Computer Vision & Cognitive Systems

SCQ5109806 - LM CS,DS,CYB,PD

Machine Learning Basics (I)

Prof. Lamberto Ballan

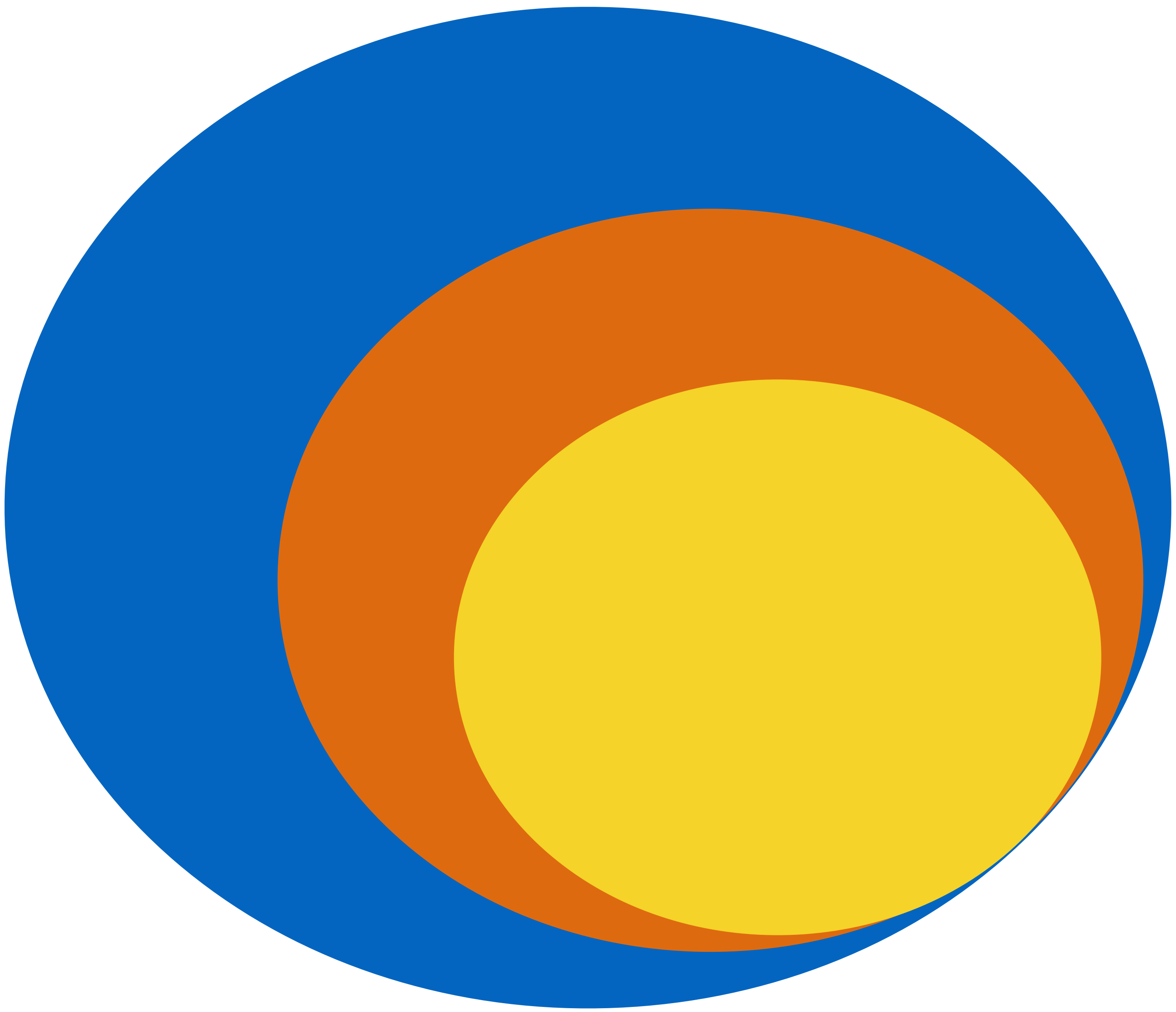
DIPARTIMENTO MATEMATICA
Department of Mathematics "Tullio Levi-Civita"

Course material & how to contact us

- Moodle: https://stem.elearning.unipd.it/course/view.php?id=13820
  - Updated on a weekly basis: slides, references, etc
  - Use this for most communication with course staff
  - Ask questions about logistics, assignments, etc
  - Submit project report

Moodle enrolment key: vls10N-25

AI, Machine and Deep Learning



Artificial Intelligence

The science to make things smart

Machine Learning

Building machines that can learn

Deep Learning

A class of ML algorithms

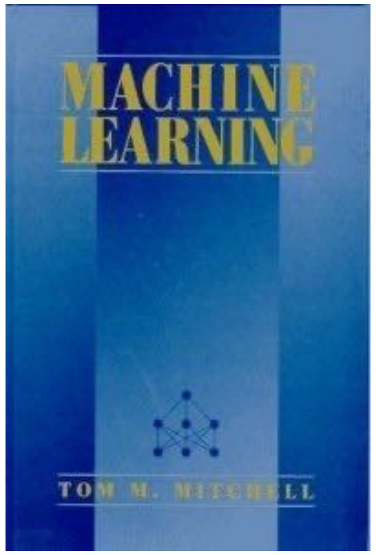
What is Machine Learning?

- "*ML is the field of study that gives computers the ability to learn without being explicitly programmed.*"
  - This is an older (informal) definition by Arthur Lee Samuel who popularised the term in 1959

A.L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers", IBM Journal of Research & Dev. 1959

What is Machine Learning?

- "*A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.*"
    - Tom Mitchell provides a more modern definition (1998)
    - Example: playing checkers
        - Experience *E* = the experience of playing many games of checkers
        - Task *T* = the task of playing checkers
        - Performance measure *P* = percent of games won against opponents

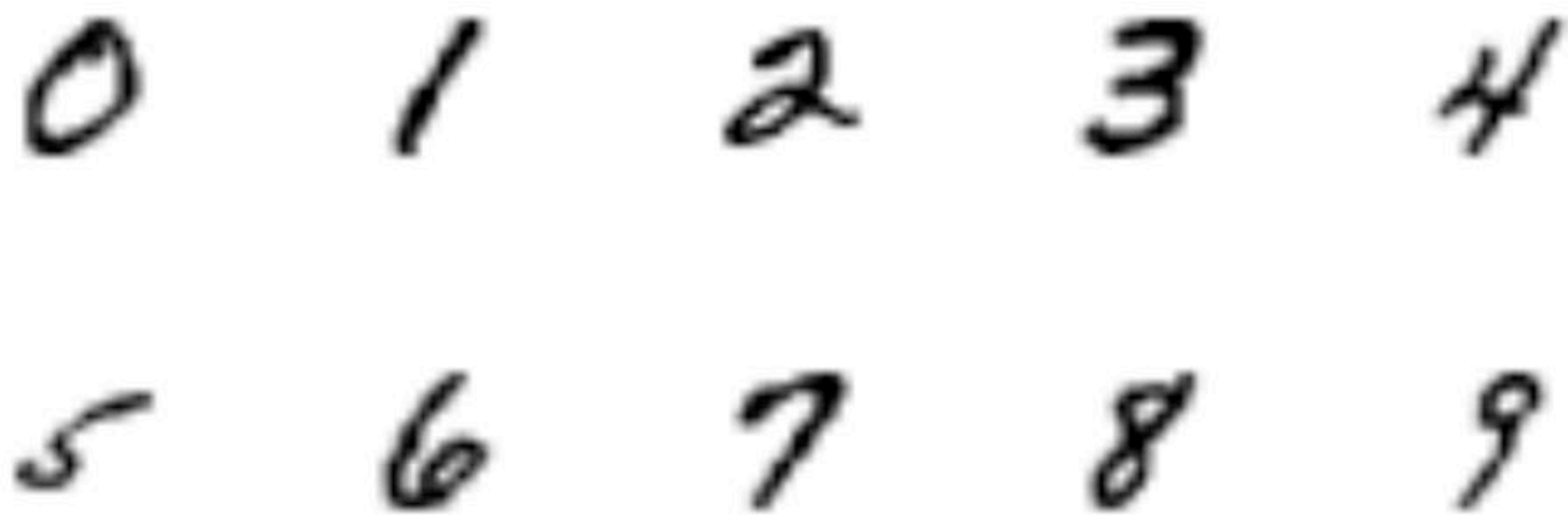


When to use Machine Learning?

- First, why not to use a traditional algorithmic approach?
  - Impossible to **exactly formalise** the problem *(and so to give an algorithmic solution)*
  - Presence of **noise** and/or **uncertainty**
  - **High complexity** in formulating a solution, i.e. it cannot be done manually
  - Lack of **compiled knowledge** with respect to the problem to be solved

When to use Machine Learning?

- Example: handwritten digit recognition



  - Very hard to exactly formalize the problem
  - Noise may be present and data may be ambiguous

When to use Machine Learning?

- **Learning algorithm**: an algorithm that is able to learn from data
- Three main ingredients:
  - The Task
  - The Performance Measure
  - The Experience

The Task

- A task is described in terms of how the machine learning algorithm should process an *example*



- How is an example represented?
    - As a collection of features *(that can be "measured")*

Features

1. Color: Red
2. Type: Fruit
3. Shape: Round
4. …

The Task

- The task is defined by the problem we want to tackle and the desired output
- Example: classification

Apple

Orange

Mango

The Performance Measure

- How good is the machine learning system?
  - We need to measure its performance, i.e. how good is the function/model returned by the learning algorithm
- The performance measures depends on the task
  - Example (classification): *accuracy* is the proportion of examples for which the model gives the correct output

O:	mango	apple	orange	mango	orange	apple	orange	orange	apple	orange
	1	1	1	1	1	0	0	1	1	0

accuracy = 70%

The Experience

- The experience is provided by the available data
- Which kind of data?
  - Real-valued features, discrete features, …
- How do we get data
  - Obtained once for all (batch), acquired incrementally by interacting with the environment (online learning)
- How can data be used?
  - Learning paradigms

Main Learning Paradigms

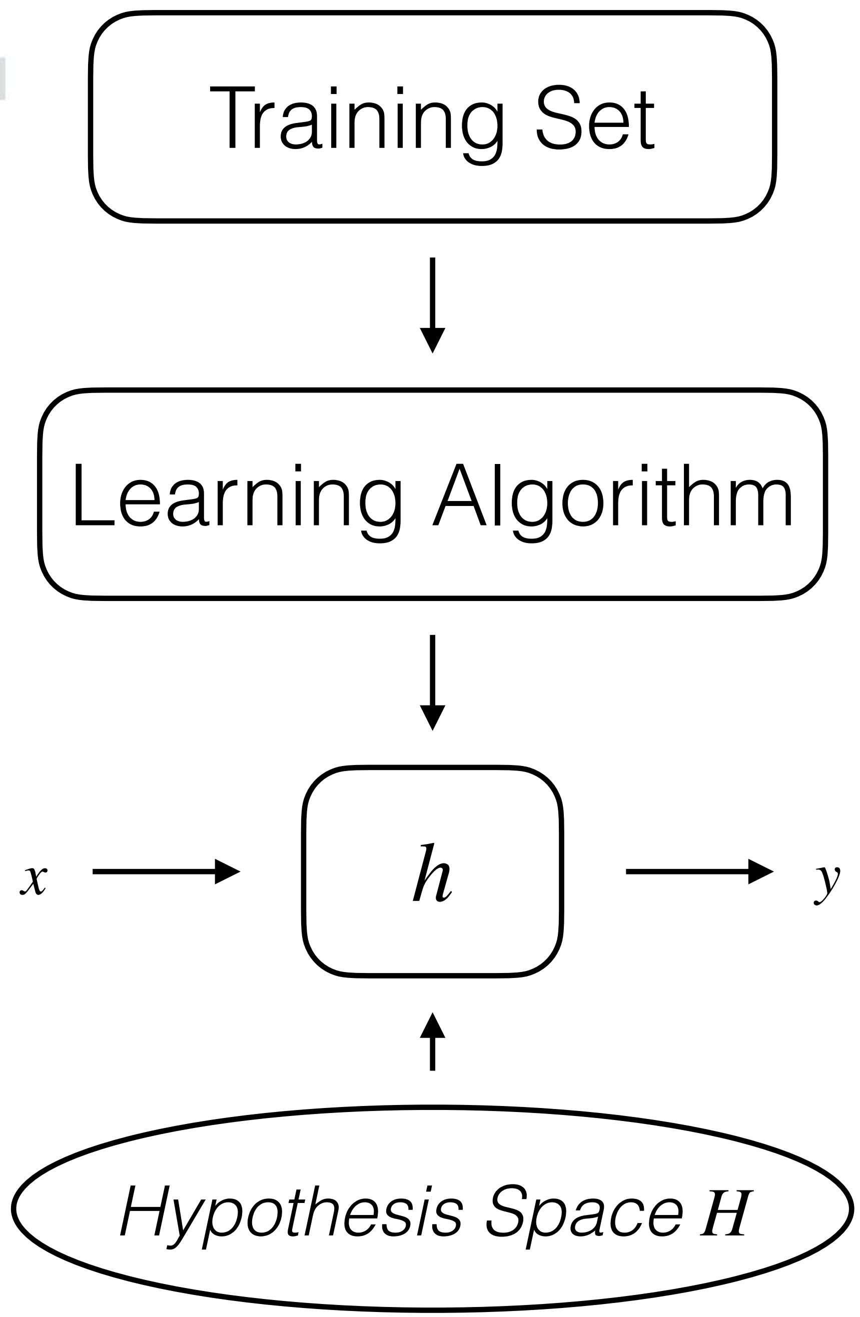
- **Supervised Learning**
  - **Goal:** give the "right answer" for each example in the data
  - Given examples $\{(x^{(i)}, y^{(i)})\}$, learn a function (description) which captures the information content of the examples
  - Basically we look for a function $h(\cdot)$ which is able to map in a predictive way $x^{(i)}$'s to $y^{(i)}$'s, *i.e* $h: X \rightarrow Y$
  - An expert (or teacher) provides the supervision *(i.e. the values of $h(\cdot)$ corresponding to the instances $x^{(i)}$)*
  - **Output:** Classification (discrete-valued) *vs* Regression (real-valued output)

Main Learning Paradigms

- **Supervised Learning**

$\{(x^{(i)}, y^{(i)})\}$

Size in feet² (x)	Price ($) in 1K's (y)
2104	460
1416	232
1534	315
852	178
…	…



h approximates the unknown target f

$$h \sim f \colon X \rightarrow Y$$

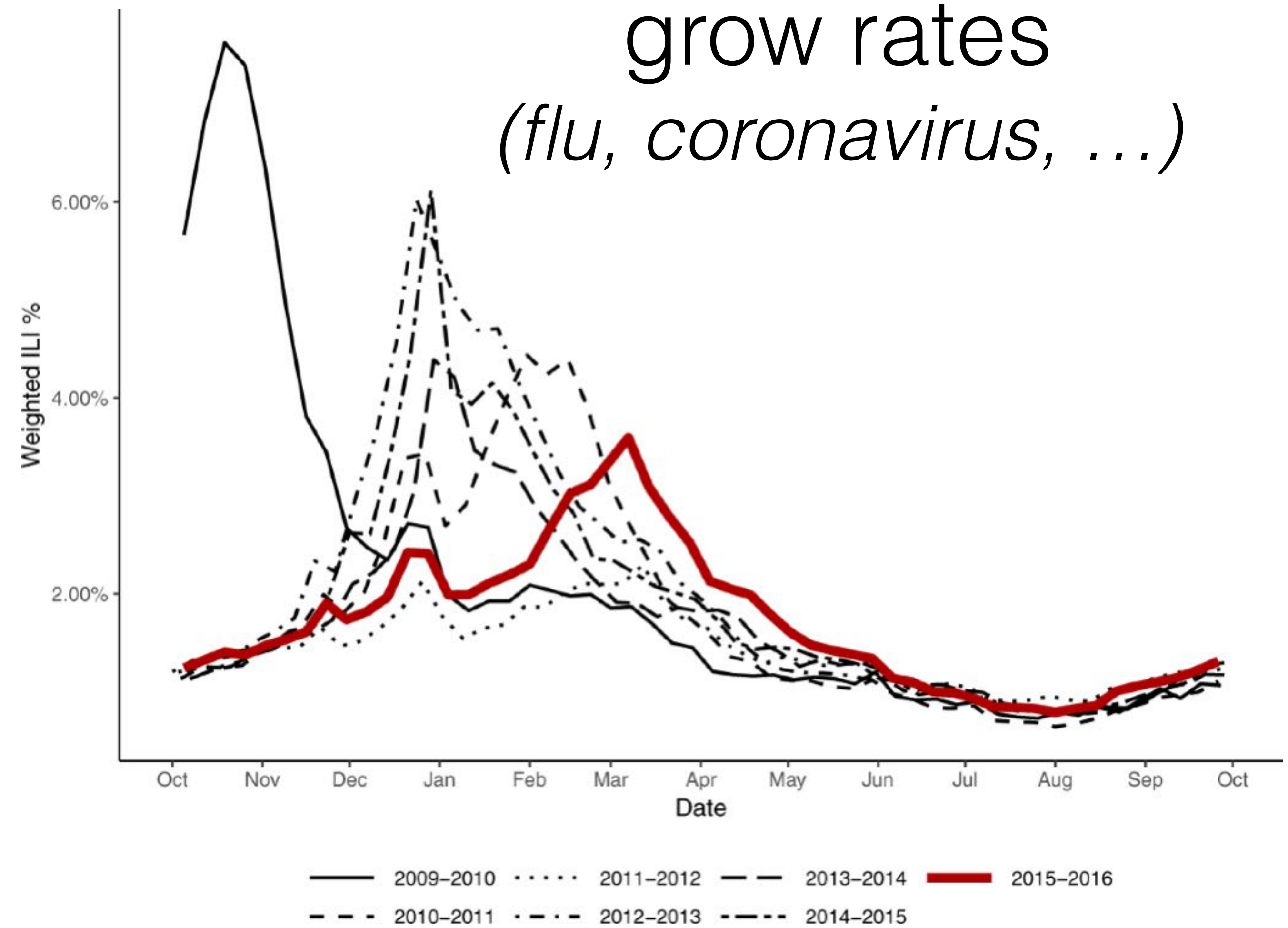
Main Learning Paradigms

- **Supervised Learning:** a few examples

Object Recognition

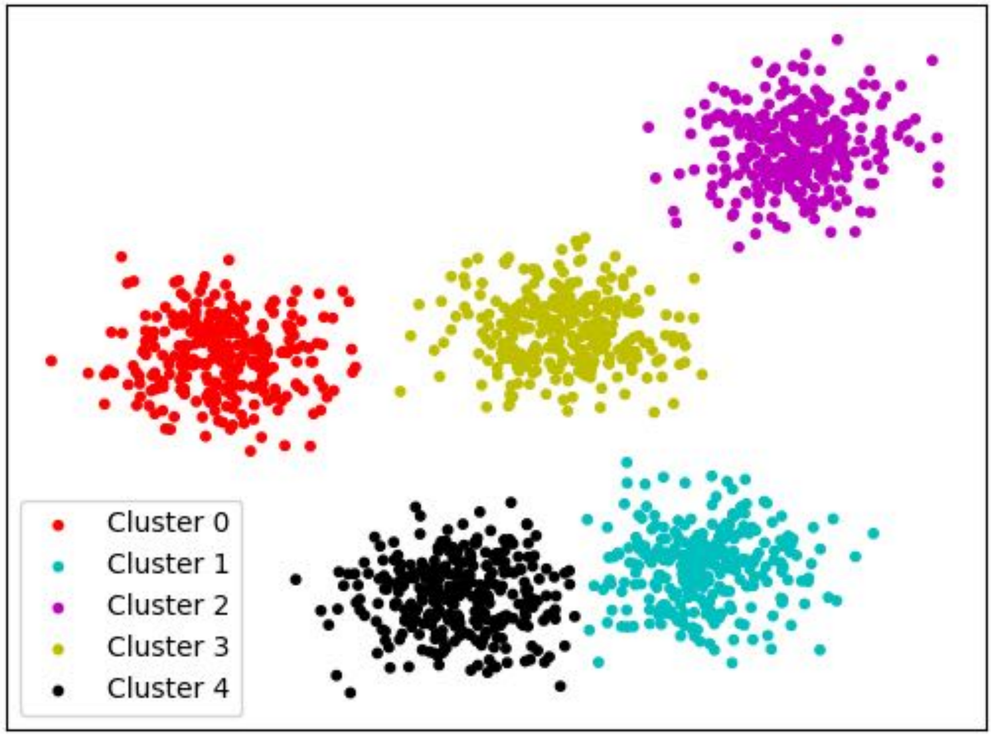
Predicting pandemic grow rates

(flu, coronavirus, …)



Main Learning Paradigms

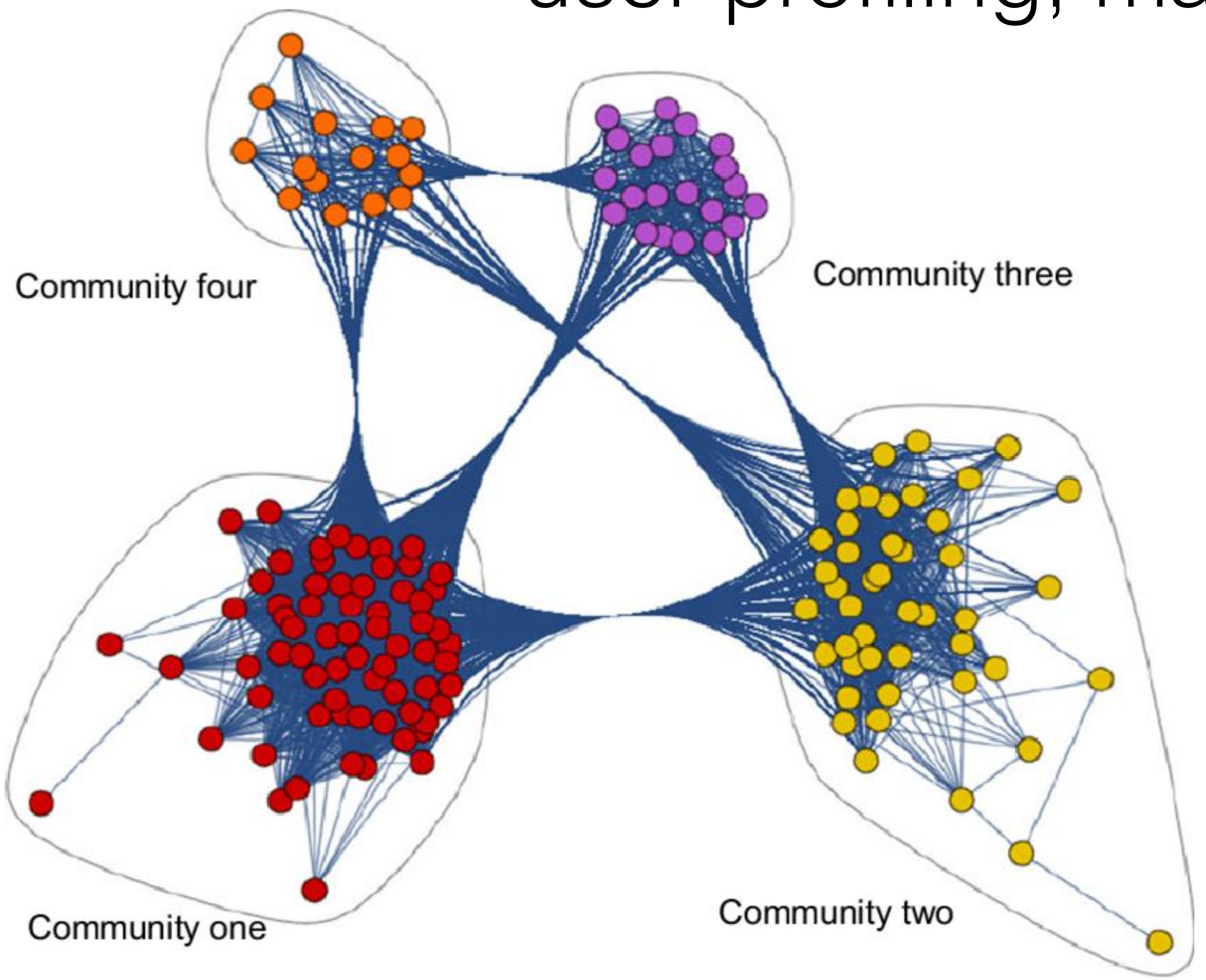
- **Unsupervised Learning**
  - **Goal:** find regularities / patterns on the data
  - Given examples $\{x^{(i)}\}$, discover regularities on the whole input domain
  - There is no expert (*i.e.* no supervision)

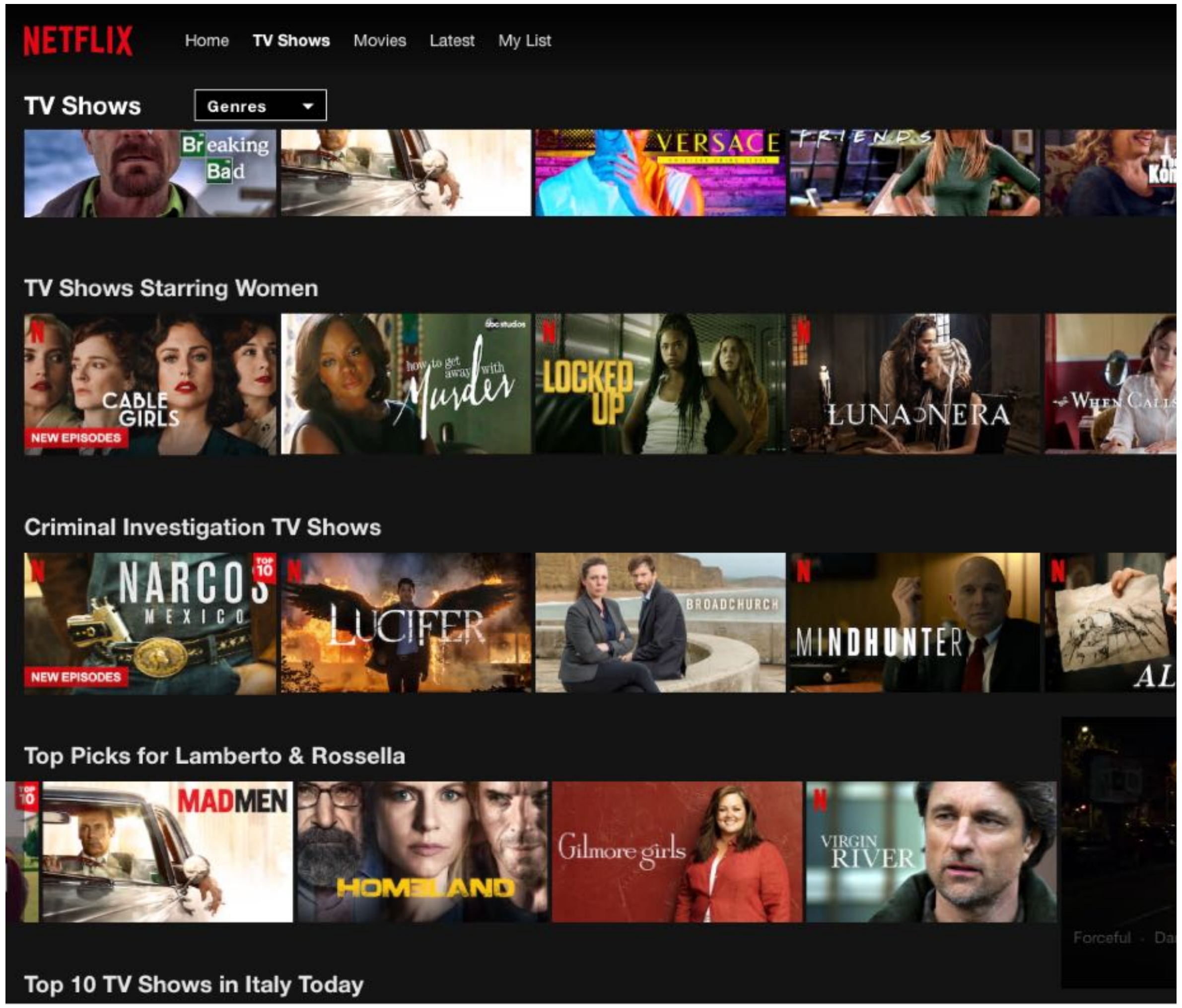


Main Learning Paradigms

- **Unsupervised Learning:** a few examples

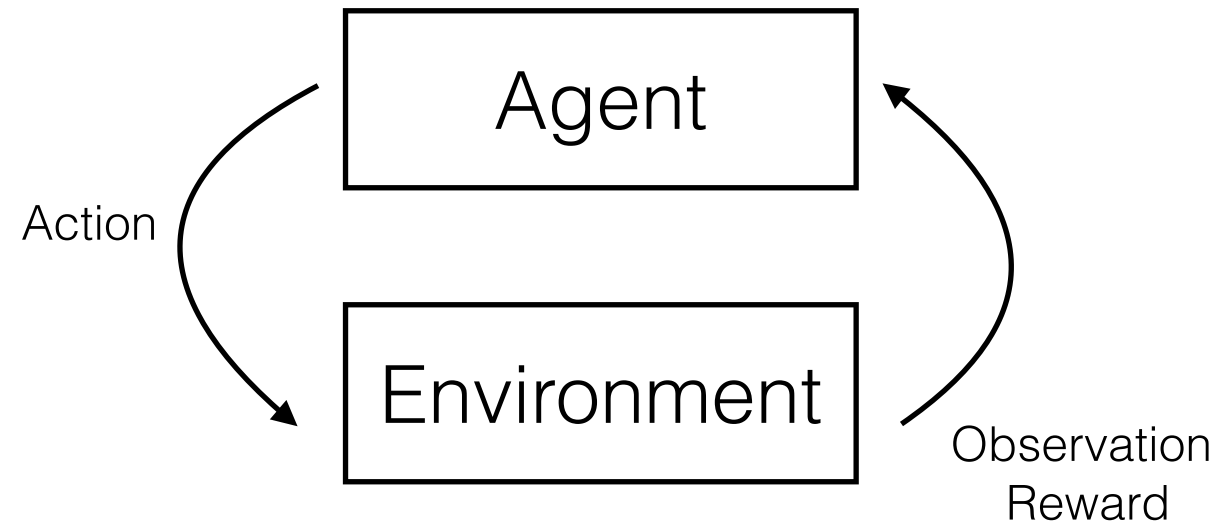
Community detection in social media,
user profiling, market analysis …





Main Learning Paradigms

- **Reinforcement Learning**
  - Agent which may
    - be in state s
    - execute action a *(among the ones admissible in state s)*



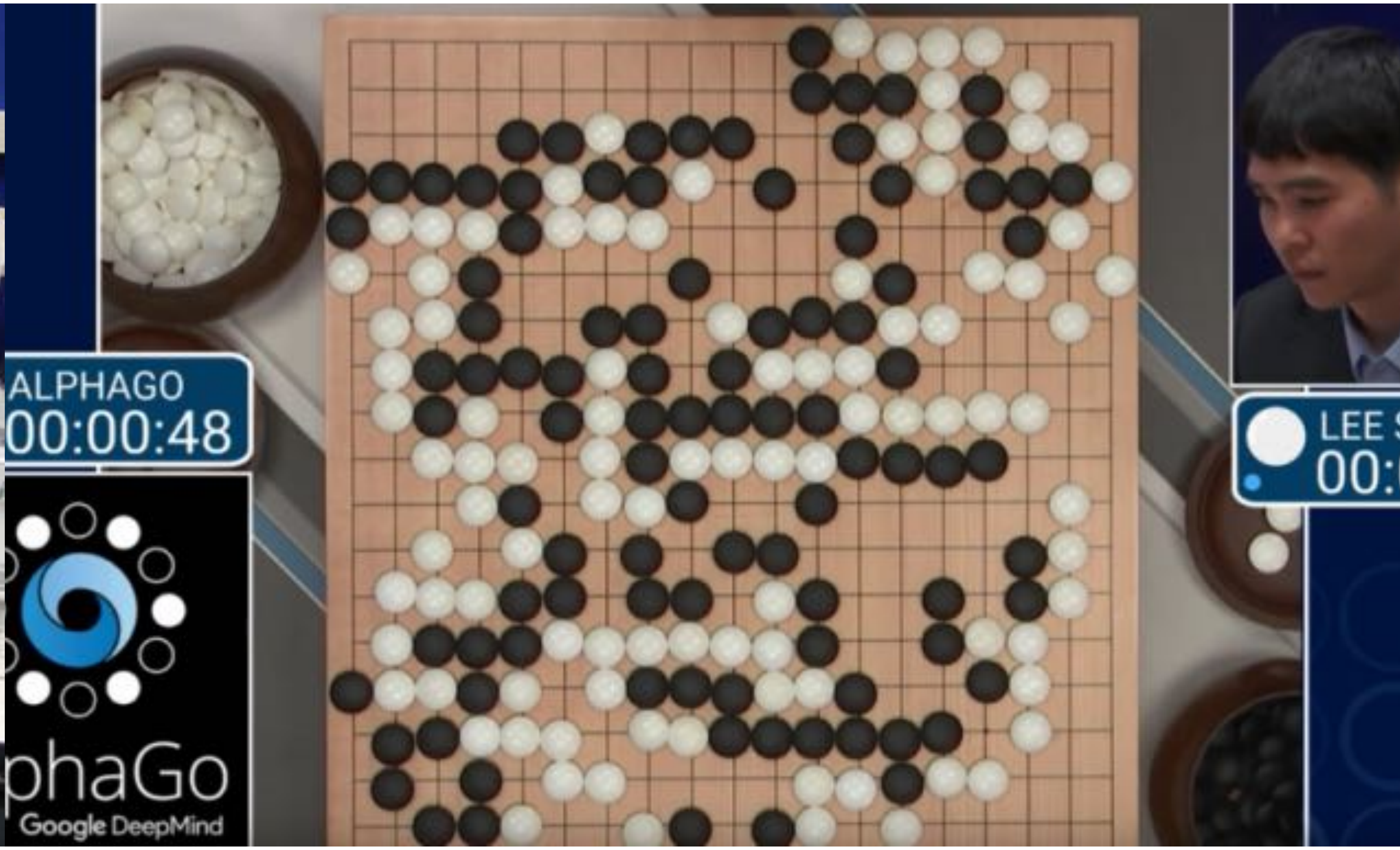
  - It operates in an environment e, which in response to action a in the state s returns
    - the next state and a reward *r* (which can be positive, negative or neutral)
  - The goal of the agent is to maximize a reward function

Main Learning Paradigms

- **Reinforcement Learning:** a few examples

Robotics

Games



Other Learning Strategies

- Active Learning
- Online Learning, Incremental & Continual Learning
- Weak-supervised Learning
- Self-supervised Learning
- Deep Learning and Representation Learning
- Federated Learning

Fundamental Ingredients

- Training data D (drawn from the instance space X)
- Hypothesis space H
  - i.e. the set of functions which can be implemented by the machine learning system
  - we assume that the function to be learned f may be represented/approximated by the hypothesis $h \in H$
- Learning algorithm
  - It can be seen as a search algorithm into H

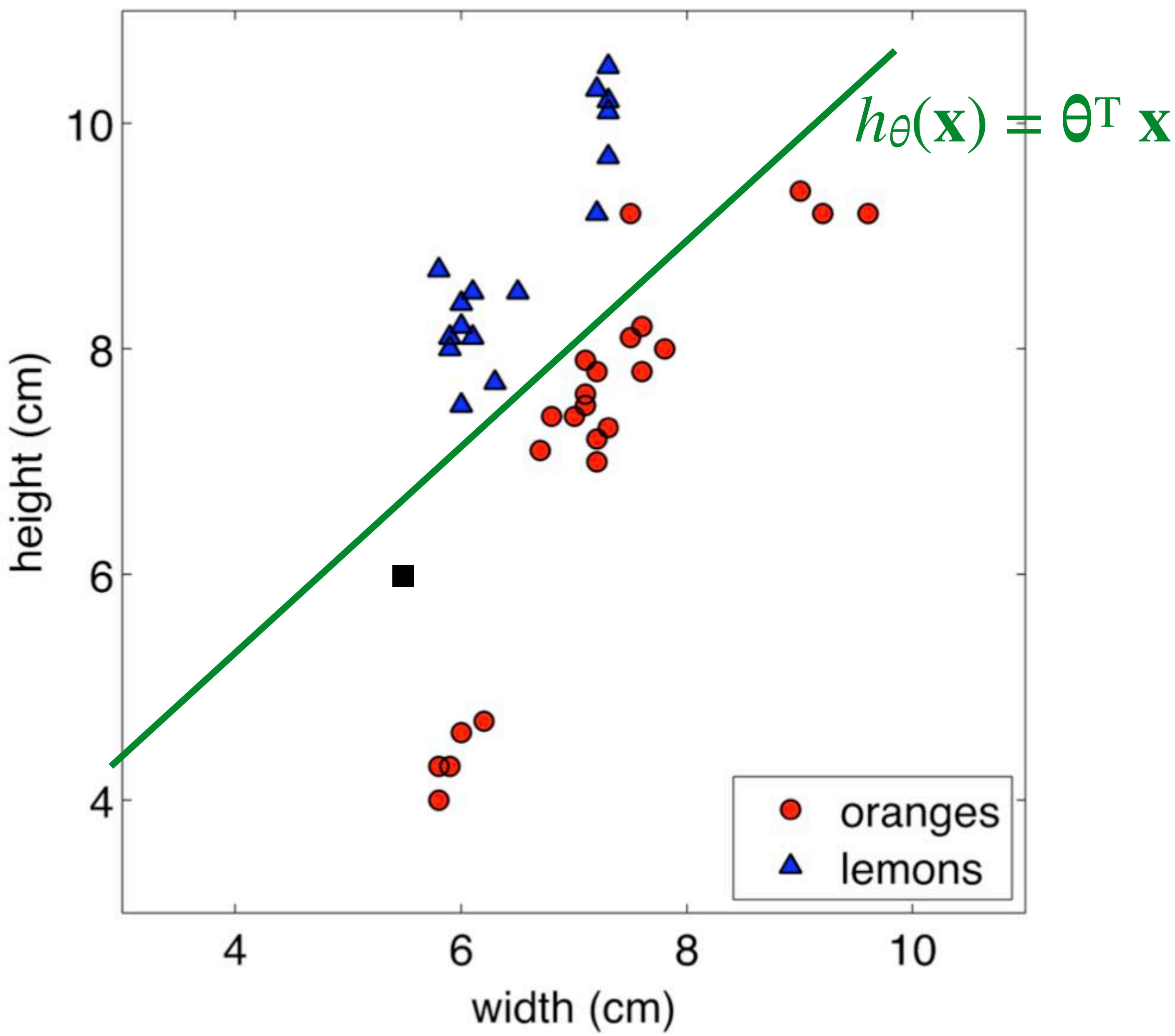
Inductive (or learning) Bias: on the representation (H) and/or on the search (learning algorithm)

Inductive Bias

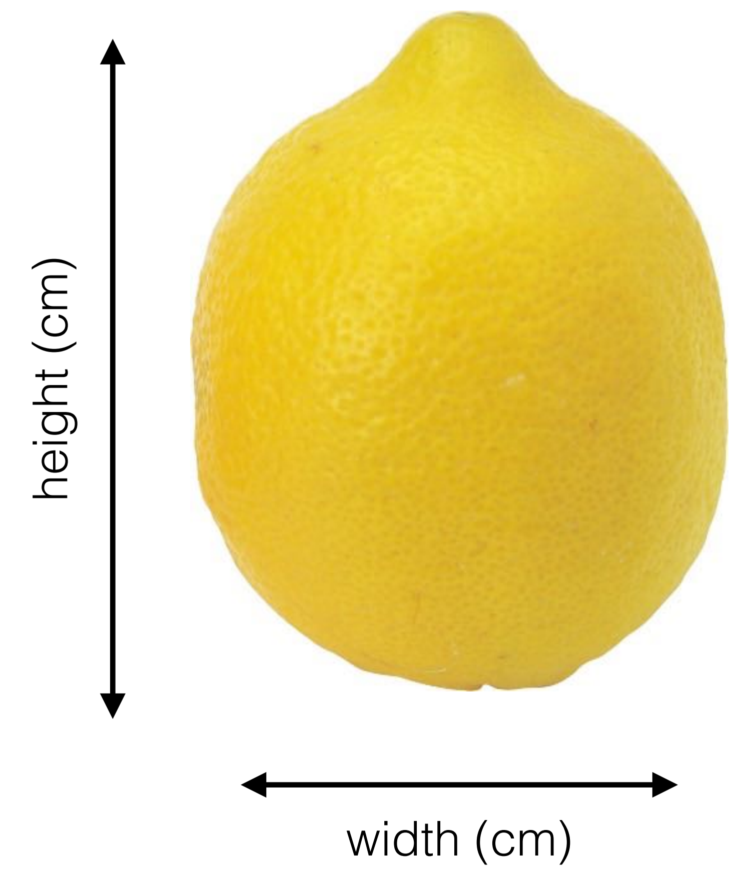
- Experience alone might not allow us to make conclusions about unseen data instances
- **Inductive Bias** = all the assumptions about the “nature” of the target function and its selection
- Two type of bias:
  - Restriction: limit the hypothesis space
  - Preference: impose ordering on hypothesis space

Example of Inductive Bias

- 1. Linear regression/classification, 2. k-NN

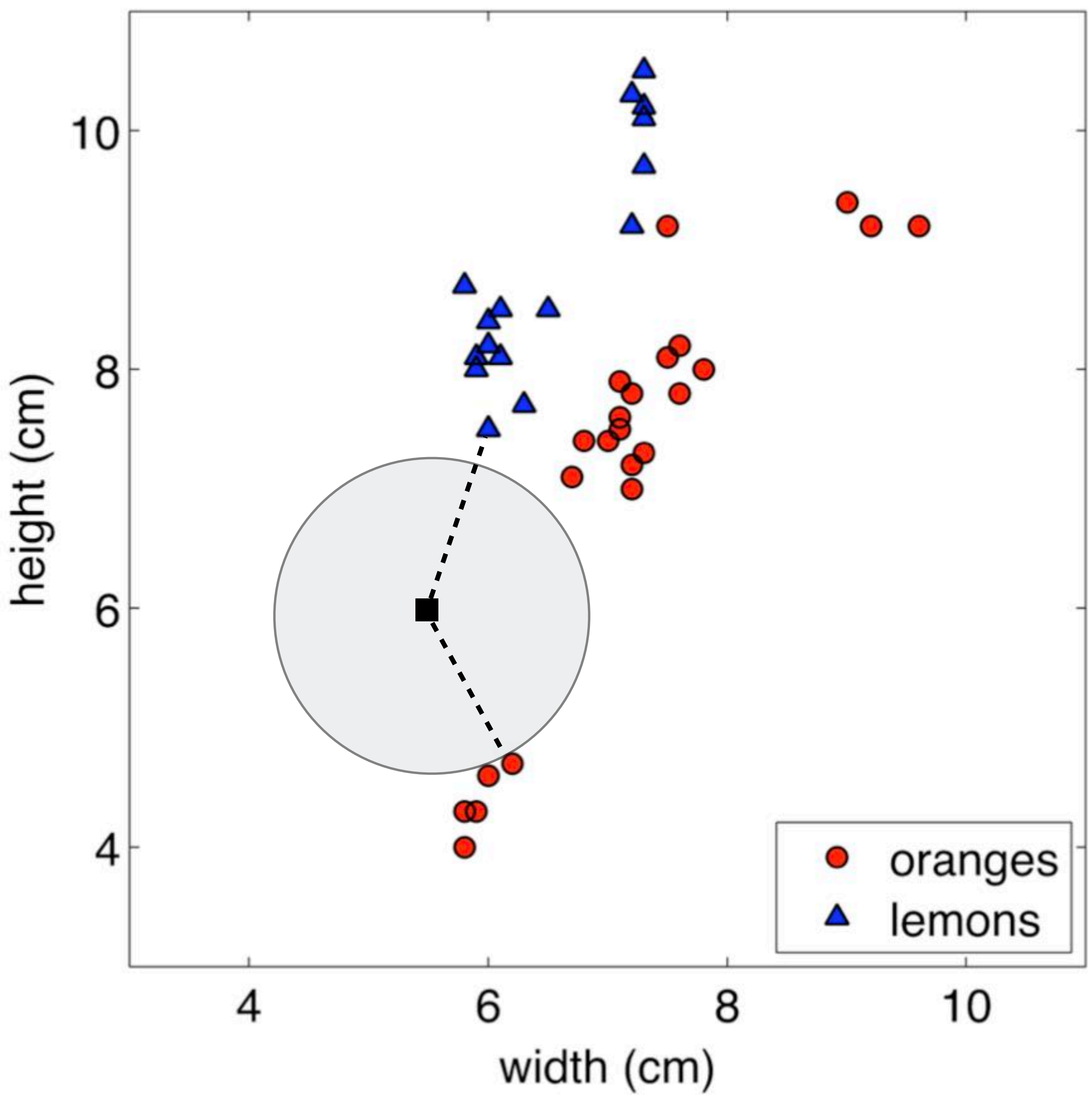


Binary classification problem
("lemon" *vs* "non-lemon")

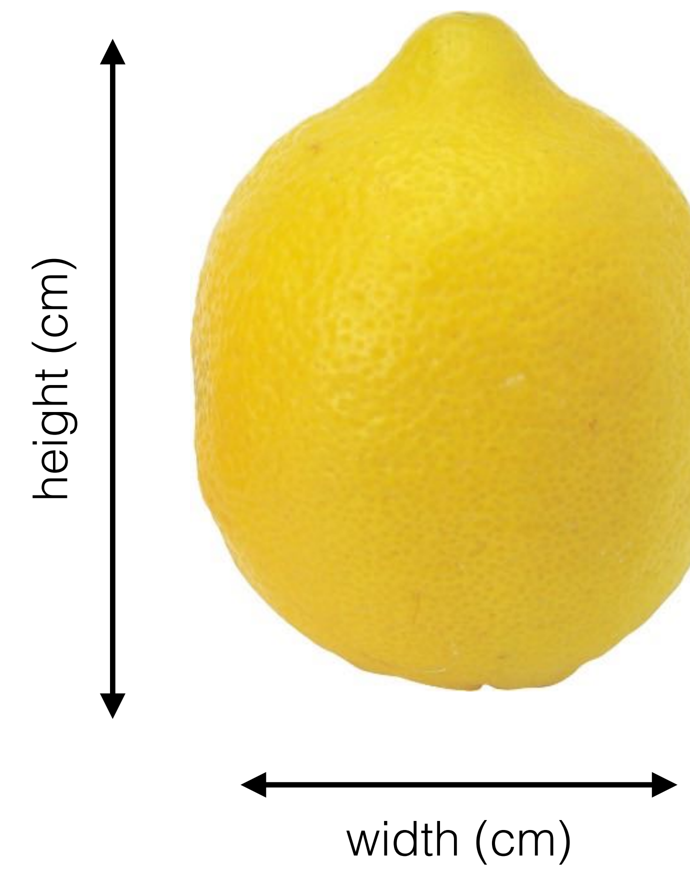


Example of Inductive Bias

- 1. Linear regression/classification, 2. k-NN



Binary classification problem
(“lemon” *vs* “non-lemon”)



Example of Inductive Bias

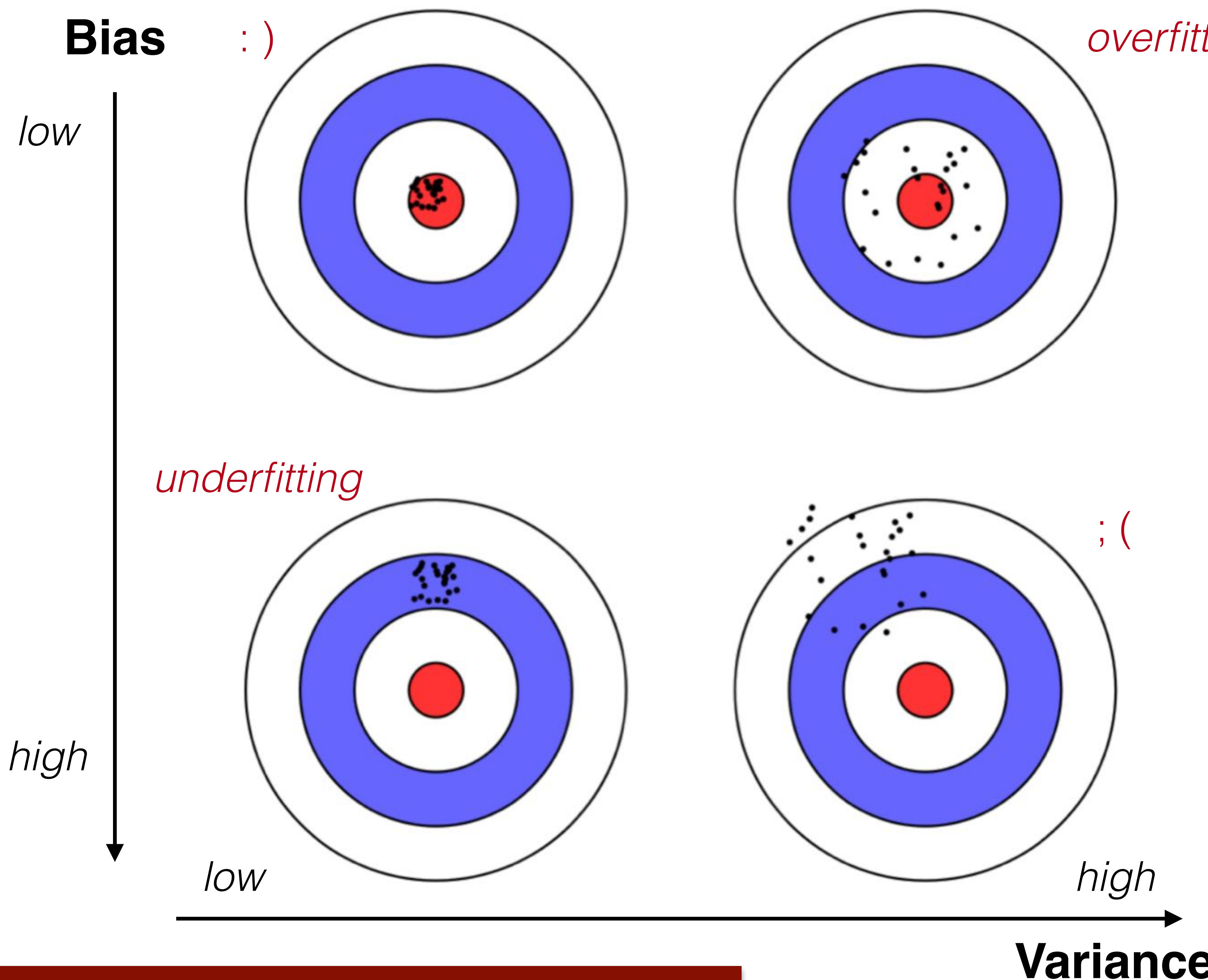
- 1. Linear regression/classification, 2. k-NN

    - **Linear regression**: assume that the output or dependent variable is related to independent variable linearly (in the weights)

    - **(k-)Nearest Neighbors**: assume that most of the cases in a small neighborhood in feature space belong to the same class

Bias-Variance Tradeoff

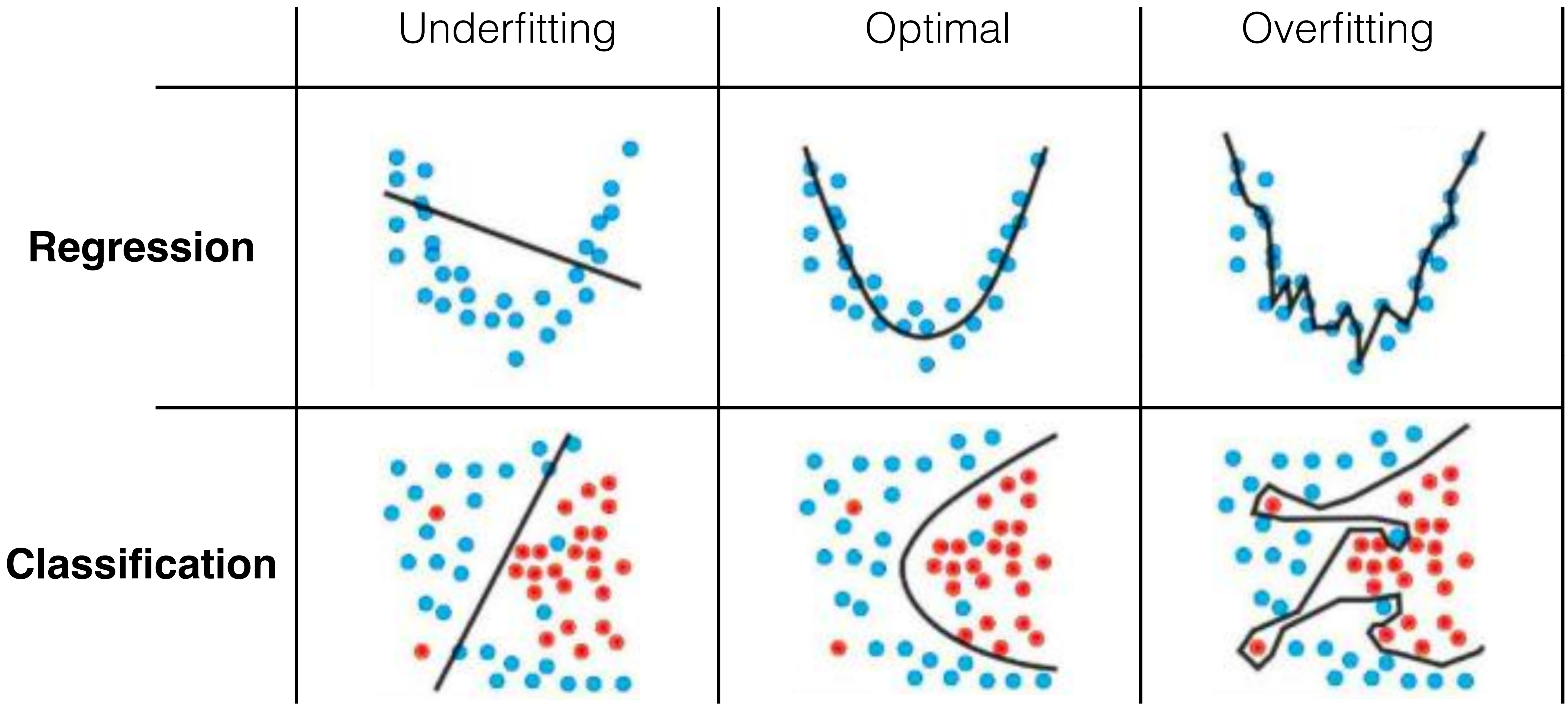
- Dartboard metaphor illustrating bias and variance:



Bias-Variance Tradeoff

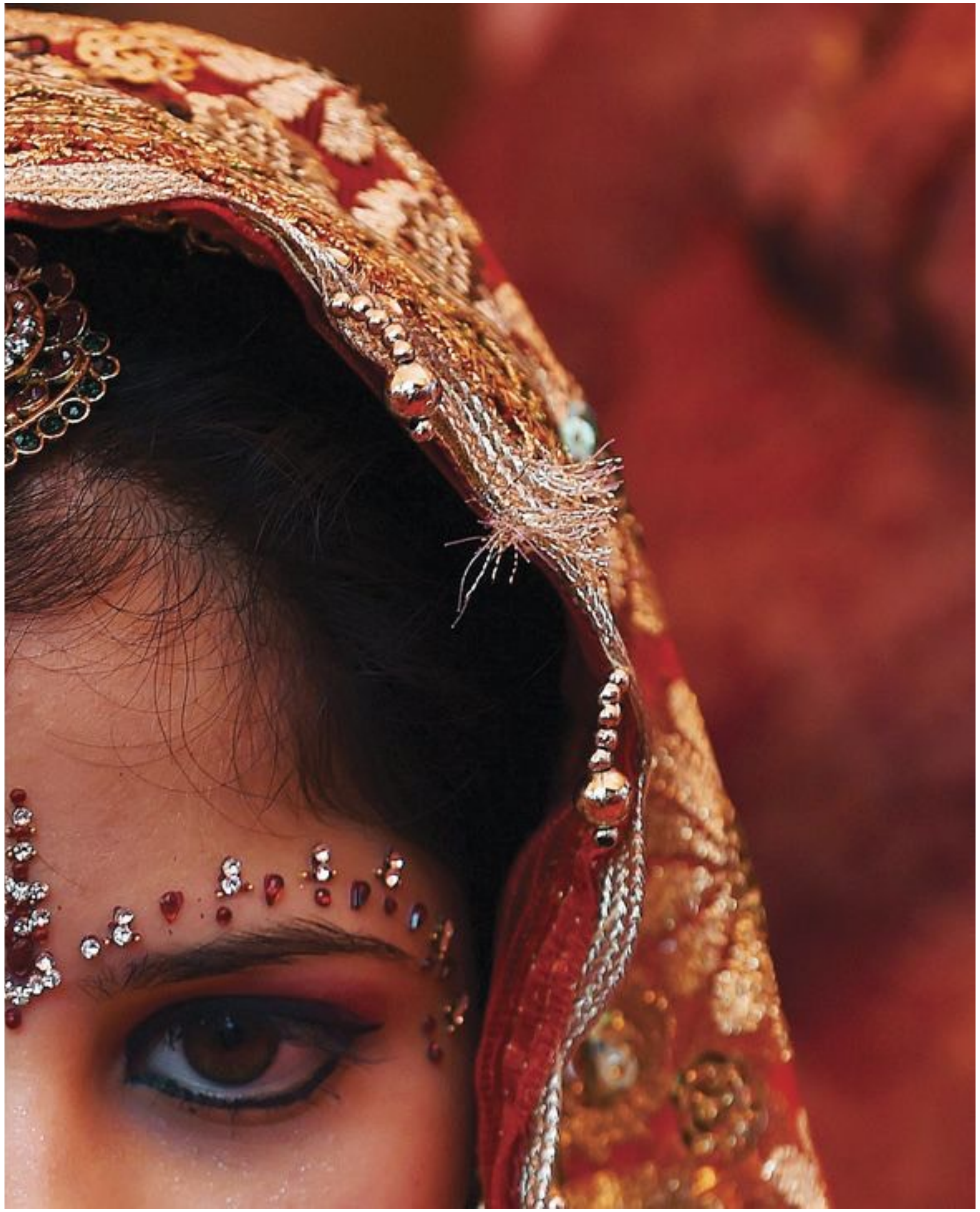
- The **bias** error is produced by weak assumptions in the learning algorithm
    - High bias can cause an algorithm to miss the relevant relations between features and target outputs (***underfitting***)
- The **variance** is an error produced by an over-sensitivity to small fluctuations in the training set
    - High variance can cause an algorithm to model the random noise in the training data, rather than the intended outputs (***overfitting***)

Bias-Variance Tradeoff



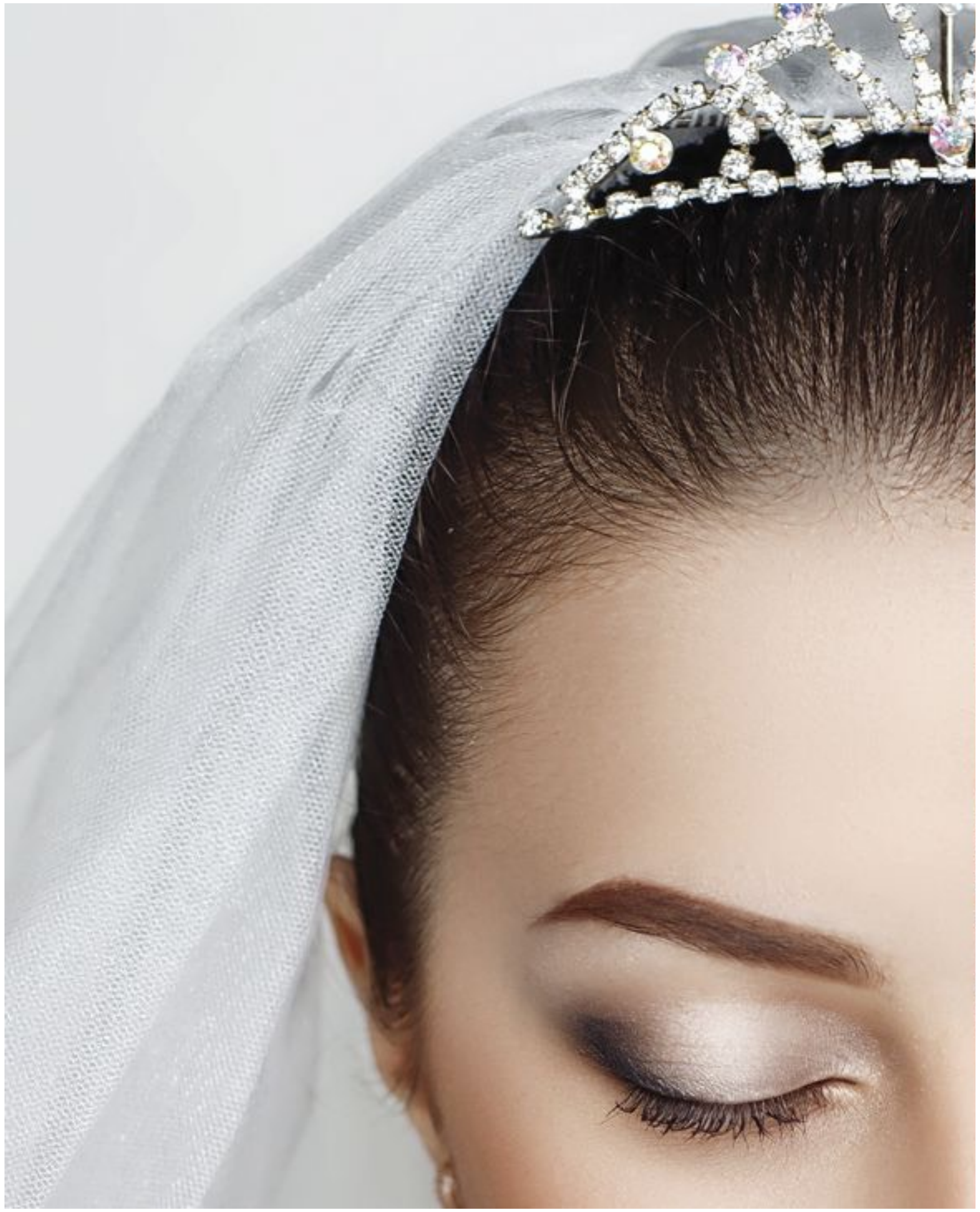
Algorithmic Bias

- How would you label this image?



Algorithmic Bias

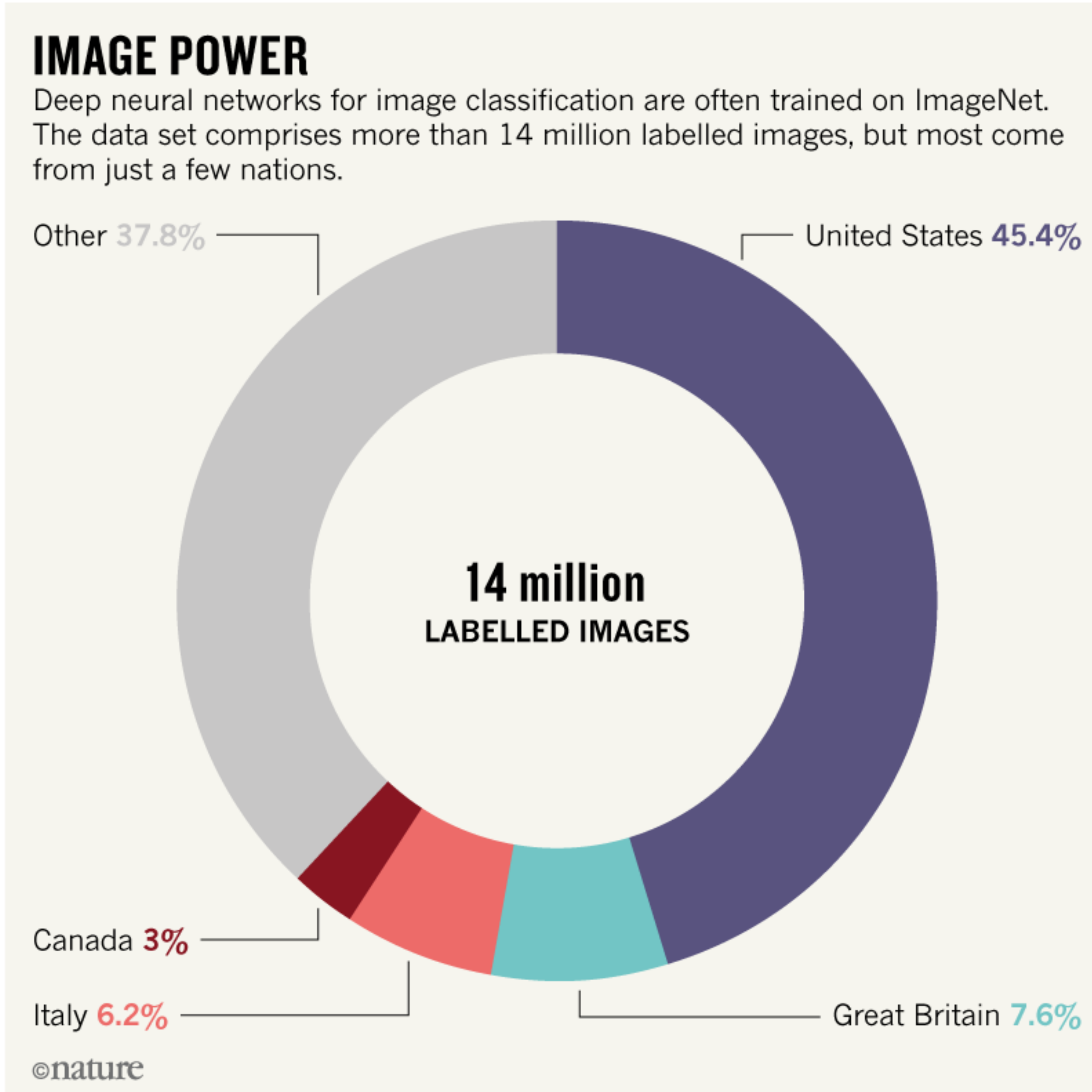
- What about this one?



Algorithmic Bias

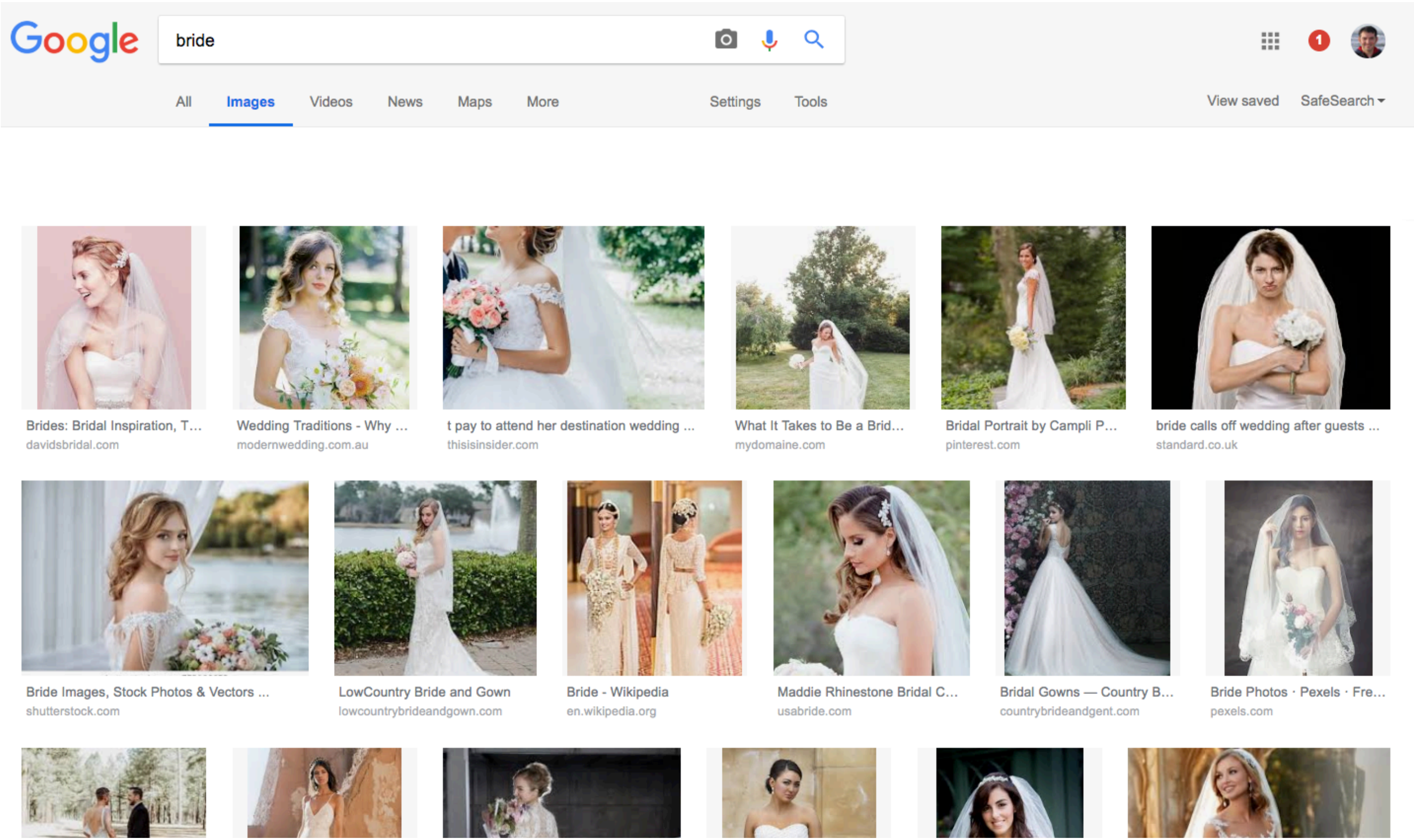
J. Zou, L. Schiebinger, *Nature 559: 324-326 (2018)*

Algorithmic Bias



J. Zou, L. Schiebinger, *Nature 559: 324-326 (2018)*

Algorithmic Bias



Algorithmic Bias

Gender error: ~1% on lighter-skinned males

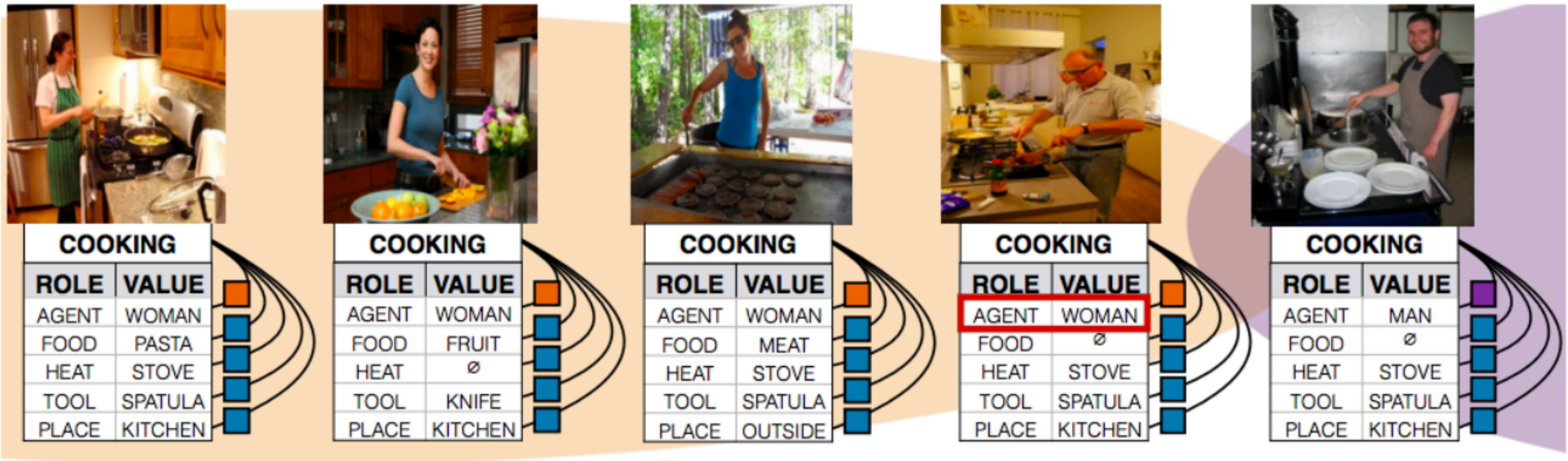
Gender error: ~7% on lighter-skinned females

Gender error: ~12% on darker-skinned males

Gender error: ~35% on darker-skinned females

J.Buolamwini, T.Gebru, *PMLR 81:1-15 (2018)*

AI models may amplify social bias



"Men also like shopping: Reducing Gender Bias Amplification using Corpus-level constraints", J. Zhao, T. Wang, M. Yatskar, V. Ordonez, K.-W. Chang, *EMNLP 2017*

AI models may amplify social bias

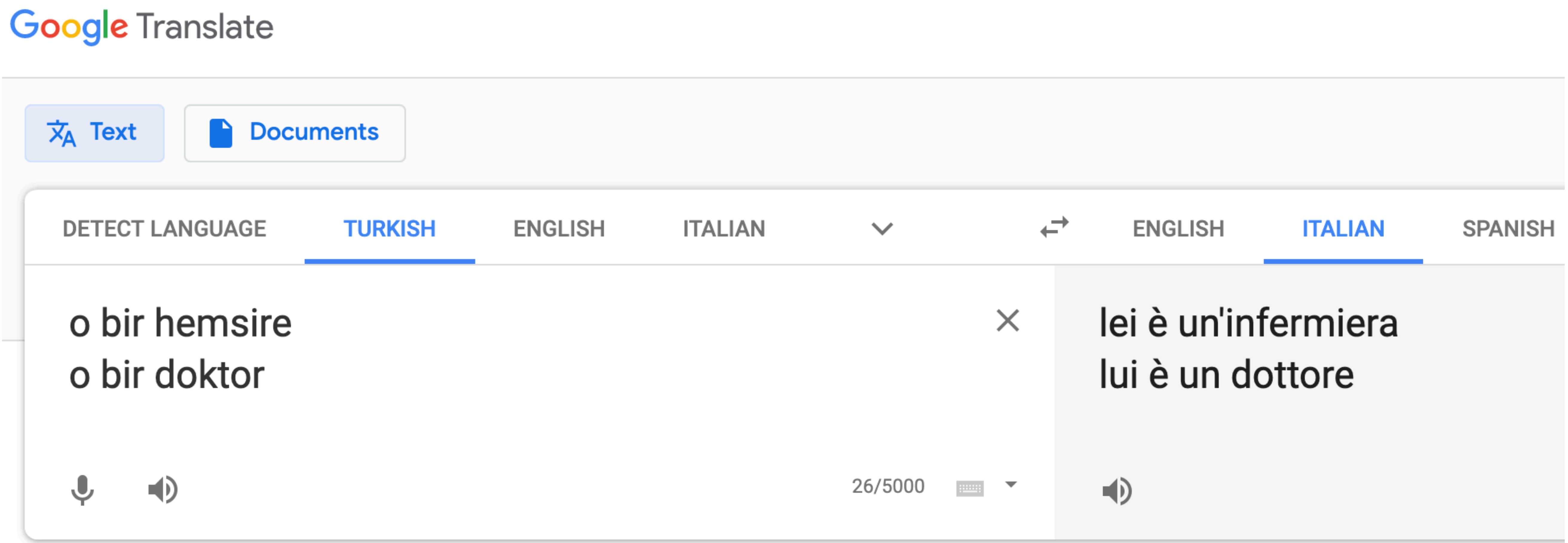
Word embeddings (w2vNEWS)

Extreme *she*	**Extreme *he***
1. homemaker	1. maestro
2. nurse	2. skipper
3. receptionist	3. protege
4. librarian	4. philosopher
5. socialite	5. captain
6. hairdresser	6. architect
7. nanny	7. financier
8. bookkeeper	8. warrior
9. stylist	9. broadcaster
10. housekeeper	10. magician

"Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings", T. Bolukbasi, K.-W. Chang, J. Zou , V. Saligrama, A. Kalai, *NPS 2016*

"Semantics derived automatically from language corpora contain human-like biases", A. Caliskan, J. J. Bryson, A. Narayanan, *Science 356-6334 (2017)*

AI models may amplify social bias

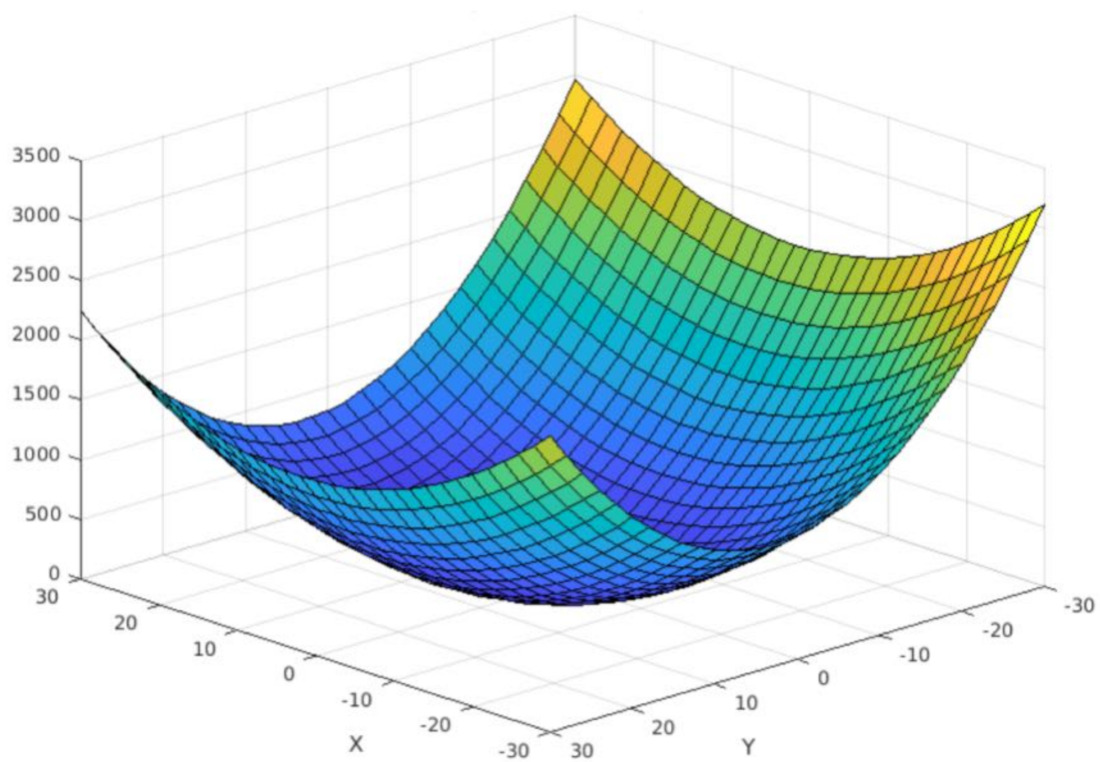


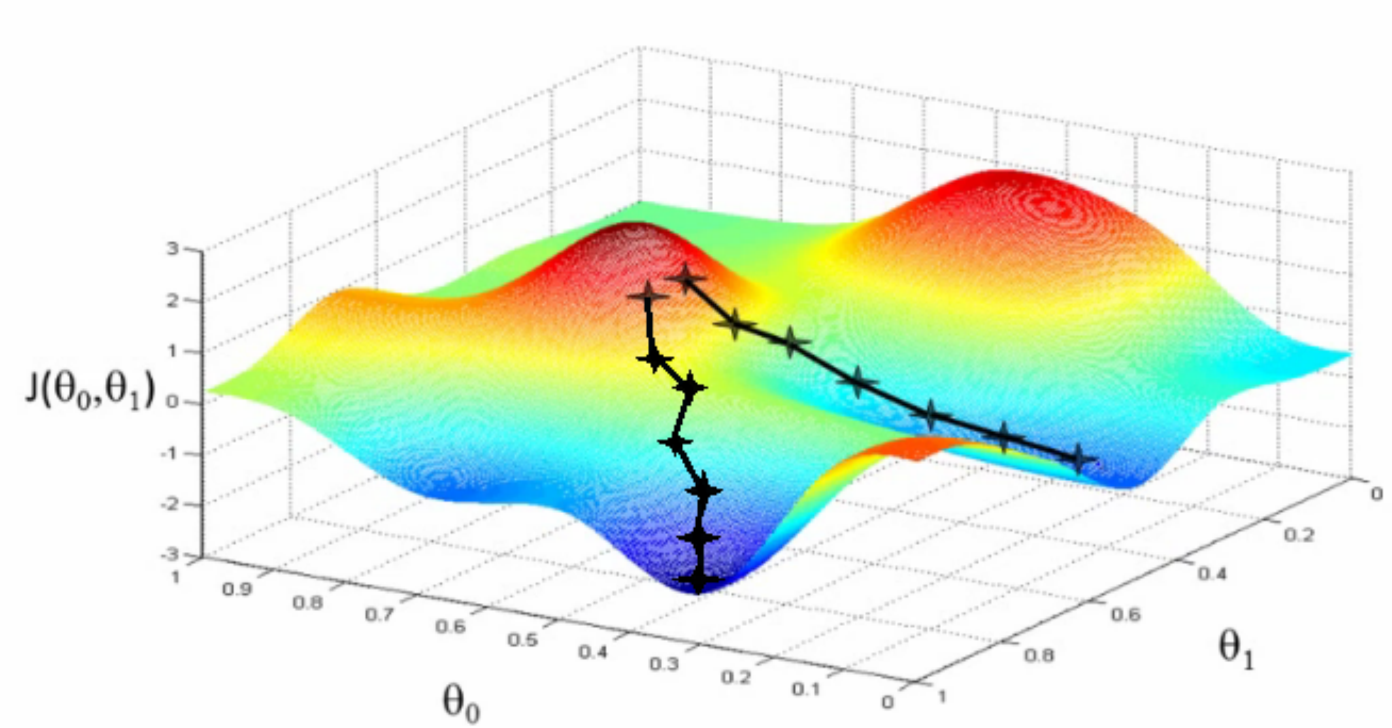
"Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings", T. Bolukbasi, K.-W. Chang, J. Zou , V. Saligrama, A. Kalai, *NPS 2016*

"Semantics derived automatically from language corpora contain human-like biases", A. Caliskan, J. J. Bryson, A. Narayanan, *Science 356-6334 (2017)*

Coming up

- Next lecture:
  - Linear Regression
  - Gradient Descent





Contact

- **Office:** Torre Archimede, room 6CD3
- **Office hours** (ricevimento)**:** Friday 9:00-11:00

✉ lamberto.ballan@unipd.it

🏠 http://www.lambertoballan.net

🏠 http://vimp.math.unipd.it

@ lambertoballan.bsky.social