INTRO TO MACHINE LEARNING

SESSION - 1

What is a shoe???

pixtastock.com - 64031702

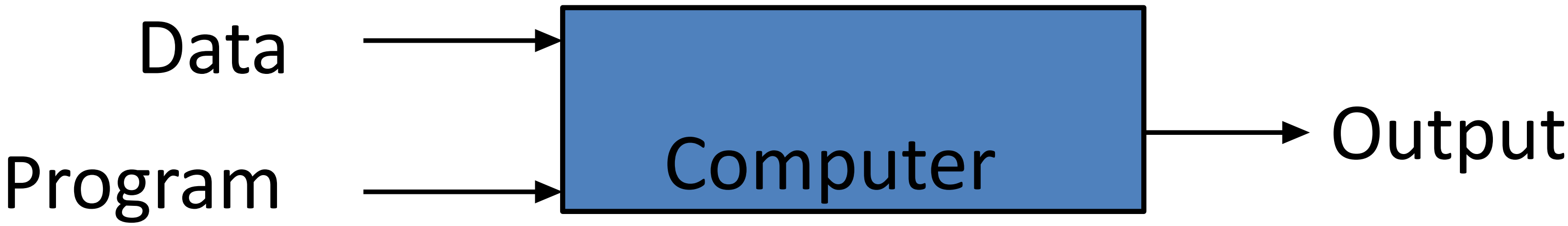
What is Machine Learning?(Defn)

“Learning is any process by which a system improves performance from experience” -- Herbert Simon

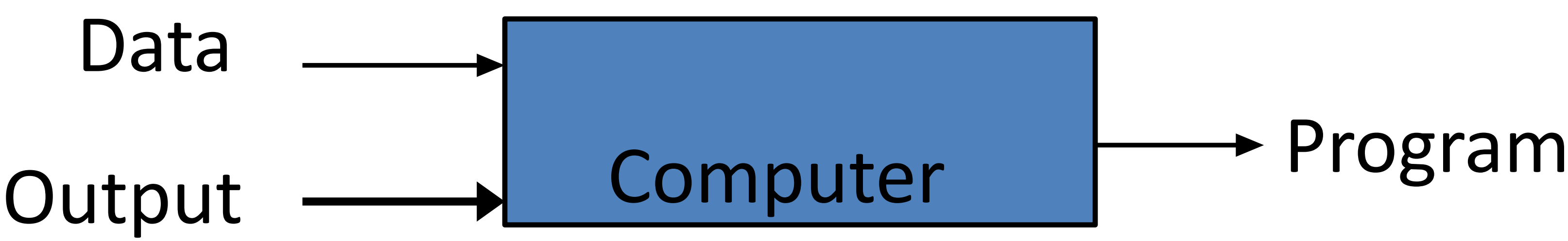
Machine Learning is the study of algorithms that
  improves their performance P
  at some task T
  with experience E.

Traditional Programming Vs Machine Learning

Traditional Programming



Machine Learning



Why now?

- Flood of available data (especially with the advent of the Internet)
- Increasing computational power
- Growing progress in available algorithms and theory developed by researchers
- Increasing support from industries

Some more examples of tasks that are best solved by using a learning algorithm

- Recognizing patterns:
  - Facial identities or facial expressions
  - Handwritten or spoken words
  - Medical images
- Generating patterns:
  - Generating images or motion sequences
- Recognizing anomalies:
  - Unusual credit card transactions
  - Unusual patterns of sensor readings in a nuclear power plant
- Prediction:
  - Future stock prices or currency exchange rates

Types of Learning

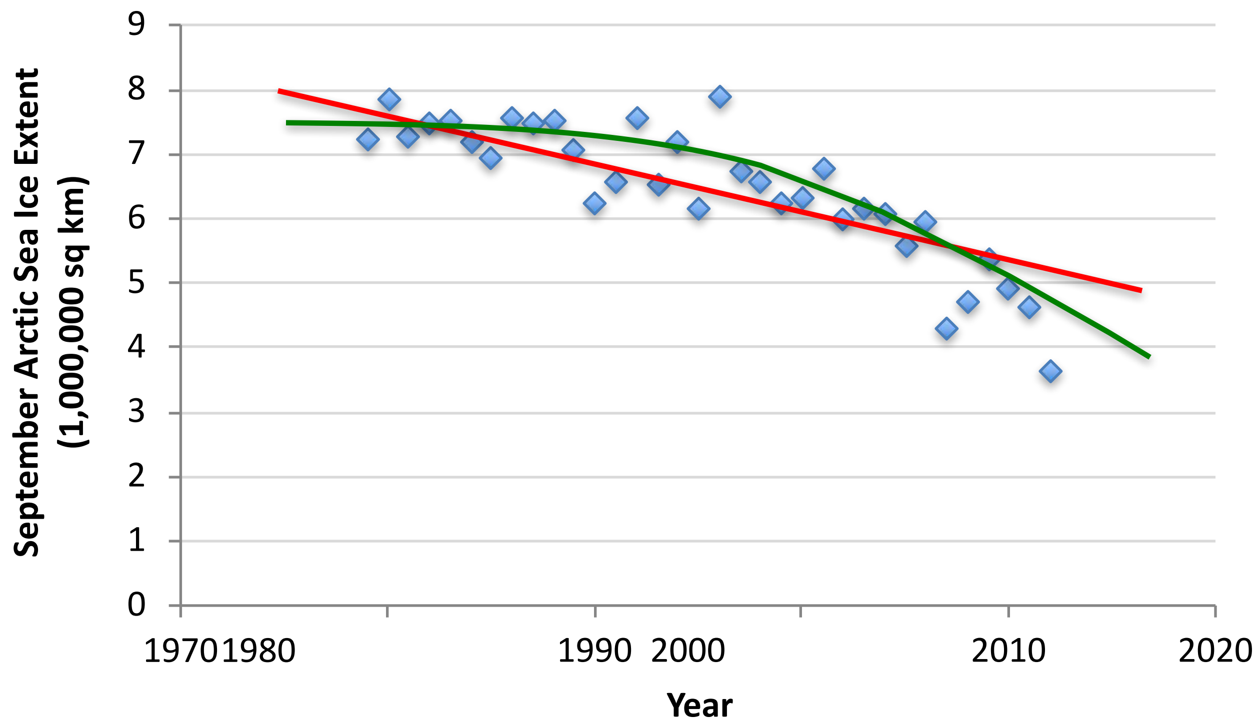
Types of Learning

- **Supervised (inductive) learning**
  - Given: training data + desired outputs (labels)
- **Unsupervised learning**
  - Given: training data (without desired outputs)
- **Semi-supervised learning**
  - Given: training data + a few desired outputs
- **Reinforcement learning**
  - Rewards from sequence of actions

Based on slide by Pedro Domingos

Supervised Learning: Regression

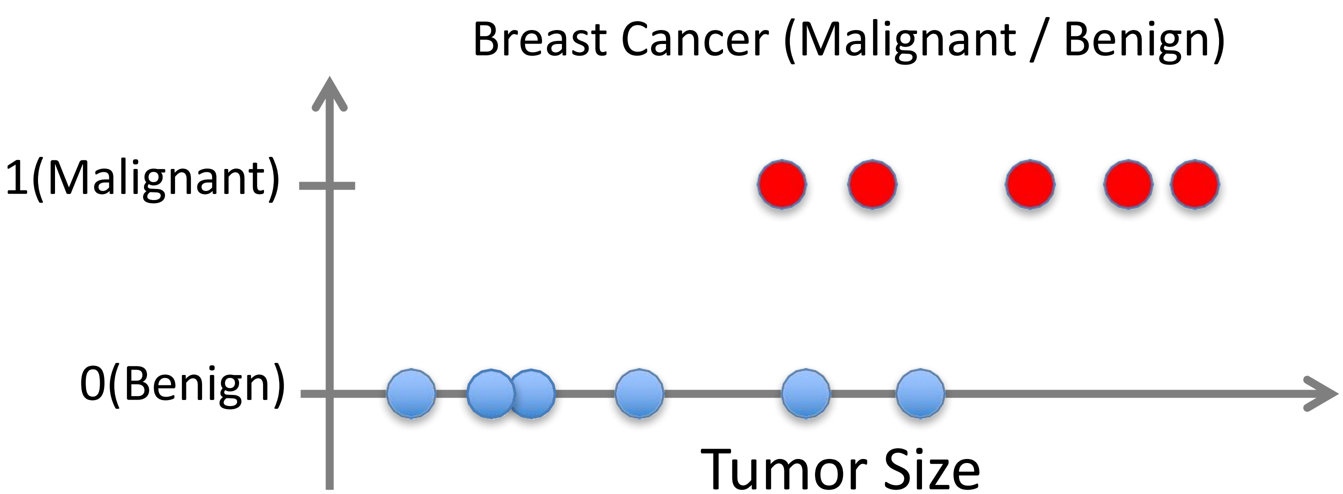
- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is real-valued == regression



Data from G. Witt. Journal of Statistics Education, Volume 21, Number 1 (2013)

Supervised Learning: Classification

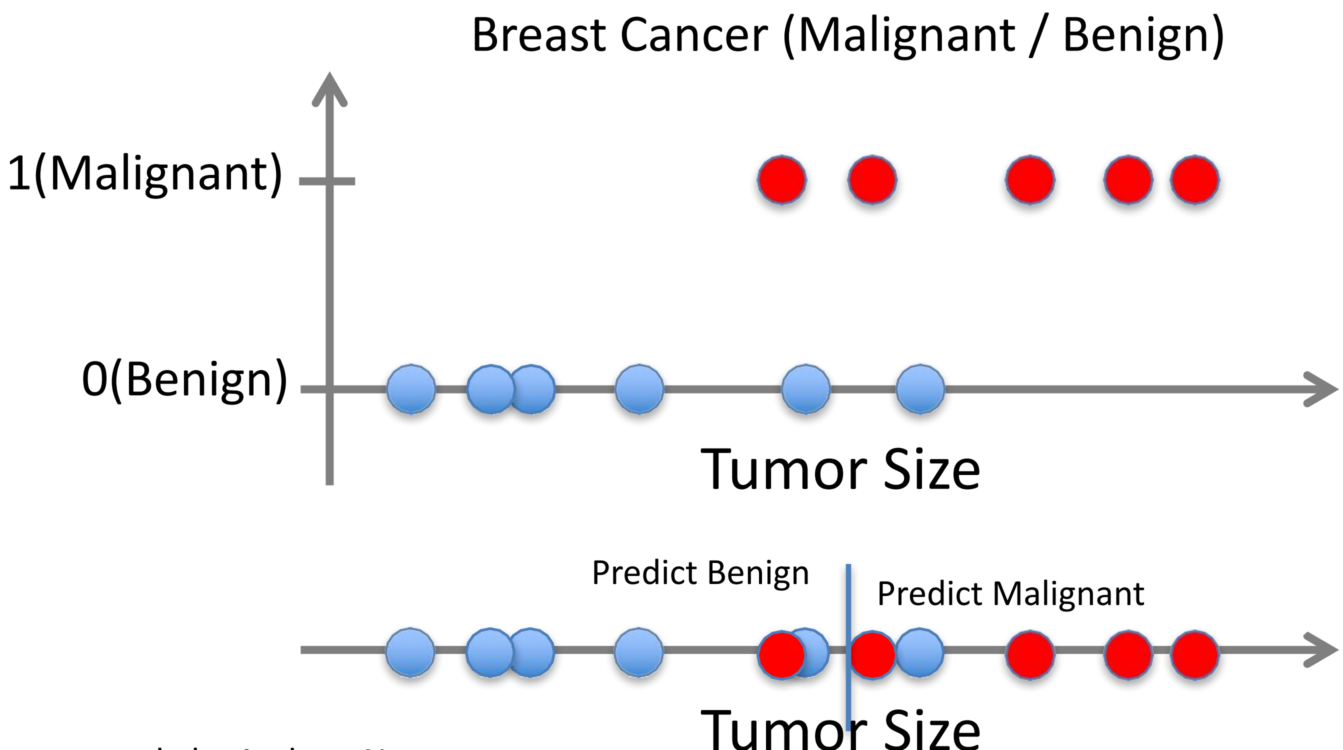
- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



Based on example by Andrew Ng

Supervised Learning: Classification

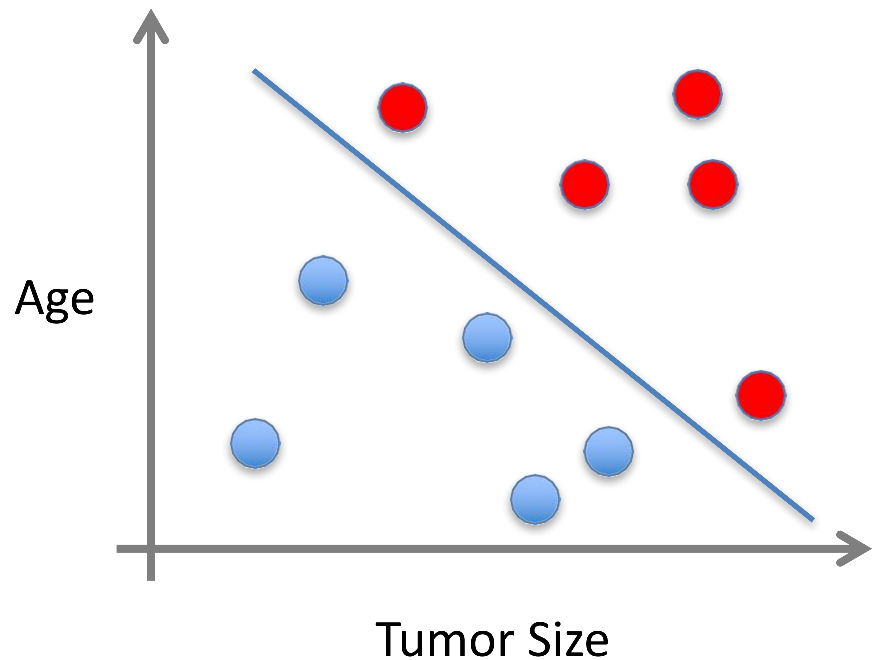
- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



Based on example by Andrew Ng

Supervised Learning

- x can be multi-dimensional
  - Each dimension corresponds to an attribute



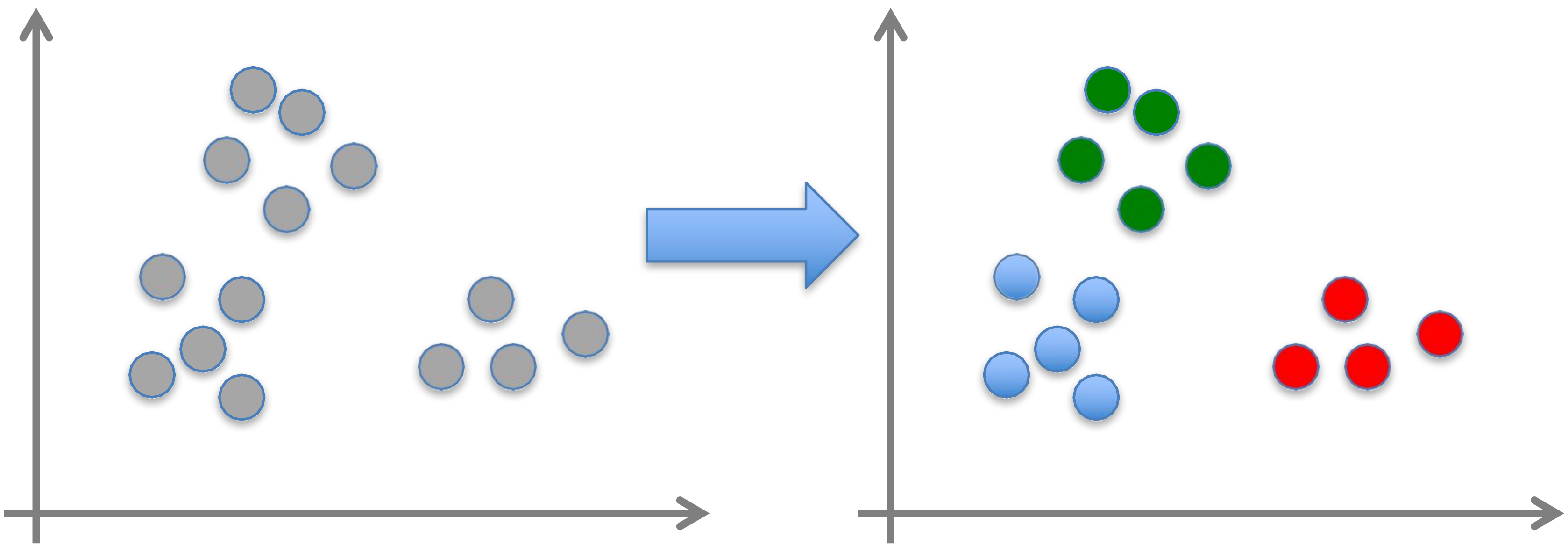
- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape

…

Based on example by Andrew Ng

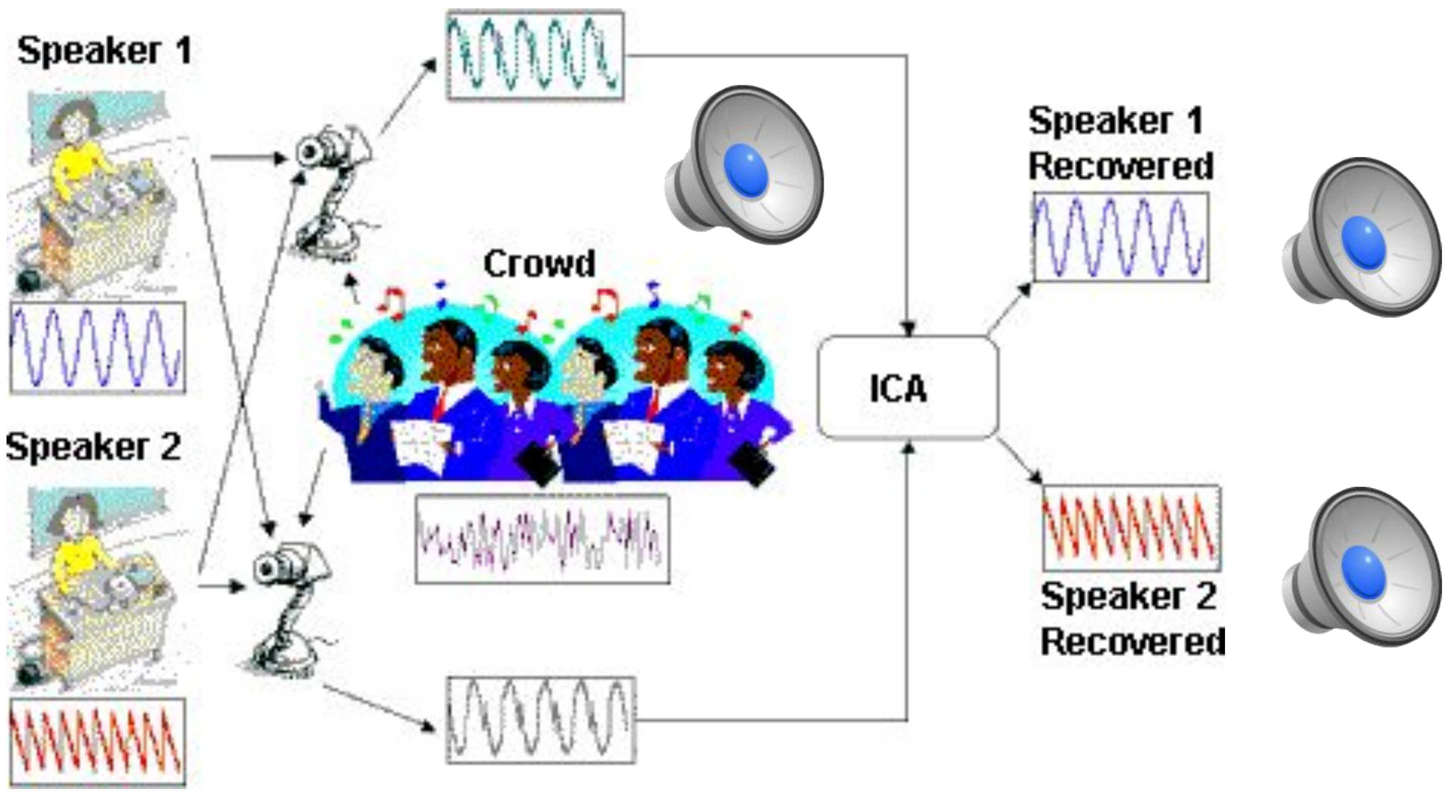
Unsupervised Learning

- Given $x_1, x_2, \ldots, x_n$ (without labels)
- Output hidden structure behind the x's
  - E.g., clustering



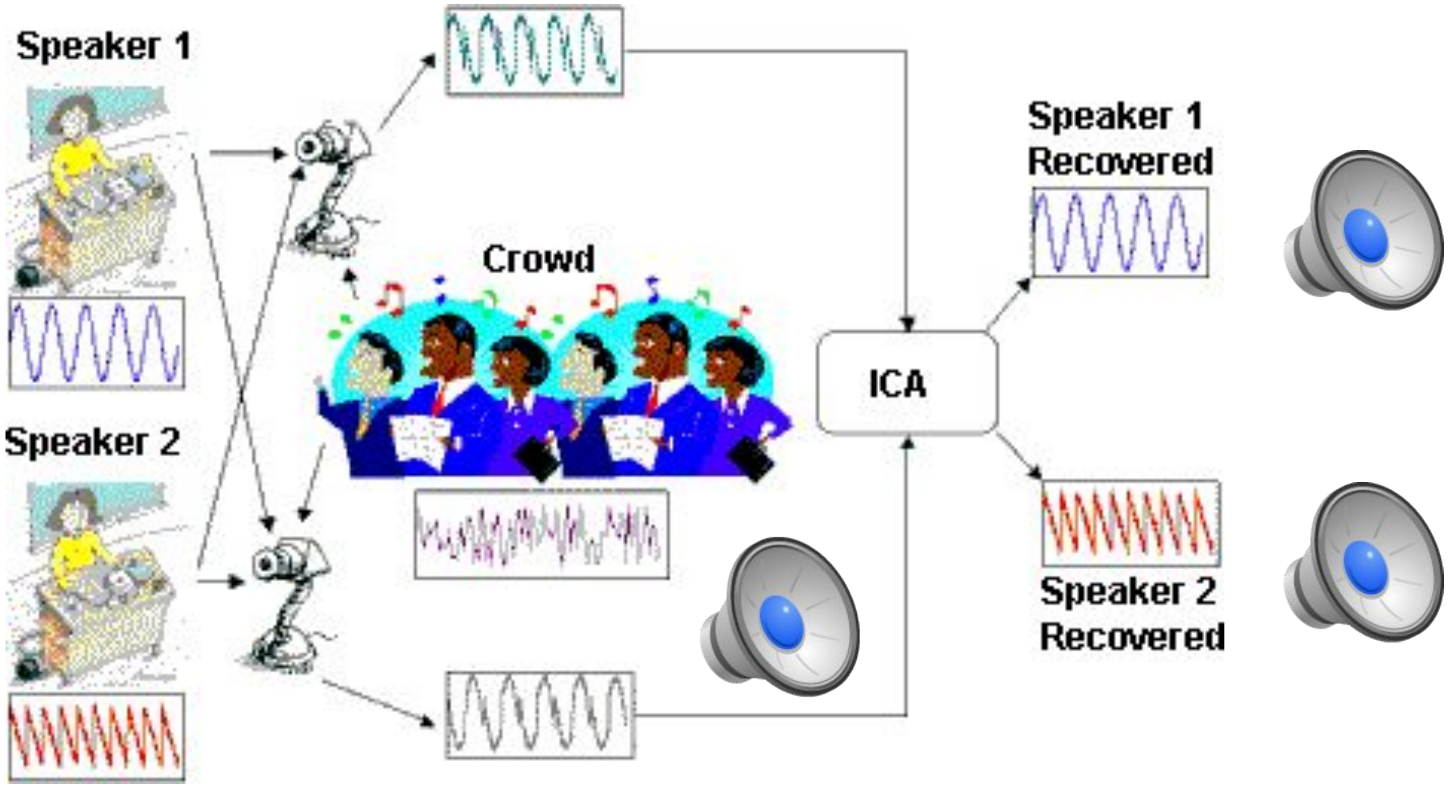
Unsupervised Learning

- Independent component analysis – separate a combined signal into its original sources



Unsupervised Learning

- Independent component analysis – separate a combined signal into its original sources

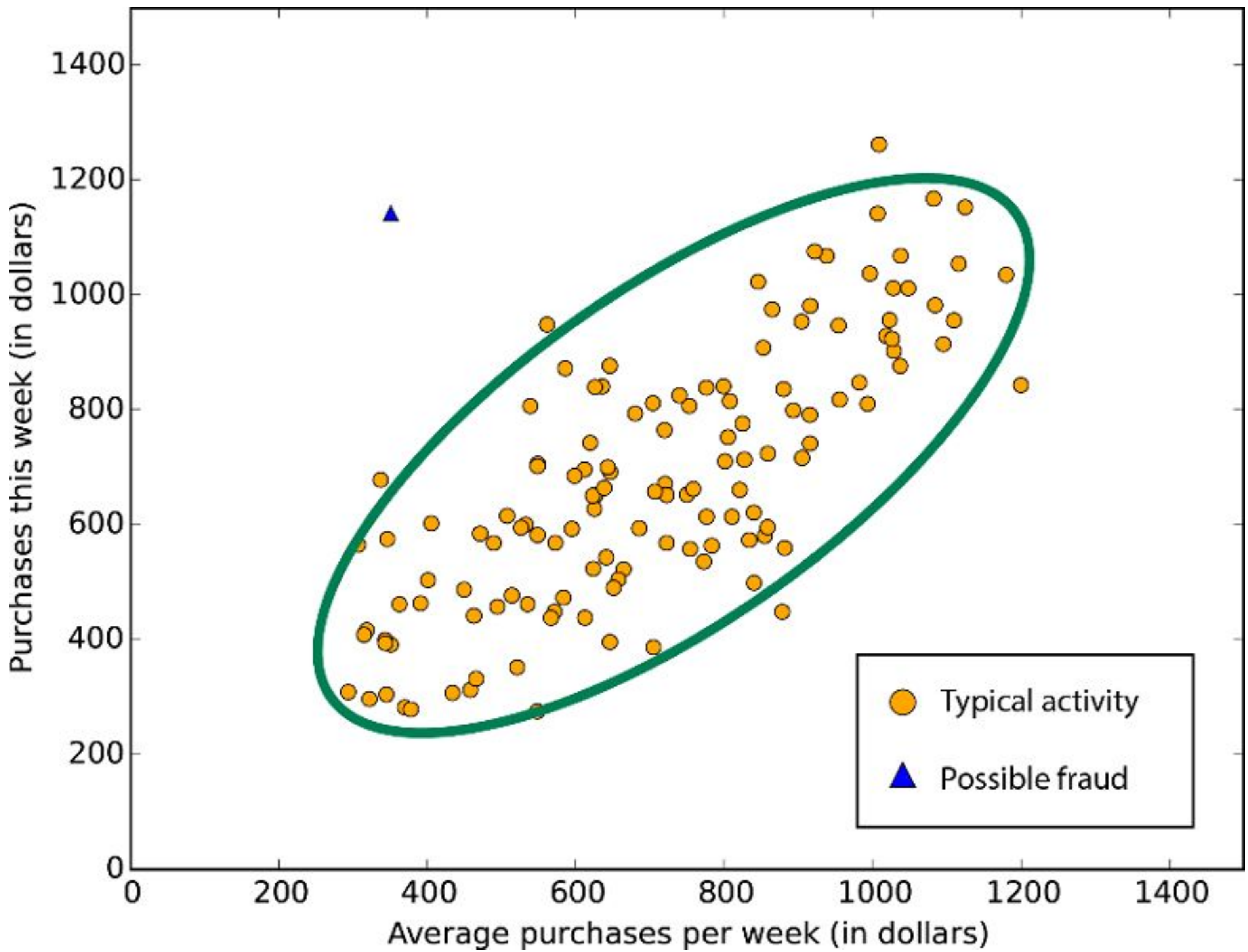


Unsupervised Learning

While **dimensionality reduction** is an important tool in machine learning/data mining, we must always be aware that it can distort the data in misleading ways.

Above is a two dimensional projection of an intrinsically three dimensional world....

Unsupervised Learning



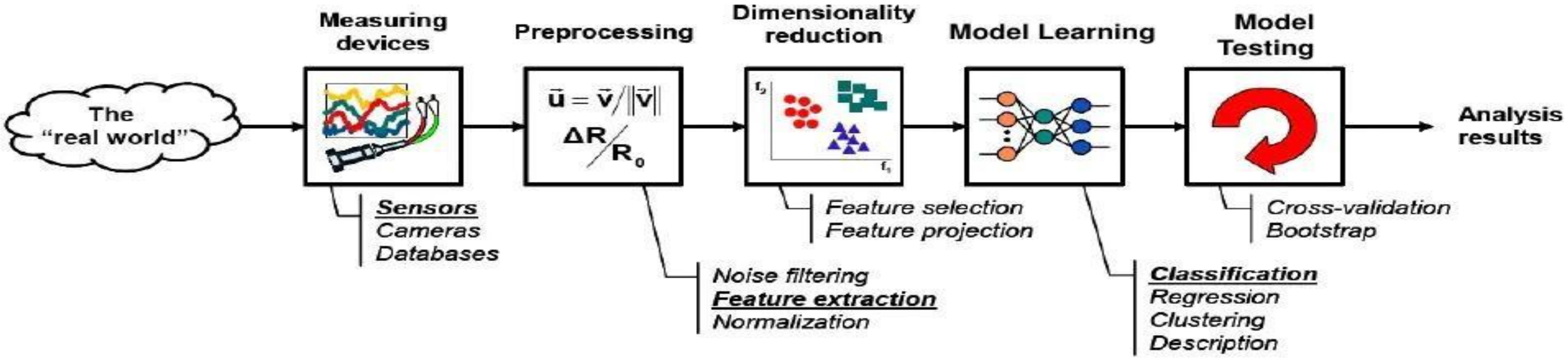
Reinforcement Learning

- Given a sequence of states and actions with (delayed) rewards, output a policy
  – Policy is a mapping from states → actions that tells you what to do in a given state
- Examples:
  – Credit assignment problem
  – Game playing
  – Robot in a maze
  – Balance a pole on your hand

Reinforcement Learning

https://www.youtube.com/watch?v=4cgWya-wjgY

The Learning Process



Evaluation metrics for an regression

Mean squared error	$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n} e_t^2$
Root mean squared error	$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$
Mean absolute error	$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n} \lvert e_t \rvert$
Mean absolute percentage error	$\text{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n} \left\lvert \frac{e_t}{y_t} \right\rvert$

Evaluation metrics for classification

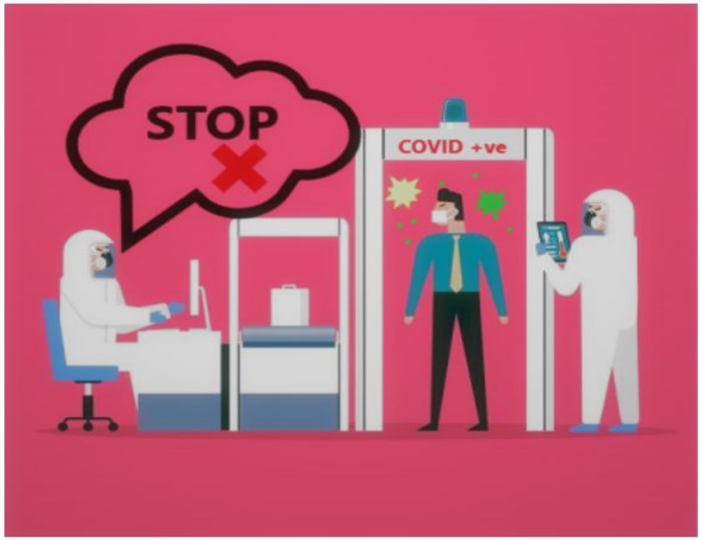


Fig.1 – True Positive

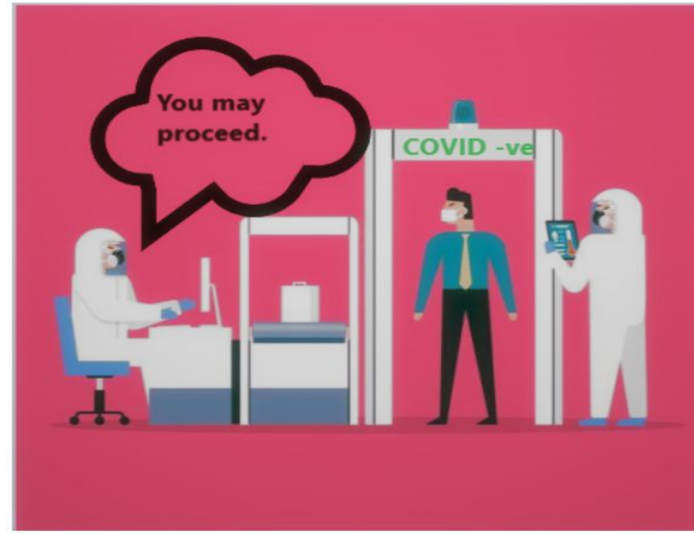


Fig.2 – True Negative

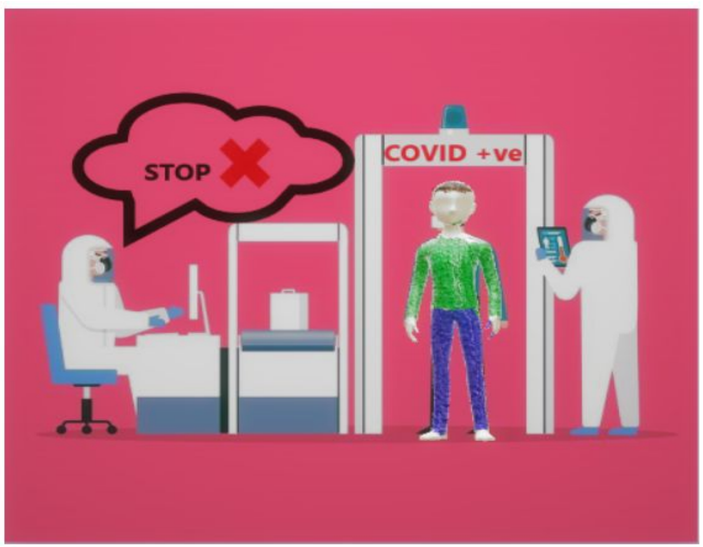
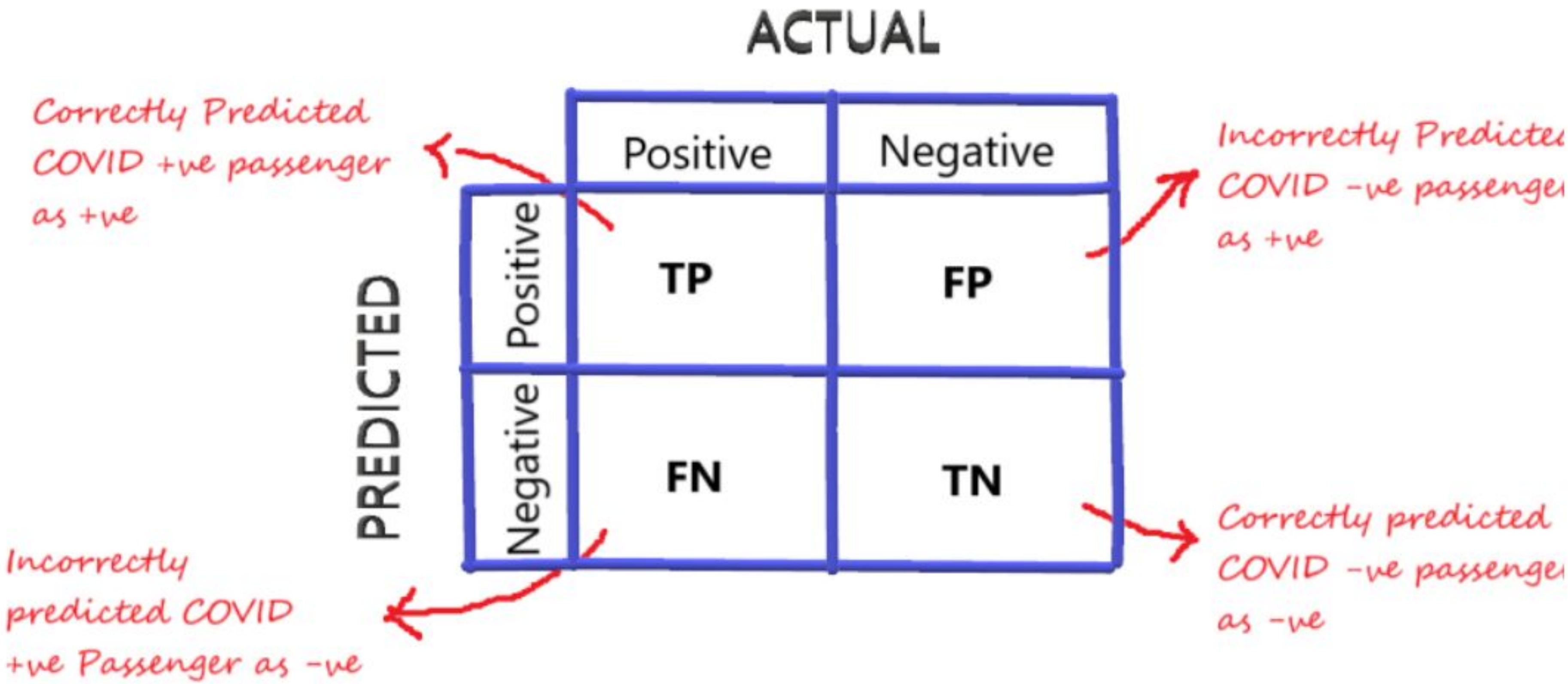


Fig.3 – False Positive

Fig.4 – False Negative

CONFUSION MATRIX



Accuracy

Accuracy = Number of correct predictions / Total number of predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TP}$$

Now, let's consider 50,000 passengers travel per day on an average. Out of which, 10 are actually COVID positive. If our learned model predicted everyone as a covid -ve. Lets check the accuracy of this model

		ACTUAL	
		Positive	Negative
PREDICTED	Positive	TP = 0	FP = 0
	Negative	FN =10	TN 50,000 - 10 = 49,990

Accuracy for this case will be:

Accuracy = 49,990/50,000 = 0.9998 or **99.98%**

Precision and Recall

Correctly predicted as COVID +ve

$$\text{Recall} = \frac{TP}{\underbrace{TP + FN}_{\text{Total COVID +ve Passengers}}}$$

Correctly Predicted as COVID +ve

$$\text{Precision} = \frac{TP}{\underbrace{TP + FP}_{\text{Total Predicted as COVID +ve}}}$$

F1 Score

		ACTUAL	
		Positive	Negative
PREDICTED	Positive	**TP** = 1	**FP** = 0
	Negative	**FN** = 9	**TN** 50,000 - 9 = 49,991

$$\textbf{F1 Score} = 2 * \frac{\textbf{Precision} * \textbf{Recall}}{\textbf{Precision} + \textbf{Recall}}$$

So Here precision is
P=1(TP)/1(TP)+0(FP)

P=1

So here Recall is
R=1(TP)/1(TP)+9(FN)
R=0.1

F1=2*(1*0.1)/(1+0.1)
F1=0.18