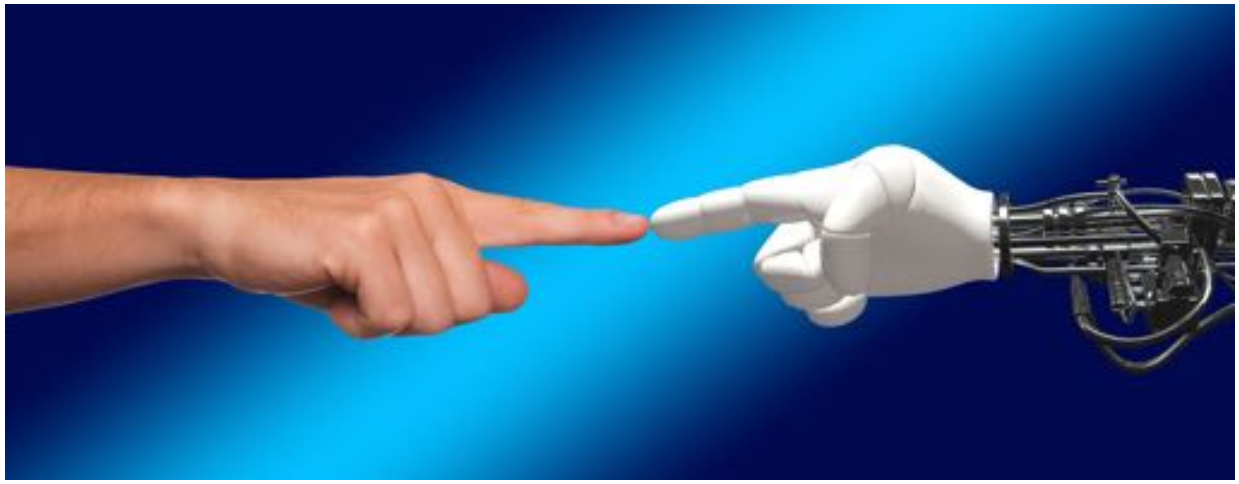
Artificial Intelligence Cognitive Computing - a practical introduction



Ansaf Salleb-Aouissi

Technovation Talks
United Nations – New York
December 14, 2017

AI beyond the movies

Definition of AI

"The science and engineering of making intelligent machines"

McCarthy.

"The study and design of intelligent agents, where an intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success."

Russel and Norvig

Artificial Intelligence: a modern approach.

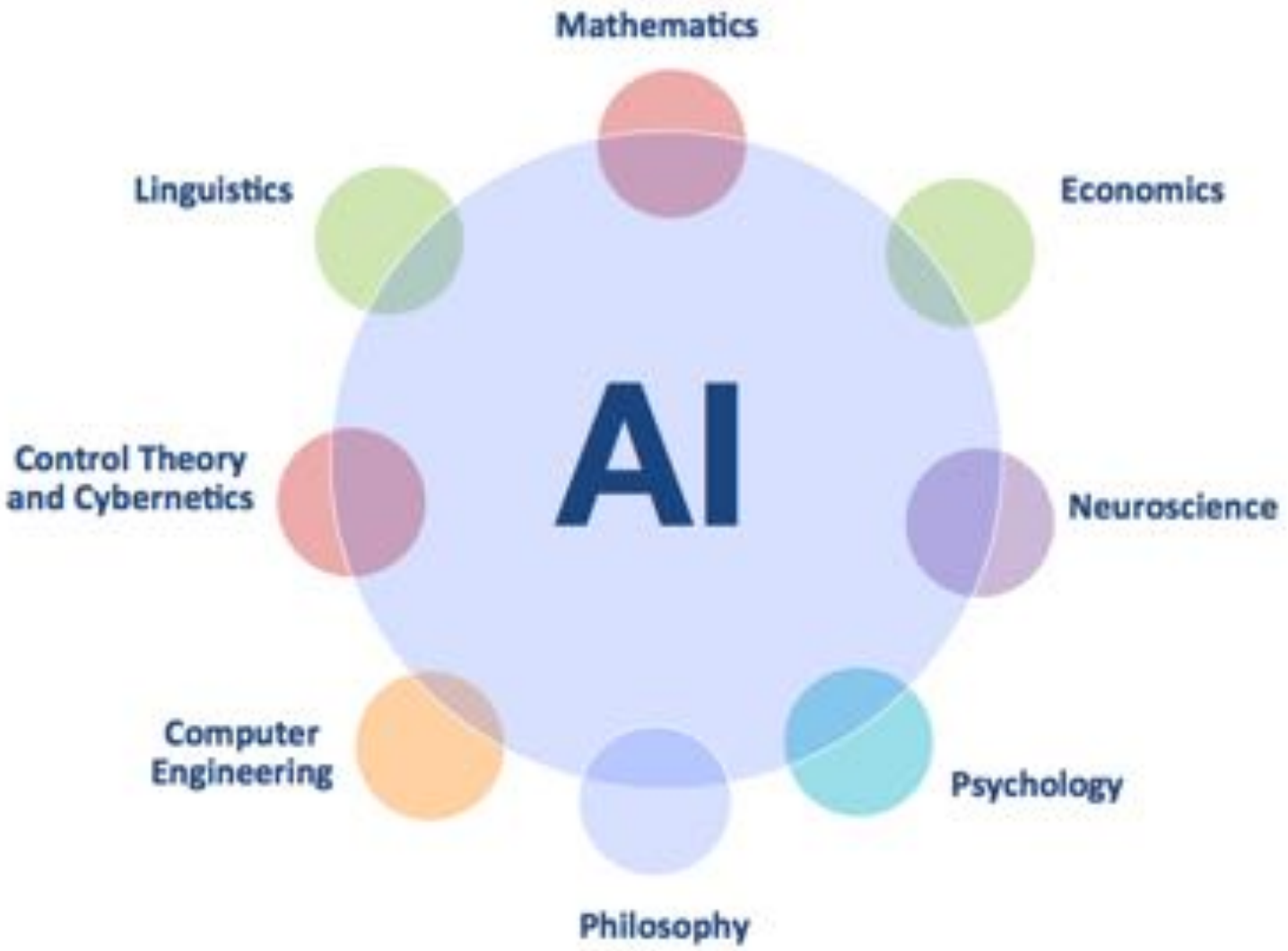
Why AI?

AI is a revolution!

"Just as the Industrial Revolution freed up a lot of humanity from physical drudgery, I think AI has the potential to free up humanity from a lot of the mental drudgery."

Andrew Ng.

Foundation of AI



Turing Test

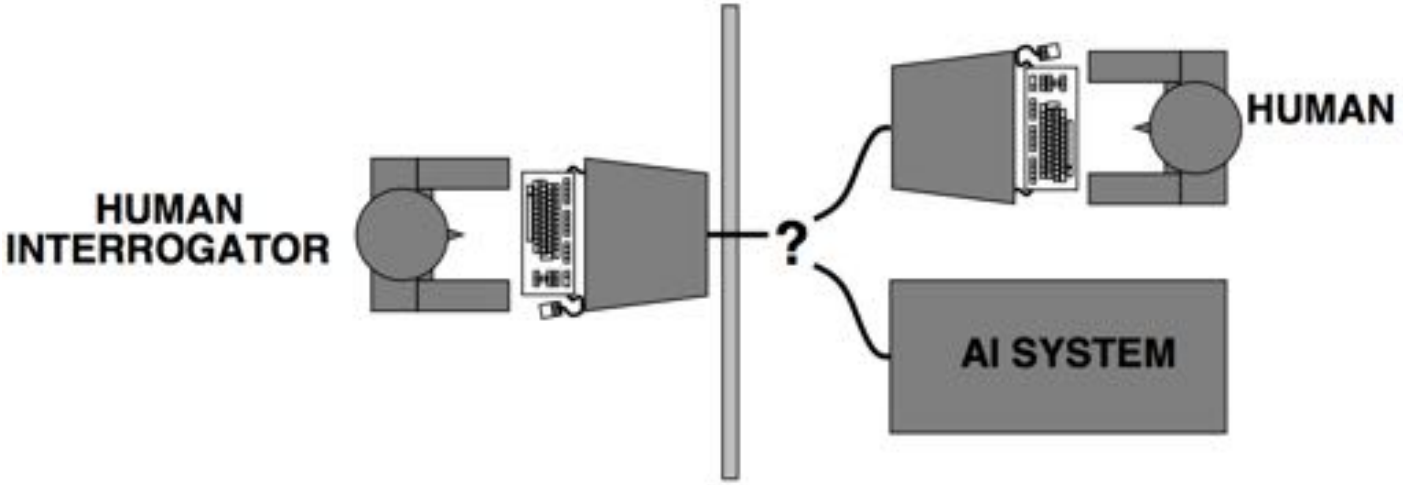
Alan Turing (1912-1954)

- Famous British mathematician.
- Code breaker during World War II.
- Proposed an operational test for intelligent behavior: The Imitation Game.
- In "Computing machinery and intelligence" (1950), he laid down AI major components:
  (language, reasoning, knowledge, learning, understanding).

http://www.turingarchive.org/browse.php/B/9

Turing Test

- **Turing test (Alan Turing 1950)**: A computer passes the test of intelligence, if it can fool a human interrogator.



Credit: From Russel and Norvig slides.

History of AI

- **1940-1950**: Gestation of AI
  - McCulloch & Pitts: Boolean circuit to model of brain
  - Turing's Computing Machinery and Intelligence
    http://www.turingarchive.org/browse.php/B/9
- **1950-1970**: Early enthusiasm, great expectations
  - Early AI programs, Samuel's checkers program
  - Birth of AI @ Dartmouth meeting 1956.
  - Check out the MIT video "The thinking Machine" on youtube
    https://www.youtube.com/watch?v=aygSMgK3BEM
- **1970-1990**: Knowledge-based AI
  - Expert systems, AI becomes an industry
  - AI winter

History of AI

- **1990-present**: Scientific approaches
  - Neural Networks: le retour
  - The emergence of intelligent agents
  - AI becomes "scientific", use of probability to model uncertainty
  - The availability of very large datasets.
    * Data will drive future discoveries and alleviate the complexity in AI.
  - AI Spring!

Applications of AI

Handwriting recognition (check, zipcode)

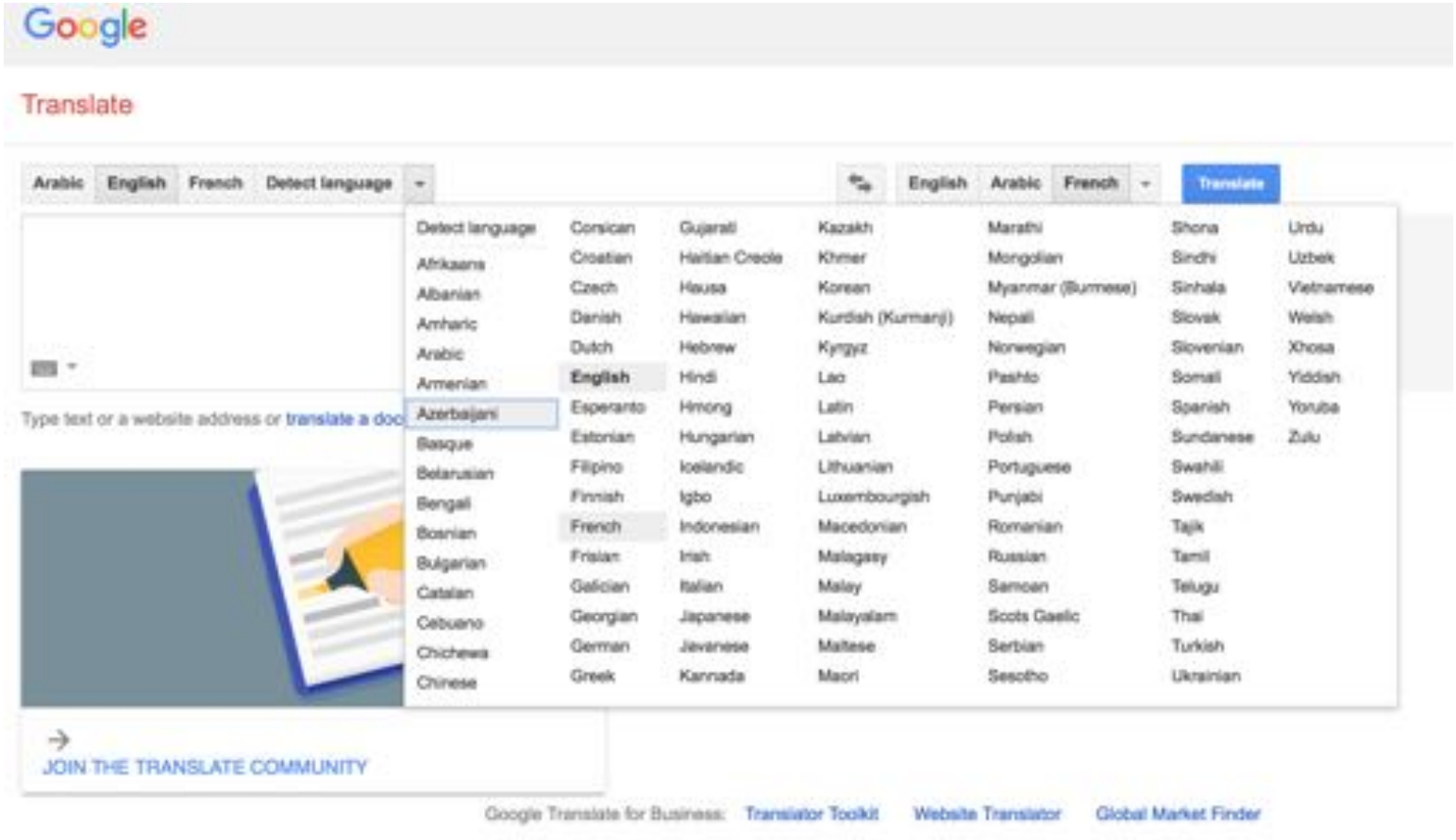
Applications of AI

Machine translation

- Historical motivation: translate Russian to English.
- MT has gone through ups and downs.
- First systems using **mechanical translation** (one-to-one correspondence) failed!
- "Out of sight, out of mind" $\Rightarrow$ "Invisible, imbecile".
- Today, **Statistical Machine Translation** leverages the vast amounts of **available translated corpuses**, e.g., Canadian Hansard, European Parliament Proceedings.

Applications of AI

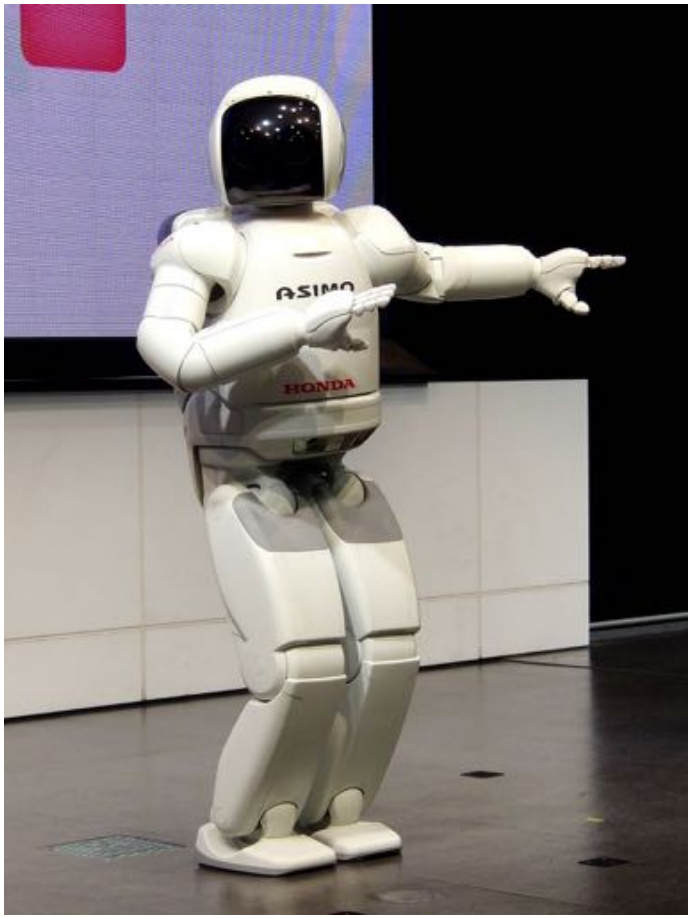
Machine translation



100+ languages

Applications of AI

Robotics: Awesome robots today! NAO, ASIMO, and more!



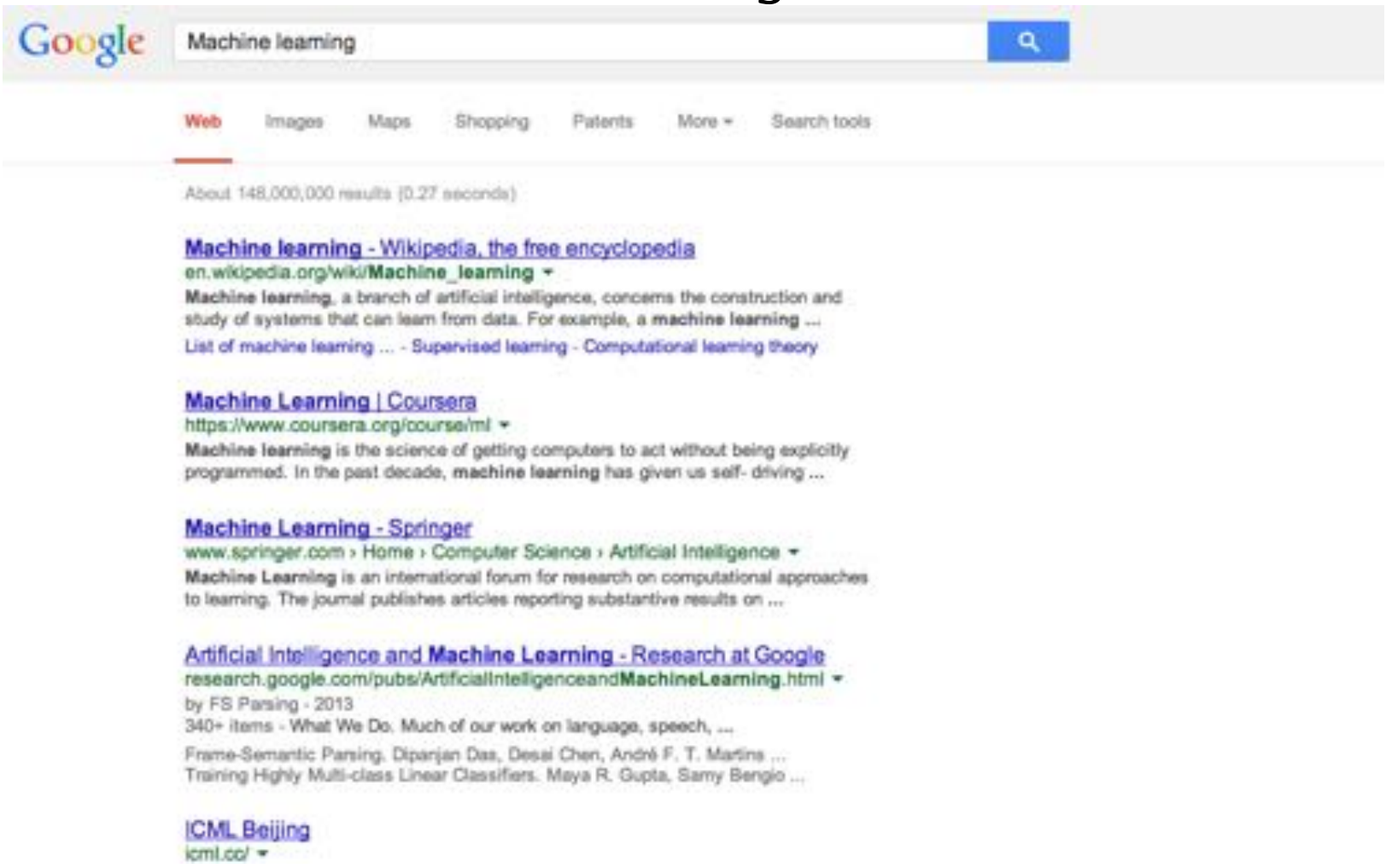
Credit: By Momotarou2012, via Wikimedia Commons.

Applications of AI

Recommendation systems (collaborative filtering)

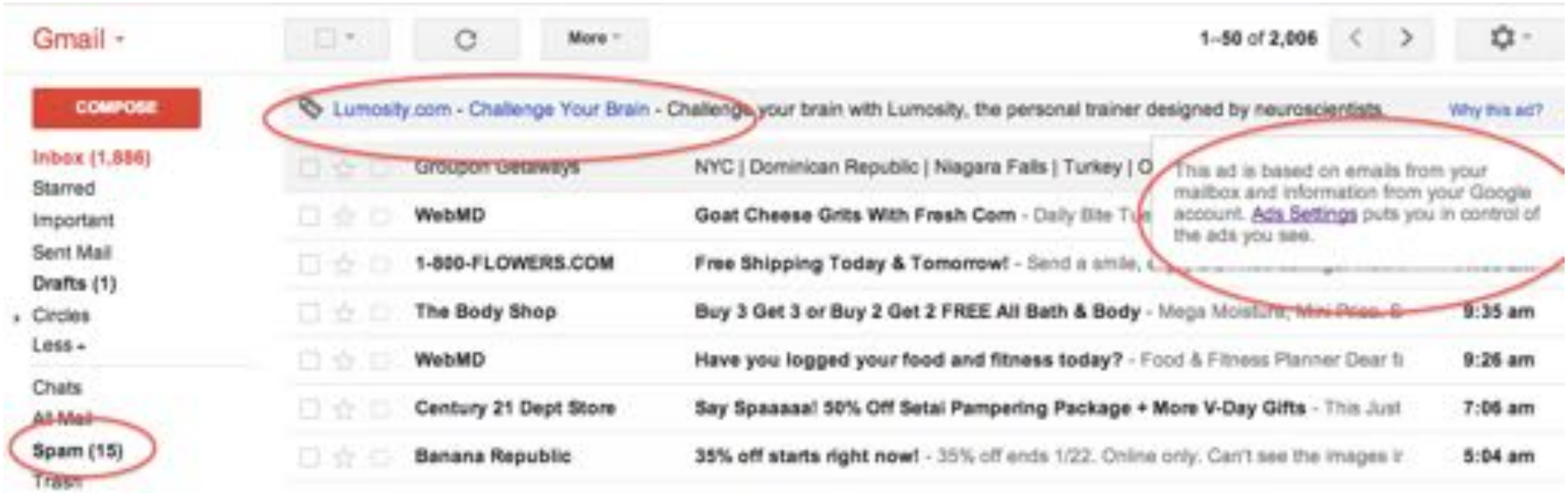
Applications of AI

Search engines



Applications of AI

Spam filtering



Applications of AI

Face detection

Viola-Jones method.

Applications of AI

Speech recognition

- Virtual assistants: Siri (Apple), Echo (Amazon), Google Now, Cortana (Microsoft).
- "They" helps get things done: send an email, make an appointment, find a restaurant, tell you the weather and more.
- Leverage deep neural networks to handle **speech recognition** and **natural language understanding**.

Applications of AI

Chess (1997): Kasparov vs. IBM Deep Blue

(Left) Copyright 2007, S.M.S.I., Inc. - Owen Williams, The Kasparov Agency, via Wikimedia Commons (Right) By James the photographer, via Wikimedia Commons

Powerful search algorithms!

Applications of AI

Jeopardy! (2011): Humans vs. IBM Watson

By Rosemaryetoufee (Own work), via Wikimedia Commons

Natural Language Understanding and information extraction!

Applications of AI

Go (2016): Lee Sedol versus Google AlphaGo

AlphaGo

(Left) By LG Electronics, via Wikimedia Commons (Right) By Google DeepMind, via Wikimedia Commons

Deep Learning, reinforcement learning, and search algorithms!

Applications of AI

Autonomous driving

By User Spaceape on en.wikipedia, via Wikimedia Commons

- DARPA Grand Challenge
  - 2005: 132 miles
  - 2007: Urban challenge
  - 2009: Google self-driving car

AI Schools

Four schools of thoughts (Russel & Norvig)

Thinking humanly	**Thinking rationally**
"The exciting new effort to make computers think... *machines with minds*, in the full and literal sense." (Haugeland, 1985)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985
Acting humanly	**Acting rationally**
"The study of how to make computers do things which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998)

AI Schools

Thinking humanly: cognitive approach

Requires to determine how humans think!

1960's "cognitive revolution".

Requires scientific theories of internal activities of the brain

- What level of abstraction? "Knowledge" or "circuits"?
- How to validate?

Today, Cognitive Science and Artificial Intelligence are distinct disciplines.

AI Schools

Acting humanly:

AI Schools

Thinking rationally: Laws of thoughts.

- Codify "right thinking" with **logic**.
- Several Greek schools developed various forms of logic: *notation* and *rules of derivation* for thoughts.
- Problems:

  1. Not all knowledge can be expressed with logical notations.
  2. Computational blow up.

AI Schools

Acting rationally:

- The right thing: that which is expected to maximize goal achievement, given the available information.
- A **rational agent** is one that acts so as to achieve the best outcome, or when there is uncertainty, the best expected outcome.
- Aristotle (Nicomachean Ethics):
  "Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good."

AI Schools

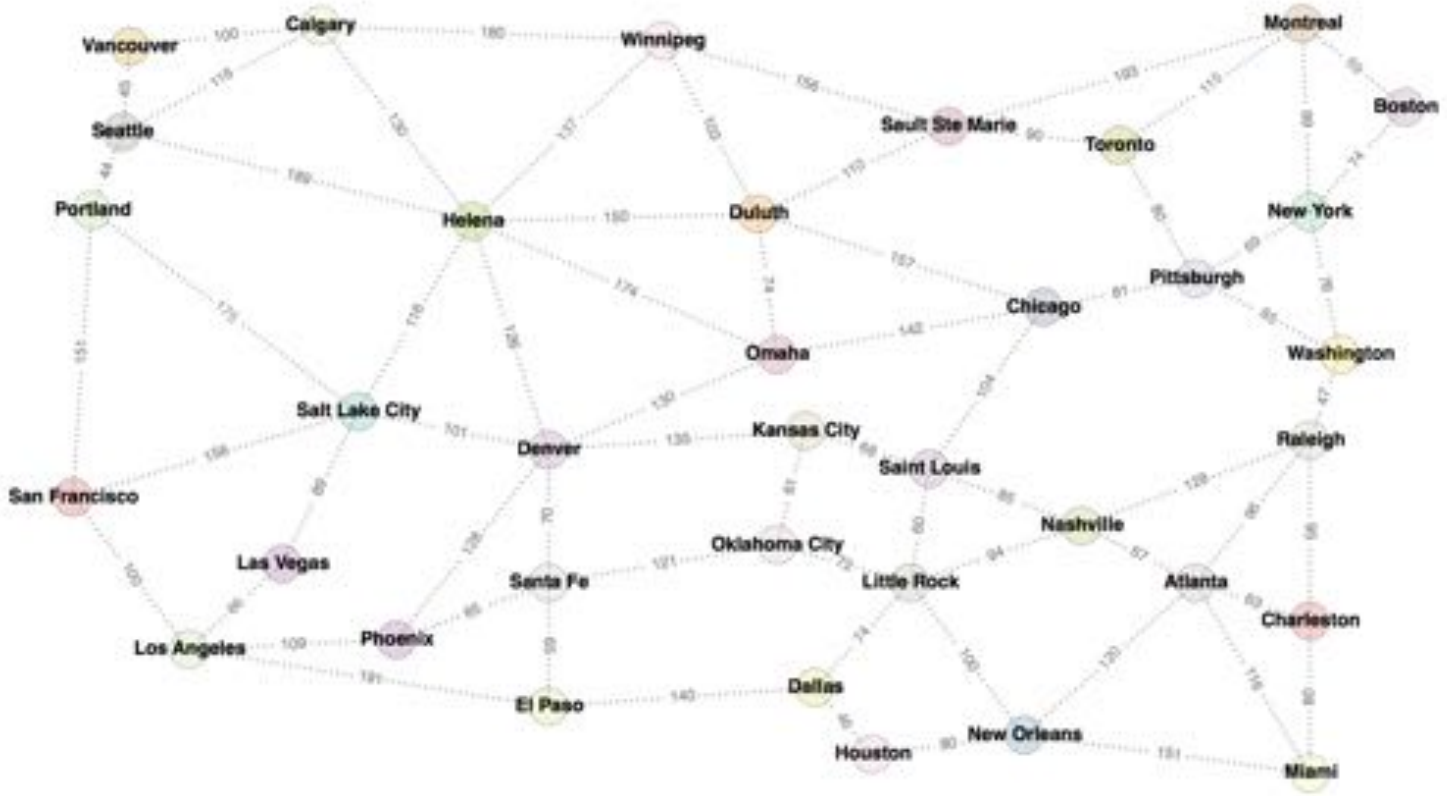
Four schools of thoughts (Russel & Norvig)

Thinking humanly	**Thinking rationally**
"The exciting new effort to make computers think... *machines with minds*, in the full and literal sense." (Haugeland, 1985)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985
Acting humanly	**Acting rationally: Our approach**
"The study of how to make computers do things which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998)

Search agents

Agents that work towards a goal.

Start: Las Vegas – Goal: Calgary



Explore + Execute

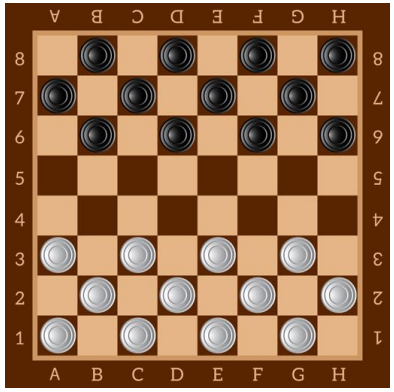
Adversarial agents

Adversarial search problems $\equiv$ game.

There is an **opponent** we can't control!

Checkers:

- Chinook ended 40-year-reign of human world champion Marion Tinsley in 1994.
- Used an endgame database defining perfect play for all positions involving 8 or fewer pieces on the board, a total of 443,748,401,247 positions.



Constraint satisfaction agent

- Agents that solve problems with constraints.
- **Find the assignment(s)** that satisfy all constraints.
- E.g., map coloring, scheduling problems, manufacturing, etc.

8		9	5		1	7	3	6
2		7		6	3			
1	6							
				9		4		7
	9		3		7		2	
7		6		8				
							6	3
			9	3		5		2
5	3	2	6		4	8		9

Constraint satisfaction agent

8	4	9	5	2	1	7	3	6
2	5	7	8	6	3	9	1	4
1	6	3	7	4	9	2	5	8
3	2	5	1	9	6	4	8	7
4	9	8	3	5	7	6	2	1
7	1	6	4	8	2	3	9	5
9	8	4	2	7	5	1	6	3
6	7	1	9	3	8	5	4	2
5	3	2	6	1	4	8	7	9

Variables: $X_{l,c}$ for $1 \leq l \leq 9$ and $1 \leq c \leq 9$.

Constraints: All 3x3 grid, row, column, must contain digits 1..9 and all of them!

Solution: Find the assignments to the variables that satisfy the constraints.

Machine learning agents

"How do we create computer programs that improve with experience?"

Tom Mitchell

Supervised vs. Unsupervised

Given: Training data: $(x_1, y_1), \ldots, (x_n, y_n)$ / $x_i \in \mathbb{R}^d$ and y_i is the label.

example $x_1 \rightarrow$	x_{11}	x_{12}	$\ldots$	x_{1d}	$y_1 \leftarrow$ label
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
example $x_i \rightarrow$	x_{i1}	x_{i2}	$\ldots$	x_{id}	$y_i \leftarrow$ label
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
example $x_n \rightarrow$	x_{n1}	x_{n2}	$\ldots$	x_{nd}	$y_n \leftarrow$ label

Supervised vs. Unsupervised

fruit	length	width	weight	label
fruit 1	165	38	172	Banana
fruit 2	218	39	230	Banana
fruit 3	76	80	145	Orange
fruit 4	145	35	150	Banana
fruit 5	90	88	160	Orange
...				
fruit n	...	...	...	...

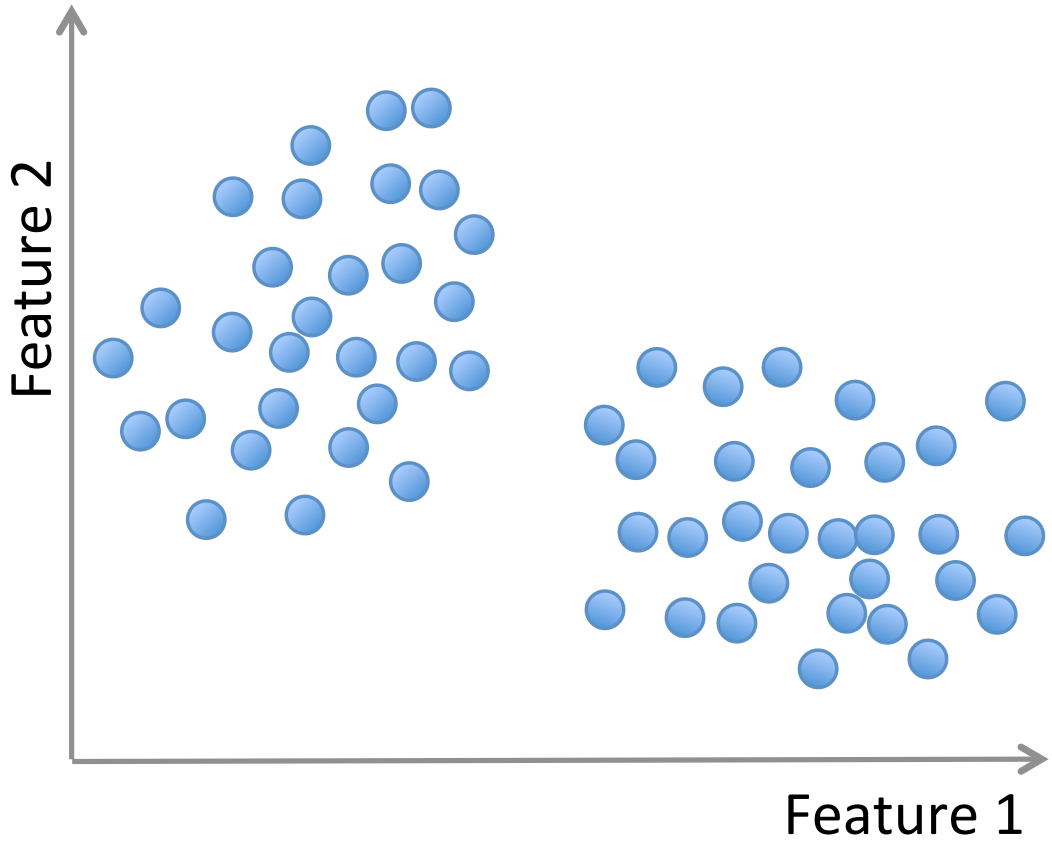
Unsupervised learning:

Learning a model from **unlabeled** data.

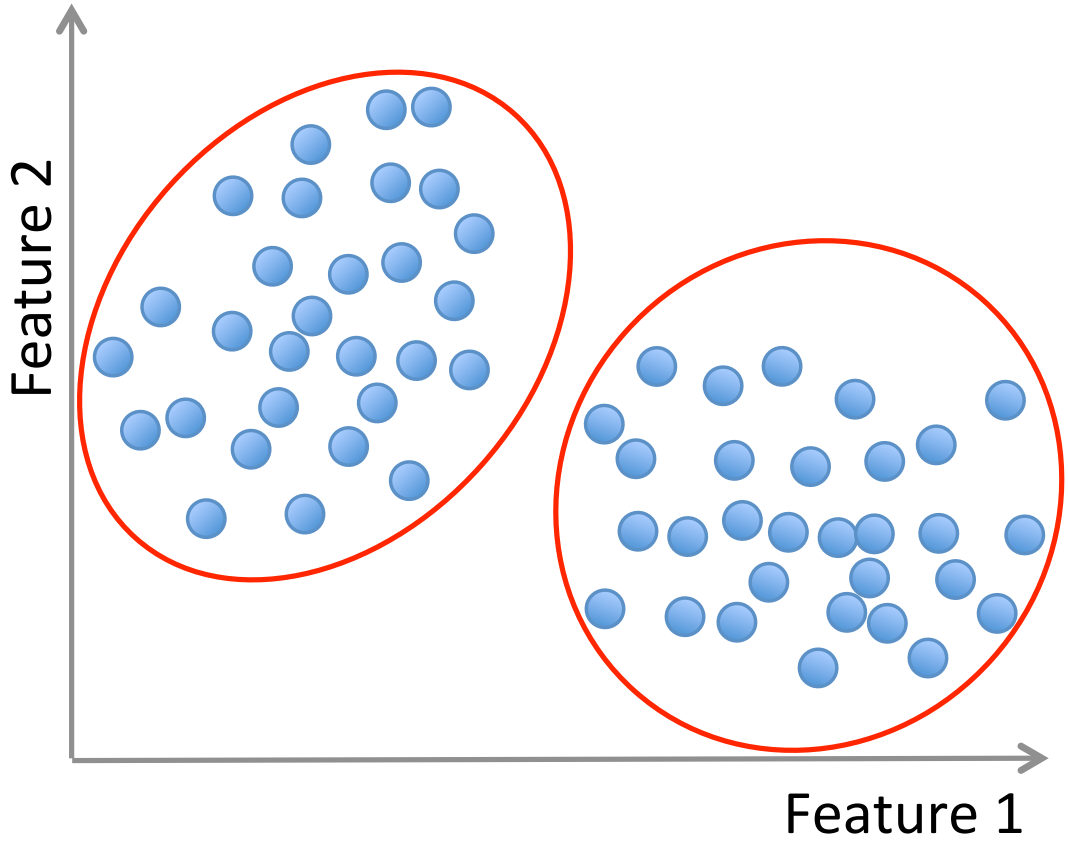
Supervised learning:

Learning a model from **labeled** data.

Unsupervised learning



Unsupervised learning



Methods: K-means, gaussian mixtures, hierarchical clustering, spectral clustering, etc.

Example: Obama 2012 campaign.

Unsupervised learning

Training data: "examples" x.

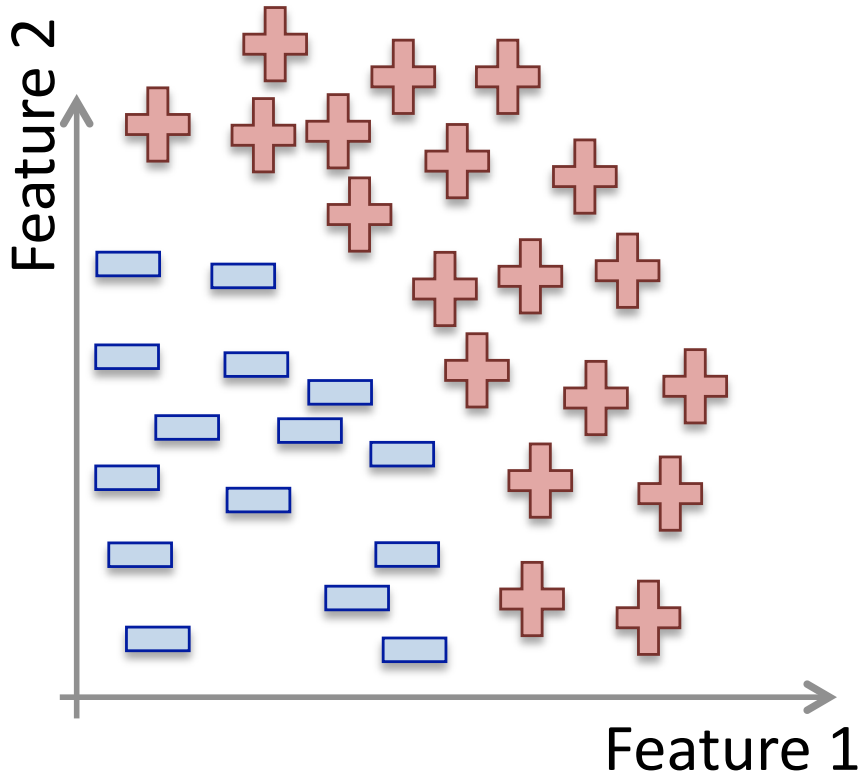
$$x_1, \ldots, x_n, \; x_i \in X \subset \mathbb{R}^n$$

- **Clustering/segmentation:**

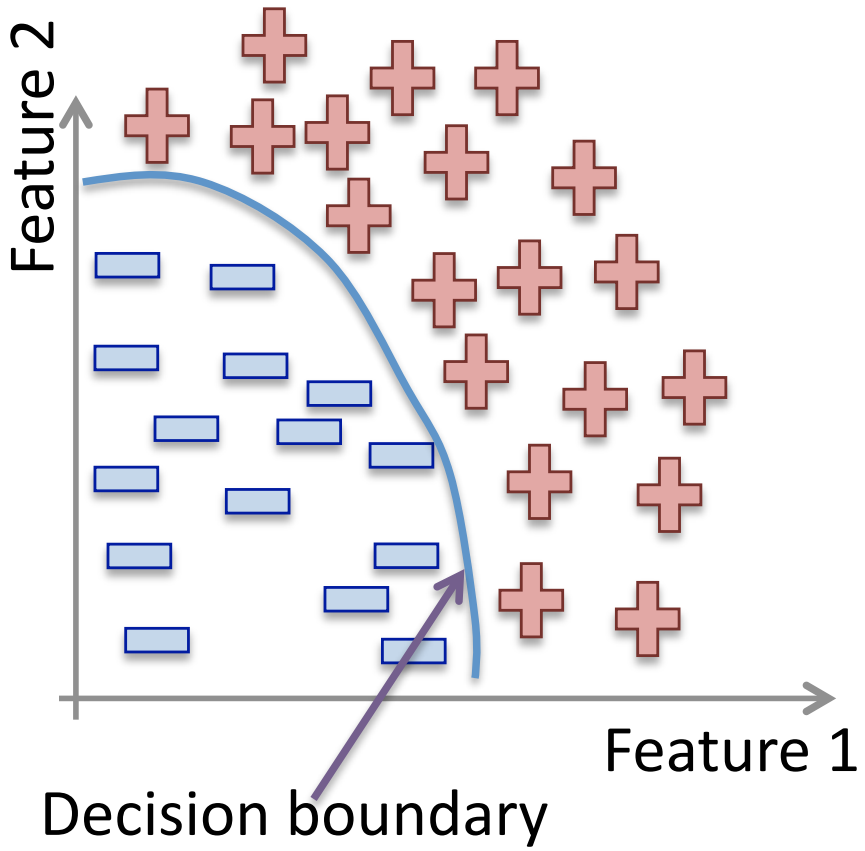
  $f : \mathbb{R}^d \longrightarrow \{C_1, \ldots C_k\}$ (set of clusters).

  Example: Find clusters in the population, fruits, species.

Supervised learning



Supervised learning



Supervised learning

Training data: "examples" x with "labels" y.

$$(x_1, y_1), \ldots, (x_n, y_n) \ / \ x_i \in \mathbb{R}^d$$

- **Classification:** y is discrete. To simplify, $y \in \{-1, +1\}$

$f : \mathbb{R}^d \longrightarrow \{-1, +1\}$ f is called a **binary classifier**.

Example: Approve credit yes/no, spam/ham, banana/orange.

Supervised learning

Training data: "examples" x with "labels" y.

$$(x_1, y_1), \ldots, (x_n, y_n) \ / \ x_i \in \mathbb{R}^d$$

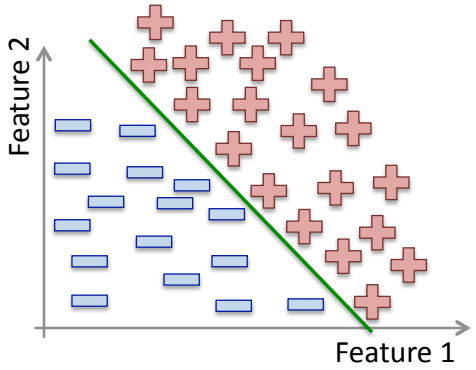
- **Regression:** y is a real value, $y \in \mathbb{R}$

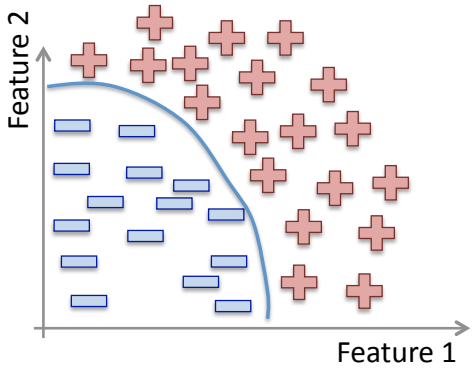
  $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ f is called a **regressor**.

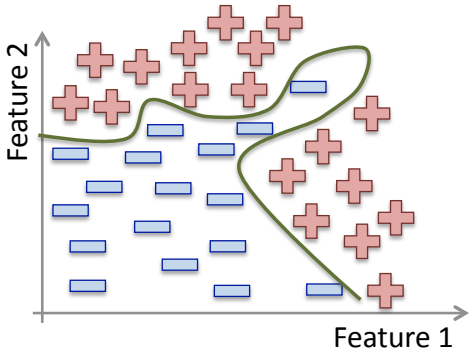
  Example: amount of credit, weight of fruit.

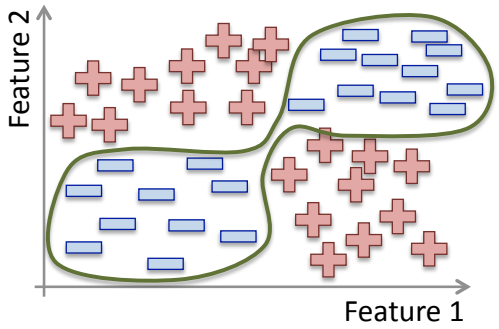
Supervised learning

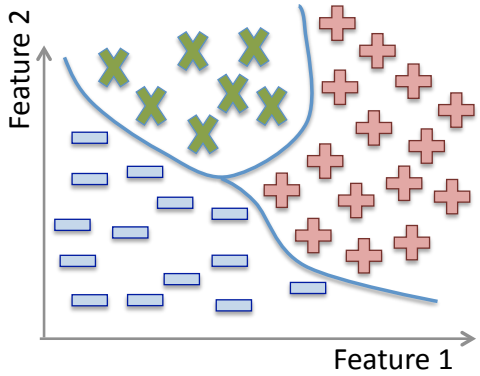
Classification:











Methods: Support Vector Machines, neural networks, decision trees, K-nearest neighbors, naive Bayes, etc.

Objective function

We want to optimize:

$$\text{Classification term} + C \times \text{Regularization term}$$

$$\sum_{i=1}^{n} \ell oss(y_i, f(x_i)) + C \times R(f)$$

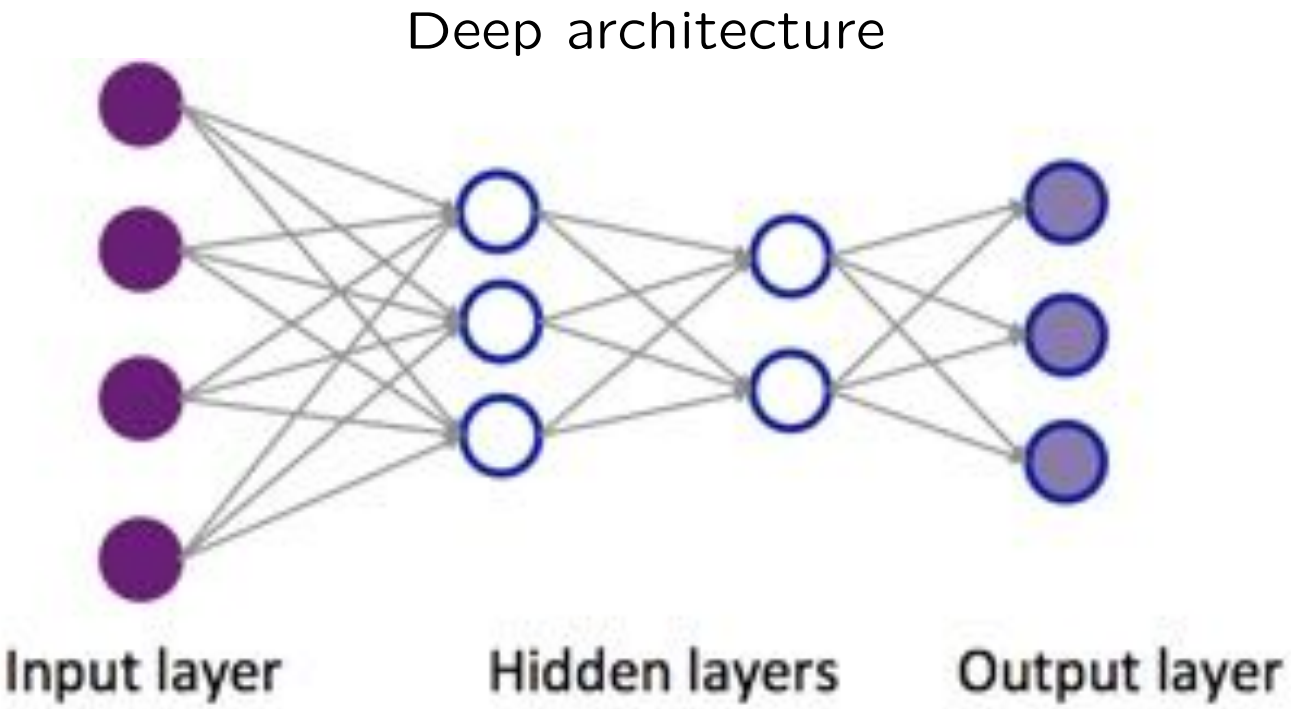
Neural Networks

- 1950-60s: Neural networks (Rosenblatt, etc.)
- 1970's: Slow progress
- 1986: Backpropagation
- 1990s: Convolutional neural networks (LeCun)
- 1990s: Recurrent neural networks (Schmidhuber)
- 2006: NN, le retour. Breakthrough: Deep belief networks (Hinton et al., 2006) and Autoencoders (Bengio et al., 2007).
- 2013: Huge industrial interest. Why now?

Lots of data and more computational power!

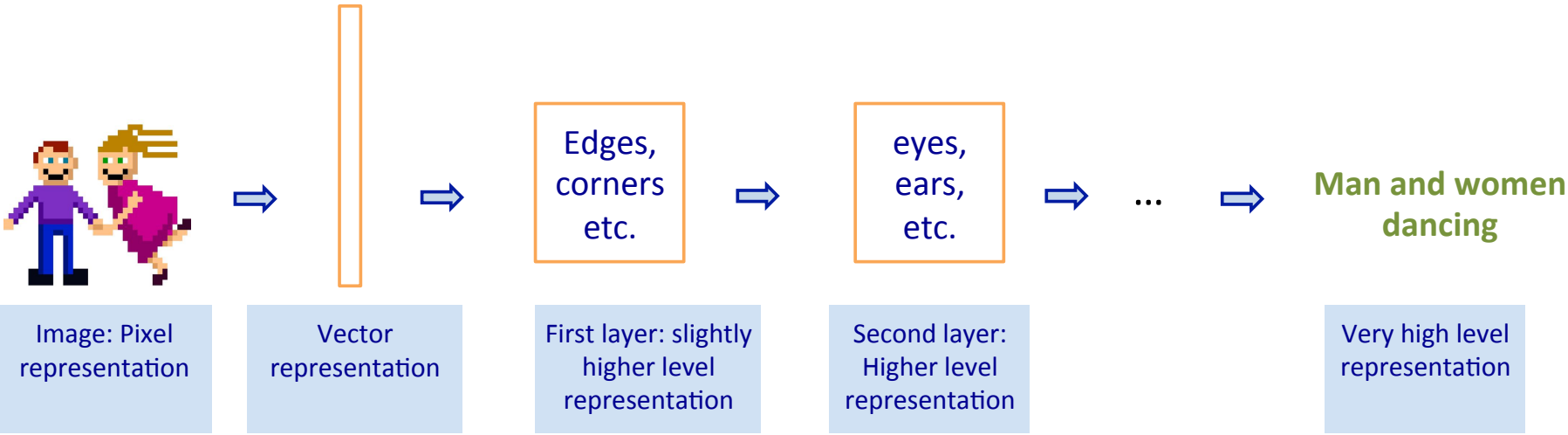
Work well, breakthrough results (vision and speech)

What is Deep Learning?



Deep learning: means using a neural network with a series of hidden layers of non-linear operations between input and output.

Why a deep architecture?



Deep architecture: The series of layers between input and output learn feature hierarchies/feature identification at different levels.

Hidden layers: Act as **feature detectors**, will leads to an *automatic abstraction of data*.

Successive layers: Learn high level features.

AI Challenges and potential

- AI is a flourishing, and a broad field shaping our world
- AI **potential**: to be applied broadly from education, health, to manufacturing, transportation and deeply impact everyday life
- AI **concerns**:
  - Is AI a threat to our humankind?
  - How will AI impact the job market?
  - How will AI transform our work, cities, politics?
  - How will AI change our regulations and laws?

AI & Inclusion

https://blogs.harvard.edu/aiandinclusionsymposium/

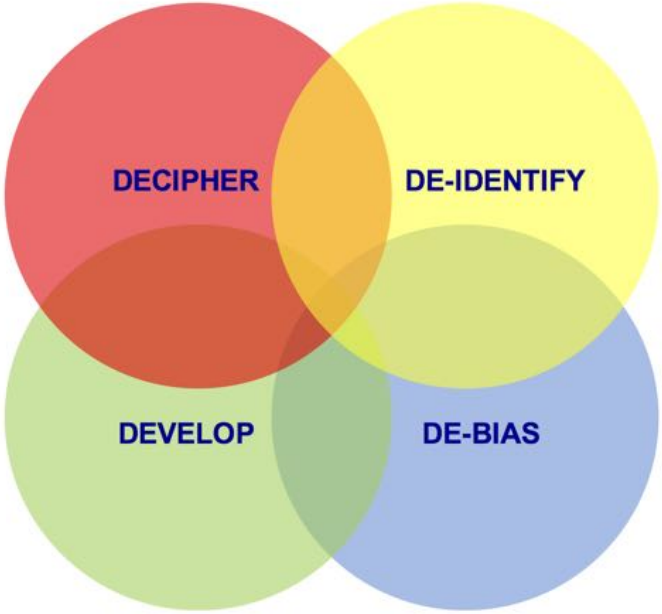
- Co-organized by the Institute for Technology and Society of Rio de Janeiro (ITS Rio) and the Berkman Klein Center for Internet & Society at Harvard University
- Goal: "Address AI opportunities and challenges of AI-based technologies through the lens of inclusion,..., identify, explore, and address the opportunities and challenges of AI as we seek to build a better, more inclusive, and diverse world together."

AI & Inclusion

How to develop **inclusive AI systems** optimized for accuracy, safety, privacy, non-discrimination, transparency?

AI & Inclusion

- AI and inclusion evolves around the **four** following dimensions.



1. **Develop**: to empower individuals worldwide with AI education and avoid "digital divide"
2. **Decipher**: to provide the right for explanation through understandable models
3. **De-identify**: to protect people privacy, and the right not to be categorized which may lead to social exclusion
4. **De-bias**: to ensure fairness and avoid digital discrimination.

Develop AI knowledge

- Quality of education, research and innovation in developing countries is a bottleneck.
- The digital divide may deepen with AI. Artificial Intelligence (AI) and the evolution of digital divides, Andres Lombana Bermudez. July 2017
- The importance of self-learning and online learning (MOOCS).
- Case study: Columbia University AI Micromasters on EdX.

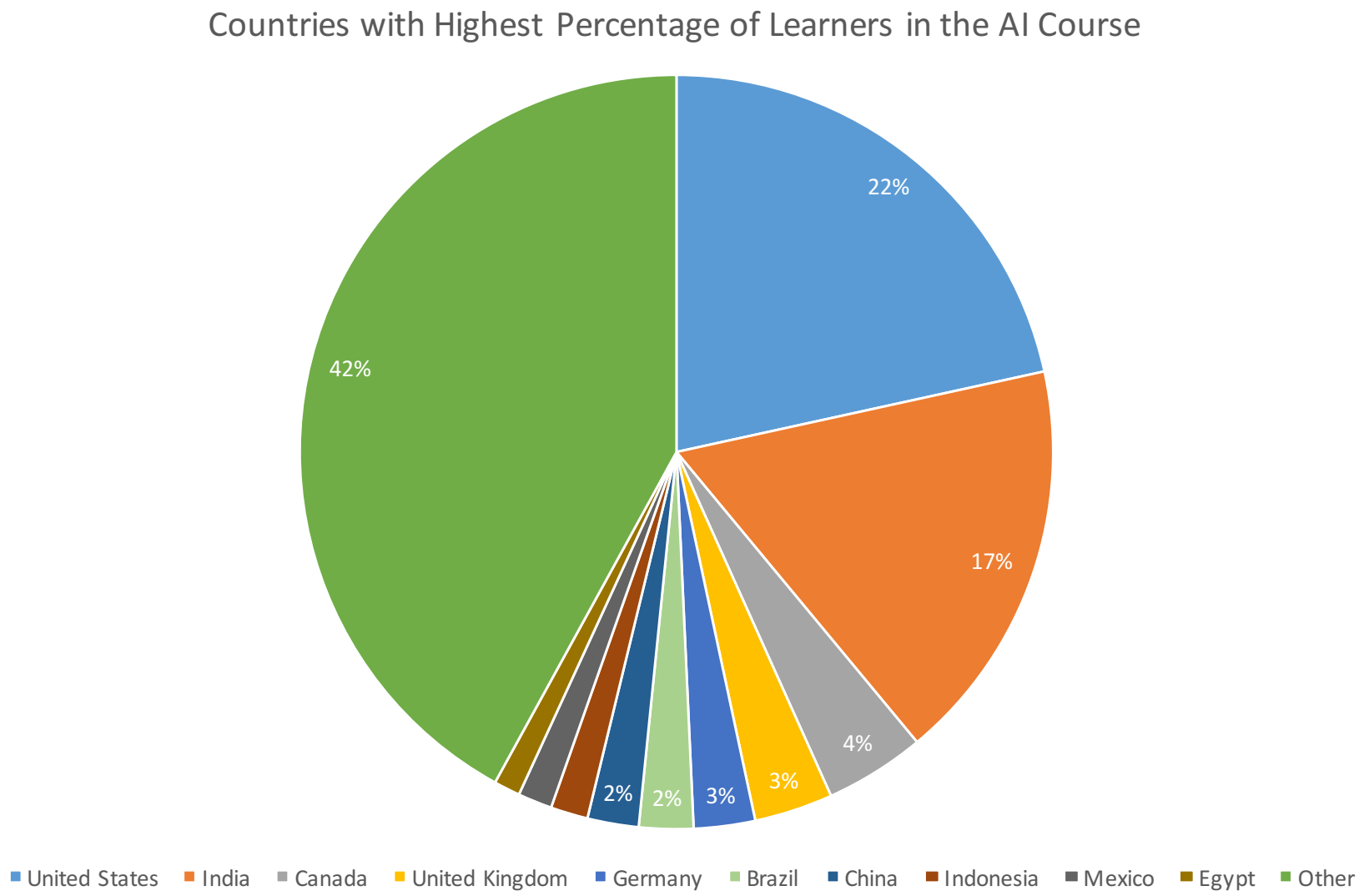
Develop AI knowledge

https://www.edx.org/micromasters/columbiax-artificial-intelligence

https://www.edx.org/course/artificial-intelligence-ai-columbiax-csmm-101x-1

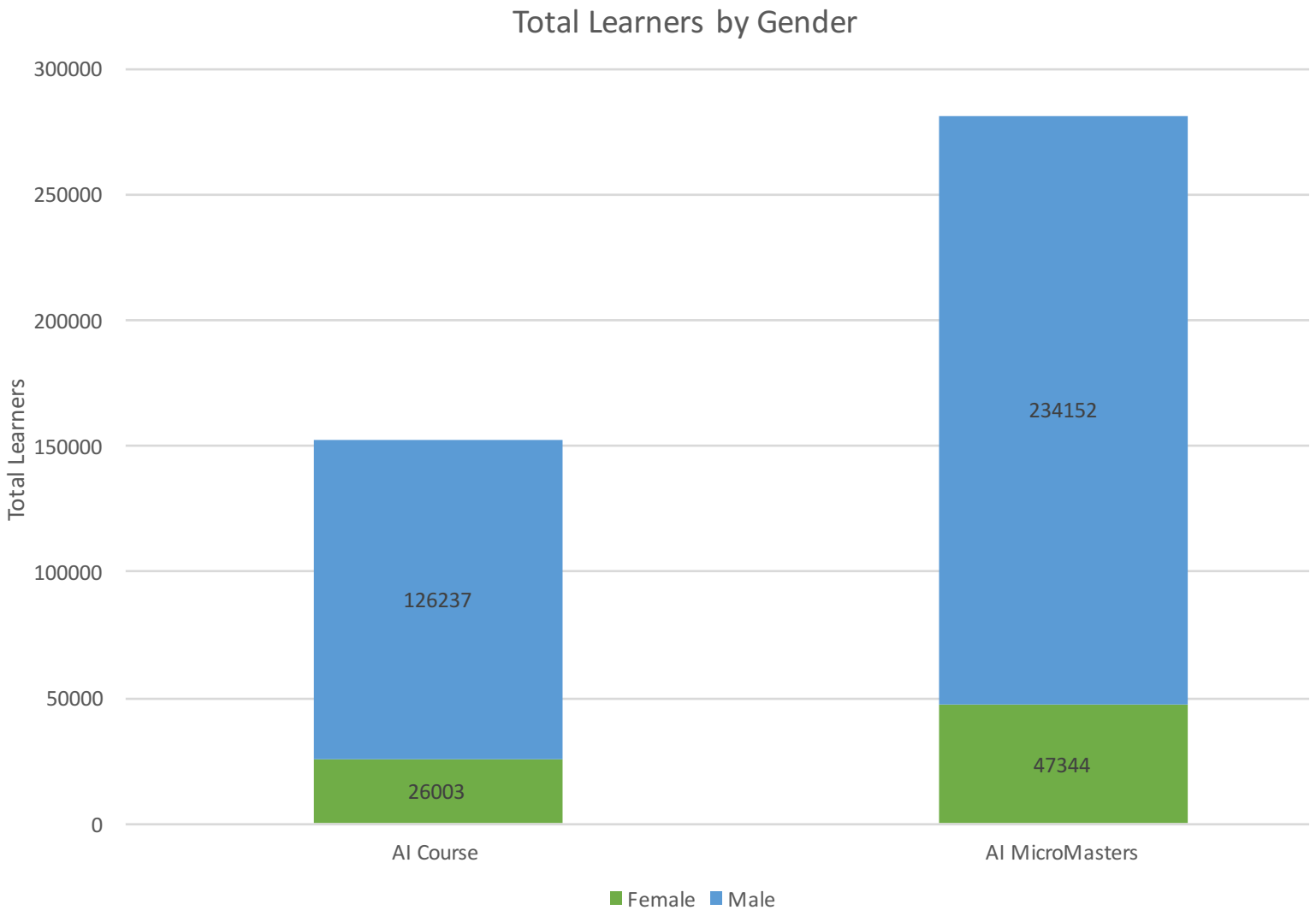
- Four courses: Artificial Intelligence, Machine Learning, Robotics and Animation and CGI Motion.
- The Micromasters attracted **285,726** learners in total.
- The AI course alone attracted **153,257**.

Develop AI knowledge



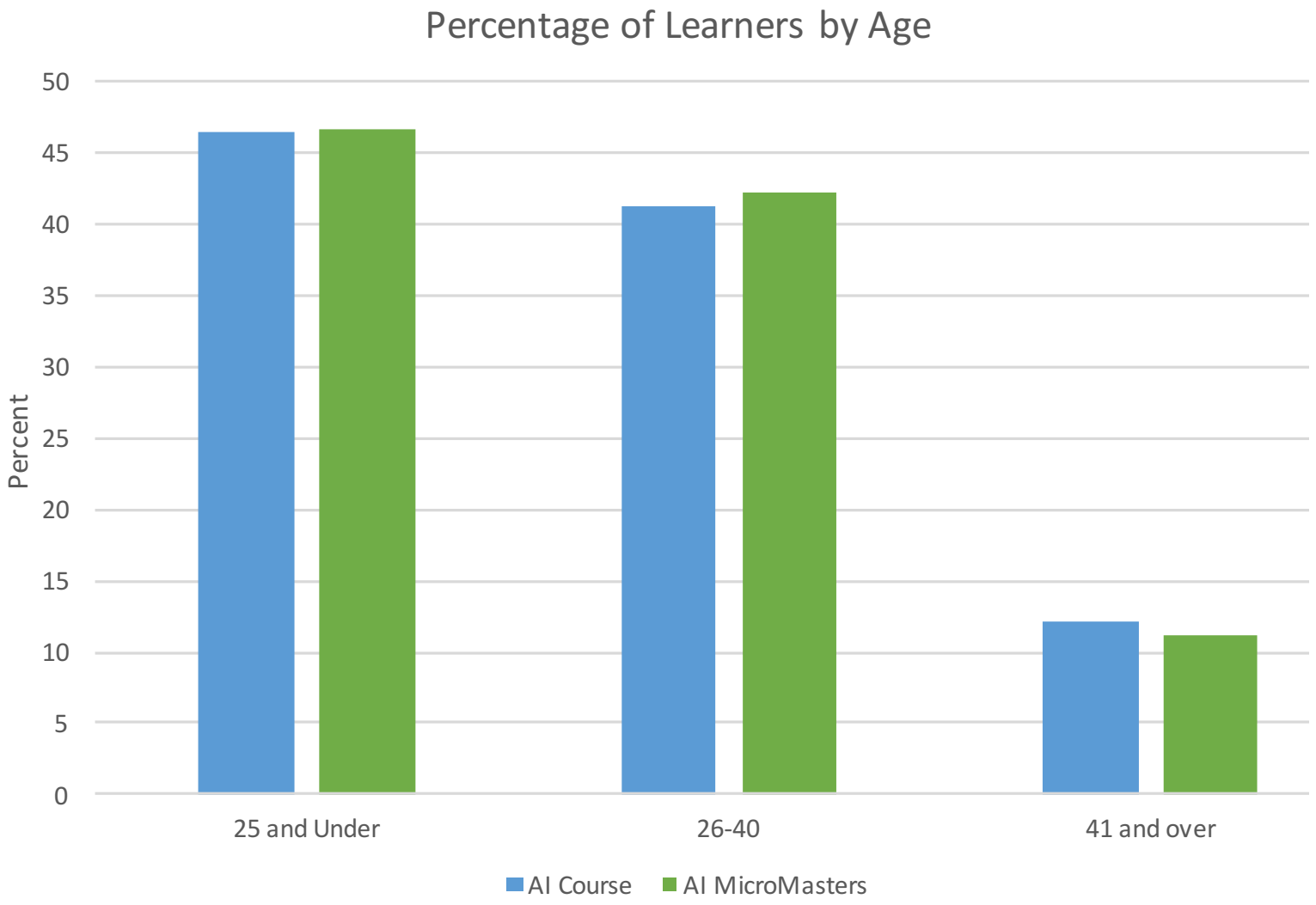
Courtesy Columbia Video Network

Develop AI knowledge



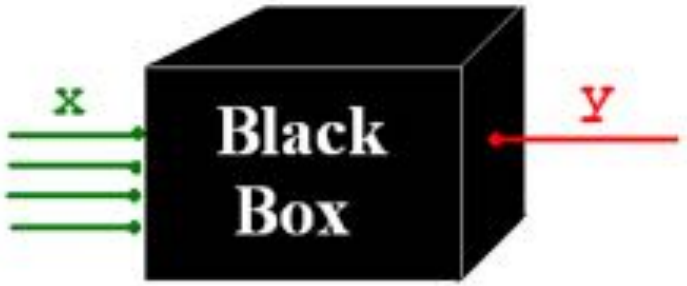
Courtesy Columbia Video Network

Develop AI knowledge



Courtesy Columbia Video Network

Decipher models



- Many of the best machine learning algorithms (e.g., SVMs, Neural Nertworks, Random Forests) produce black box models

- Being able to decipher models, or devise intelligible, interpretable, transparent, understandable models can help:
  - detect bias and fix the model
  - understand decisions
  - communicate/explain predictions to other concerned parties
  - bridge the gap between AI practitioners and consumers

Decipher models

- Explainability or interpretability represents a research opportunity for machine learning
- An emerging research topic in machine learning but it is hard to quantify the criteria of interpretability
  - *Rationalizing Neural Predictions*, Lei, Barzilkay and Jaakola 2016
  - *Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission*, Caruana et al., 2015
  - *Discovering Characterization Rules from Rankings*, Salleb-Aouissi et al. 2009

Decipher models

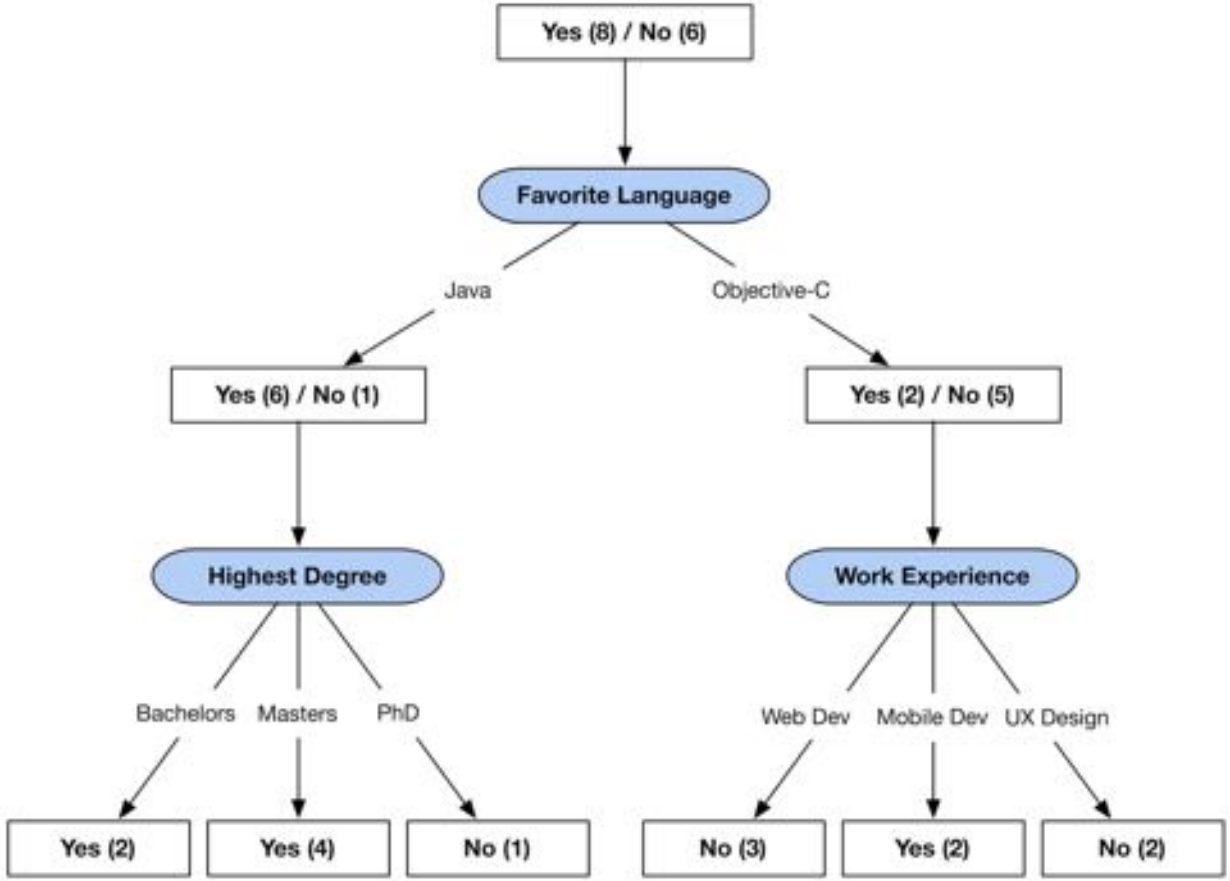
- "European Union regulations on algorithmic decision-making and a *right to explanation*" Goodman and Flaxman, 2016.
- The General Data Protection Regulation (GDPR), agreed upon by the European Parliament and Council in April 2016 includes the right of citizens to receive an explanation for algorithmic decisions will take effect in **Mid 2018**.
- Despite the growing literature there is no rigorous framework of interpretability. "Towards a Rigorous Science of Interpretable Machine Learning", Doshi-Velez and Kim, 2017.

Decipher models

- What an explication should look like? How complex should it be?

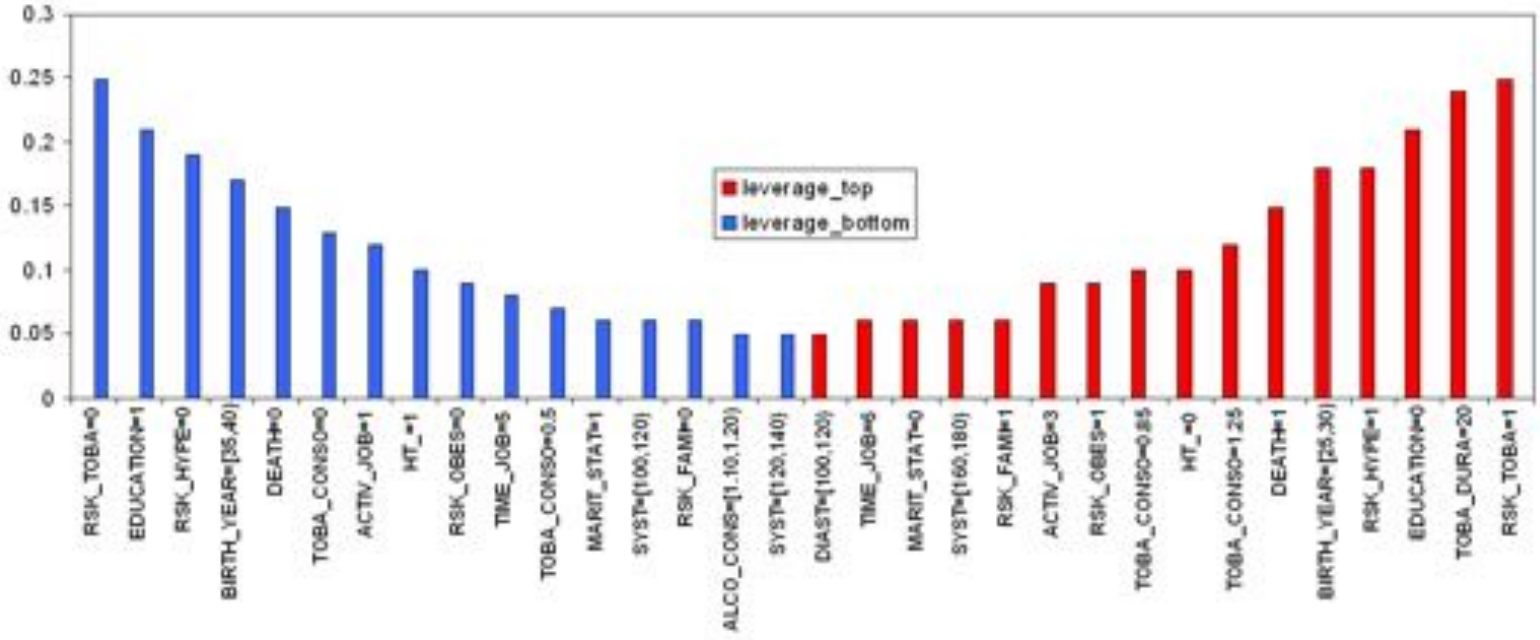
Decipher models

- What an explication should look like? How complex should it be?



Decipher models

- What an explication should look like? How complex should it be?



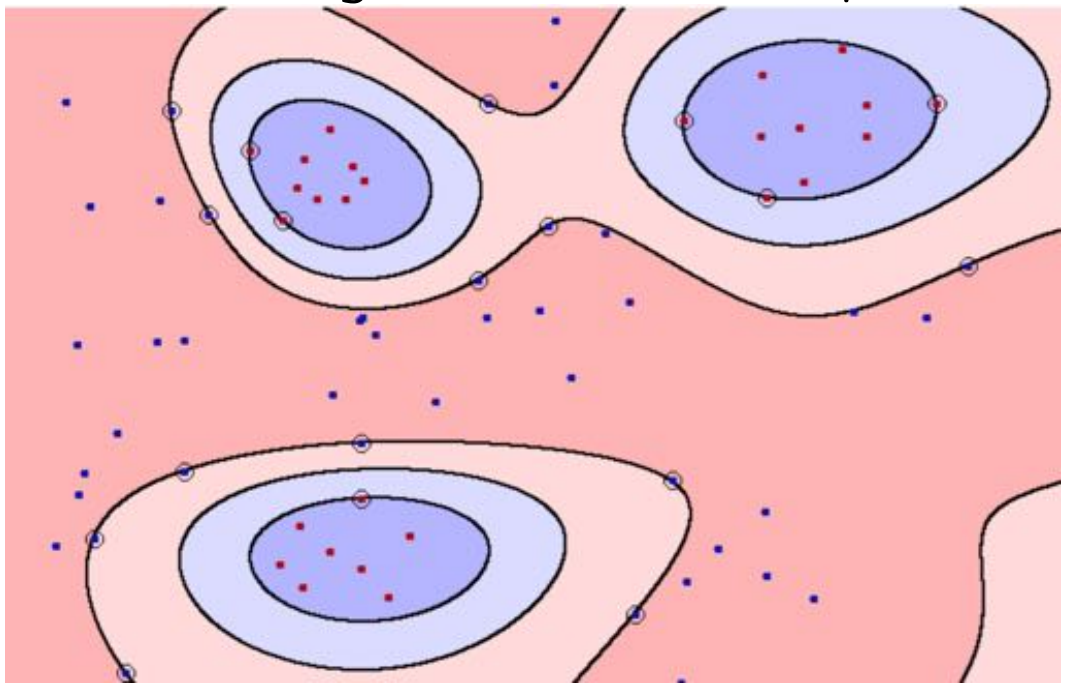
1. *support = 49 (32%), confidence = 96 % : Class = Iris-setosa --> petal_length in [1.1; 1.9]*
2. *support = 48 (32%), confidence = 94 % : Class = Iris-setosa --> petal_width in [0.1; 0.4]*
3. *support = 45 (30%), confidence = 90 % : Class = Iris-versicolor --> petal_length in [3.3; 4.8]*
4. *support = 48 (32%), confidence = 96 % : Class = Iris-versicolor --> petal_width in [1.0; 1.6]*
5. *support = 48 (32%), confidence = 94 % : Class = Iris-setosa --> petal_length in [1.0; 1.9] AND petal_width in [0.1; 0.4]*
6. *support = 41 (27%), confidence = 82 % : Class = Iris-versicolor --> petal_length in [3.3; 4.7] AND petal_width in [1.0; 1.5]*

Decipher models

- What an explication should look like? How complex should it be?
- What machine learning method for interpretable models?

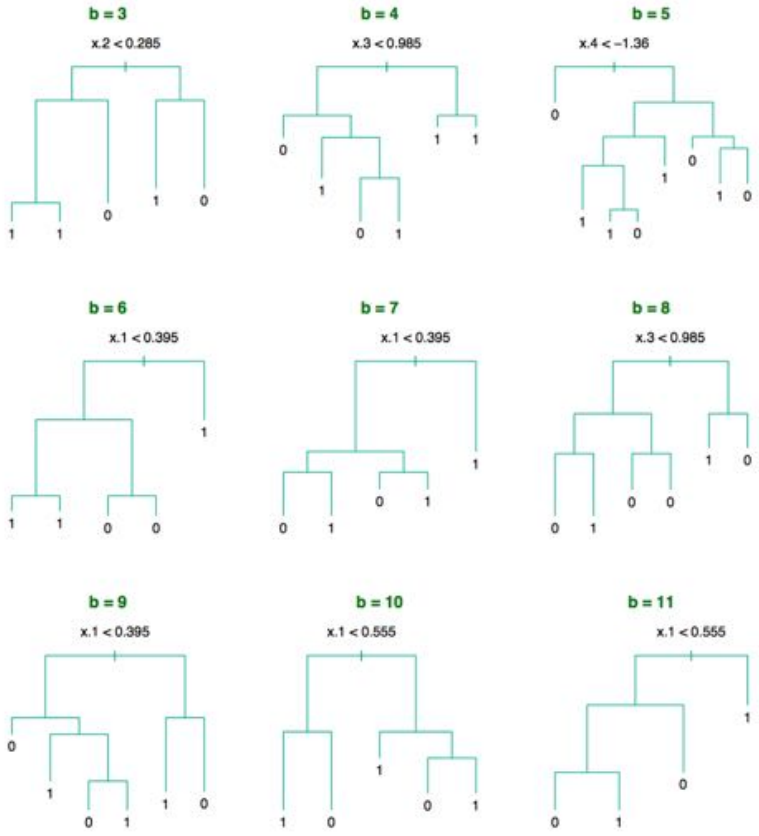
Decipher models

- What an explication should look like? How complex should it be?
- What machine learning method for interpretable models?



Decipher models

- What an explication should look like? How complex should it be?
- What machine learning method for interpretable models?



Decipher models

- What an explication should look like? How complex should it be?
- What machine learning method for interpretable models?
- Should interpretability come at the cost of accuracy? Will interpretability prevent the use of complex models?
- Should interpretability be learned at the same time the model is built, or should we build a model and then decipher it?
- Interpretability now versus long term (e.g., reason for refusing a loan vs. advancing medical research and science)

De-identify

- Do we have control of our own data?
- "The right to be forgotten" as mentioned in the GDPR.
- Avoid profiling, labeling and social exclusion.
- Protect people's privacy.
- Challenging with the web, and different data types.

De-identify

- Protected features (e.g., race, age, gender) can be revealed by all kind of data:
  - Facebook "likes" reveal personal attributes
  - Facial recognition can detect private information
  - Writing can reveal your gender, ethnicity.
- This means deleting personal identifiers is not enough. The information is embedded in other forms and revealed to the world!
- De-identifying is a complex task.

De-bias models

- Automated decision making is common in recommendation systems, credit scoring, job hiring, etc.
- Decisions rely on predictive models that are as fair and unbiased as the data they were trained on.
- Data can be biased, incomplete and even include past discrimination decisions and ML will reproduce it.
- Leads to the **digital** discrimination (Wihbey, 2015) of members of underrepresented groups.

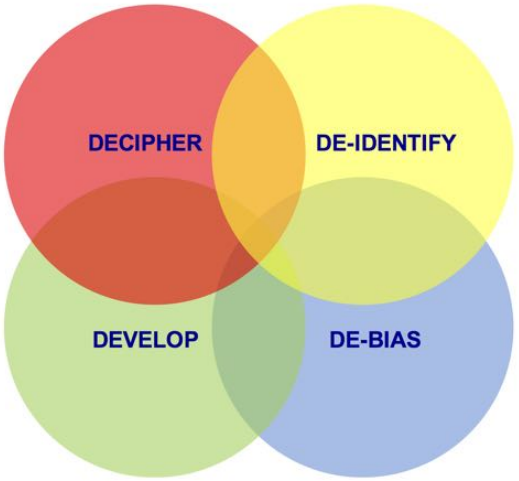
De-bias models

- What is being protected? race, ethnicity, disability, age, gender, religion, sexual orientation, nationality, obesity, etc.
- Discriminatory decisions can occur in access to employment, education, social protection, services.
- Discrimination-aware machine learning models aim to detect bias and prevent it.
  - *The possibilities of digital discrimination*, Wihbey, 2015
  - *A survey on measuring indirect discrimination in machine learning*. Zliobaite, 2015
- Split the features into regular and protected
- Deploys statistical tests to determine the presence of discrimination
- Use discrimination measures like mean difference, mutual information to indicate the magnitude/spread of the discrimination.

Summary

- AI is a flourishing, exciting and broad field with high impact on humanity and society.
- Trend today: Machine Learning, deep learning, reinforcement learning, complex models, and natural language understanding.
- The potential of AI is amazing but challenging from an inclusion perspective.

Summary



- AI and inclusion: Lot more work to do to include the four dimension in the learning process.
  - **Methods** are so different and vary from linear to non linear, from discriminative to probailitstic methods.
  - **Data** is different: structured, images, text, or all of them.
  - There is a lack of **consensus** on how to quantify the criteria of inclusion and how to optimize ML models including those.

Credit

- Artificial Intelligence, A Modern Approach. Stuart Russell and Peter Norvig. Third Edition. Pearson Education. http://aima.cs.berkeley.edu/
- Preparing for the Future of Artificial Intelligence. Executive Office of the President, National Science and Technology, Council Committee on Technology. October 2016.
- Computing Machinery and Intelligence. Alan Turing, 1950. (available here http://loebner.net/Prizef/TuringArticle.html)
- AI and Inclusion: Global Symposium, An evolving reading list. 2017.
- There is a blind spot in AI research. 2016.
- European Union regulations on algorithmic decision-making and a *right to explanation*" Goodman and Flaxman, 2016.
- Artificial Intelligence. The road ahead in low and middle-income countries, June 2017.

Credit

- Artificial Intelligence (AI) and the evolution of digital divides, Andres Lombana Bermudez, July 2017.
- WEF on economic inclusion, Artificial intelligence could help reverse latin america's economic slowdown
- Malavika Jayaram discussing geographic challenges/opportunities of AI
- The possibilities of digital discrimination, Wihbey, 2015
- A survey on measuring indirect discrimination in machine learning. Zliobaite, 2015
- Big data's disparate impact, Barocas, and Selbst, 2016
- Discovering Characterization Rules from Rankings, Salleb-Aouissi et al. 2009
- Rationalizing Neural Predictions, Lei, Barzilkay and Jaakola 2016
- Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission, Caruana et al., 2015