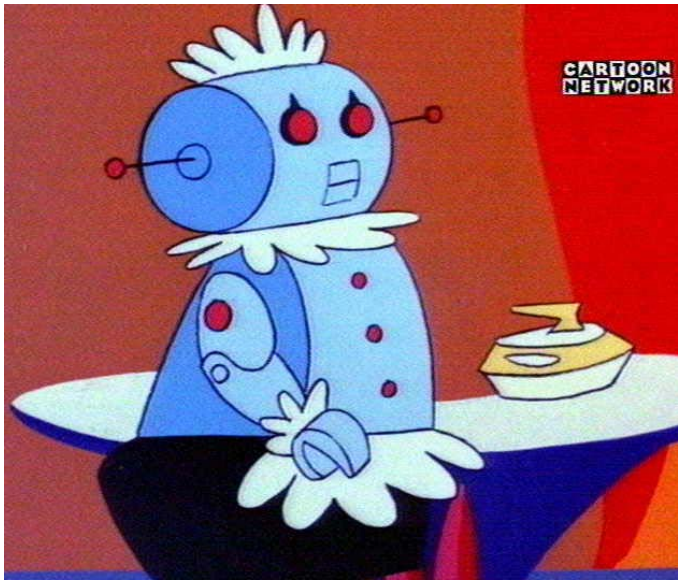
Artificial Intelligence
&
Machine Learning

ICS 90 Guest Lecture

Prof. Alex Ihler

What is AI?


CARTOON NETWORK

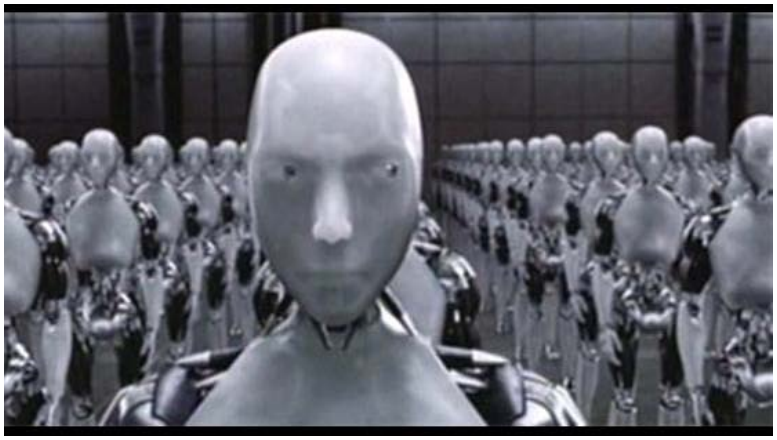
?

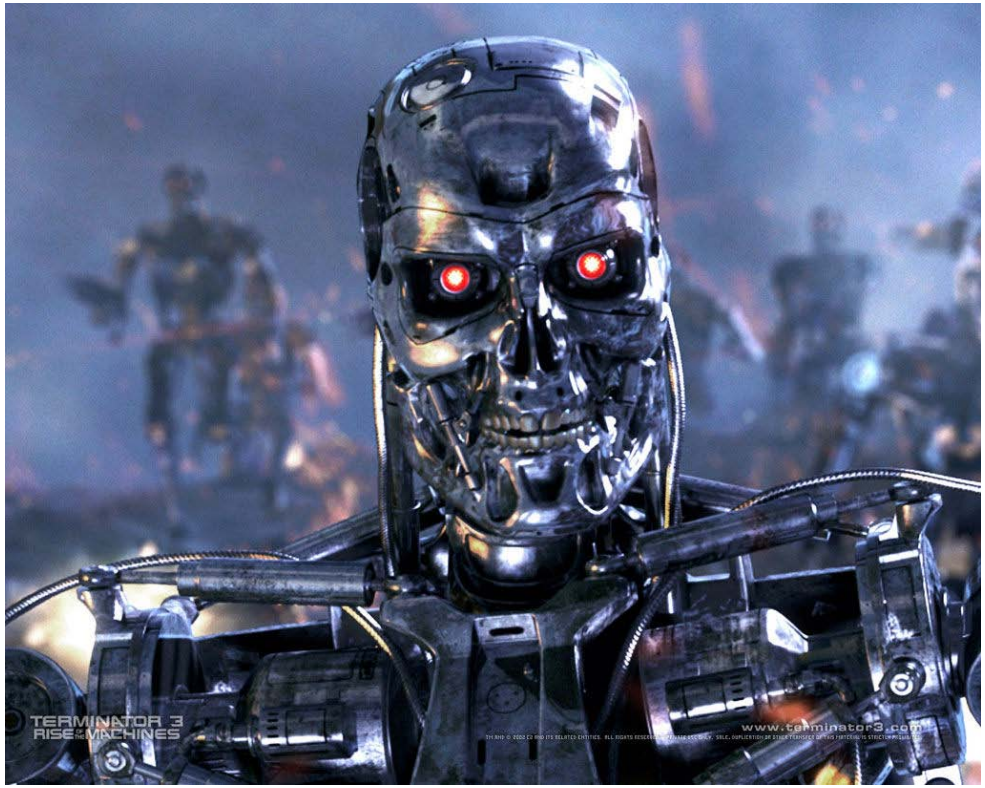

Roomba
SCHEDULER
iRobot

?


JVC
D-VHS
D-THEATER

What is AI?





What is AI?

- Not (just) killer robots
- Solving complex problems

- Some examples
  - Mathematics, games
  - Reasoning about the world
  - Expert advice: medicine, law, taxes
  - Automated decisions and actions

History of AI

Some successes:

Chess (Deep Blue v. Kasparov)

RoboCup

Darpa GC (Stanley)

Watson (Jeopardy)

Agents

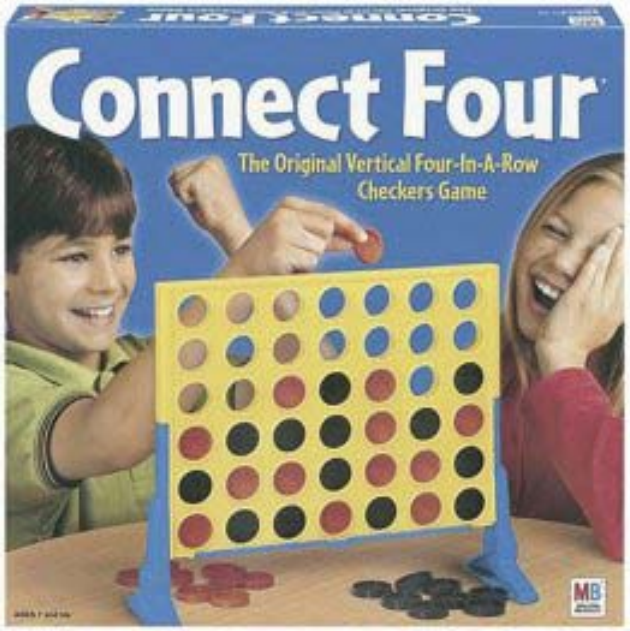
- An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators

- Human agent:
  - Sensors: eyes, ears, …
  - Actuators: hands, legs, mouth…

- Robotic agent
  - Sensors: cameras, range finders, …
  - Actuators: motors

Agents and Games

- AI as an opponent
  - Responsive & “realistic” behavior
- AI as an assistant
  - Simplify interface, “do what you want”



What is ML?

- How can a computer “learn” from experience (observed data)?
- Less than the whole of AI?
  - Just one part of intelligence…
- More than just AI?
  - Applicable to many “practical” problems
  - Making sense of data automatically
  - Found in
    - Data mining & information retrieval
    - Computational biology
    - Signal processing
    - Image processing & computer vision
    - Data compression and coding

Classification

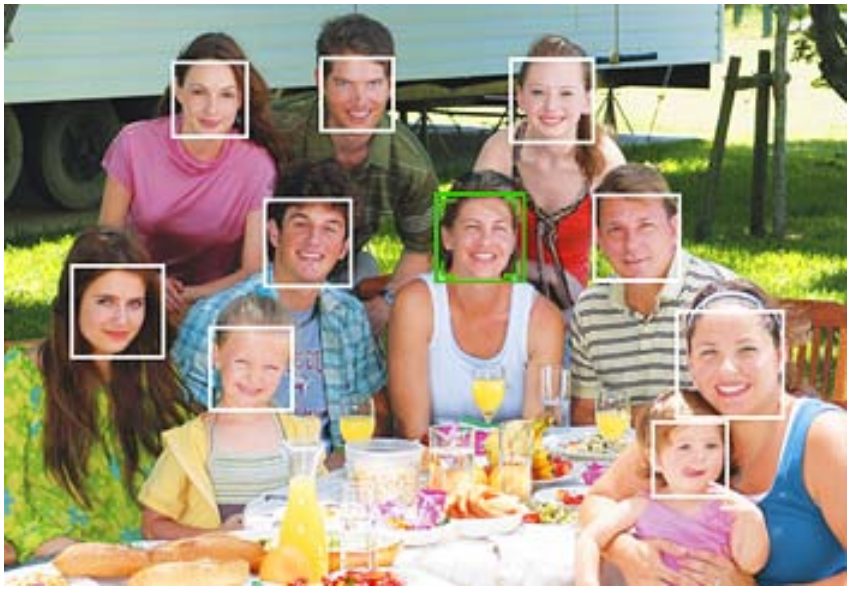
- Discriminating between two (or more) types of data
- Example: Spam filtering

Cures fast and effective! - Canadian *** Pharmacy #1 Internet Inline Drugstore Viagra Our price $1.15 Cialis Our price $1.99 …

Interested in your research on graphical models - Dear Prof. Ihler, I have read some of your papers on probabilistic graphical models. Because I …

Classification

- Example: face detection



Prediction

- Based on past history, predict future outcomes

Wall Street

Netflix

Data Mining & Understanding

- Massive volumes of data available
  - Webpages, Google books, …
  - Too large to hand-curate or organize

- How does Google decide the “most relevant” documents?
- How can we look for text documents “about” law, medicine, etc?
- What makes a document “similar”?

Tools for Machine Learning

- Probability and Statistics
  - Allows computing with / about uncertainty
  - Combine multiple sources of (uncertain) information
  - Search for “simple” explanations
- Also: optimization, information theory, …

- Classes
  - CS171, Intro to AI
  - CS178, Machine Learning
  - CS177, Probability in Computer Science
  - CS175, Projects in AI
  - CS179, Graphical models

Spam Filtering

- How can we tell spam from real email?

Cures fast and effective! - Canadian *** Pharmacy #1 Internet Inline Drugstore Viagra Our price $1.15 Cialis Our price $1.99 …

Interested in your research on graphical models - Dear Prof. Ihler, I have read some of your papers on probabilistic graphical models. …

Spam Filtering

- How can we tell spam from real email?

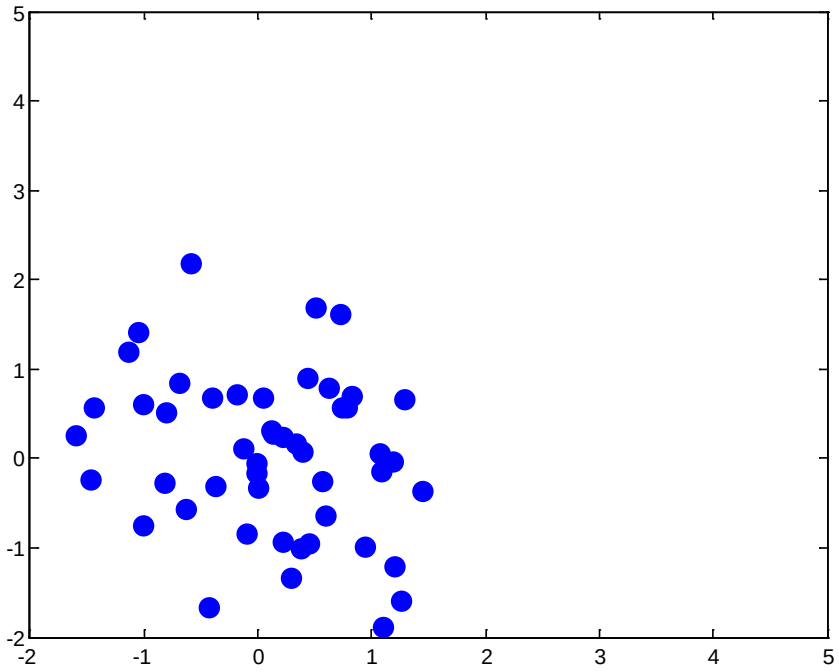
  Cures fast and effective! - Canadian *** Pharmacy #1 Internet Inline Drugstore Viagra Our price $1.15 Cialis Our price $1.99 …

  Interested in your research on graphical models - Dear Prof. Ihler, I have read some of your papers on probabilistic graphical models. …

- Word distributions
  - Spam: “winner”, “v1agra”, “$1,000,000”, …
  - Non-spam: “class”, “grade”, “research”, …
- Each word we see modifies our belief that the email is real
- Assign each word a weight
- Sum the weights of all the words that appear
- How can we pick the weights?
  - Use past data (labeled spam/real email) to estimate

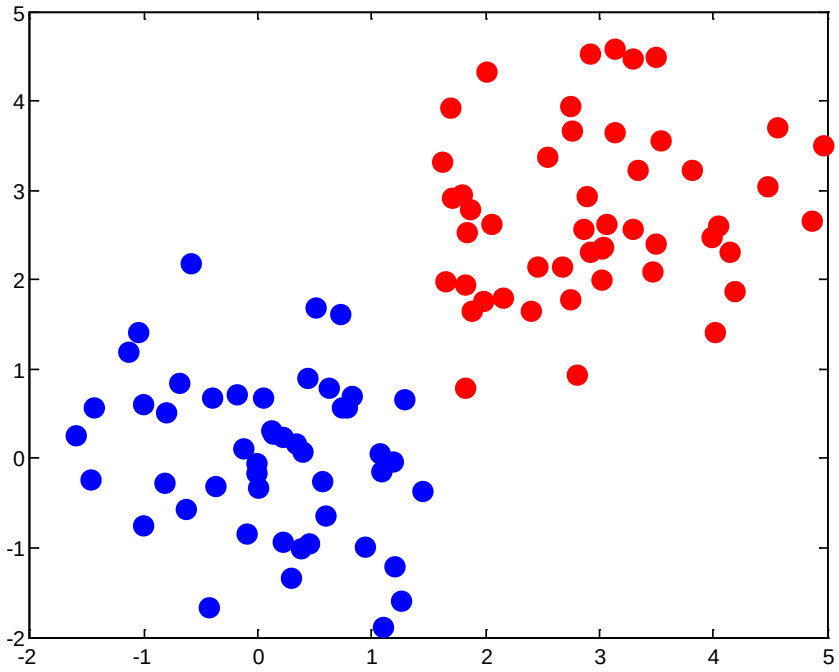
Visualization

- Feature space
  - Each point represents a value of all features



Visualization

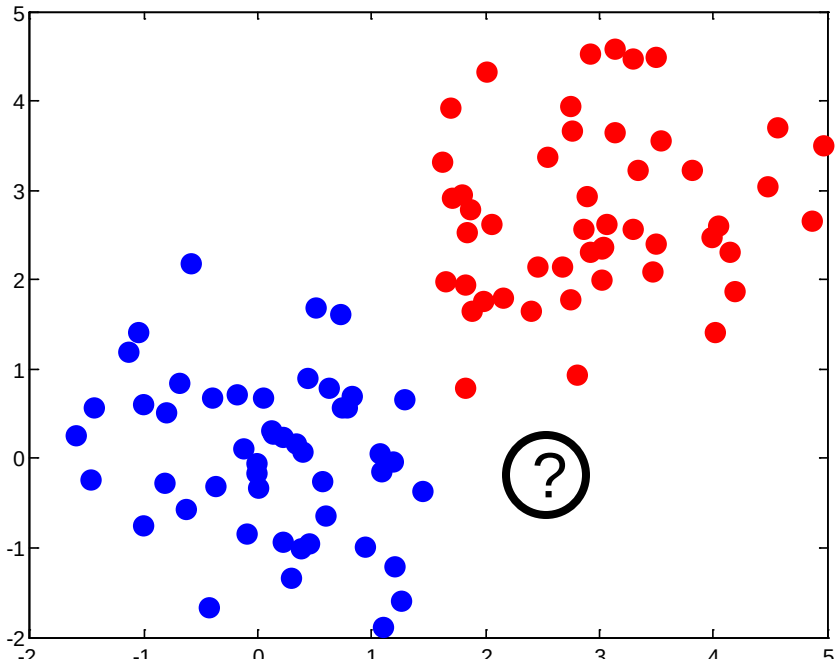
- Feature space
  - Each point represents a value of all features



- Multiple classes = multiple scatterplots

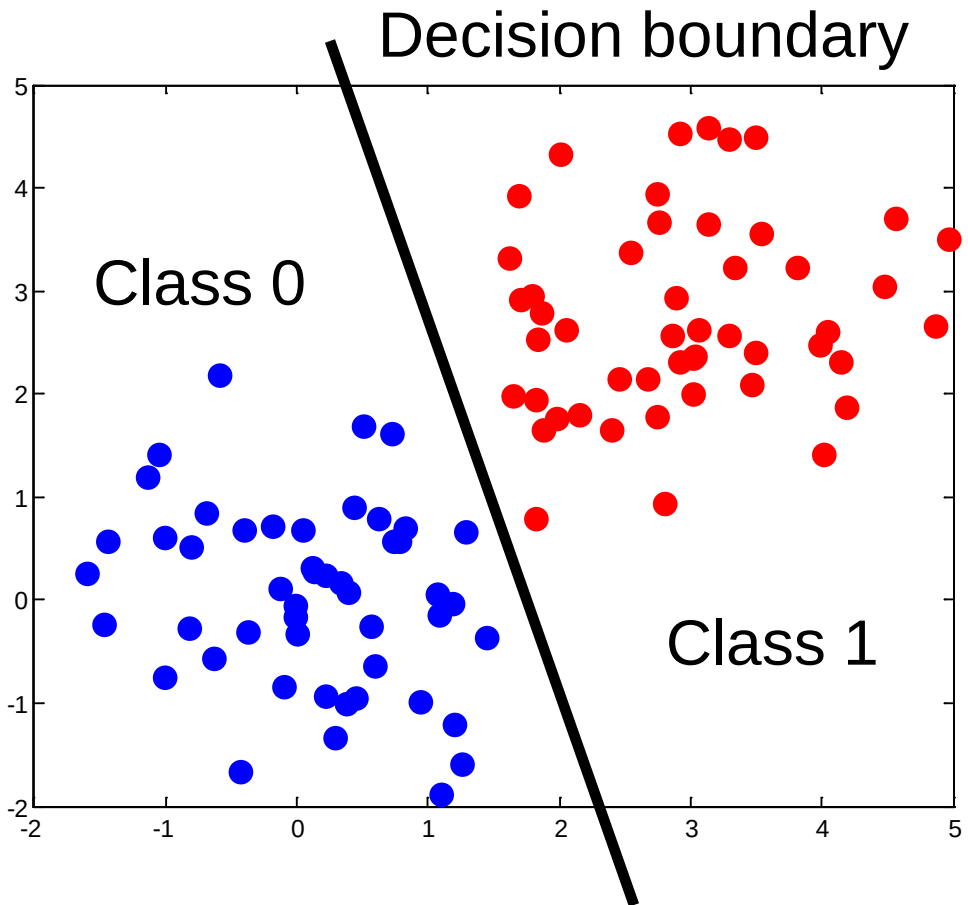
Classifiers

- Classifier: function from x to c
  - Each value of x maps to some class c
  - Partitions the feature space into sets
- Decision boundary
  - The boundary between two sets with different class estimates
  - Can describe function in terms of boundary



Linear decision boundaries

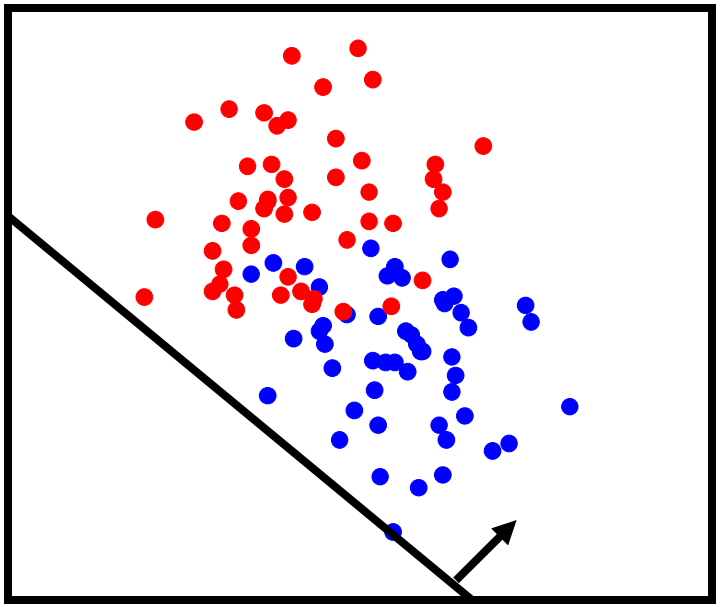
- Linear decision boundaries
  - Simplest form of decision boundary

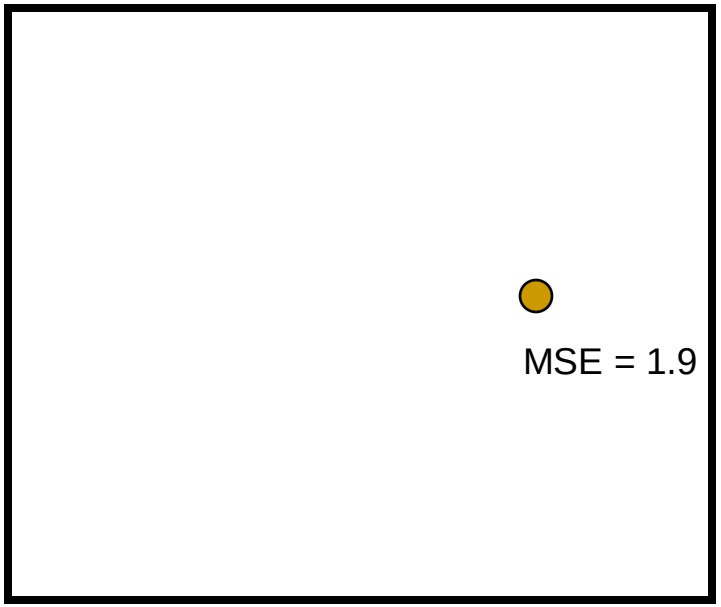


Training the Classifier

- Once we have a smooth measure of quality, we can find the “best” settings for the parameters of
  f(X1,X2) = a*X1 + b*X2 + c

- Example: 2D feature space ⇔ parameter space

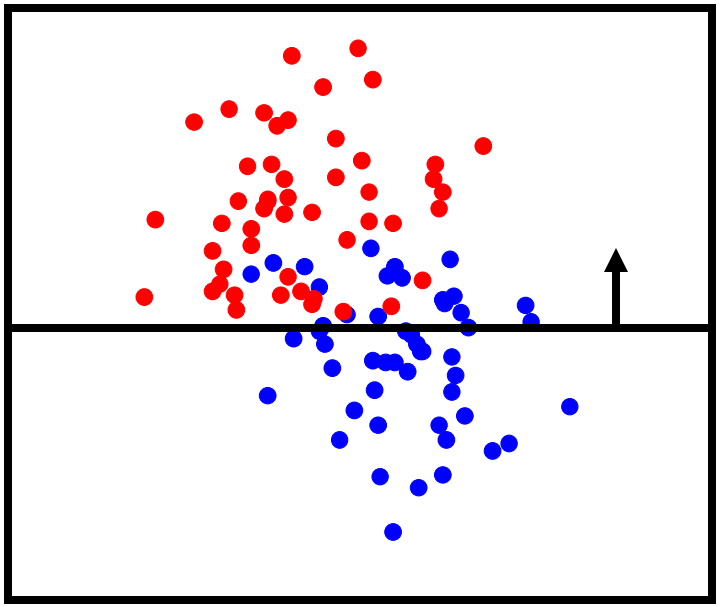


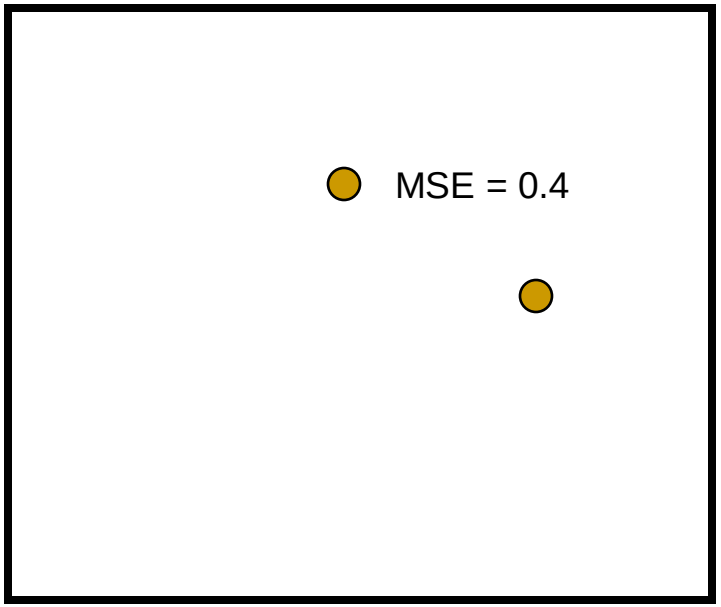


Training the Classifier

- Once we have a smooth measure of quality, we can find the “best” settings for the parameters of
  f(X1,X2) = a*X1 + b*X2 + c

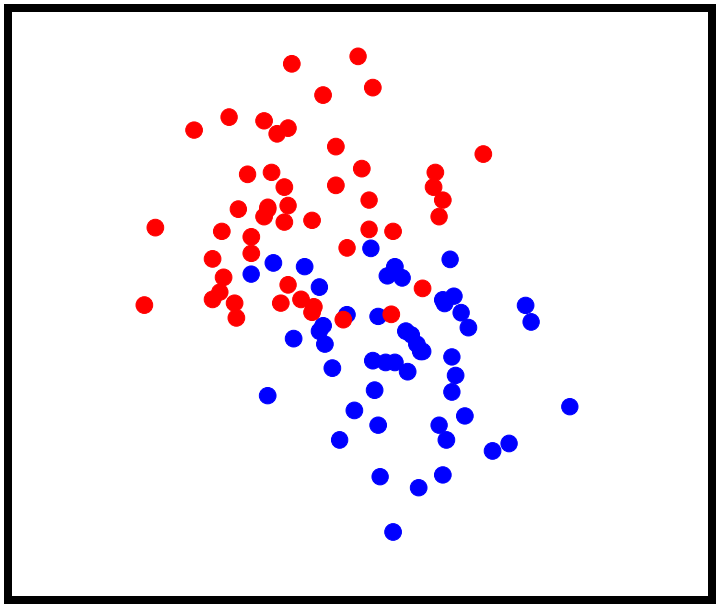
- Example: 2D feature space ⇔ parameter space

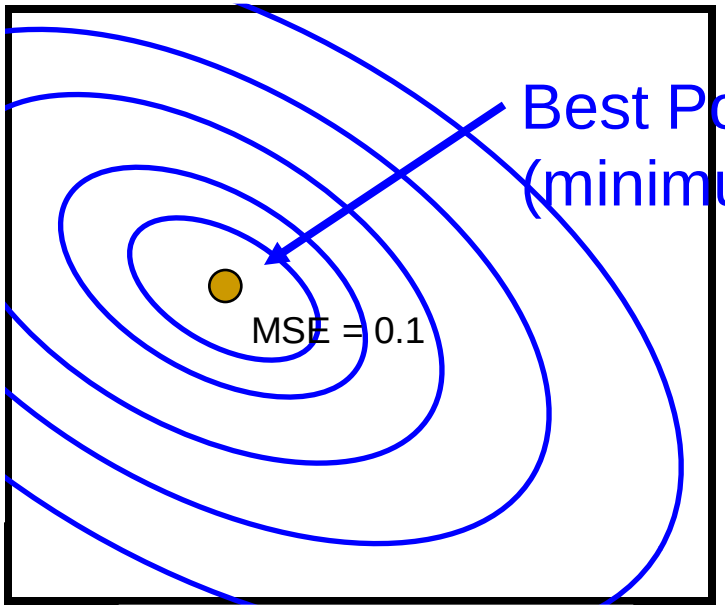




Training the Classifier

- Once we have a smooth measure of quality, we can find the "best" settings for the parameters of
  f(X1,X2) = a*X1 + b*X2 + c
- Finding the minimum MSE in parameter space…



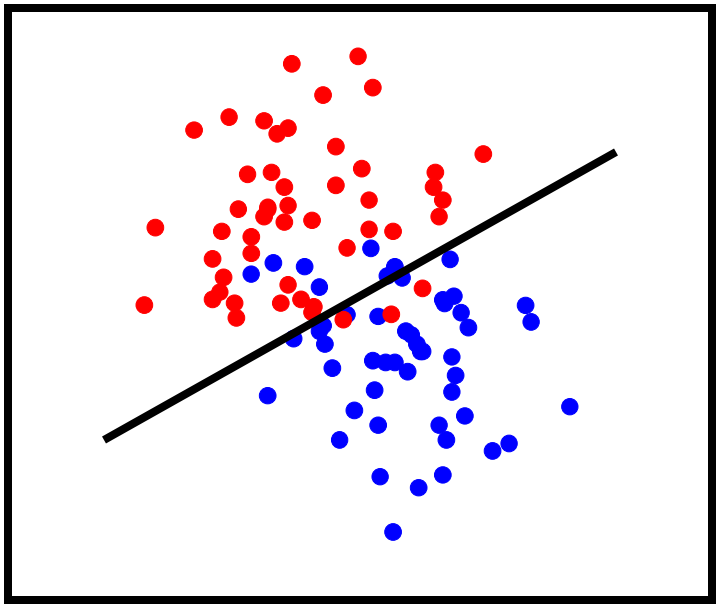


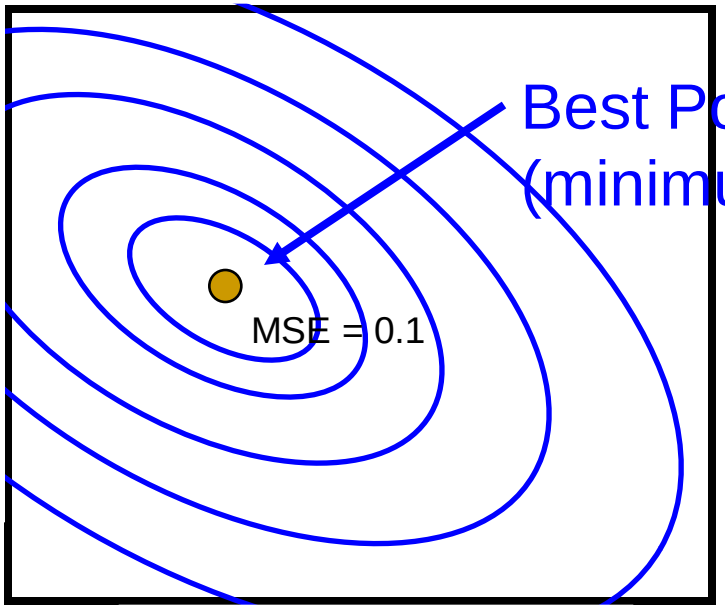
- [a b c] = ?

[arctan(A/B), c] = [-pi/4, 1]

Training the Classifier

- Once we have a smooth measure of quality, we can find the "best" settings for the parameters of
  f(X1,X2) = a*X1 + b*X2 + c
- Finding the minimum MSE in parameter space…





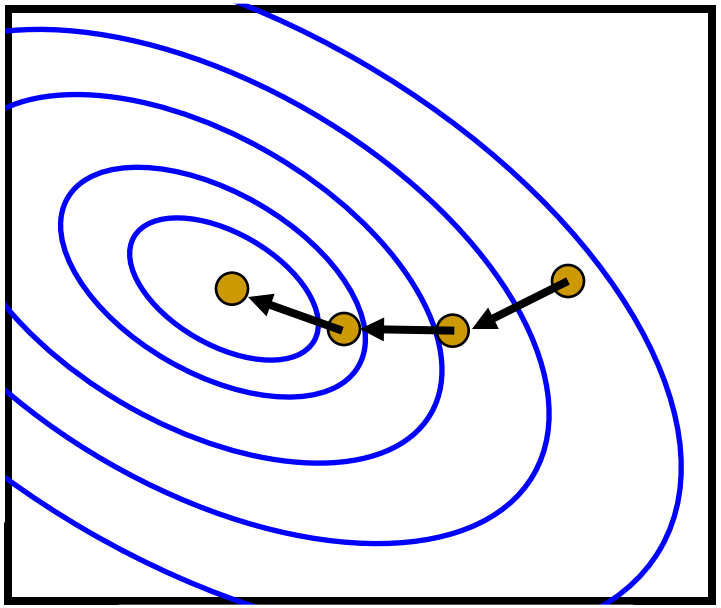
- [a b c] = ?

[arctan(A/B), c] = [-pi/4, 1]

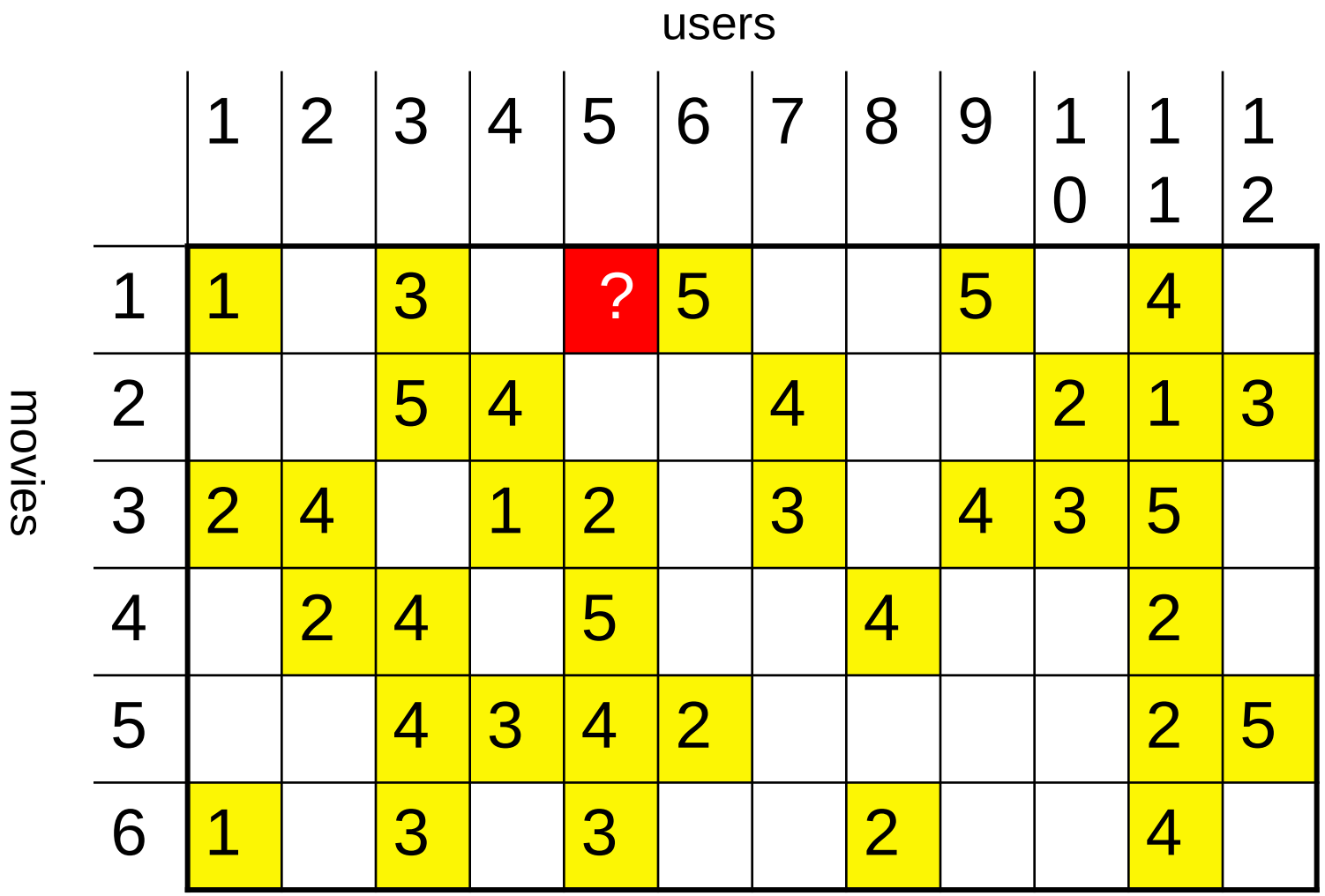
Finding the Best MSE

- As in linear regression, this is now just optimization
- Methods:
  - Gradient descent
    - Improve MSE by small changes in parameters ("small" = learning rate)
  - Or, substitute your favorite optimization algorithm…
    - Coordinate descent
    - Stochastic search
    - Genetic algorithms

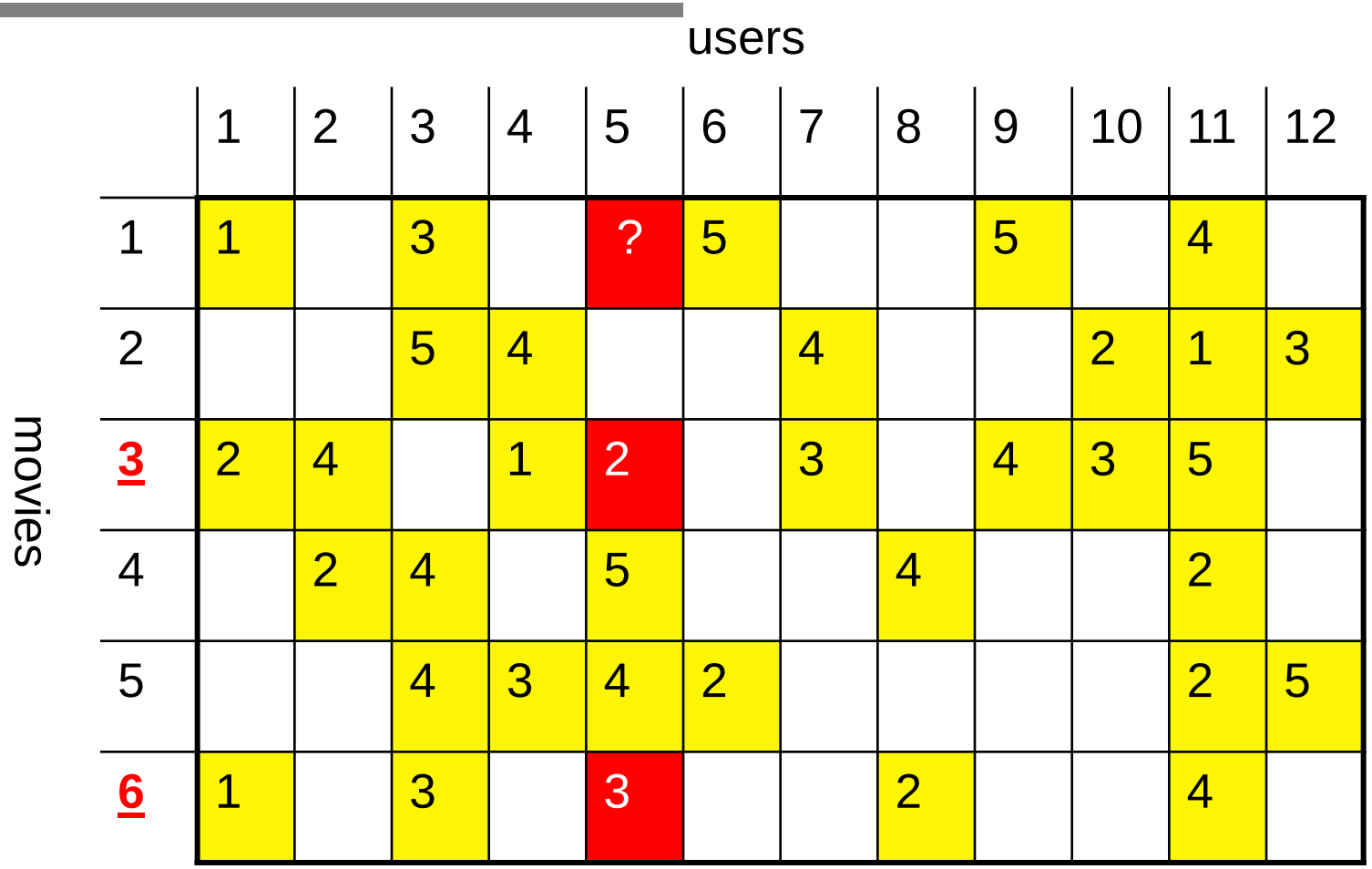
Gradient Descent



Collaborative Filtering (Netflix)



Nearest-Neighbor methods



Neighbor selection:
Identify movies similar to 1, rated by user 5

Nearest-Neighbor methods

users

movies	1	2	3	4	5	6	7	8	9	10	11	12
1	1		3		?	5			5		4	
2			5	4			4			2	1	3
3	2	4		1	2		3		4	3	5	
4		2	4		5			4			2	
5			4	3	4	2					2	5
6	1		3		3			2			4	

Compute similarity weights:

$s_{13}=0.2$, $s_{16}=0.3$

Nearest-Neighbor methods

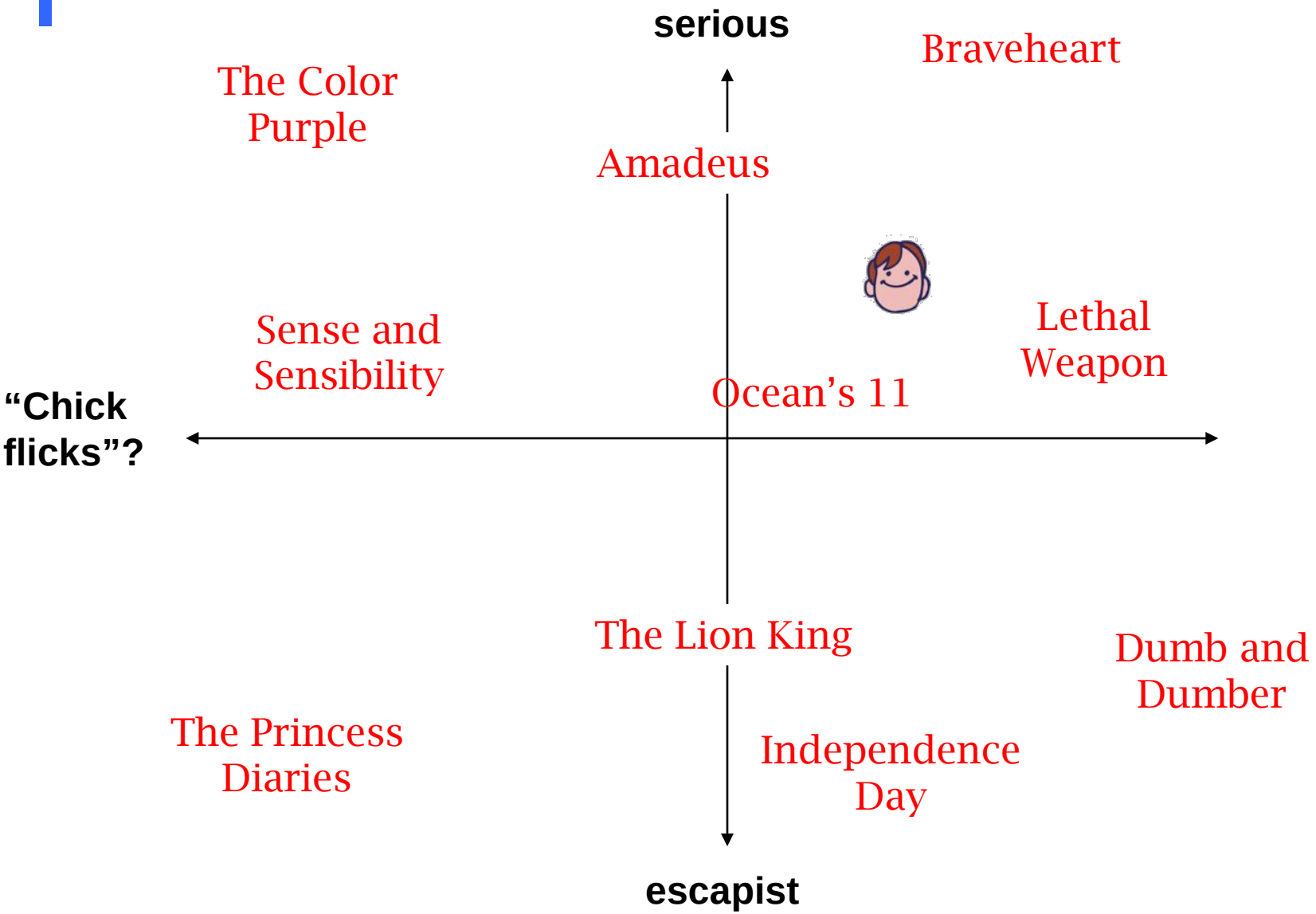
users

movies	1	2	3	4	5	6	7	8	9	10	11	12
1	1		3		**2.6**	5			5		4	
2			5	4			4			2	1	3
3	2	4		1	2		3		4	3	5	
4		2	4		5			4			2	
5			4	3	4	2					2	5
6	1		3		3			2			4	

Predict by taking weighted average:

(0.2*2+0.3*3)/(0.2+0.3)=2.6

Latent Space Models



Latent Space Models

Model ratings matrix as “user” and “movie” positions

Infer values from known ratings

Extrapolate to unranked

users

items

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

items

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

●

users

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

Graphical models

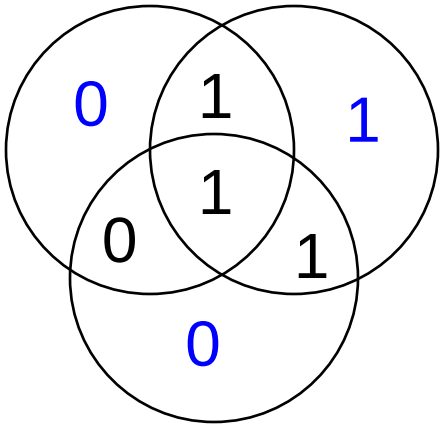
- Constraint networks
  - Satisfaction problems, coding, …
- Probabilistic networks
  - Represent & reason about uncertainty
  - Learning from data

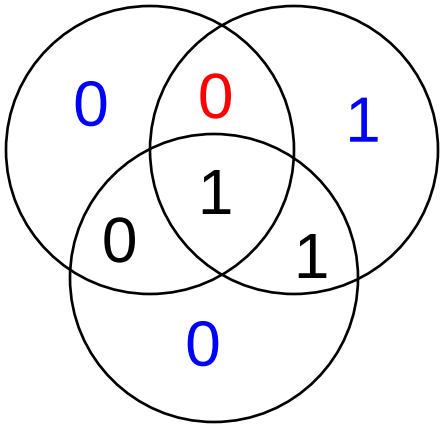
Coding

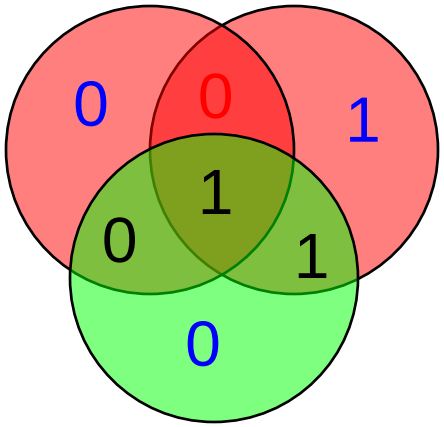
- One more example…
- Suppose we want to correct for missing or erroneous data

  Send: 1101 …. Receive 0101

- We can add extra “check bits” to help
  - Called a Hamming code
  - Works very well for small codes

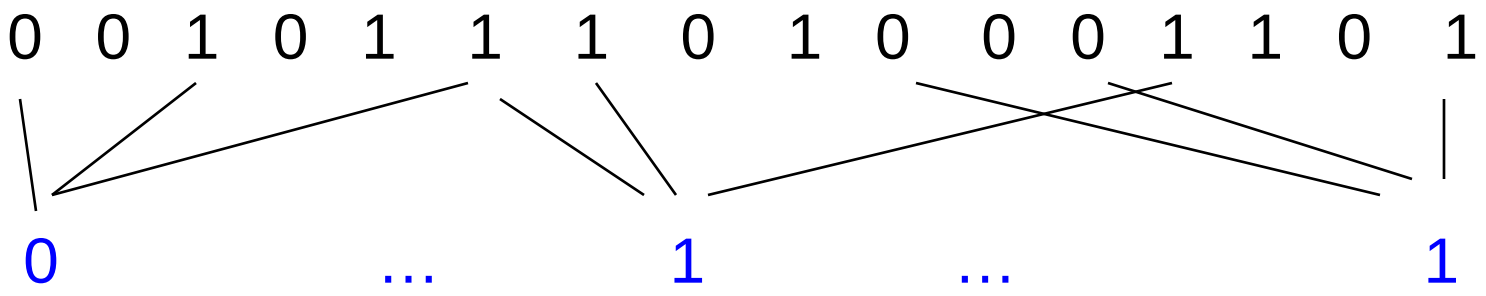






Coding

- Large codes: “Low density parity check”
- Similar principle, but just use random check bits



- This is a graphical model
  - Decoding: try to find a few changes that satisfy all constraints
  - Uses “belief propagation”, an algorithm from machine learning
- LDPC codes are used in satellite communications, “raptor codes” (used by torrents), and many more…

LDPC demo

- Data bits (top)
- Check bits (bottom)
  - 50% overhead

- Add noise (~7.5% errors)
- Correct with iterative algorithm

- Example from David Mackay’s website

Conclusions

- Artificial Intelligence
  - Search, planning, making decisions
- Machine learning
  - Organizing and understanding observed data
  - Finding simple representations
  - Making predictions in complex systems
  - Approximate computations for difficult problems
- Related disciplines
  - Computer vision
  - Computational biology

- Lots of great classes: 171, 177, 178, 175, …
- Lots of great faculty here at UCI