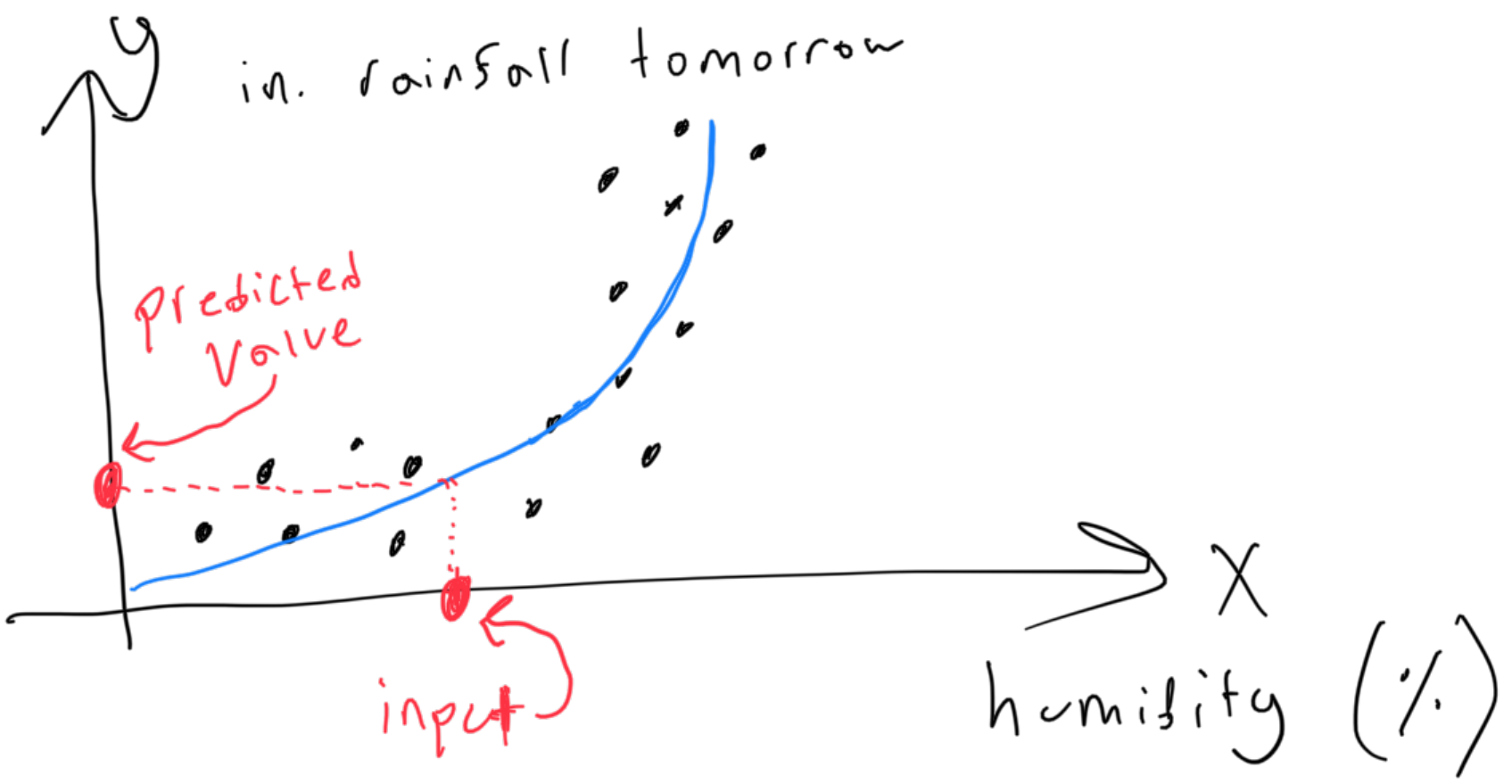
Day 1: ML Overview

ICS 435 / ICS 635 / DATA 435

Supervised Learning

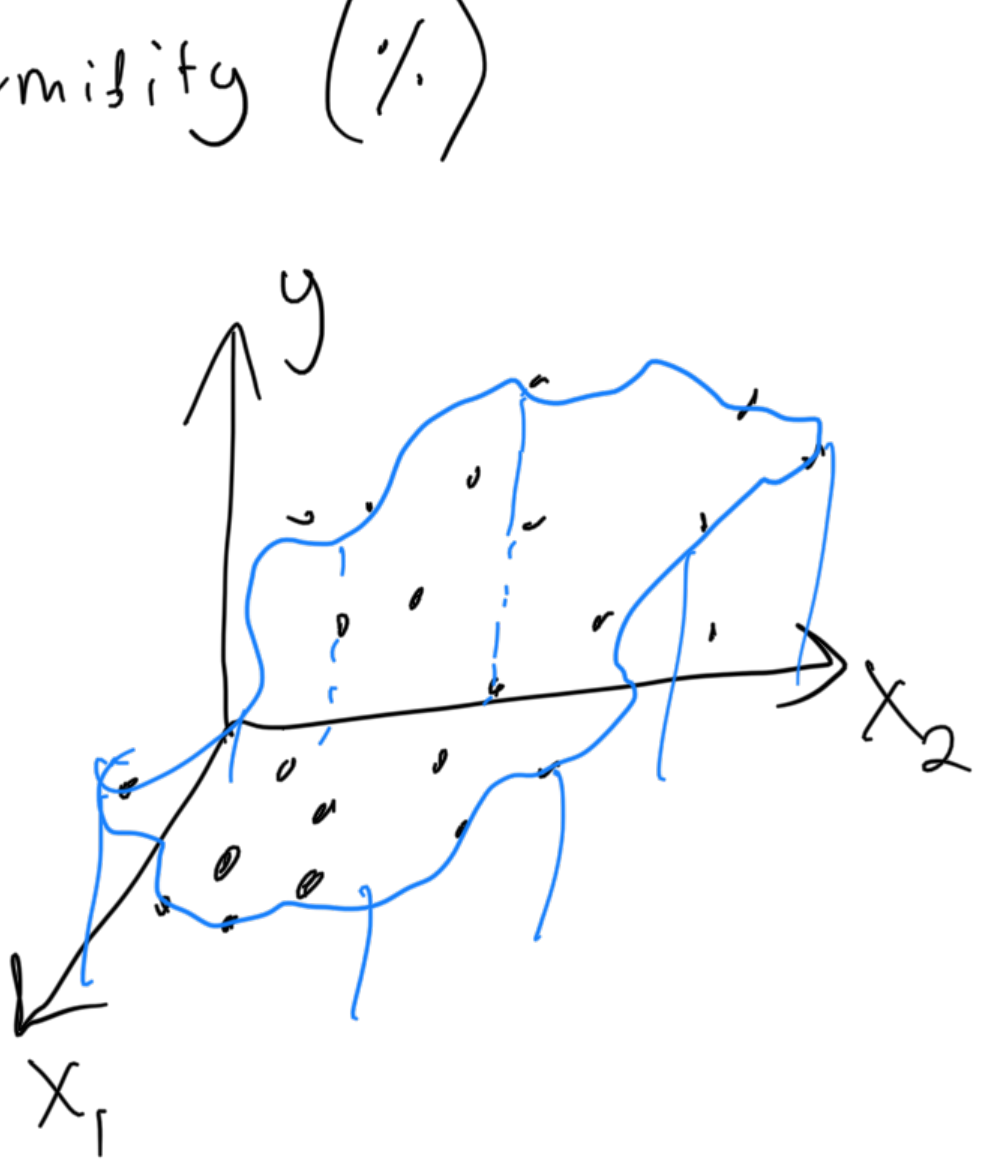
Predict from data

Regression: predict a number



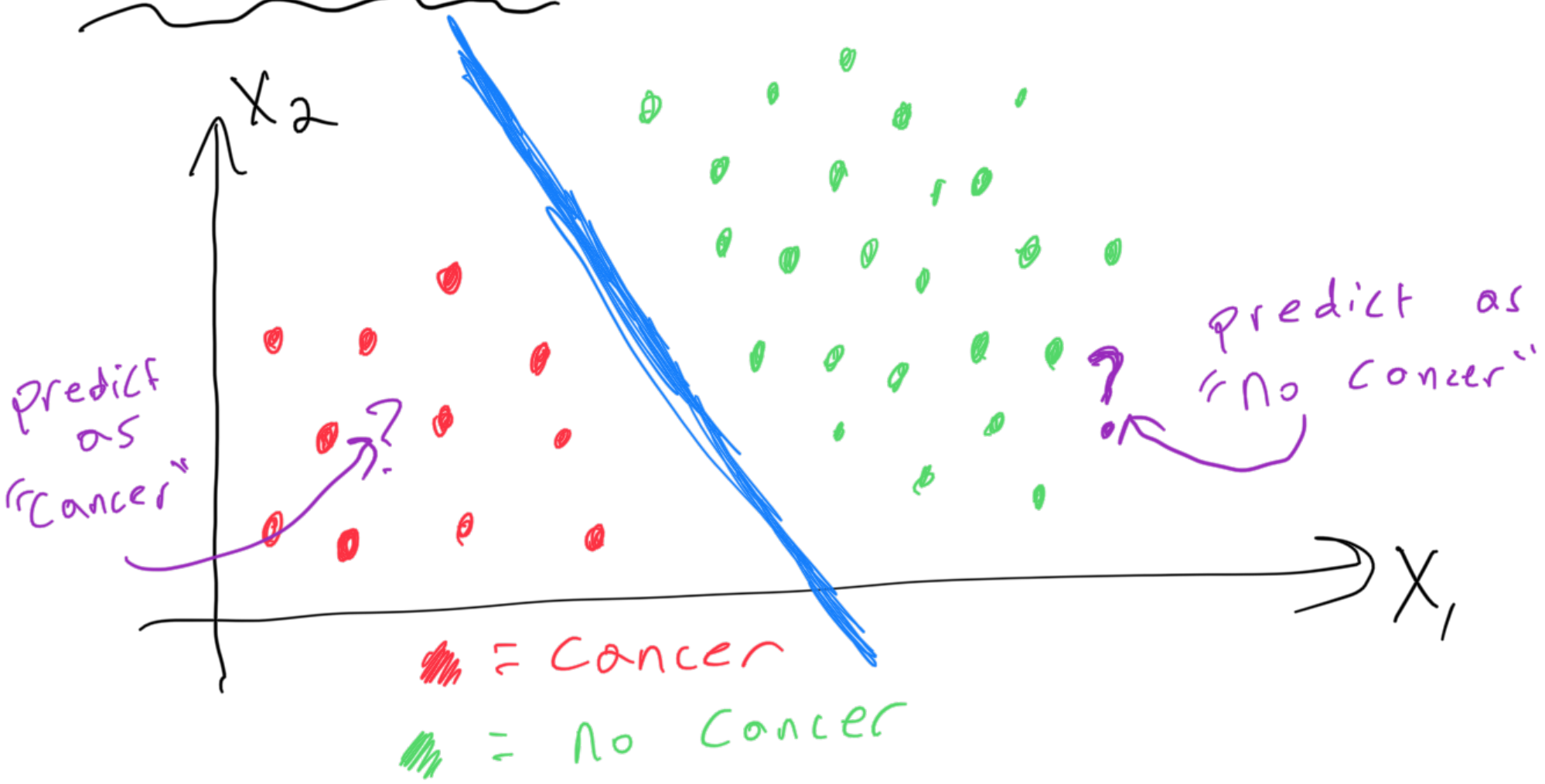
Examples:

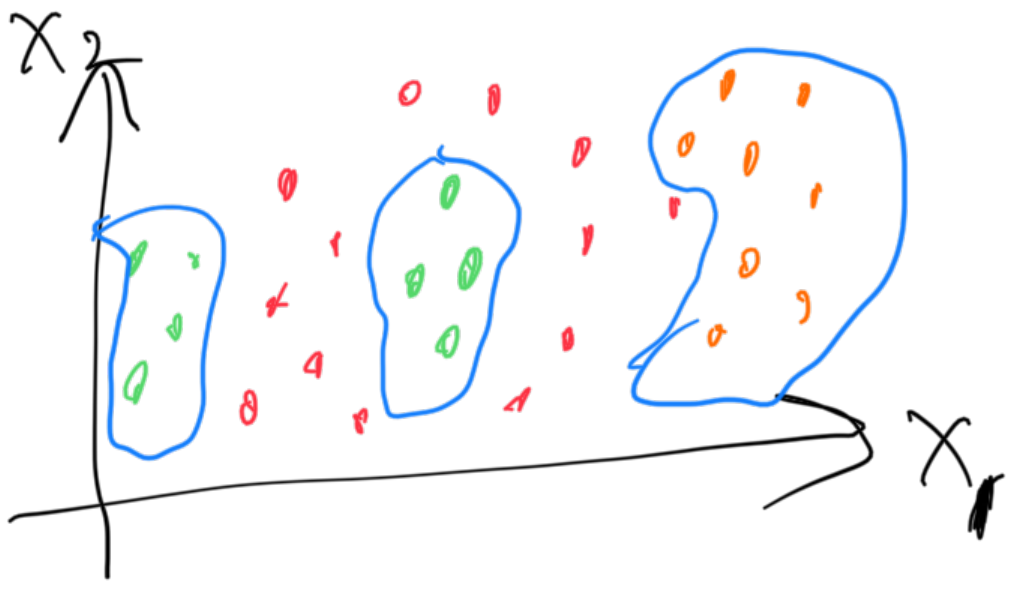
1. Input: sq ft, ...
   Output: $



Categories

Classification: predict categorical





Examples

Classify proteins

Name Pokemon from image

CAPTCHA - filter Classify artist

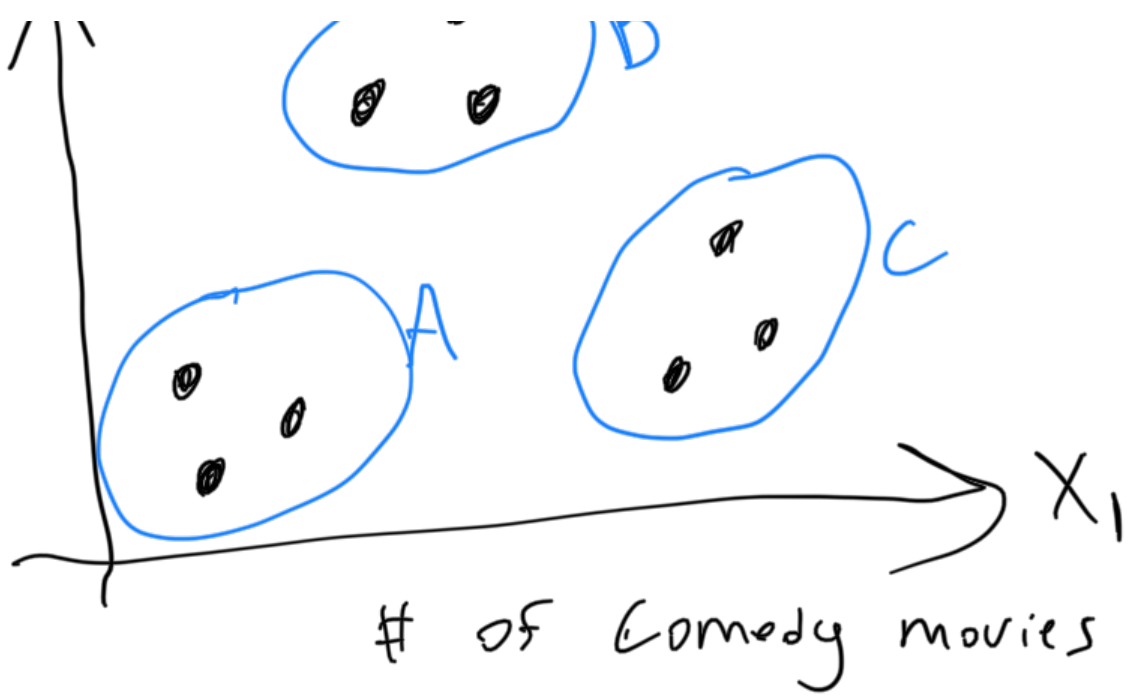
TSA filtering

Unsupervised Learning

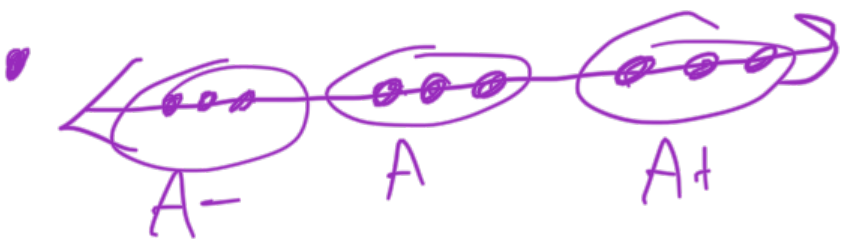
Learn patterns from data

Clustering

Examples:

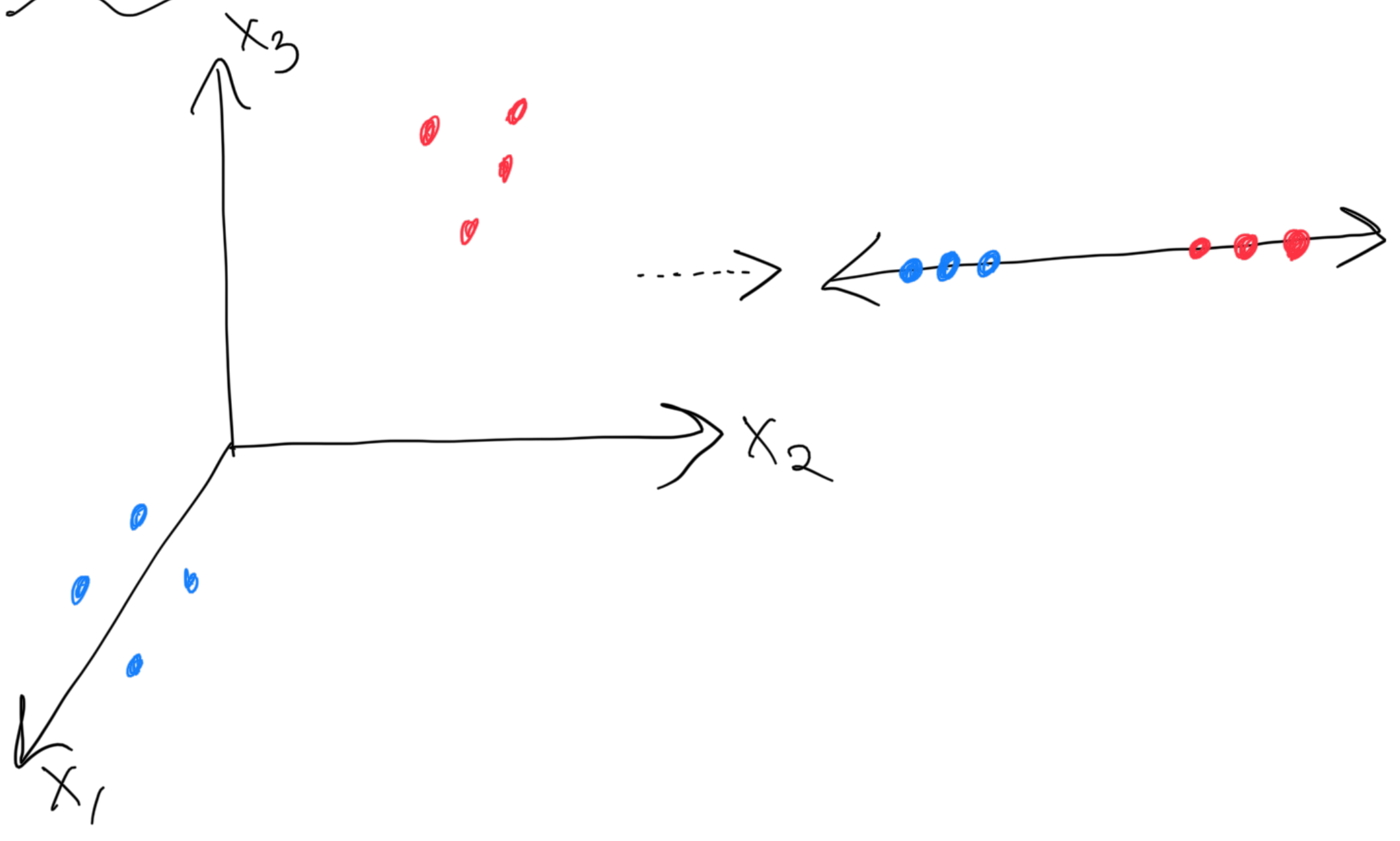


- Clustering ... items
- Recommender Systems

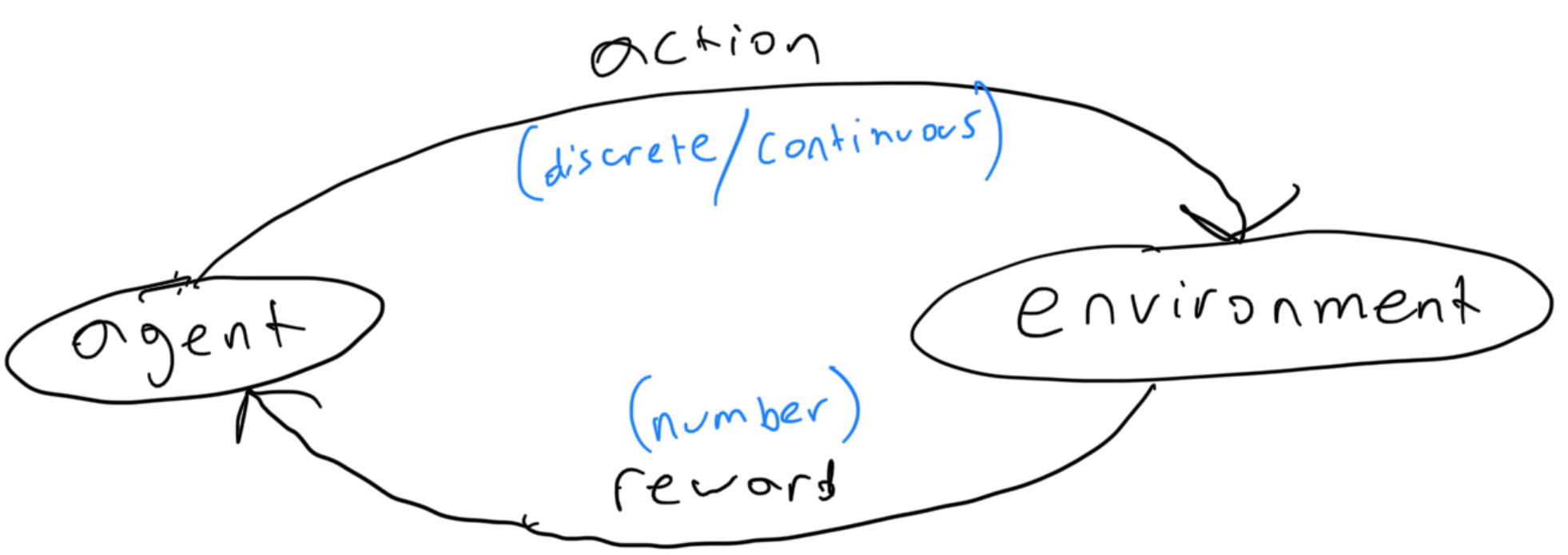


Anomaly Detection

Dimensionality Reduction



Reinforcement Learning



$S_1 \xrightarrow{a_1} S_2 \xrightarrow{a_2} S_3 \longrightarrow \cdots$

(collection of #s, e.g., vector or matrix)

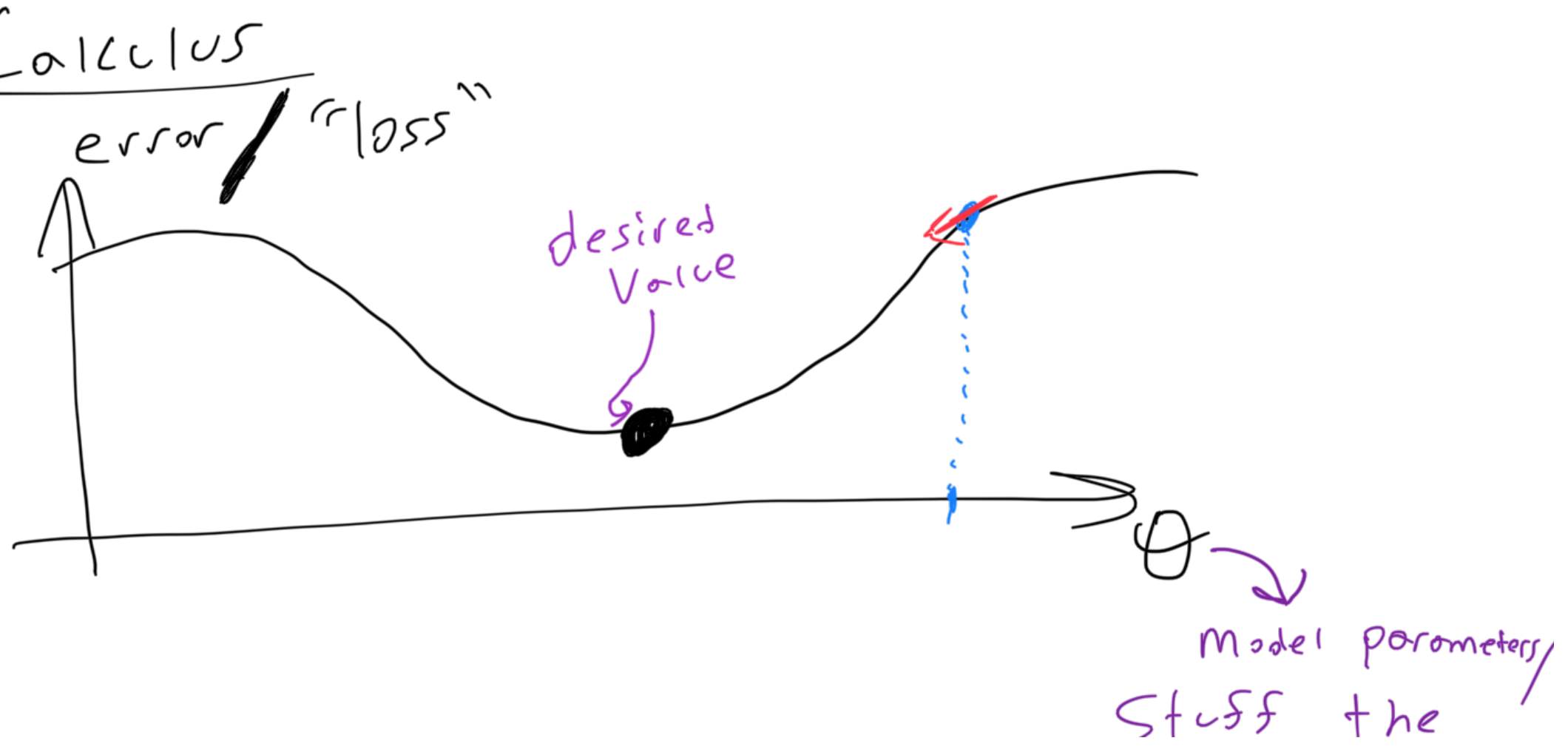
Learn POLICY function $P(S, a) =$

prob. of taking action a in state S

$$\begin{bmatrix} 0.1 & 0.8 & 0.1 \\ a_1 & a_2 & a_3 \end{bmatrix}$$

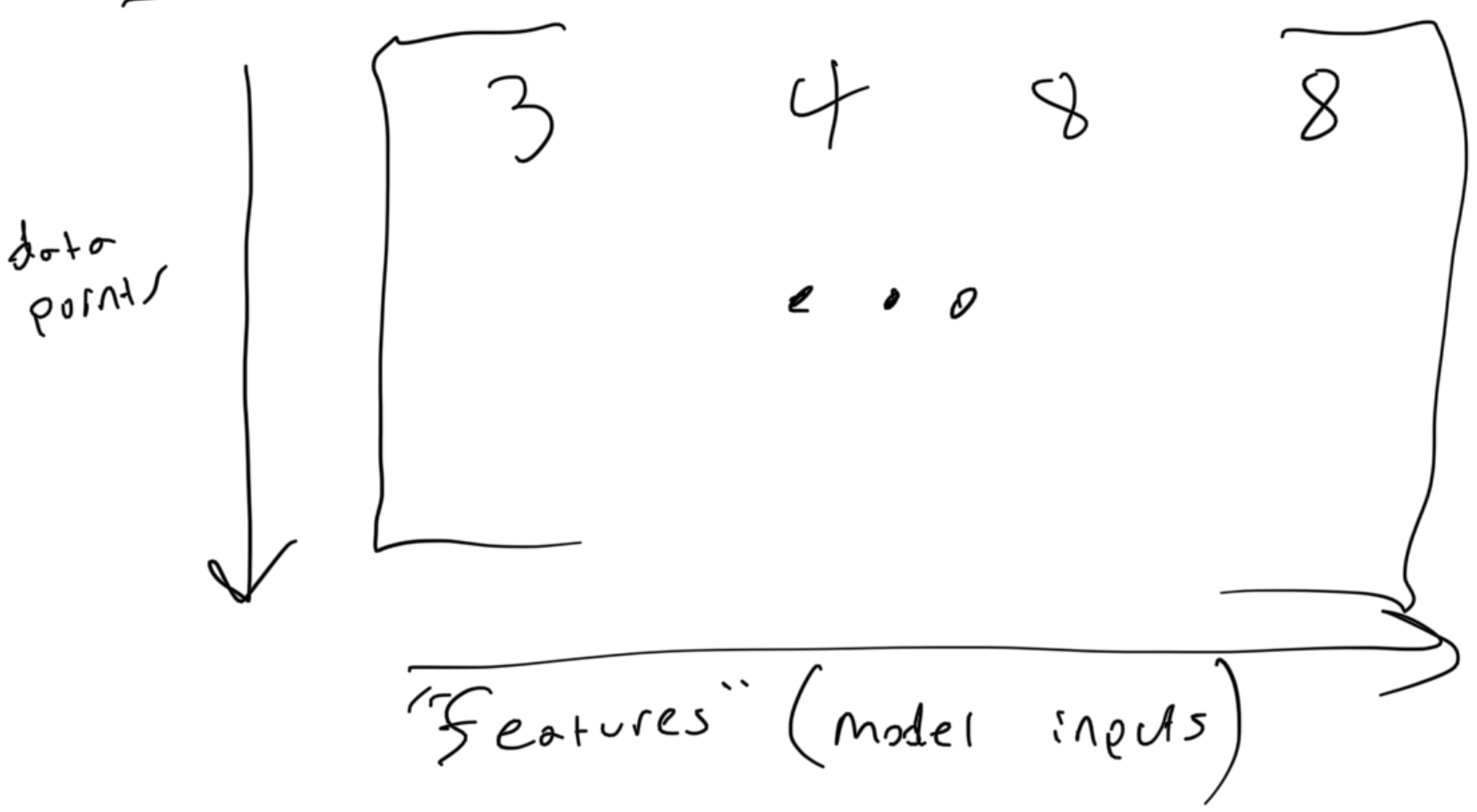
Math

Calculus

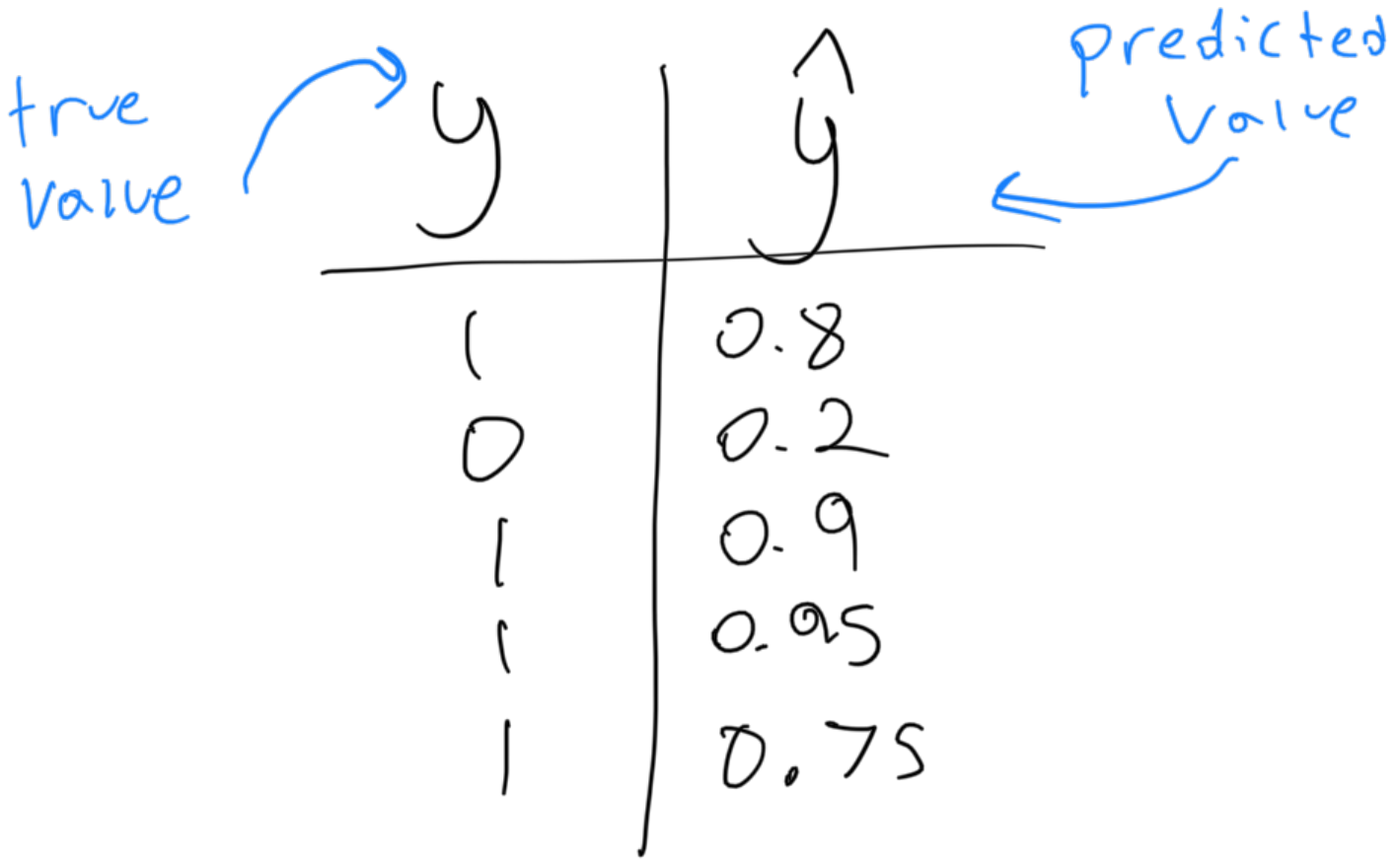


model learns

Linear Algebra



Probability



Representing Data

$$y = f(x)$$

output data

input data

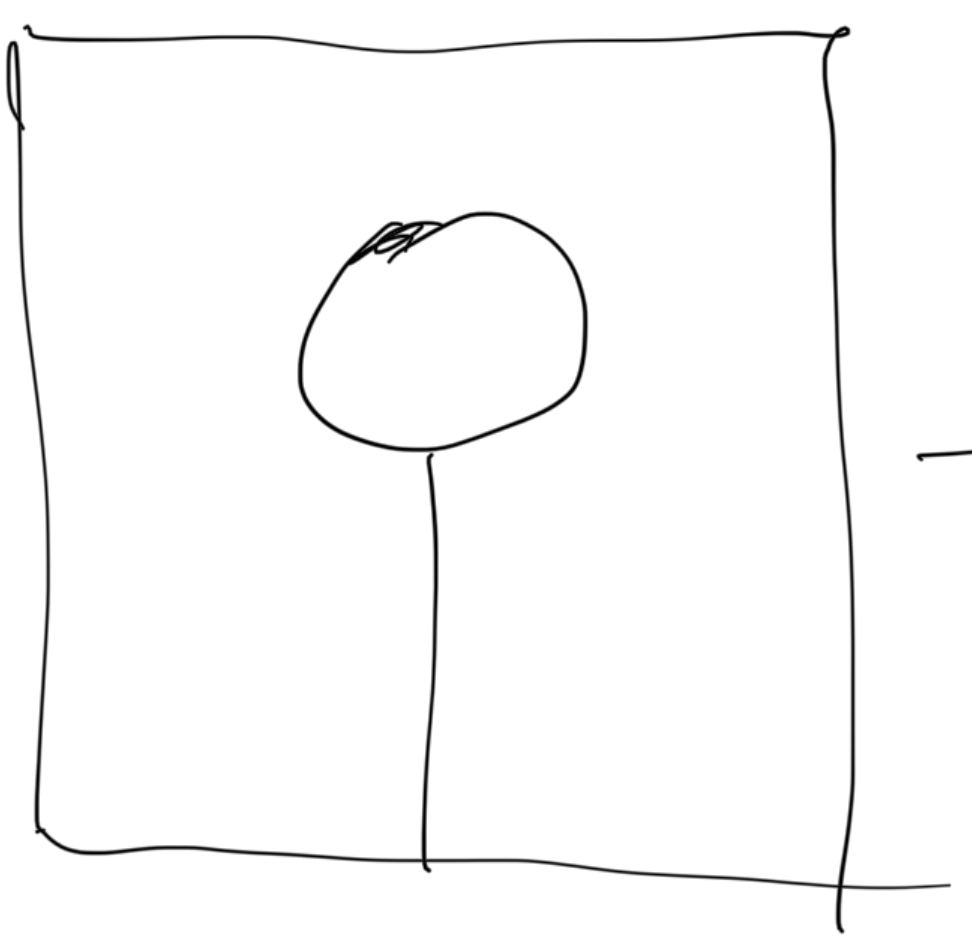
Tabular data (e.g., CSV, Excel, etc.)

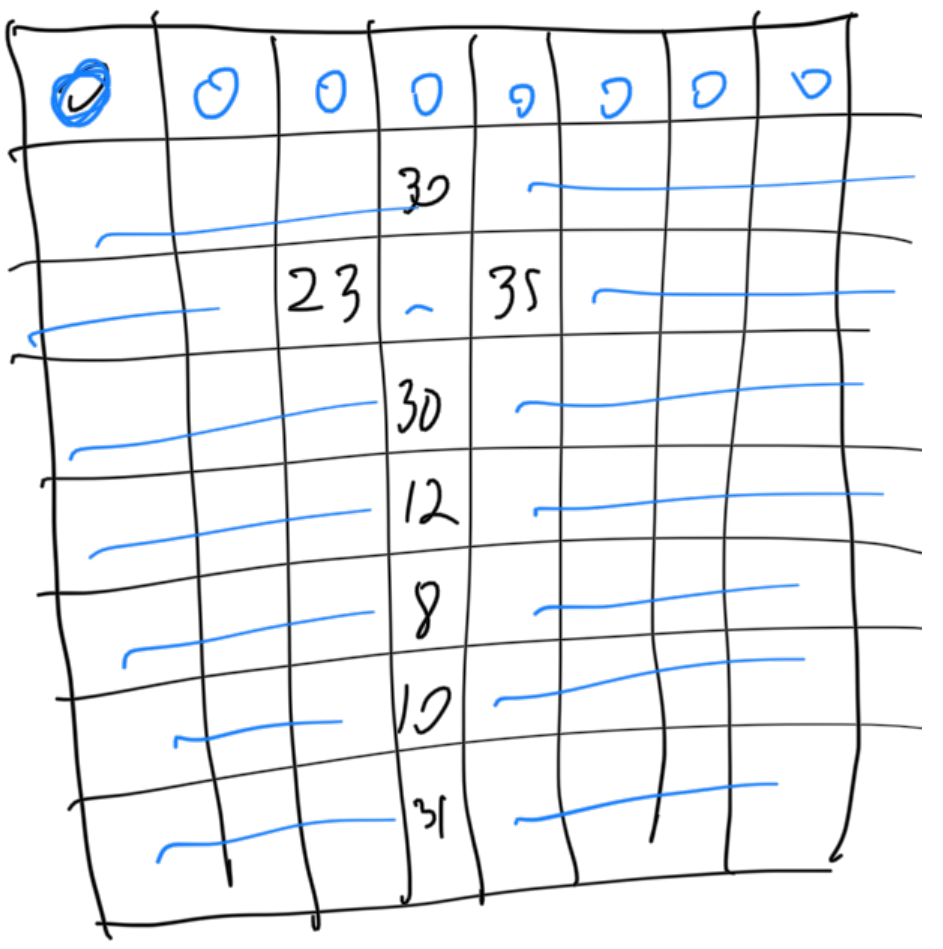
	# tweets per day	avg len of tweet				should we block account?
user 1	34	8	0	3	...	1
user 2						0
...						1
						0
						1

data points are rows

"features"/model inputs are columns

X y





"I love ICS 435/635 DATA 435" → [-3.1 4.3 8.3 .....]