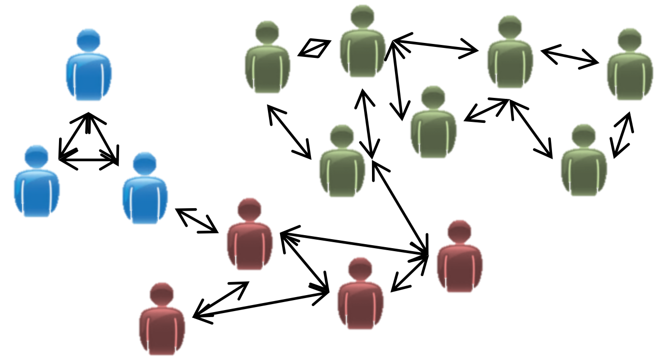


Training set: $\{x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(m)}\}$

slide credit: Andrew Ng

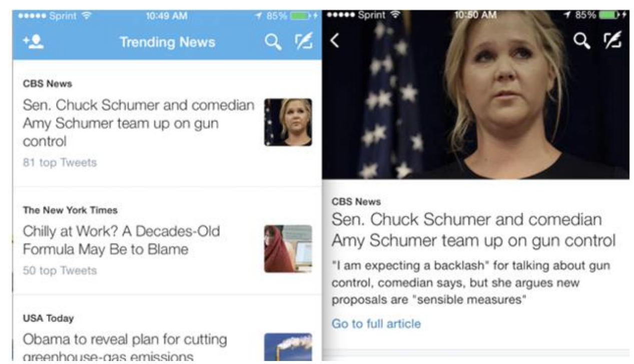
Clustering

Gene analysis

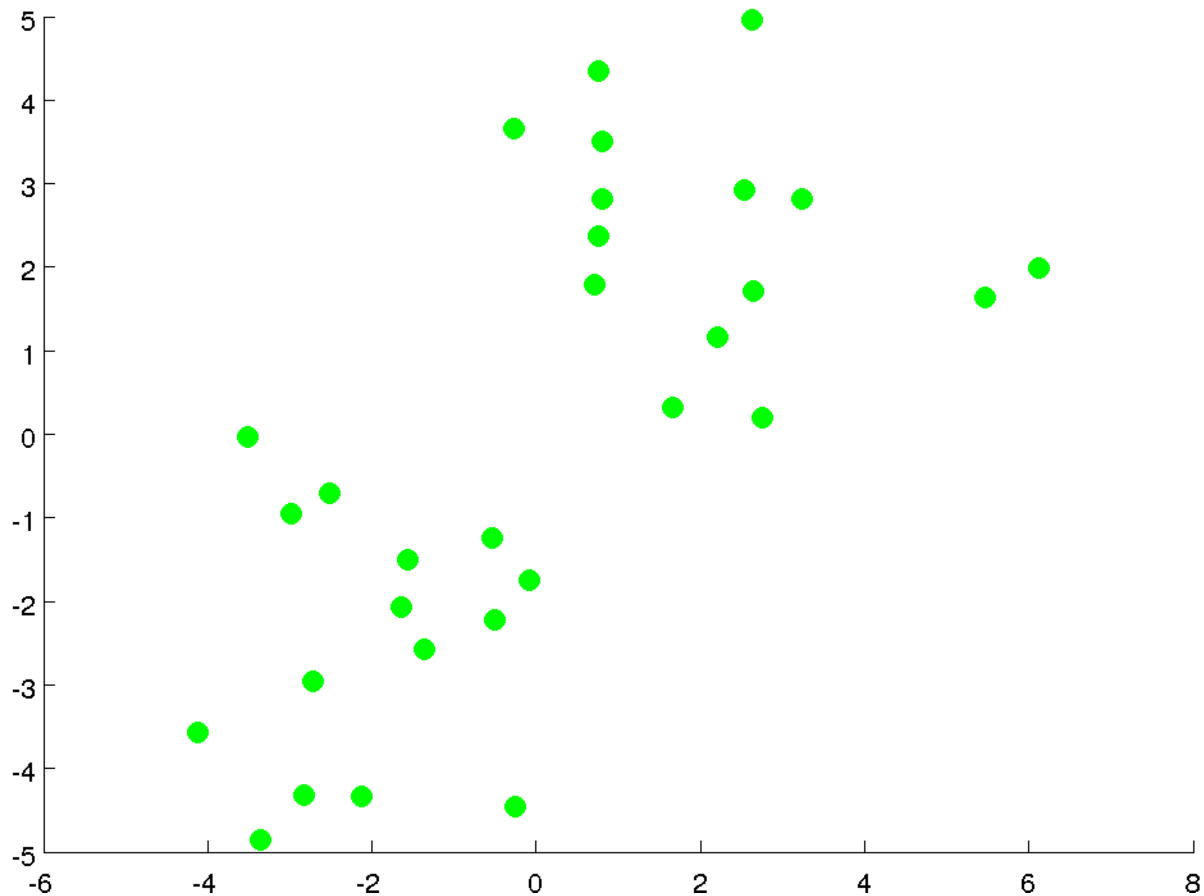


Social network analysis

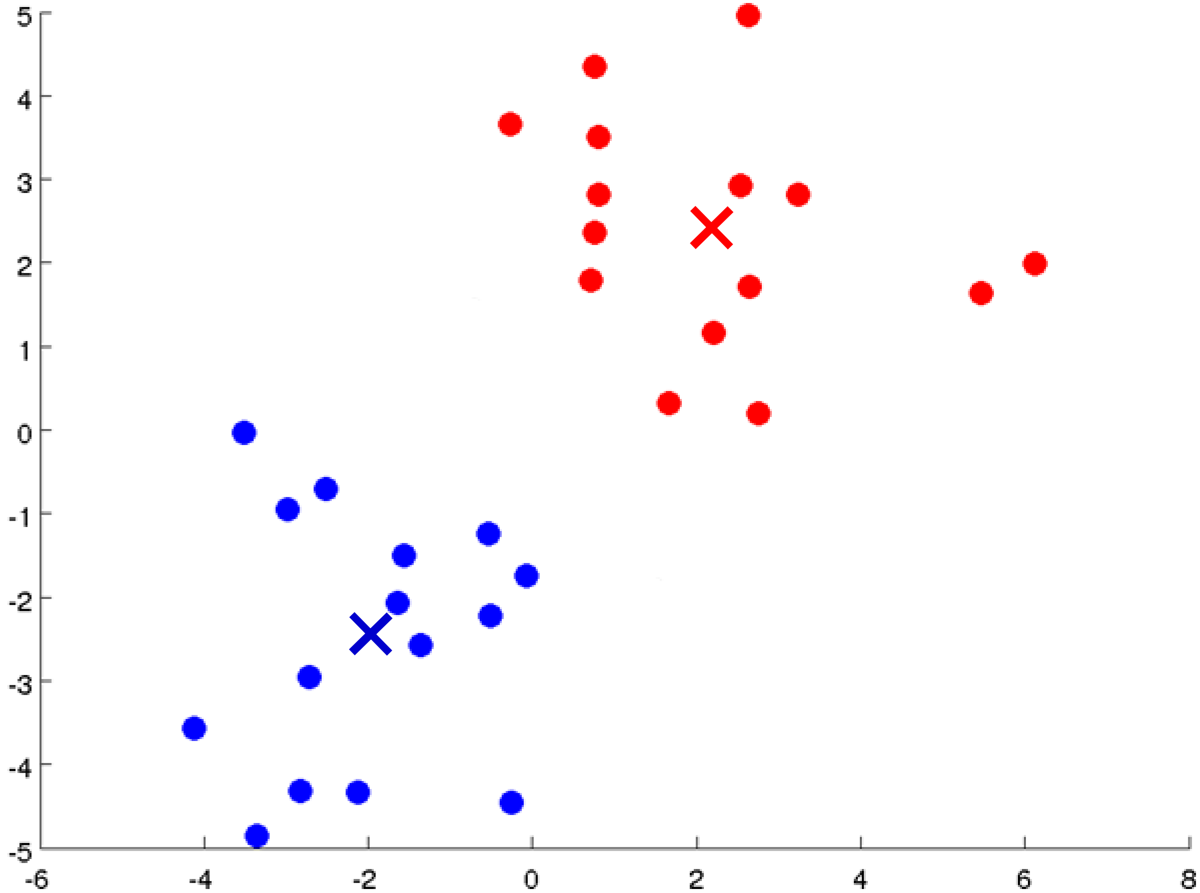
Types of voters



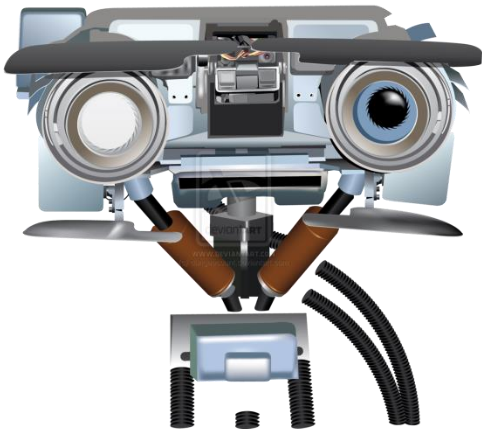
Trending news



slide credit: Andrew Ng



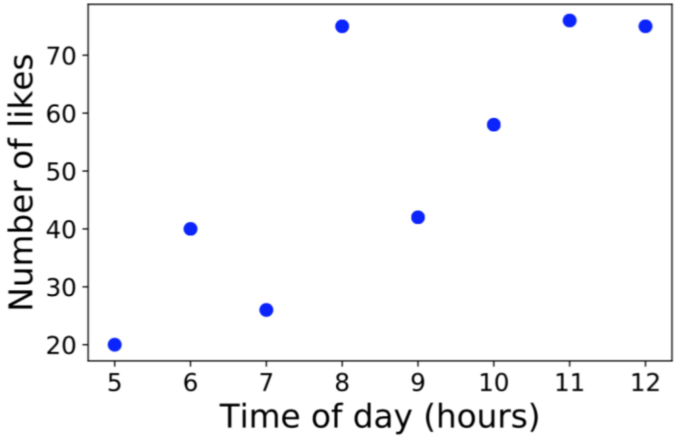
slide credit: Andrew Ng

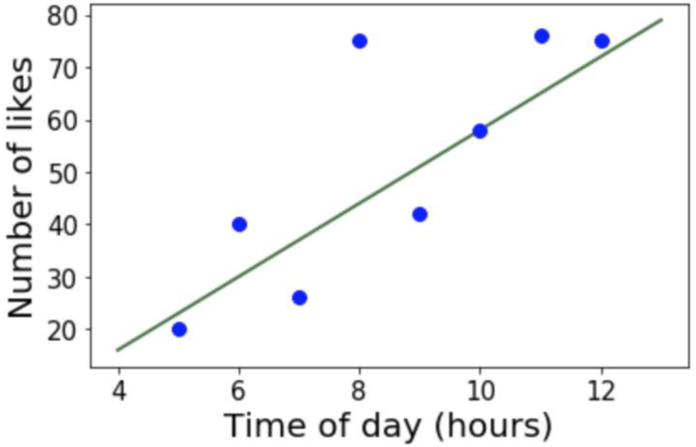


Supervised Learning: Linear Regression

Example of Supervised Learning: Linear Regression

- When the label is a real number
- Training a **model** to find a relationship between input and output values
- Learning a **line of best fit**



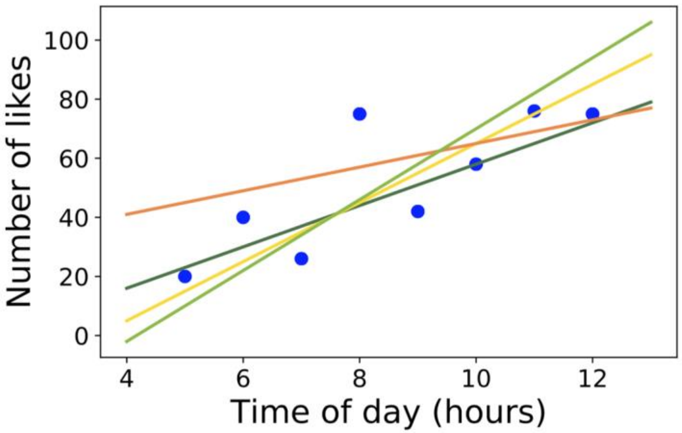


Linear Regression: Model Parameters

- Learning a best fit line means learning parameters (or weights) for our model.

$$h_\theta(x) = \theta_0 + \theta_1 x$$

θ_i's: Parameters



Linear Regression: Cost Function

- How do we know that the model parameters result in a “best fit” line? If parameter values minimize our cost function.

Cost Function:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

SSD = sum of squared differences, also
SSE = sum of squared errors

Multivariate Linear Regression

Hypothesis:

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

For convenience of notation, define $x_0 = 1$.

θ_i 's: Parameters

Cost Function:

$$J(\theta_0, \theta_1, \ldots, \theta_n) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Goal: $\underset{\theta_0, \theta_1, \ldots, \theta_n}{\text{minimize}} \; J(\theta_0, \theta_1, \ldots, \theta_n)$ **How??**

Two potential solutions

$$\min_{\theta} J(\theta; x^{(1)}, y^{(1)}, \ldots, x^{(m)}, y^{(m)})$$

Gradient descent (or other iterative algorithm)

- Start with a guess for θ
- Change θ to decrease $J(\theta)$
- Until reach minimum

Direct minimization

- Take derivative, set to zero
- Sufficient condition for minima
- Not possible for most "interesting" cost functions