ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

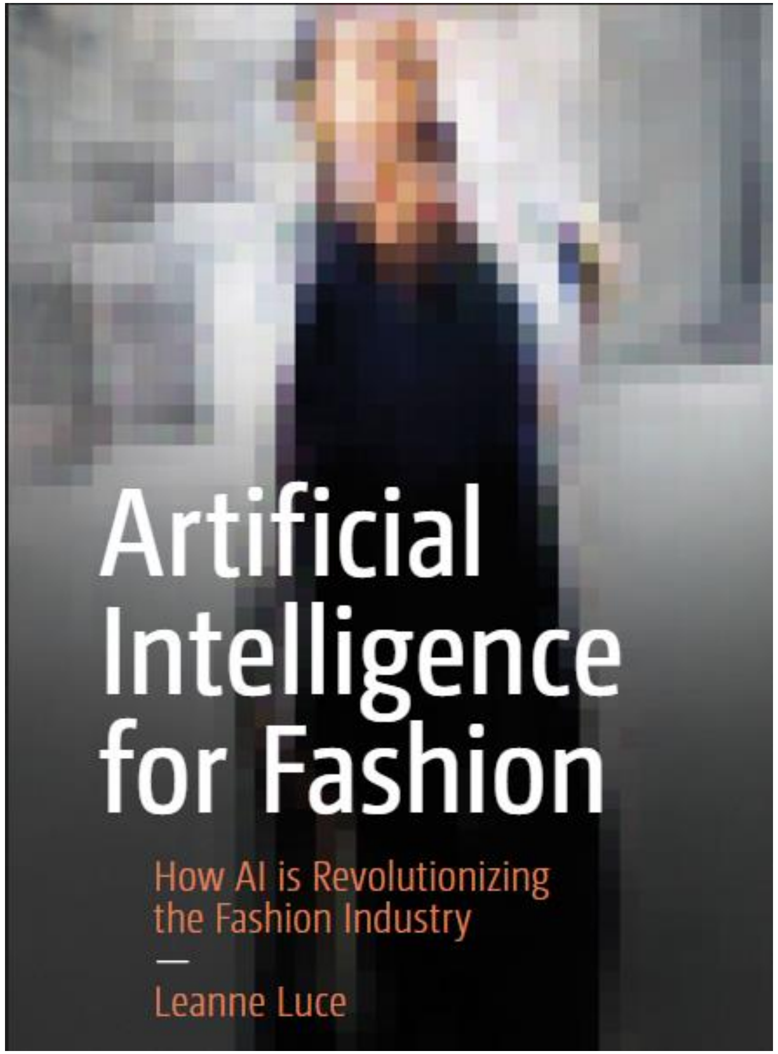
AI for fashion

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

Main textbook



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Syllabus

- Ai in fashion: an impact perspective
- Definitions of AI
- History of AI;
- AI technologies in the fashion field;
- Where we are and where we are going

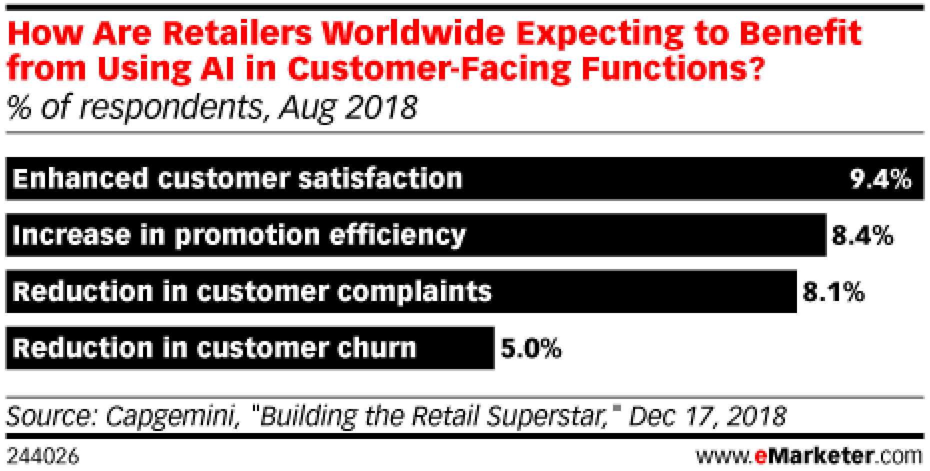
An impact perspective

- Most people grew up thinking that artificial intelligence is something we would only see in movies;
- As it turns out, AI helps more than we could have ever thought;
- We as humans make mistakes in our day-to-day tasks, also in the fashion industry.
- This could impact several Fashion activities: **Manufacturing, Design, Merchandising and Inventory**;
- Fashion and artificial intelligence is a new partnership that is sure to do nothing but succeed for years to come.

Example of garments retrieval [1]

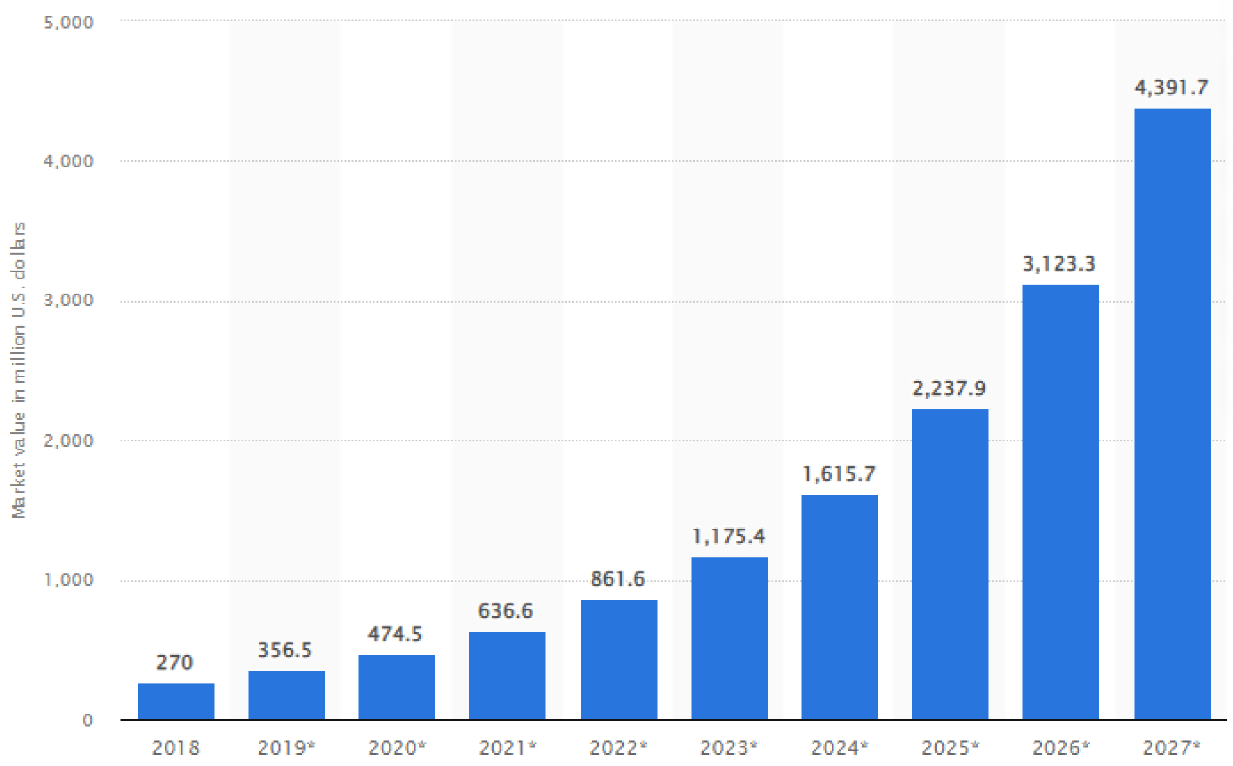
Example of virtrual try-on [1]

- AI will for sure impact the way consumers relate to fashion;
- From garments retrieval to virtual try-on, AI could speed-up the buying and selling process;
- Adopting Extended Reality Techniques, this process could also be not only faster but also more effective;



Market value of artificial intelligence in fashion worldwide from 2018 to 2027

(in million U.S. dollars)



Source: https://www.statista.com/statistics/1070736/global-artificial-intelligence-fashion-market-size/

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

- According to AI in Fashion Market Research Report 2021 [2], under the cumulative impact of COVID-19, global spending on AI in the fashion market is expected to grow from USD 229 million in 2019 to USD 1,260 million by 2024;

- Global expenditure on AI in the fashion market is expected to grow from USD 352.58 million in 2020 to USD 825.19 million by the end of 2025;

- Those economical phenomena have also impacted the number of publications in AI for fashion [1].

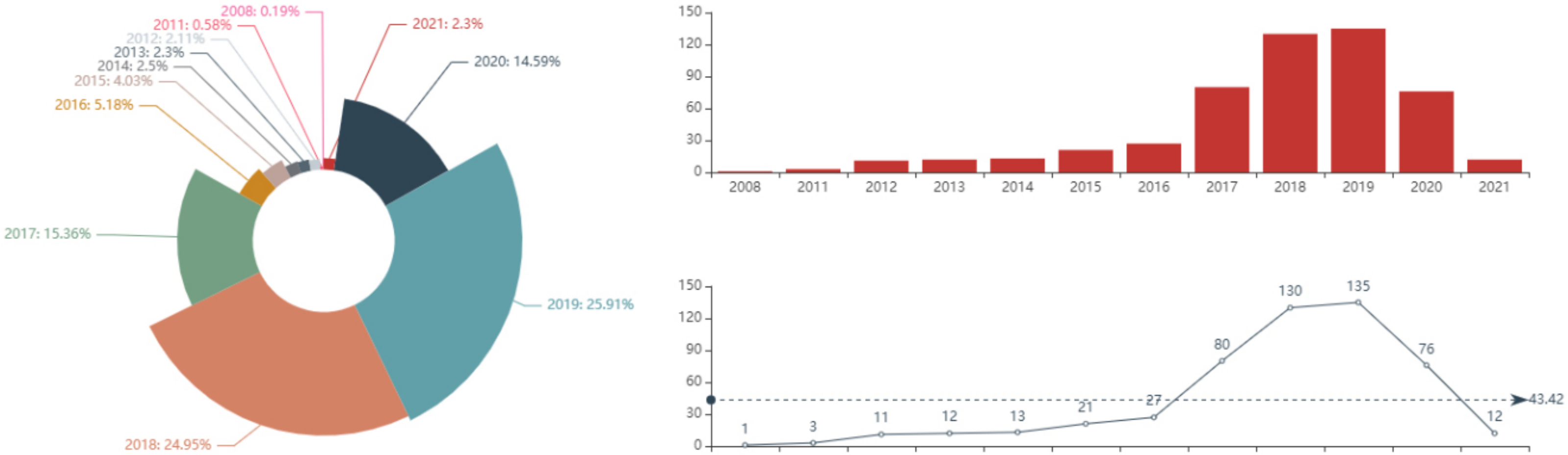


Figure 2: Distribution of published *fAshIon* papers by year. A search identified 521 papers published in the field. Apart from the 12 papers published in 2021 (which is not a final count) and the 76 papers published in 2020, we can see that the overall trend of these papers is growing.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Why AI?

- Simply put, AI allows companies and consumers to make better decisions, improving core processes by increasing strategic decision-making speed and accuracy.

- AI can:
    - (**Automation**) Automate a repetitive task that was previously done manually, without feeling any fatigue or having to take breaks like a human employee would need to do;

    - (**Enhancement**) Make products and services smarter and more effective, improving experiences for end-users, via capabilities like optimizing conversation bots or customer service menus, and delivering better product recommendations;

    - (**Analysis**) Analyze data at a much faster rate than humans, allowing it to find patterns much more quickly, and it can also analyze much larger datasets than humans, allowing it to uncover patterns humans would simply miss;

    - (**Accuracy)** Become more accurate than humans, utilizing its ability to harvest and interpret data to provide better decisions for many tasks (also in fashion).

Why now?

- **More computing power**: AI requires a lot of computing power, recently many advances have been made and complex deep learning models can be deployed and one of the great technology that made this possible are GPUs;

- **More data**: In order to have a useful AI-agent to make smart decisions like telling which item to recommend you next when shopping online or classifying an object from an image, AI are trained on large data sets and big data enables us to do this more efficiently;

- **Better algorithms**: State of the art algorithms, most of them based on Neural networks (NN) are constantly getting better results;

- **Broad investment**: universities, governments, startups, and tech giants (Google, Amazon, Facebook, Baidu, Microsoft) are all investing heavily in AI;

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

What is Artifical Intelligence?

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

What is Artificial intelligence?

- In this Figure, taken from [3], we see eight definitions of AI, two of each lies in a particular approach at the intersection of two dimensions;

	Fidelty to Human performance	Idea performance measure
Reasoning	**Thinking Humanly** "The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	**Thinking Rationally** "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Behaviour	**Acting Humanly** "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	**Acting Rationally** "Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

- Historically, all four approaches to AI have been followed, each by different people with different methods;

- A human-centered approach must be in part an empirical science, involving observations and hypotheses about human behavior;

- A rationalist approach involves a combination of mathematics and engineering.

Acting humanly: The Turing Test approach

- **The Turing Test**, proposed by Alan Turing (1950), was designed to provide a satisfactory operational definition of intelligence. A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer;

- The computer would need to possess the following capabilities:

  - **Natural language processing** (NLP) to enable it to communicate successfully in English;
  - **Knowledge representation** to store what it knows or hears;
  - **Automated reasoning** to use the stored information to answer questions and to draw new conclusions;
  - **Machine learning** (ML) to adapt to new circumstances and to detect and extrapolate patterns;

- However, the total Turing Test includes physical simulation that so need:

  - **Computer Vision** (CV) to perceive objects;
  - **Robotics** to manipulate objects;

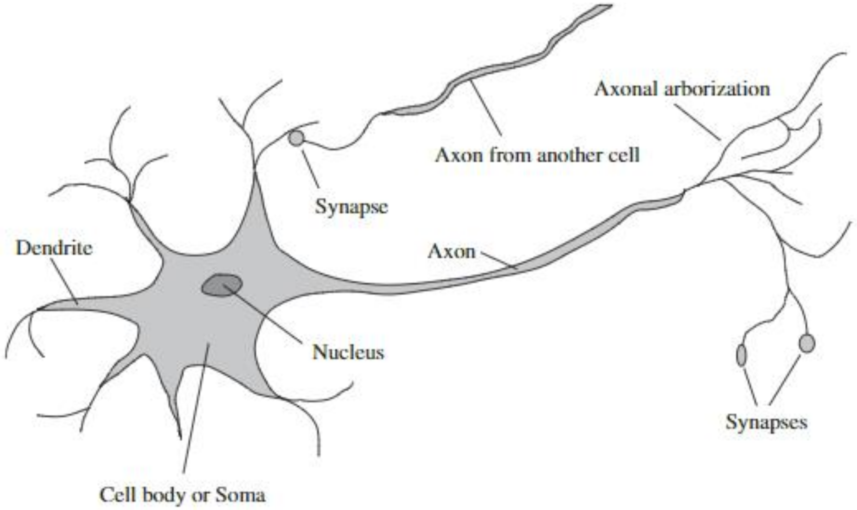
Thinking humanly: The cognitive modeling approach

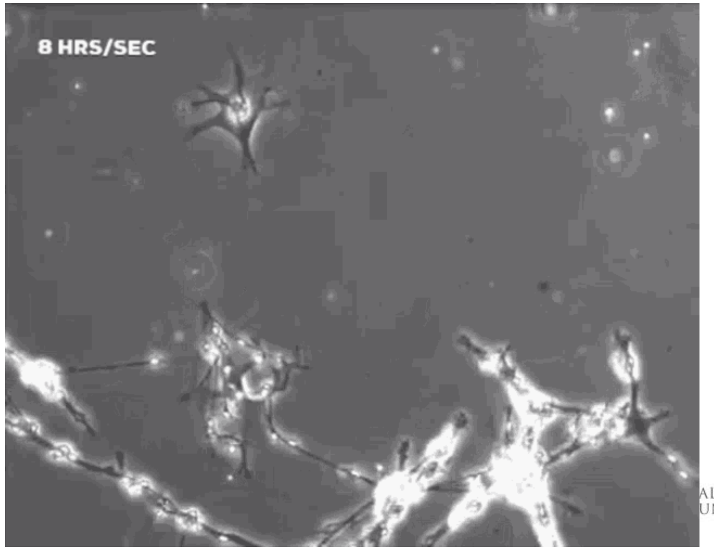
- **If we are going to say that a given program thinks like a human, we must have some way of determining how humans think;**

- The interdisciplinary field of **cognitive science** brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind;

- Clearly, a lot of AI is based on human cognition, which mechanisms were and are revealed by **cognitive/neuroscientists**;

- For example, Neural networks, were formalized taking inspiration from real neuron circuits!





Thinking rationally: The “laws of thought” approach

- The Greek philosopher Aristotle was one of the first to attempt to codify “right thinking,” that is, irrefutable reasoning processes;

- His syllogisms provided patterns for argument structures that always yielded correct conclusions when given correct premises—for example, “Socrates is a man; all men are mortal; therefore, Socrates is mortal.” These laws of thought were supposed to govern the operation of the mind; their study initiated the **field called logic;**

- Logicians Computer Scientist from the 20^{th} formalized a programming paradigm to solve any solvable problem described in logical notation;

- However, it is not easy to take informal knowledge and state it in the formal terms required by logical notation, particularly when the knowledge is less than 100% certain;

- Moreover, there is a big difference between solving a problem “in principle” and solving it in practice. Even problems with just a few hundred facts can exhaust the computational resources of any computer unless it has some guidance as to which reasoning steps to try first;

- Nevertheless, the logic programming was one of the first of real-world use of AI;

Acting rationally: The rational agent approach

- **An agent is just something that acts;**

- Of course, all computer programs do something, but computer agents are expected to do more: operate autonomously, perceive their environment, persist over a prolonged period, adapt to change, and create and pursue goals.

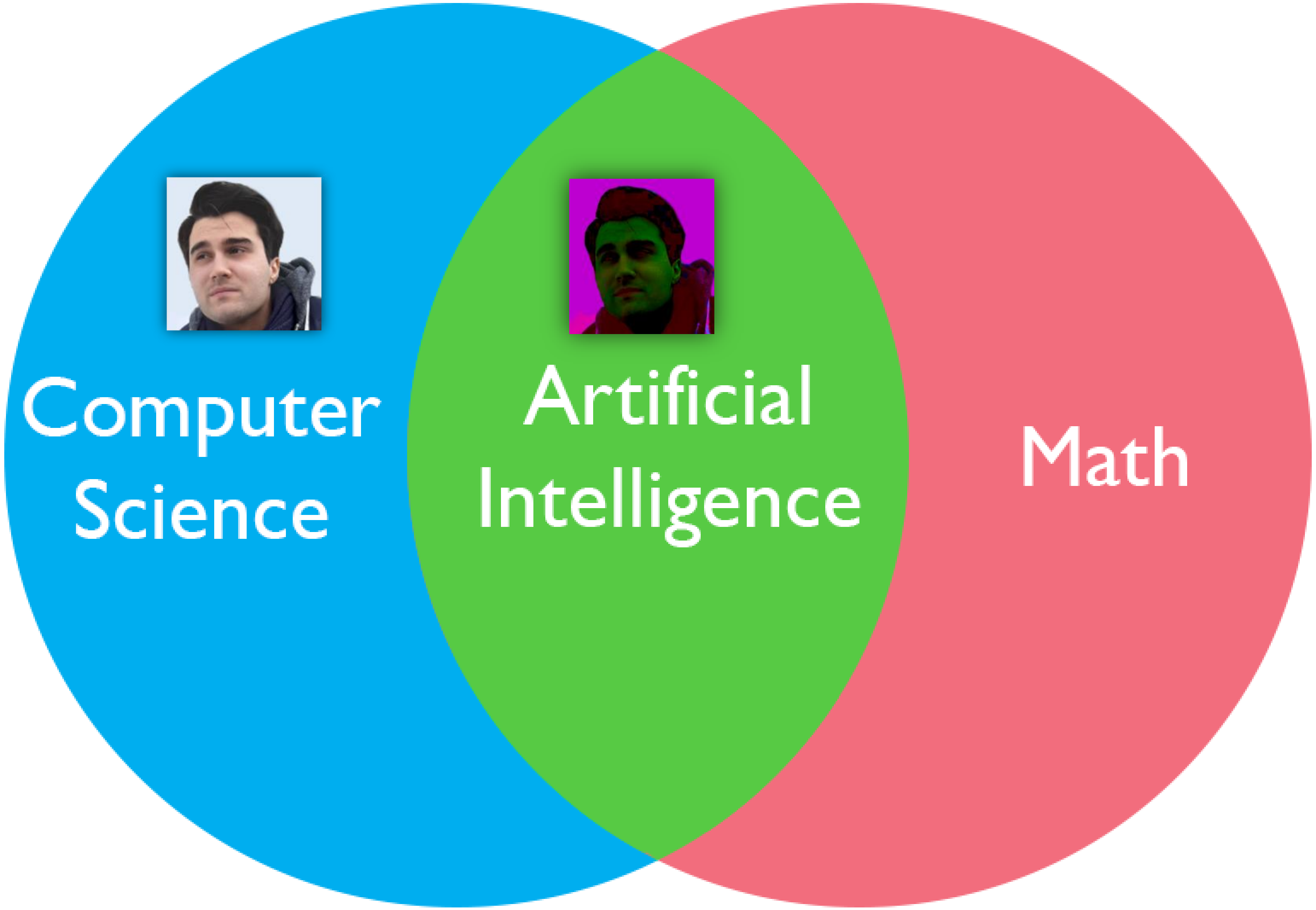
- The rational agent acts to achieve the best outcome or when there is uncertainty, the best-expected outcome;

- 1st advantage: All the skills needed for the Turing Test also allow an agent to act rationally;

- 2nd advantage: Easier to develop from a scientific perspective with respect to generalize processes like human behavior;

	Fidelty to Human performance	Idea performance measure
Reasoning	**Thinking Humanly** “The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense.” (Haugeland, 1985) “[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . .” (Bellman, 1978)	**Thinking Rationally** “The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985) “The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)
Behaviour	**Acting Humanly** “The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990) “The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)	**Acting Rationally** “Computational Intelligence is the study of the design of intelligent agents.” (Poole *et al.*, 1998) “AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)

What is Artificial intelligence in practice?



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

History of AI and Intelligent agents

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

Brief History of AI

- AI was a term first coined at Dartmouth College in 1956 when the cognitive scientist Marvin Minsky was optimistic about this technology's future;

- From 1957 to 1974, AI flourished. Computers could store more information and become faster, cheaper, and more accessible. Machine learning algorithms also improved, and people got better at knowing which algorithm to apply to their problem;

- Breaching the initial fog of AI revealed a mountain of obstacles. The biggest was the lack of computational power to do anything substantial: computers simply couldn’t store enough information or process it fast enough;

Source: Harvard Review

- In the 1980s, AI was reignited by two sources: an expansion of the algorithmic toolkit, and a boost of funds.

- John Hopfield and David Rumelhart popularized “**deep learning**” techniques which allowed computers to learn using experience;

- On the other hand, Edward Feigenbaum introduced **expert systems** which mimicked the decision-making process of a human expert. The program would ask an expert in a field how to respond in each situation, and once this was learned for virtually every situation, non-experts could receive advice from that program.

- The Japanese government heavily funded expert systems and other AI-related endeavors with $400 million dollars with the goals of revolutionizing computer processing, implementing logic programming, and improving artificial intelligence.

- Despite the absence of government funding and public hype, AI keeps thriving. In 1997, reigning world chess champion and grandmaster Gary Kasparov was defeated by IBM’s Deep Blue [4], a chess-playing computer program.

- We haven’t gotten any smarter about how we are coding artificial intelligence, so what changed?

- In the video, the developers of DeepBlue called the combination testing “brute force”… what is brute force?

- In computer science, **brute-force search or exhaustive search**, also known as generate and test, is a very general problem-solving technique and algorithmic paradigm that consists of systematically enumerating all possible candidates for the solution and checking whether each candidate satisfies the problem's statement;

- It turns out, the fundamental limit of computer storage that was holding us back 30 years ago was no longer a problem;

- Moore’s Law, which estimates that the memory and speed of computers double every year, had finally caught up and in many cases, surpassed our needs;

- This is precisely how Deep Blue was able to defeat Gary Kasparov in 1997, and how Google’s Alpha Go was able to defeat Chinese Go champion, Ke Jie in 2017;


ALPHAGO

Nowadays… What else changed? All started from an agent

- DeepBlue and DeepMind were able to succeed because hardware performance have speeded up… but not only;
- Indeed, these AI models developed as a model and utility-based intelligent agents;
- An intelligent agent could be viewed as an entity that: "**Operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change, and create and pursue goals**";
- A model-based agent considers an internal state;
- A utility-based agent consider a utility function which it tries to maximize (i.e., a goal to reach);

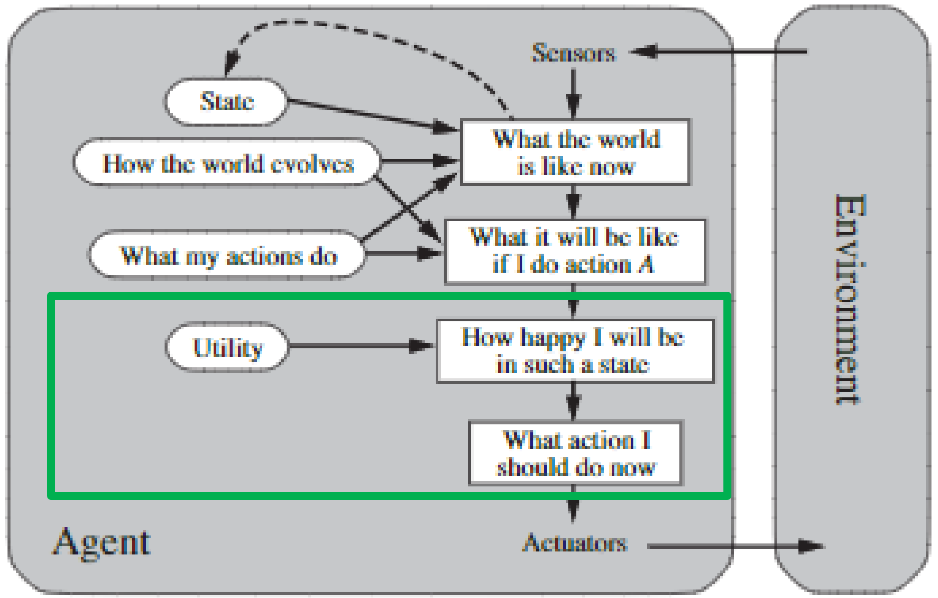


Figure 2.14 A model-based, utility-based agent. It uses a model of the world, along with a utility function that measures its preferences among states of the world. Then it chooses the action that leads to the best expected utility, where expected utility is computed by averaging over all possible outcome states, weighted by the probability of the outcome.

Example of intelligent agents in our everyday lives?

AI assistants, like Alexa, are examples of intelligent agents as they use sensors to perceive a request made by the user and automatically collect data from the internet without the user's help.

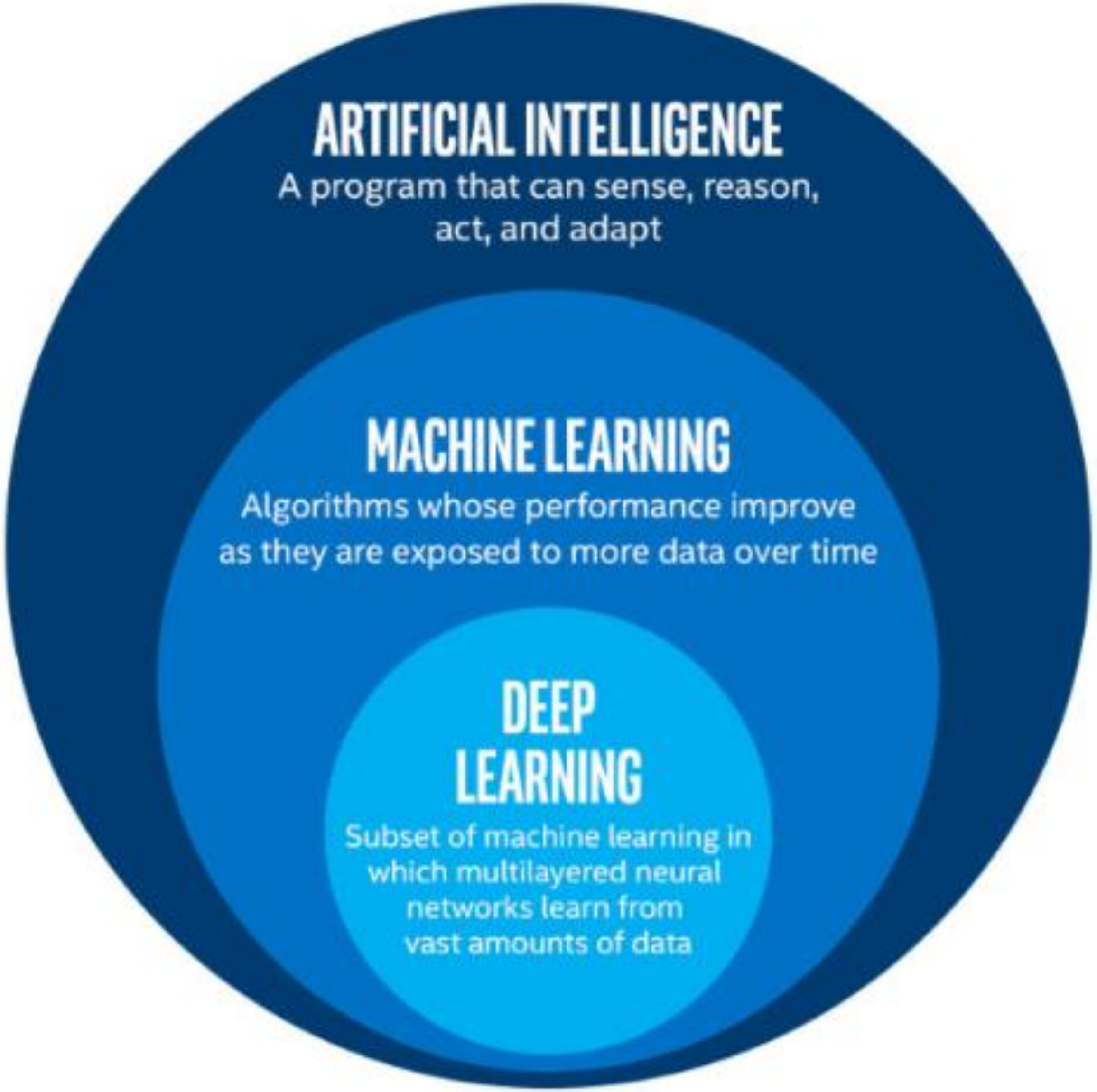
Autonomous vehicles could also be considered intelligent agents as they use sensors, GPS and cameras to make reactive decisions based on the environment to maneuver through traffic.

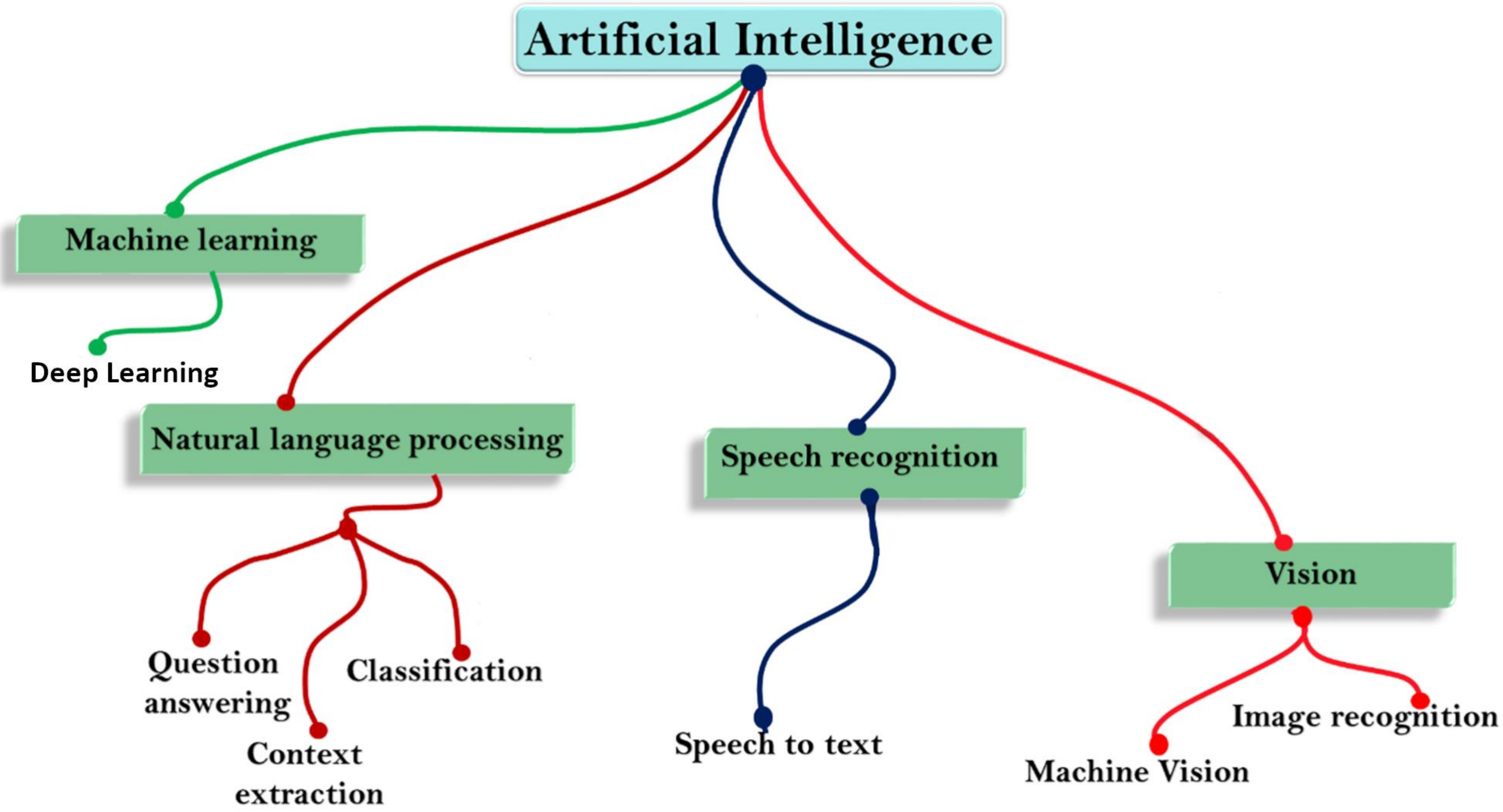
Can you think of an intelligent agent for fashion?

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

- DeepBlue and AlphaGo were able to perform at best since they could explore all the problem-space;
- Think about a chessboard: you know all the positions of different pieces now, and you could simulate thousand moves and the resulting configuration… this because you already know all the rules of the game!
- But what if you don't have rules? What if you have just a set of pieces without know what they correspond to?
- Well … the agent has to learn them from experience (i.e., data)! This is one of the basic concept behind Machine Learning (**pattern recognition**);
- Alexa is an **intelligent** and **learning** agent;
- We will see in few slides why a machine learning algorithm is what makes an agent «learning and intelligent» for real-world task;
- Before that, it is important to distinguish terms like AI and Machine Learning;

Artificial Intelligence vs Machine Learning




Artificial Intelligence
Machine learning
Deep Learning
Natural language processing
Question answering
Context extraction
Classification
Speech recognition
Speech to text
Vision
Image recognition
Machine Vision

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Why is Big Data mainstream?

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

What is Big data? *[5]*

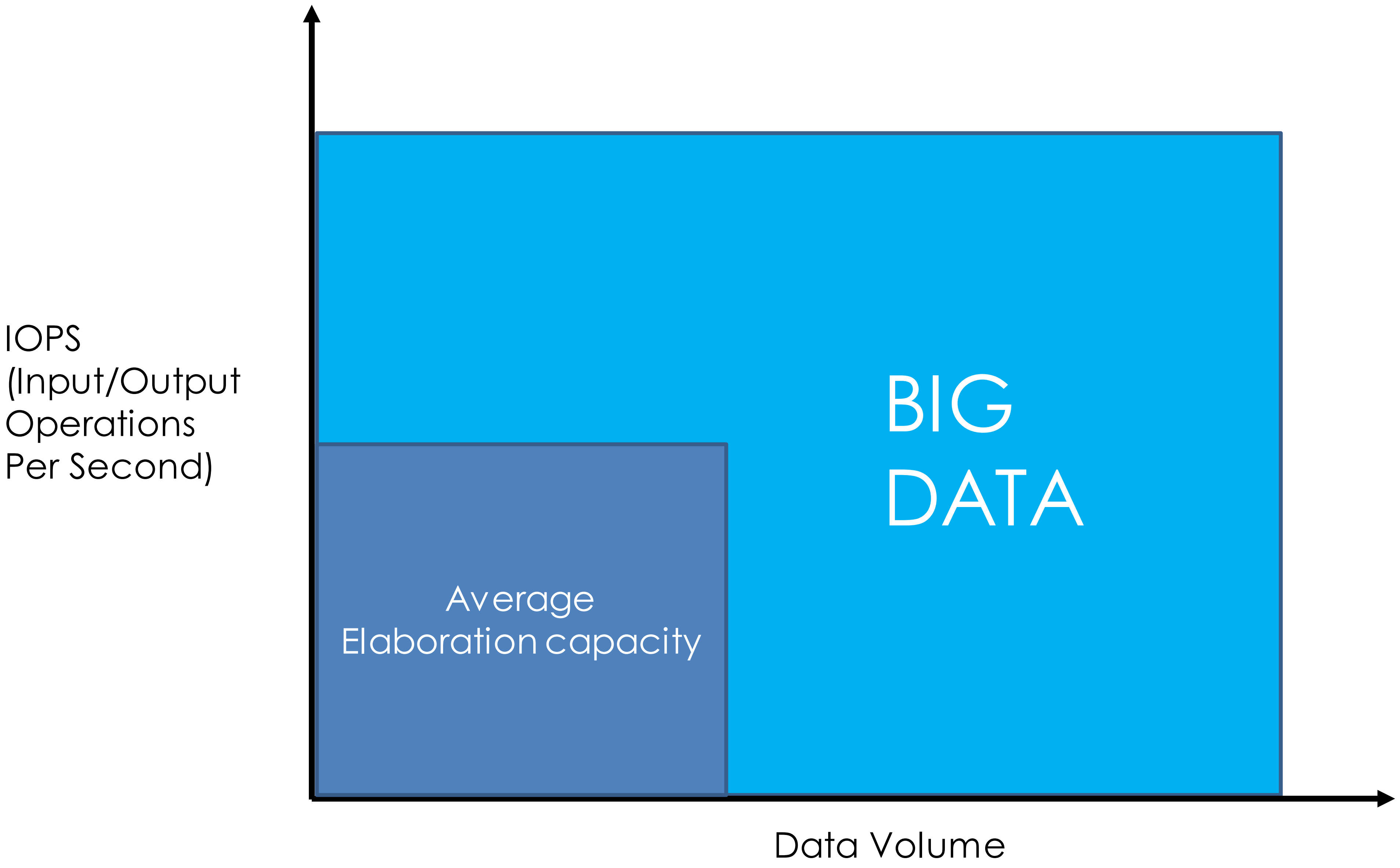
- The term Big data refers to a massive data set with a larger, more varied, and more complex structure than normal data.

- This level of bulk produces storage, analysis, and visualization difficulties for processes running on them.

- Researching huge amounts of data and discovering patterns and correlations between them is often called big data analysis.

- The data and results that we can draw from them are fundamental both for companies (advantage over the competition) but also for the advancement of research in various fields of science (biology, chemistry, art etc.)

Quando i data diventano big?



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Le 4 V dei big data

- **Volume:**
  - Big data is huge!
- **Speed:**
  - Big data is produced fast and needs to be processed fast!
- **Variety:**
  - Big data has a different nature: structured, unstructured, multimedia data.
- **Veracity:**
  - Big data could contain unreliable data sources.
  - For example, people's feelings expressed in social media are uncertain by nature, since they involve human judgment (yet they contain valuable information).

Le 8V dei big data

Short history of Big Data

- In 2005 Roger Mougalas from O'Reilly Media coined the term Big Data for the first time, only a year after they created the term Web 2.0;

- The same year of the creation of Hadoop by Yahoo, created on top of one of the most famous paradigms to analyze big data: Google MapReduce;

- In 2010 Eric Schmidt speaks at the Techonomy conference in Lake Tahoe in California and he states that "there were 5 exabytes of information created by the entire world between the dawn of civilization and 2003. Now that same amount is created every two days;

- Today, we create roughly 2.5 quintillion bytes of data (IBM).

How much is 2.5 quintillion of bytes?

- 1 kilobyte is 1000 bytes (a normal text file is measured in kilobytes);
- 1 Megabyte is 1000 kilobytes (an image file is measured in megabyte);
- 1 Gigabyte is 1000 megabytes (your phone should have at least 32GB of memory);
- 1 Terabyte is 1000 gigabytes (size of a pc memory disk);
- 2.5 quintillions of bytes are equal to 2.500.000 di Terabyte?

Where does this data come from?

- **Social data** comes from the Likes, Tweets & Retweets, Comments, Video Uploads, and general media that are uploaded and shared via the world's favorite social media platforms. This kind of data provides invaluable insights into consumer behavior and sentiment and can be enormously influential in marketing analytics.

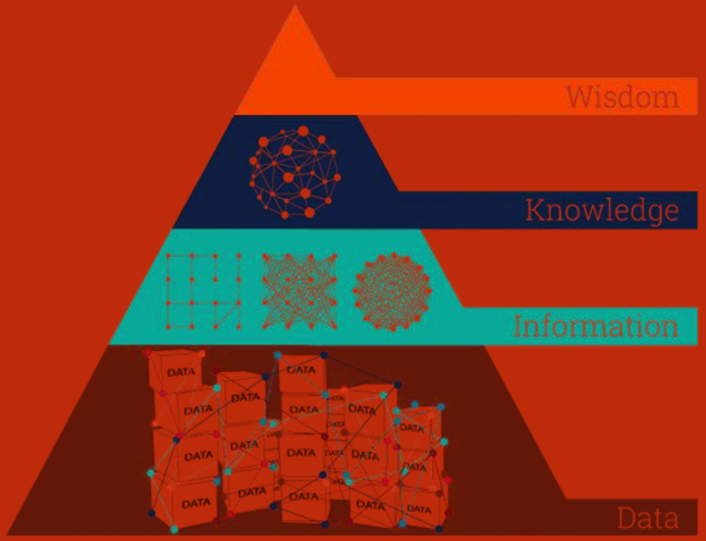
- **The web also generates a lot of data: streaming, user clicks, user subscriptions, user search.**

- **Machine data** is defined as information generated by industrial equipment, sensors installed in machinery, and even weblogs that track user behavior. This type of data is expected to grow exponentially as the **internet of things** grows ever more pervasive and expands around the world. Sensors such as medical devices, smart meters but also games nowadays deliver high velocity, volume, and variety of data very shortly.

- **Transactional data** is generated from all the daily transactions that take place both online and offline. Invoices, payment orders, storage records, and delivery receipts are characterized as transactional data, yet data alone is almost meaningless.

How can we exploit such vawe?

- We talked about intelligent agents;
- We know that, if there aren’t clear rules, the agent itself should learn them;
- A learning agent learns by means of Machine learning algorithms;
- So?
- We must learn how to tame this huge amount of complex data to let our algorithms to learn and improve!



Data, information, knowledge and wisdom

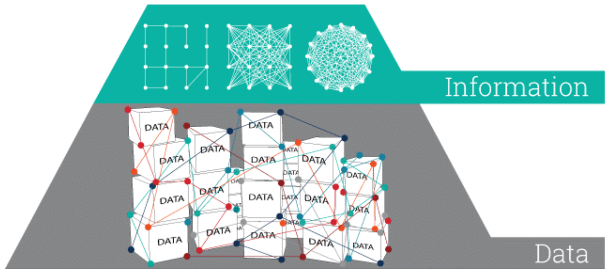
Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

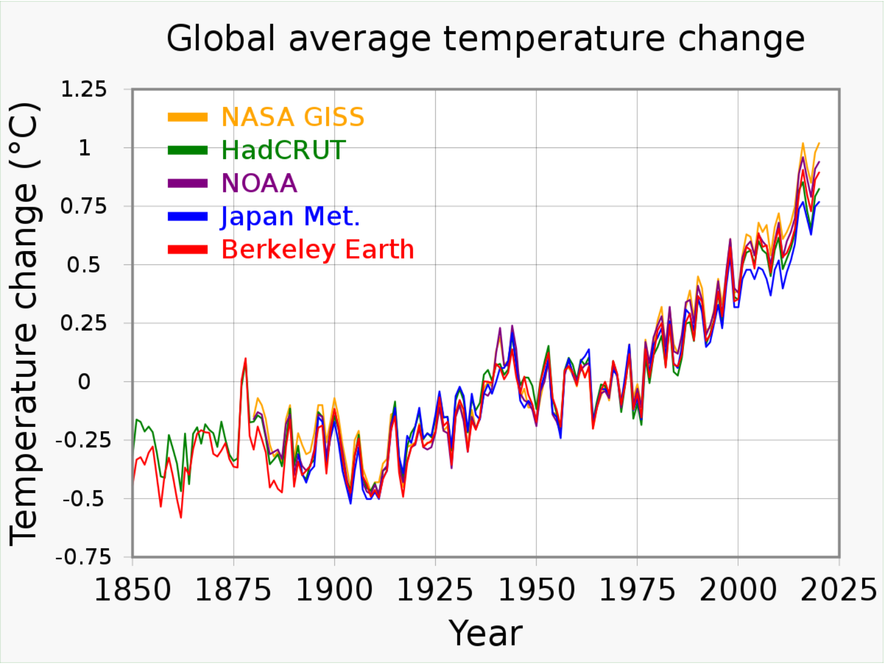
Data, Information, Knowledge and Wisdom



- **Data** is defined as a collection of rough and unorganized facts, such as numbers and characters. Without context, the data makes little sense. For example, 1 is a number, but we can see it as a temperature degree (1 °). We have therefore transformed the data into information!

- **Information** is a "clean and processed" set of data in a way that makes manipulation, analysis, and visualization easier; For example, we could make a graph to analyze the global mean temperature values over the past 170 years!

- By asking pertinent questions about "who", "what", "when", "where", etc., we can derive valuable insights from information by making it more useful. But when we come to the "how" question, we are forced to make the leap from information to knowledge!

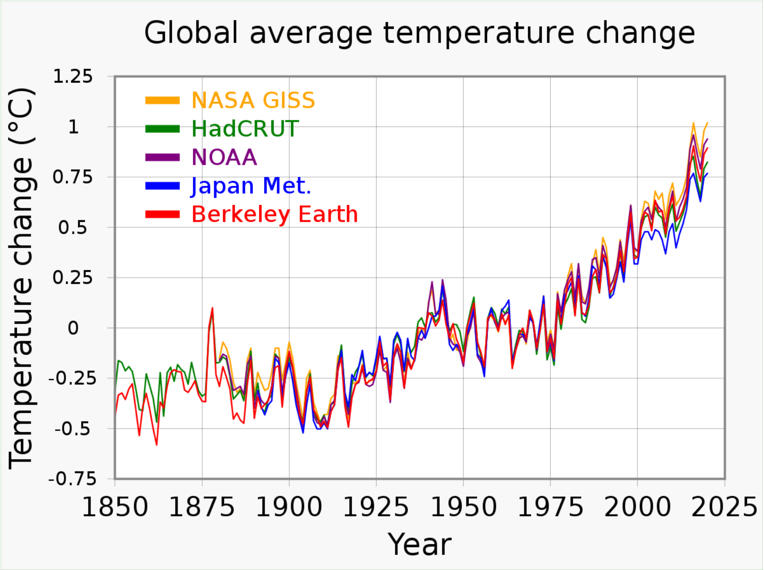


- How is the information derived from the collected data relevant to our objectives? "How" are the pieces of this information linked to other pieces to add more meaning and value? And most importantly, "how" can we apply the information to achieve our goal?

- When we understand how to apply information to achieve our goals, we turn it into knowledge.

- When we use the knowledge and insights gained from information to make proactive decisions, we can say that we have reached the final step of the pyramid: wisdom.

- Wisdom is at the top of the hierarchy and is nothing more than applied knowledge to make the best possible decision!



- Quali sono le cause del riscaldamento globale? **(conoscenza)**

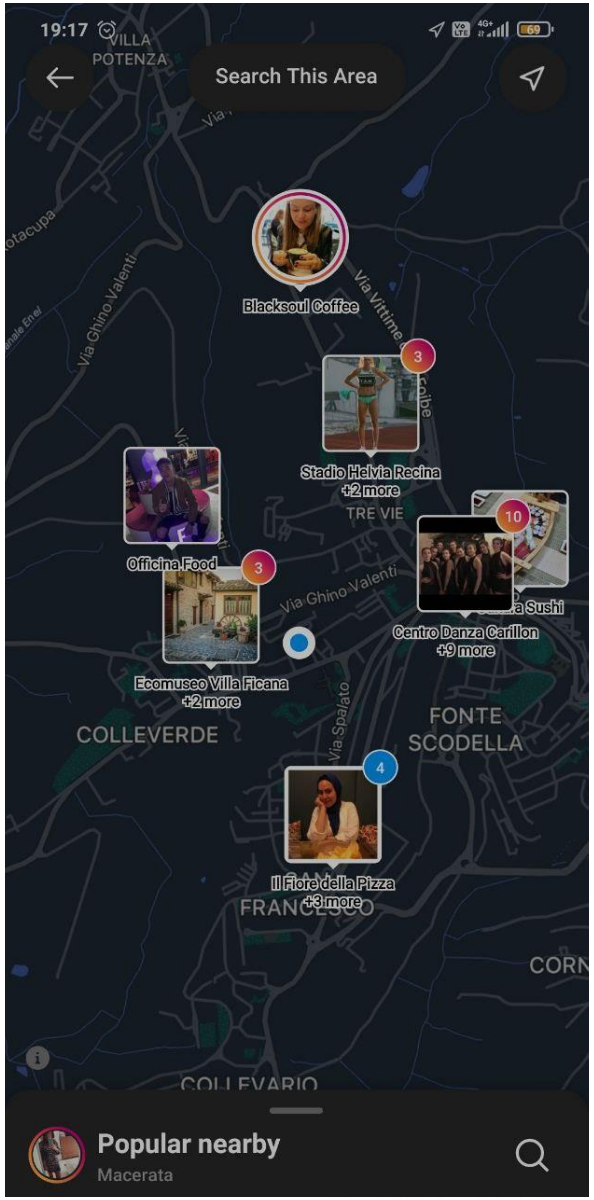
- Cosa possiamo fare per fermarlo? **(saggezza)**

Software manipulate data… and (data) science allows to produce knowledge!

- Any software can be seen as a data manipulator;
- The applications we use every day are data-driven:
  - Social network (instagram);
  - Shopping on amazon;
  - Banking transactions;
  - Digital medical records;
  - …
- Some of these applications process data and information to create knowledge!

Instagram

- A social created to share your best shots!
- When you upload a photo to instagram, dear Zuckerberg not only saves the data related to the photo file but also:
    - Who uploaded the photo;
    - Who has been tagged;
    - When the photo was uploaded;
    - The caption;
    - The hashtags;
    - The place;
    - ...
- With this info, he is able to understand what the user might like, based on where he is! (knowledge)



Amazon

amazon

- Your amazon profile is associated with thousands (if not millions) of information:
    - Personal data (username, password etc ..);
        - Order history:
        - All purchased products;
        - Who sent them;
        - Shipping company;
        - Shipping Times;
    - Do you have the prime? Access files to prime videos, music etc ...
- All this information is used to: predict interesting articles for you, next TV series to watch, next song to listen to ... (knowledge)

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

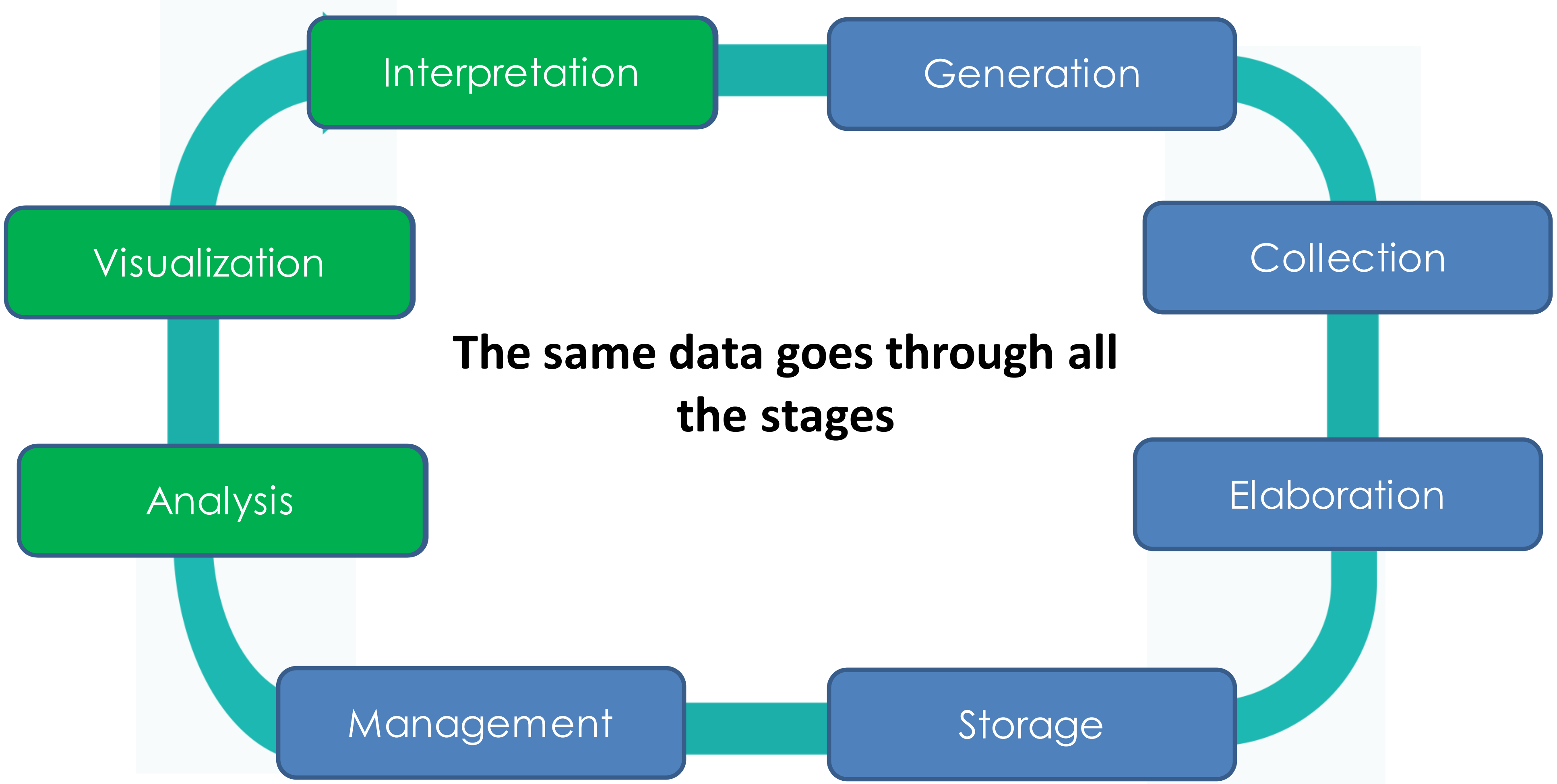
How tame data to reach our goal?

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

The (big) data life cycle



Generation

- For the data lifecycle to begin, the data must first be generated;
- Data generation occurs regardless of whether you are aware of it;
- Some of this data is generated by your school, university, company and / or third parties that you may or may not be aware of;
- Every sale, purchase, rental, communication, interaction: everything generates data;
- The generated data is called raw.

Collection

- However, not all the data generated every day are collected or used;
- Depending on the purpose we want to achieve, we would decide what information to keep;
- Data can be collected in various ways, such as:
  - Automatic collection software;
  - Surveys and Interviews: Surveys and interviews can be an effective way to gather large amounts of information from many subjects;
  - Observation and measurement: observe natural events with the human eye and label them (e.g., normal cells vs cancer cells) or use sensors suitable for the purpose (e.g., measuring temperature and humidity)!
- Understanding which variables to collect becomes essential!

Elaboration

- Once the data has been collected, it needs to be processed. Data processing may concern various activities, including:
    - Data wrangling: process of cleaning and transforming a data from its raw form into something more accessible and usable;
    - Data compression, in which data is transformed into a format that can be stored more efficiently;
    - (possibly) Data encryption, in which data is translated into an encrypted form to protect it.

Storage

- After the data has been collected and processed, it must be archived for future use;
- This process is implemented through the use of databases or data sets;
- These data sets can then be stored in the cloud, on a server or using a custom physical storage solution;
- When determining how best to store data for your organization, it is important to create some level of redundancy to ensure that a copy of the data is protected and accessible, even if the original source is damaged or compromised;

Management

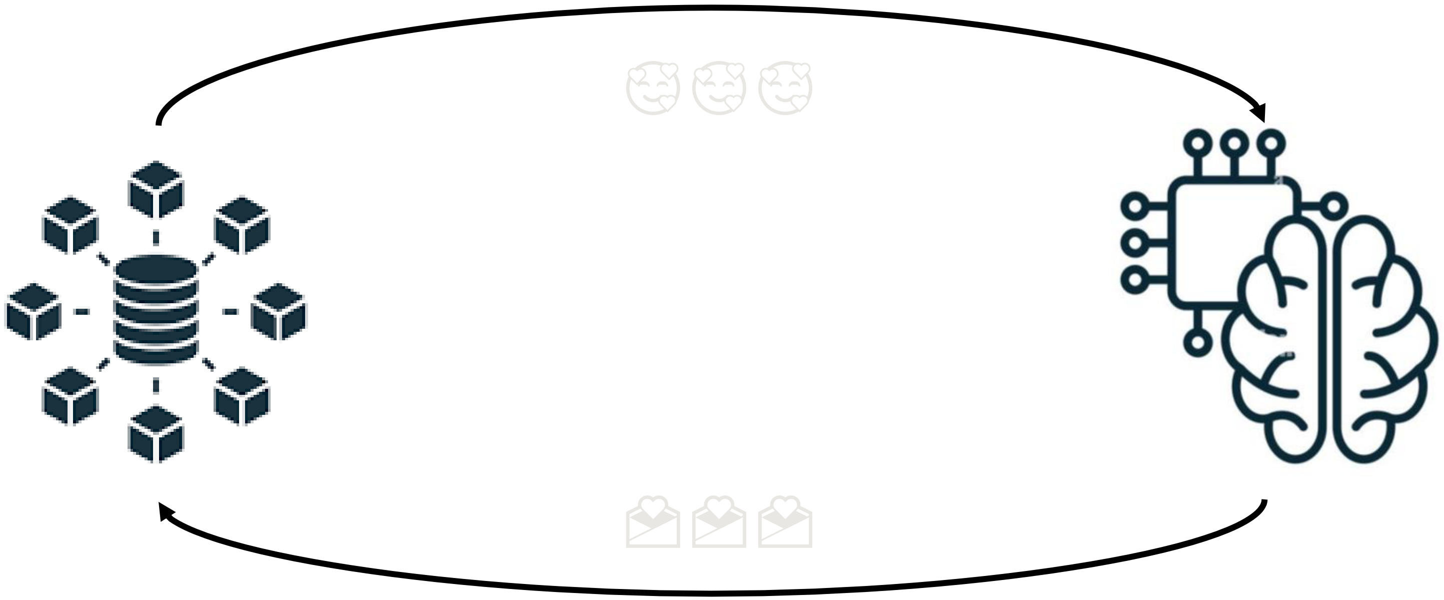
- Data management (database management) involves organizing, storing and retrieving data as needed during the life of a data project.
- Although referred to as a "step" here, it is an ongoing process that takes place from the beginning to the end of the big data lifecycle.
- Data management includes everything:
  - Archiving and (possibly) encryption;
  - Access logs and changes that keep track of who accessed the data and the changes made (so-called log files);
  - ...

Analysis

- Data analysis refers to processes that attempt to gather meaningful information and knowledge from raw data;
- Analysts and data scientists use different tools and strategies to conduct these analyzes;
- Some of the more commonly used methods include statistical models, data mining algorithms, and machine learning;
- The actions to be taken at this stage depend on the challenge we are facing!
- For example, if we wanted to classify the type of flower based on a photo, I certainly wouldn't start studying an Ikea catalog!

(or maybe, yes?)

A love that will probably never stop



- Machine learning algorithms love Big Data;
- This couple is now acclaimed as the leading foundations of the world of tomorrow;
- You will hear about them for a long time.

Visualization

- Data visualization refers to the process of creating graphical representations of information, generally using one or more visualization tools;
- Data visualization makes it easy to quickly communicate your analysis to a wider audience;
- The form that the visualization takes depends on the data you are working with and the story you want to communicate;
- While technically not a mandatory step for all data projects, data visualization has become an increasingly important part of the data lifecycle!

Interpretation

- Finally, the interpretation phase of the data life cycle offers the opportunity to make sense of the analysis and visualization.

- In addition to simply presenting the data, this step concerns the investigation of all the previous steps through the lens of one's own experience (domain knowledge);

- **Example**: if a Machine Learning algorithm validates a molecule as active against some cancer cells, our dear Valentina (pharmacologist) could synthesize it and test it in the laboratory!

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

How machines learn?

Lorenzo Stacchio

PhD student in Computer Science

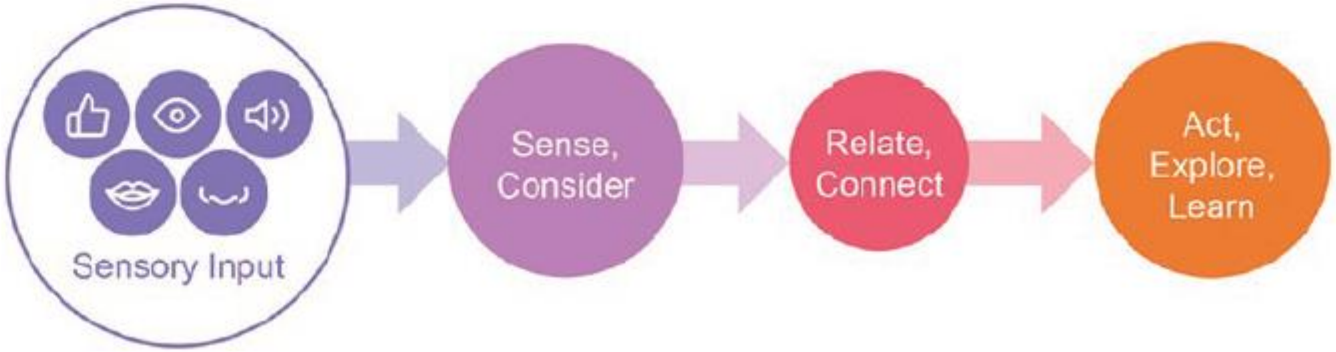
Department for Life Quality studies

How machines learn?

- Until now, we described a Machine learning algorithm as the component that makes an "agent" a "learning agent";
- We have cleared that machines can learn from data (and so experiences); However, we didn't talk about how machines learn from data;
- Citing Marylyn Ferguson, Humanistic Psychologist «*Making mental connections is the most crucial learning tool, the essence of human intelligence; to forge links; to go beyond the given; to see patterns, relationships, context.*»
- We can improve a machine's ability to mimic human behavior by searching for patterns;
- These patterns help to discover and define trends;
- By analyzing these trends and modeling them with algorithms, machines can mimic human responses to certain inputs.

What is learning?

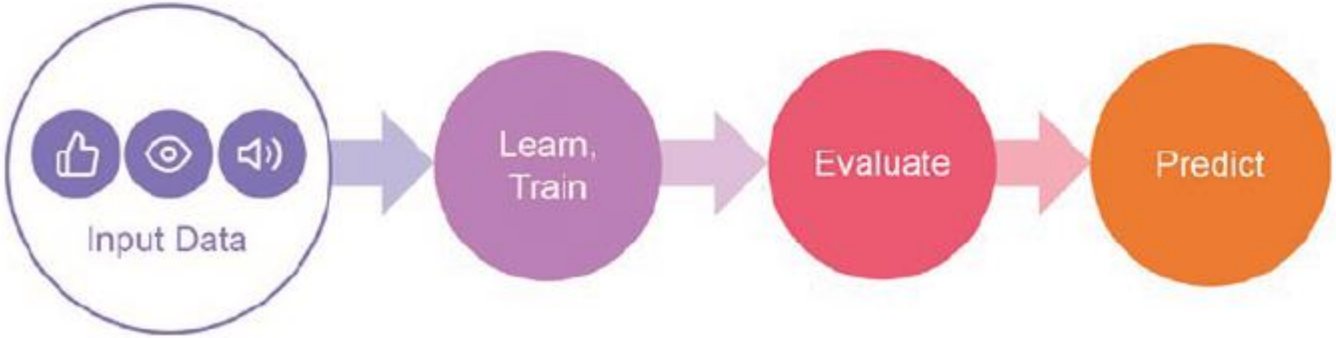
- If we could simplify human learning, we might say that humans take information from their environment, relate it to something, and then learn or act;
- These inputs could be something they see, smell, taste, hear, feel, or even their interpretation of a mood or tone. That information is related to prior knowledge a person has about the world, making a connection. From there, a human might act on their new knowledge, explore, or innovate;



What is learning for machines?

- Machines are given input in the form of data;

- The machine interprets that data and learns from it. Then the machine evaluates that data before outputting the information that has been defined as useful to a human to interpret.



- N.B. a particular sub-field of machine learning, i.e., Reinforcement learning, also exploits a mechanism called **reinforce** to learn, but we will not explore it;

How machines perceive?

- Machines can perceive the environment through sight, feeling, and hearing via sensors.
- Sensors are a part of a machine's hardware system. They measure physical events like temperature, pressure, force, acceleration, sound, and light.
- Also, cameras are sensors, used to perceive what the machines are watching and where they must act;
- Finally, Machine perceived language in code. Originally, this code was only a series of 0s and 1s, or binary. Different combinations of 0s and 1s encode different information to machines;
- Over time, humans have created programming languages that interface between human language and machine language to make the job of coding easier (we will see python).
- We will also talk about how machines understand human language!

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Use cases of Machine learning in fashion

- Natural language processing:
    - How machines read, hear and comprehend;
    - Conversational commerce;
- Computer Vision:
    - Smart mirrors (augmented reality);
    - AI designers;
    - Fashion items retrieval;
- Virtual style assistant:
    - Virtual assistant (Alexa);
    - Virtual and augmented reality role;
- Fashion trend/demand forecasting:
    - Data mining;

Machine Learning

Machine and Deep Learning

Lorenzo Stacchio

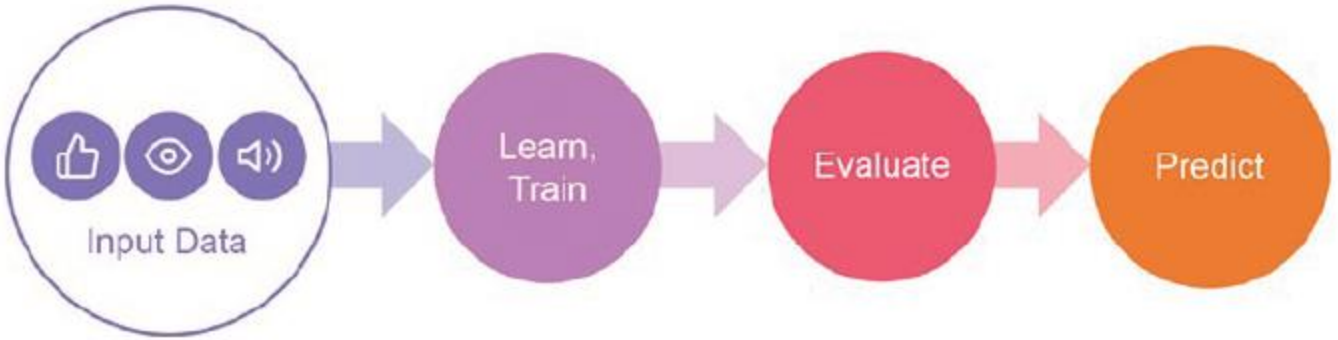
PhD student in Computer Science

Department for Life Quality studies

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

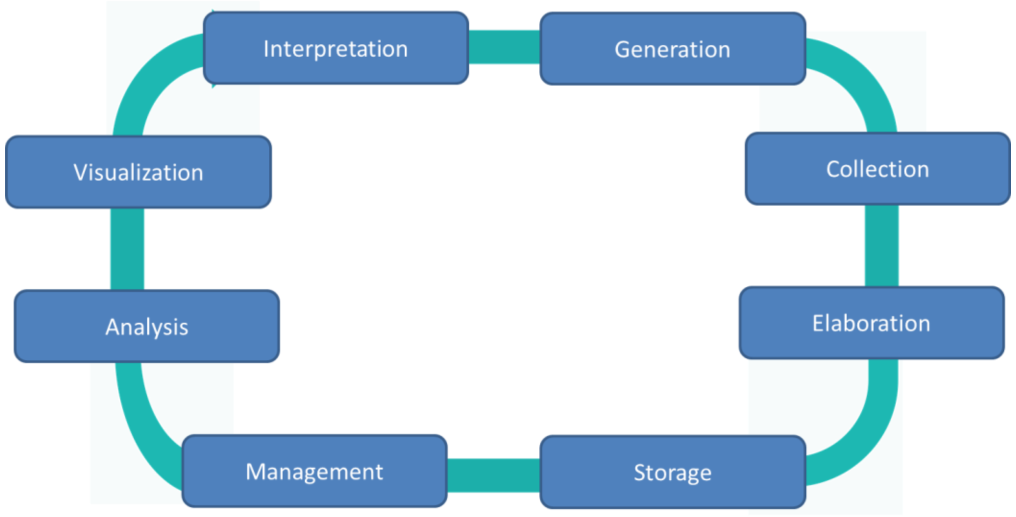
What is learning for machines?

- As we said, machines interpret data and learns from them.



- However, machines can learn and act in different ways;
- But first, we must look on how this data are formed;

Dataset in Machine learning



- From the data life cycle, we have learnt that all the steps are important;
- Even if we concentrate just on the last three, all the previous steps are responsible for the creation of what is defined as a **dataset**;
- **A dataset is a collection of data pieces that can be treated by a computer for analytic and prediction purposes;**
- There is an entire field of scientists that works on the dataset importance in machine learning and methods to ameliorate problems with them (you can check https://datacentricai.org/).

Dataset structure

- The structure of the dataset is fundamental: our machine learning algorithm will learn from them!
- We can initially see a dataset as a table, **where every column of the table represents a particular variable, and each row corresponds to a given member of the data set in question;**
- For example, we can take into consideration a table that describe university students and their carriers:

Name	Surname	Age	Machine learning	Data Science
Lorenzo	Stacchio	25	30/30	27/30
Alessia	Angeli	---	28/30	29/30
Valentina	Pellicioni	---	25/30	28/30

Name	Surname	Age	Machine learning	Data Science
Lorenzo	Stacchio	25	30/30	27/30
Alessia	Angeli	---	28/30	29/30
Valentina	Pellicioni	---	25/30	28/30

- Name, Surname, Age, Machine Learning and Data Science are all **variables and so the features of our dataset!**
- A **variable** is any characteristics, number, or quantity that can be measured or counted;
- Lorenzo, Stacchio, 25,30/30 and 27/30 compose a record and so a **sample of our dataset**!

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dataset and target variable

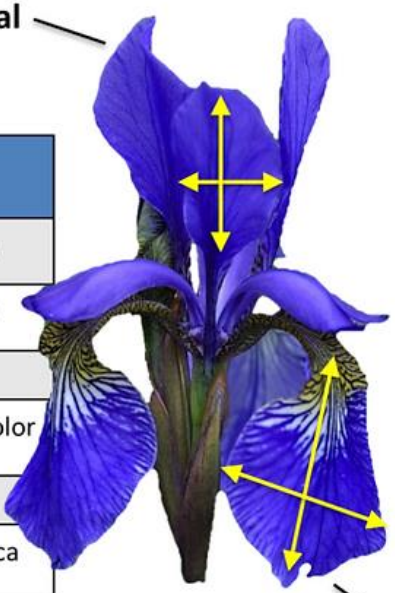
- We will see that, in the most common cases of machine learning, the dataset possess a special variable called **target;**
- The target variable is the feature of a dataset about which you want to gain a deeper understanding.
- In other words, the target variable described what you want your model learns!
- Example of target variable:

Name	Surname	Age	Machine learning	Data Science	Final mark
Lorenzo	Stacchio	25	30	27	108
Alessia	Angeli	---	28	29	107
Valentina	Pellicioni	---	25	28	106

Same task, different approaches: use case of iris recognition

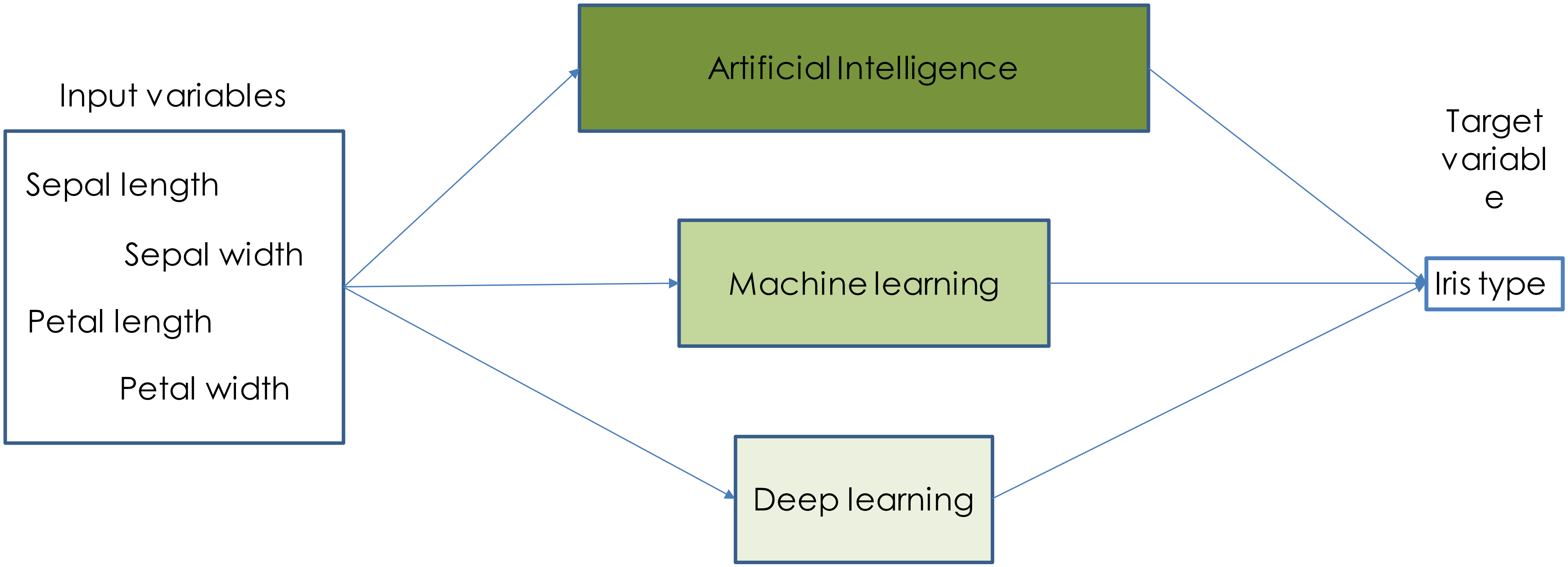
Samples
(instances, observations)

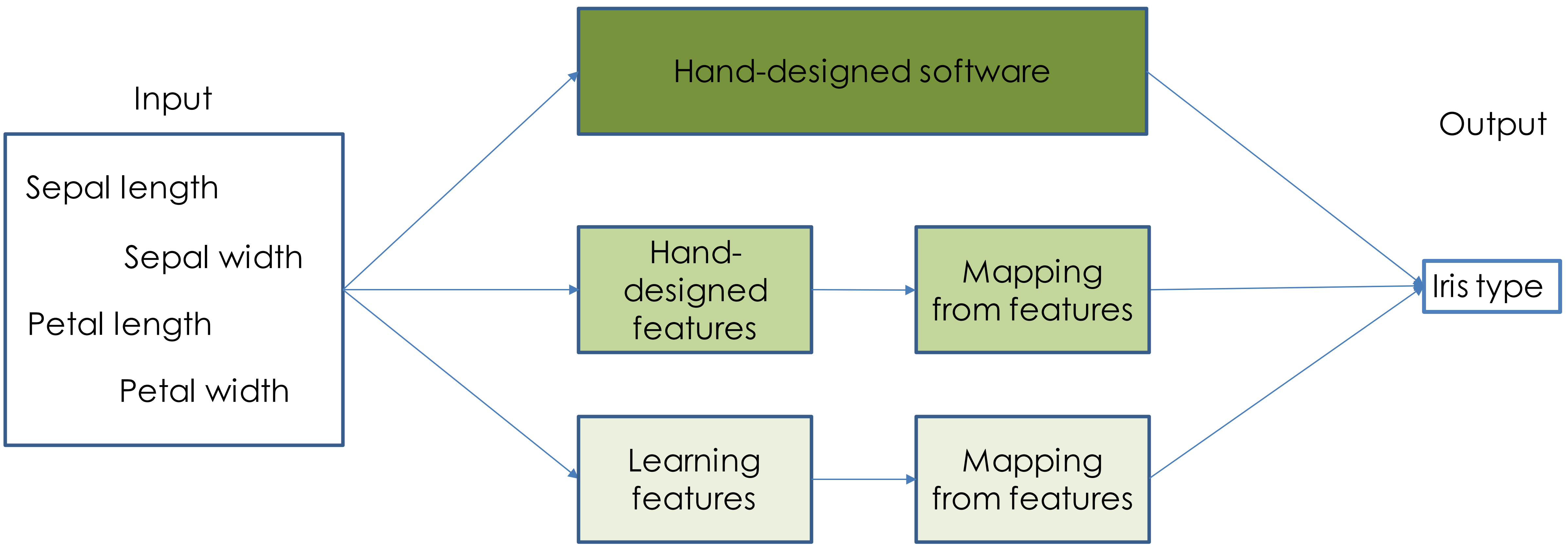
	Sepal length	Sepal width	Petal length	Petal width	Class label
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
			...		
50	6.4	3.5	4.5	1.2	Versicolor
			...		
150	5.9	3.0	5.0	1.8	Virginica



Features
(attributes, measurements, dimensions)

Class labels
(targets)

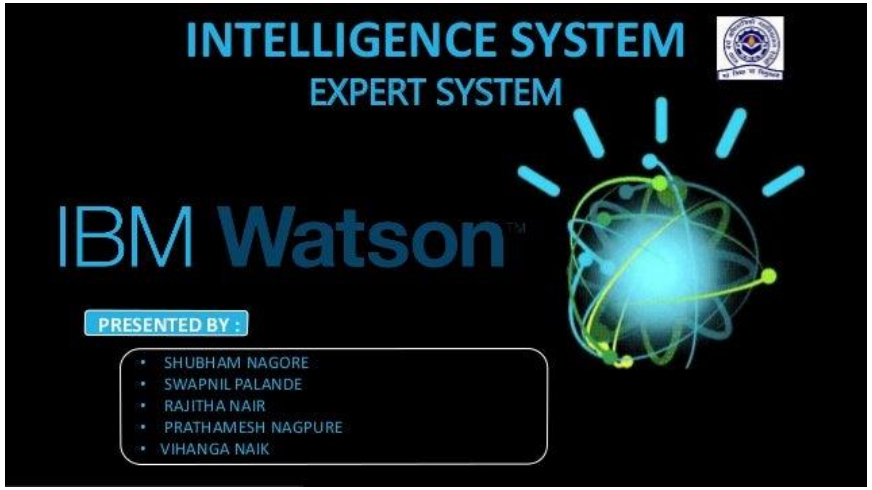

Input variables
Sepal length
Sepal width
Petal length
Petal width
Artificial Intelligence
Machine learning
Deep learning
Target variabl e
Iris type


Input
Sepal length
Sepal width
Petal length
Petal width
Hand-designed software
Hand-designed features
Mapping from features
Learning features
Mapping from features
Output
Iris type

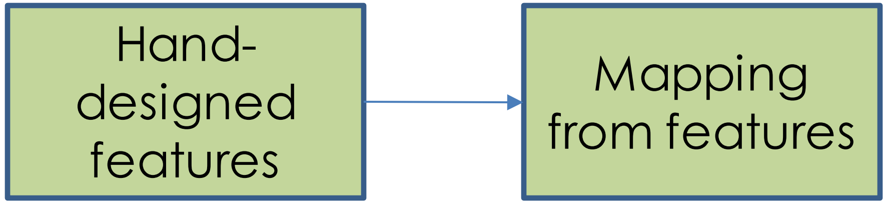
Artificial Intelligence

Hand-designed software

- Programmers write code to reach the desired goal:
- For example, In case of iris, I could write a program that checks for the values of my variable: “if petal length is 5, petal width is 4 … it is a setosa!”.
- This approach was adopted during all the ’90 and also first years of 2000’s;
- It was used to build expert systems like IBM Watson;



Machine learning



- Data scientists design machine learning algorithm to learn from data;

- However, they also carry out an analysis to manipulate and create features in a dataset;

- Taking as example the iris dataset, we can create a new feature called **petal area**, which describe the area occupied by a certain type of flower;

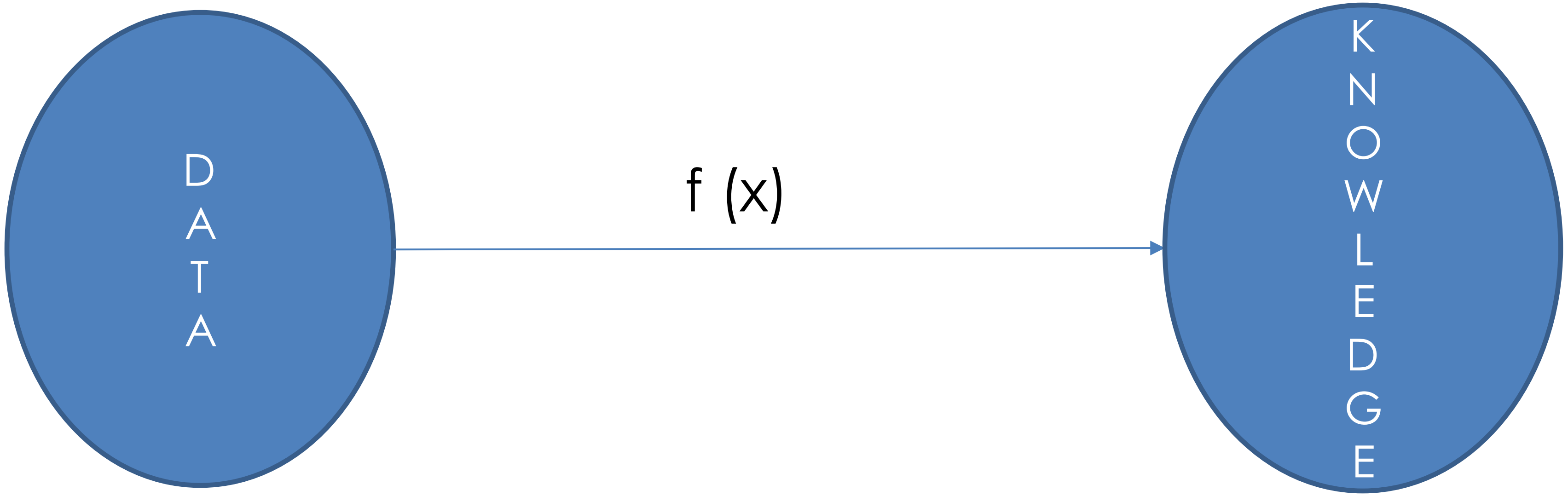
- There are many techniques to manipulate and building new features, we will talk about it in the next classes;

- What is important is that in Machine learning, algorithms can learn from data… but what they learn? What is the meaning behind “mapping from features”?

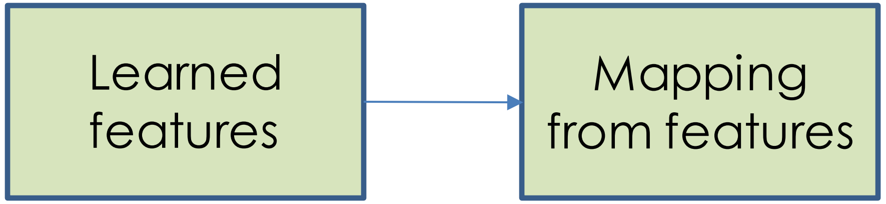
Machine learning algorithms…are instead models!

- Machine learning is about exploring and developing mathematical models for learning from data;

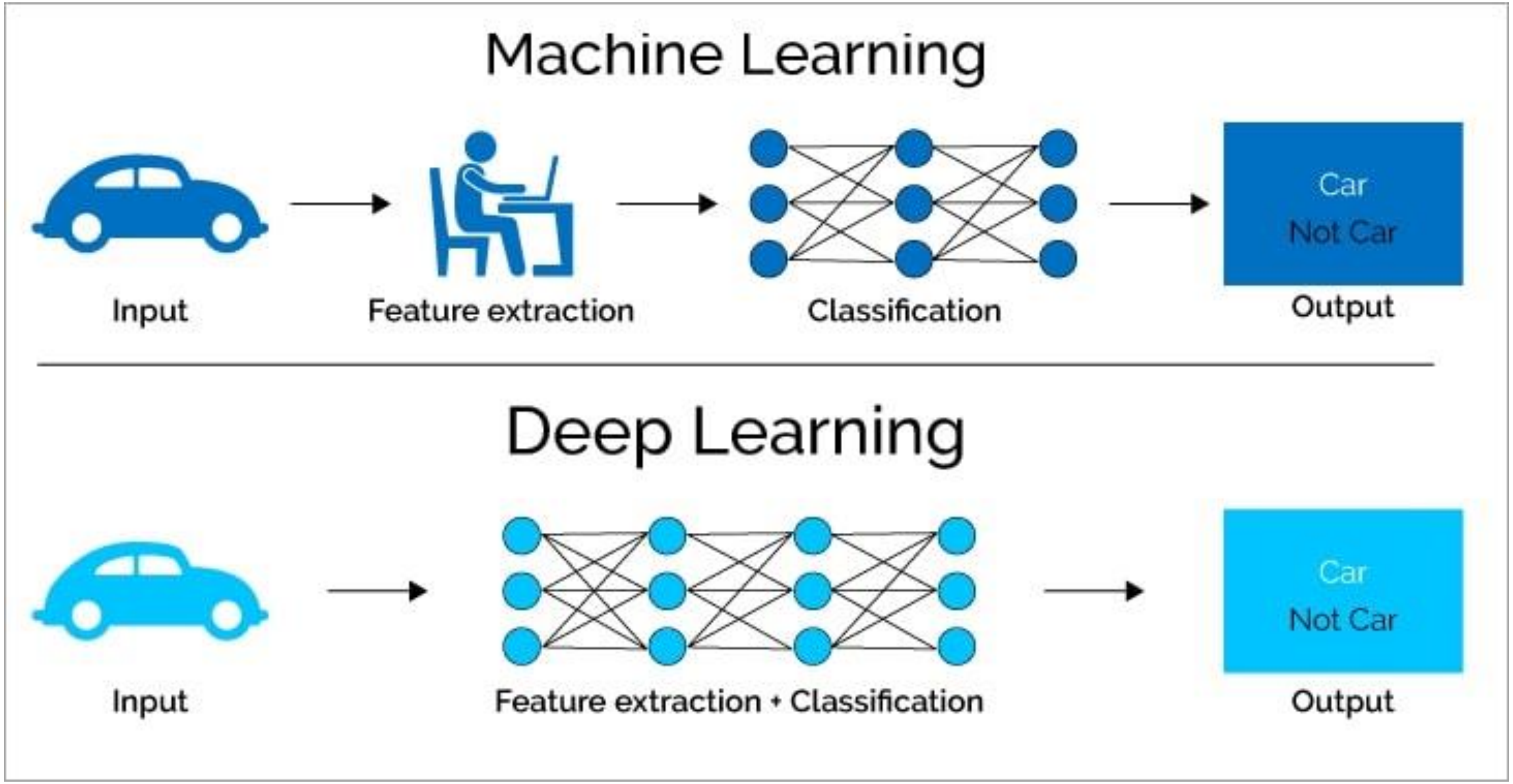
- In practice, each machine learning algorithm models an **optimal** (parametric) **mapping function** between the data domain and something we want to learn;



Deep learning



- Data scientists design deep learning algorithms to learn from data;
- Deep learning differs from machine learning, since it allows to directly learn the features, used to train!
- Deep learning was a revolution in all the fields touched by machine learning: medicine, economy … and also fashion;
- The fact that the algorithm itself could calculate the best features to achieve some goal its incredible.


Machine Learning
Car
Not Car
Input
Feature extraction
Classification
Output
Deep Learning
Car
Not Car
Input
Feature extraction + Classification
Output

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Supervised and unsupervised learning

Lorenzo Stacchio

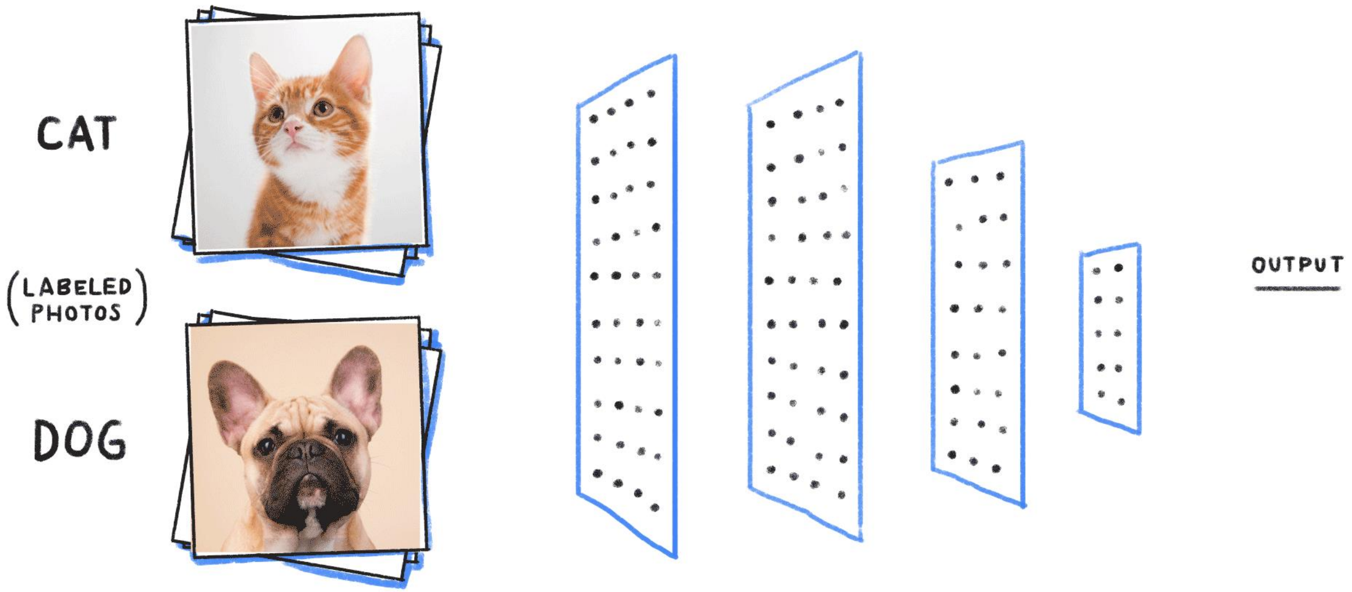
PhD student in Computer Science

Department for Life Quality studies

Supervised machine learning

- Supervised learning, is a subcategory of machine learning and it is defined by its use of labelled datasets to train algorithms that to classify data or predict outcomes accurately.

- Supervised learning helps organizations solve for a variety of real-world problems at scale, such as classifying spam in a separate folder from your inbox or recognize cats and dogs.

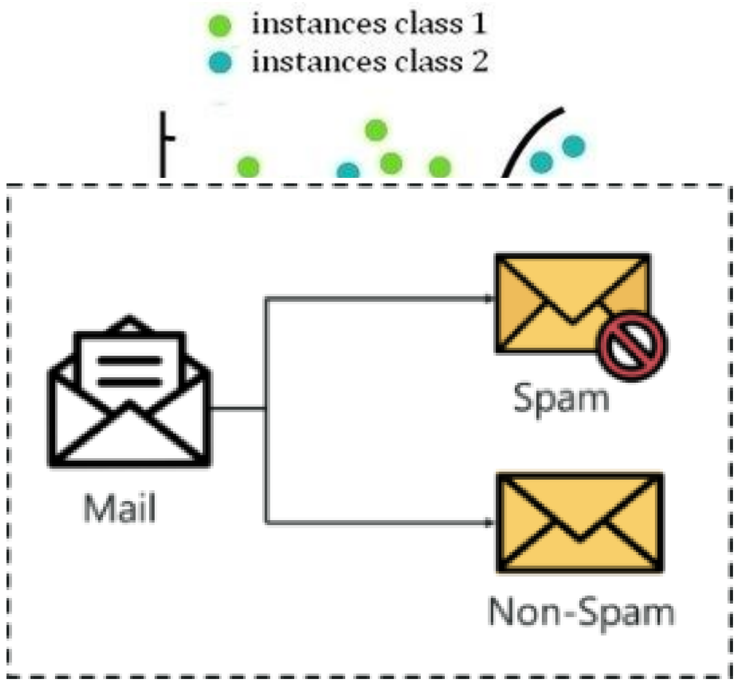


- Contextualizing in fashion, an image of a dress is used as an input, and the word “dress” is used as an output;
- With the already labelled data, our machine learning models can learn and evaluate its performance.

Supervised machine learning: regression vs classification

- A **classification** algorithm may predict a discrete value, but the discrete value is in the form of a probability for a class label;

- A **regression** algorithm may predict a continuous value, but the continuous value in the form of an integer quantity.



Unsupervised machine learning

- Another strategy used in training is **unsupervised learning;**

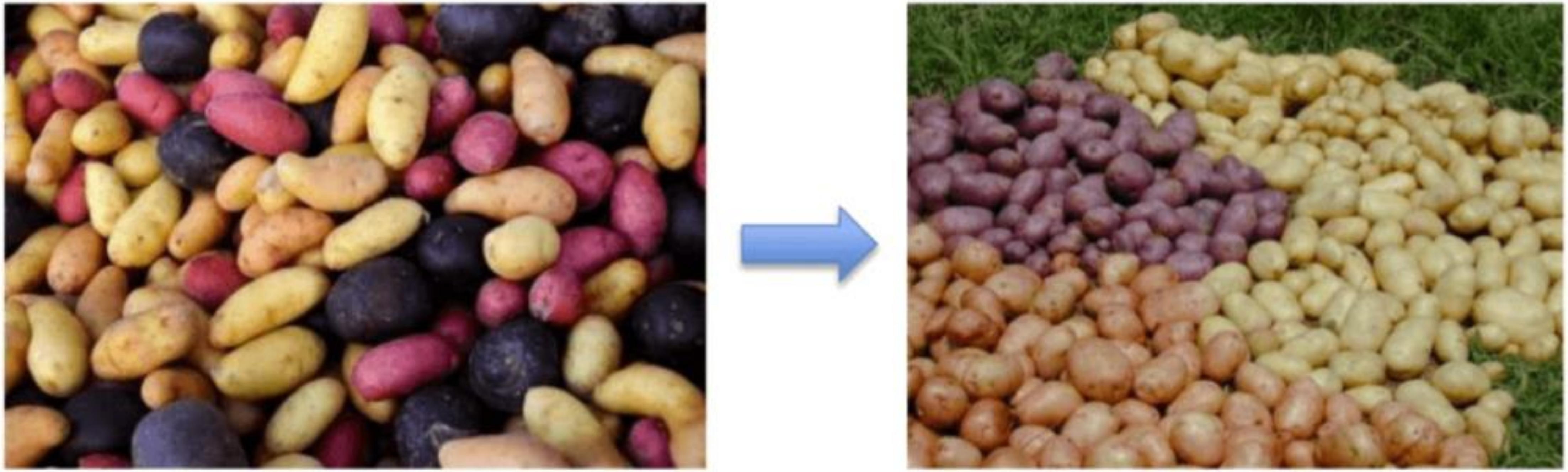
- In this case, the machine learning algorithm is not told the correct or desired output and must decide for itself how to use features to classify data and self-organize.

- This behaviour is commonly called **adaption**. The reason that unsupervised learning is a goal is because there is an ever-increasing amount of easily accessible unlabelled data, whereas creating labelled data can be a time-consuming and costly human task. However, it is a much more challenging approach.

- We won't discuss unsupervised learning techniques in great detail in this course, but it's important to know that this is a field of study that has the potential to address a major pain point in machine learning: manually labelling huge datasets.

- Think of your exam with the professor Daniela Calanca and the IMAGO dataset (we will talk about it as a use case in the next classes);

- The models find themselves the rules to cluster potatoes groups! Without knowing anything except the numerical values!



sample

Cluster/group

How machine learning algorithms learn?

We teach them by minimizing or maximizing functions

- We will not discuss the gradient descent and in general optimization algorithms, because they require a strong mathematical foundation (derivative, partial derivatives, space functions);

- However, you must know that most of the machine and deep learning models are optimized to reach the minimum possible of a loss function;

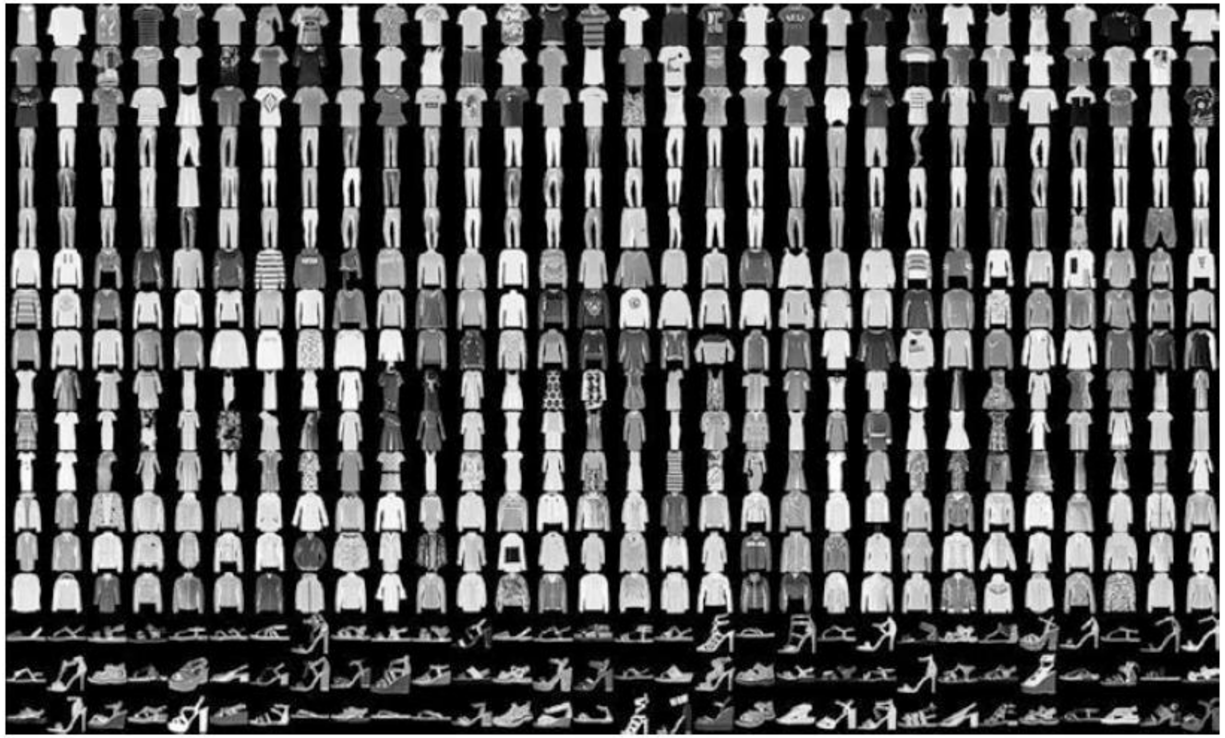
- Loss functions are used to determine the error (aka "the loss") between the output of our algorithms and the given target value. **The loss function expresses how far off the mark our computed output is.**

- However, there are also models that instead maximize other factors, such as the **entropy**;

What is the role of the data?

- The training process requires large datasets in order for the neural network to find patterns across images;
- A number of datasets have been created and made publicly available in the research community;
- We will se some of them in practical sessions;
- While in many industries, information is closed, machine learning has evolved so rapidly and effectively in part because of the sharing of valuable information and resources like these training libraries;

Zalando Fashion MNIST



- Zalando released an effective dataset for fashion called Fashion-MNIST. The Fashion-MNIST dataset contains 60,000 garment images;

- Fashion-MNIST contains ten categories of images: T-Shirt/Top, Trouser, Pullover, Dress, Coat, Sandals, Shirt, Sneaker, Bag, and Ankle Boots;

- The fashion dataset represents complex visual features (necklines, sleeves, and much more).

Train, validation and test set

- Concluding, dataset **are partitioned in three parts**: training, validation and test set.
- The first is used to train the machine learning algorithm and the other to evaluate its performance;

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Machine Learning in Fashion

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Natural Language in Fashion

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

Natural language processing

- Machine language and human language meet in natural language processing (NLP);
- NLP is a way for computers to comprehend human languages. Every day, our interactions on the Web (things we post on social media, text messages we write, and so forth) contribute to an ever-expanding mass of data;
- The data produced are often unstructured: is written in free form, unorganized, and historically hard to parse;
- **We can use NLP to understand the content and context of this unstructured data, unlocking a rich treasure trove of information about ourselves;**
- Natural language processing is applied in fashion in **conversational shopping, AI customer service chatbots, virtual assistants, and stylists**.

ELIZA

- Computer scientist Alan Turing thought the ability to use human language to be an important determinant of intelligence in machines (recall the Turing test seen before);
- At that time and throughout the 1960s, the first chatterbots were created, exemplifying the power of natural language-based interfaces;
- One of the most famous examples during that time was ELIZA, that passed the restricted form of the Turing test [6];
- It/she was created by Joseph Weizenbaum at the MIT Artificial Intelligence Laboratory.
- The ELIZA bot consists of a long list of possible responses and complex rules to determine which responses are used in conversation. In the 1980s, the architecture of these bots all changed because of machine learning;
- Chatbots today are capable of more complex interactions because of the algorithms that control them.

Chatbots

- Most chatbots can be placed into two basic categories: **scripted and artificially intelligent**.
- Scripted chatbots can follow only a predefined set of rules. This set of rules means the kinds of questions the chatbot can answer and the responses it can create are limited to the scripts it was programmed with.
- Artificially intelligent chatbots are built to interpret the natural language used by humans and can produce relevant responses to inputs that are not exactly pre-defined.
- More recently, chatbots can use images as part of a conversation in addition to text.
- Some companies create nowadays specialized chatbots for a certain domain.
- In fashion retail, such types of chatbots are usually adopted to automatize the search of a garment of a certain brand, with a certain style… this gave birth to **Conversational commerce**.

Chatbots: demo?

- One of the most powerful demos I found is the one provided by [Hugging Face];

- Link to the chatbot demo: https://huggingface.co/spaces/gradio/chatbot

- This chatbot is built on top of one of the most powerful Neural Networks in the field of Natural Language process: GPT-2 (Generative Pre-trained Transformer);

- GPT-2 was trained simply to predict the next word in 40GB of Internet text;

- GPT-2 is trained with a simple objective: predict the next word, given all of the previous words within some text;

- This provides GPT-2 a broad set of capabilities, including the ability to generate conditional synthetic text samples of unprecedented quality, where we prime the model with input and have it generate a lengthy continuation.

Conversational Commerce

"I don't know anyone who likes calling a business. And no one wants to have to install a new app for every business or service that they interact with. We think you should be able to message a business, in the same way, you would message a friend."

Mark Zuckerberg, in 2016

- Chris Messina gave this first definition: "Conversational commerce is the intersection between messaging apps and purchases, the tendency to interact with companies through messaging applications, chat or through voice technology";

- Samples of Conversational technologies are Live chats, Chatbots, and Voice assistants;

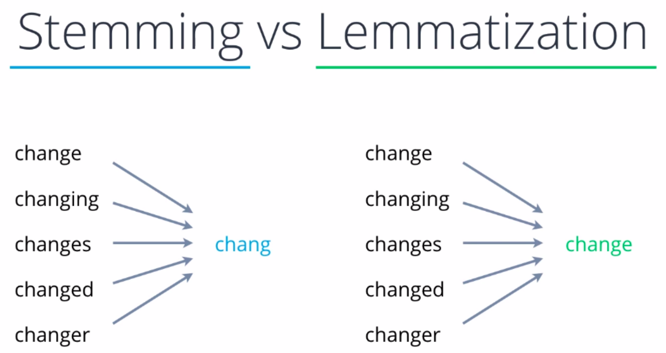
- The main idea behind conversational commerce is to reduce the number of clicks that a user must go through to reach the desired product. Rather than selecting a half dozen filters, a user can type what they're looking for in a natural language query:

Find Women's Sandals with 3-4˝ Heels in Black under $50.

Natural Language queries

Find Women's Sandals with 3-4˝ Heels in Black under $50.

- This statement is what we define as **Natural Language query**;
- A natural language query is an input that consists solely of terms or phrases spoken normally or entered as they might be spoken, without any non-language characters, such as the plus symbol or the asterisk, and without any special format or alteration of syntax;
- So, there is an error in the statement? Yes. We can avoid it with stop wording (remove useless words);
- In NLP there are pre-processing such as stemming and lemmatization that allows you to remove undesired part of the words or get their syntactic radix.

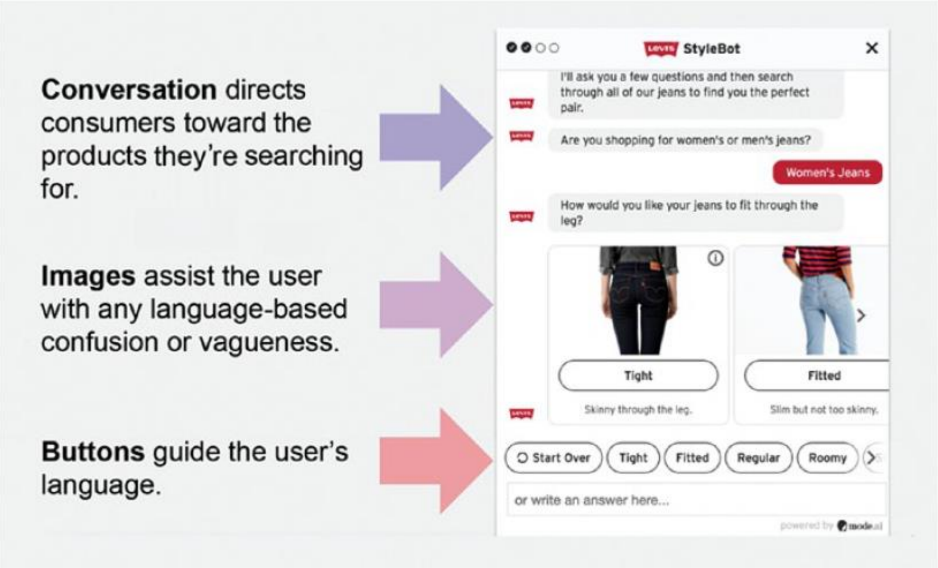


How to apply conversational commerce: Shopping and messaging

- Conversational shopping, or conversational commerce, seems to have emerged as a concept in 2015;
- The AI bots behind conversational commerce interfaces act as **live** agents on the other side of a messenger conversation with a consumer (**it can also consider past experiences**).
- By asking questions to the bot, a consumer can receive personalized recommendations, product care instructions, customer service, and even purchase products in one click;
- The Bot messaging is integrated with buttons, web views, images, and other simplified graphical user interface (GUI) components. These components can help guide the conversation between the human user and the machine by providing possible outcomes to the specific context.

- To mimic the experience of talking with an in-store sales associate, companies like Levi Strauss & Co. have partnered with AI companies like mode.ai.

- In late 2017, Levi's and mode.ai released a conversational commerce bot that helps consumers discover their perfect jeans;



- This is the conversational shopping interface by mode.ai that appears on Levi's Facebook Messenger and web site.

- This example uses a mixed UI relying not only on the user's improvised inputs but allowing the user to select common options by clicking pre-existing buttons.

How to apply conversational commerce: other examples?

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Natural Language in Fashion: how machines read

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

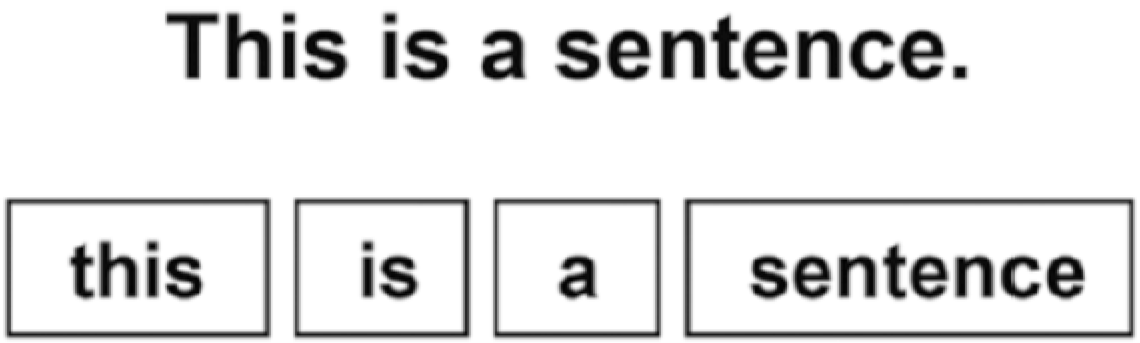
How machines read

- Natural language processing plays a critical role in many applications:
- Not only conversational shopping, but also analysing sentiments of your customers on social media [7], [8] (i.e., what they think about your last fashion item);
- But what is the computer doing when it dissects words? It usually starts by processing text to transform it into data the machine can understand;
- It usually start with correcting misspelled words, then it follows through a process of breaking the words apart, analysing them, and relating the words to each other in order to extract meaning;
- Four commonly addressed methods happening behind the scenes in NLP are as follows: Tokenization, Word embeddings, Part-of-speech tagging, Named Entity Recognition;

Tokenization

- In the process of tokenization, a machine breaks apart pieces of a sentence or phrase into tokens, usually words or terms;

- This process is also sometimes referred to as **lexical analysis**. These linguistic units make up words, punctuation, numbers, and so forth.



- Whitespace tokenization is just one of many ways to tokenize text;

- Other methods are used for more complex sentence structure or languages that don’t use whitespace to separate words;

Word embeddings: convert words to numbers

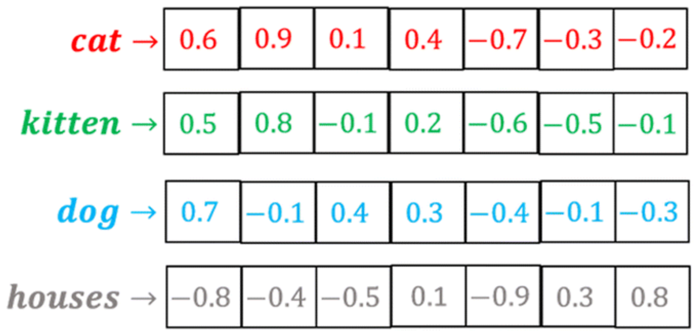
- How can a machine understand or compare words? One way is to associate each word with a set of number values, or vectors, so that a machine can compare them.

- These vectors are known as word embeddings, or word vectors;

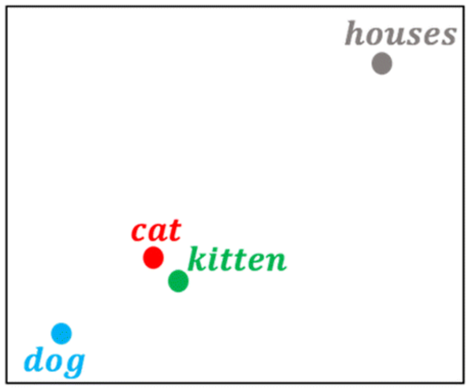
- Nowadays, this process, know as word2vec, is done with Deep Learning Models, such as BERT;

- Word embeddings make words accessible to certain types of machine learning models, by making it easy to do mathematical calculations with them;

- **They can help define relationships between words and create visualizations of these relationships.**



7-d vectors



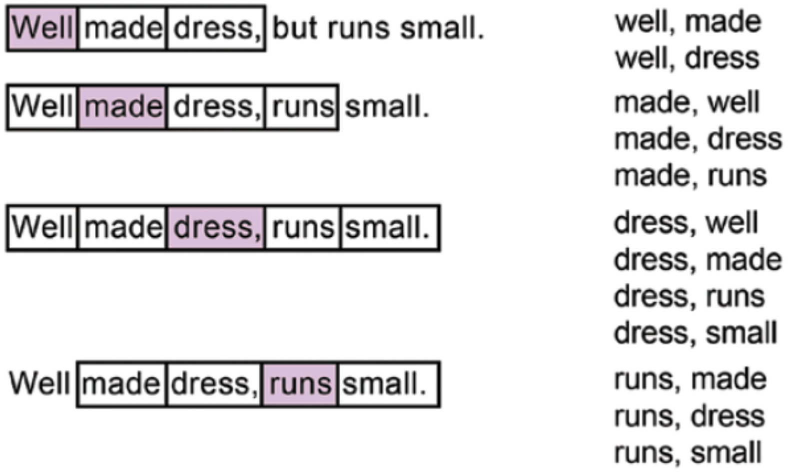
2-d representation


Initial top search results
Refined top search results
Query:
"black shoes"
Feedback:
"more formal than these"
Feedback:
"shinier than these"

How to analyze such relationships? Skip grams

- **One way of analysing these relationships and their proximity is by using skip grams;**

- The skip gram model allows us to analyse how likely two words are to be co-occurring near each other in a text;



- The input word is highlighted in pink. Looking at the nearest two words, the skip gram word pairs from this sentence are shown on the right. What this means is that for this item, the word pair "dress, small" is more probable than "dress, large."

- It is fundamental for relation extraction;

How to analyze such relationships? Part-of-speech tagging

- Part-of-speech (POS) tagging is the process of defining the part of speech of a given word based not only on the word's definition, but also its context;

- The English language contains eight parts of speech: nouns, pronouns, adjectives, verbs, adverbs, prepositions, conjunctions, and interjections;

- Knowing the part of speech of a word reveals a lot of information about its neighbors and helps to comprehend the sentence;

- Often, tokenization is done as a precursor to this task in order to separate the words to be tagged. The process of tagging is a disambiguation task, making the ambiguity of words clearer within their context.

- It is fundamental for relation extraction and sentiment analysis;

Relation extraction

- Relation extraction is a way of extracting specific pieces of information and their relationship to the subject from a text.
- There is a long list of methods used to achieve relation extraction, but the goals are approximately the same for all methods.
- Relation extraction looks to comprehend things such as the subject, relation, and object of a given sentence. Less specifically, it is looking for relationships in unstructured text.
- This is useful because as a machine reads through text, it can learn from the sentences in that text, store that knowledge to a database, and recall that information when prompted.

Table 2-1. *Extracting Relations from Natural Language into Structured Data for Later Use*

Subject	Relation	Object
PVH	Location	New York, NY
PVH	Incorporated	April 8, 1976
PVH	Is A(n)	Apparel Company

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

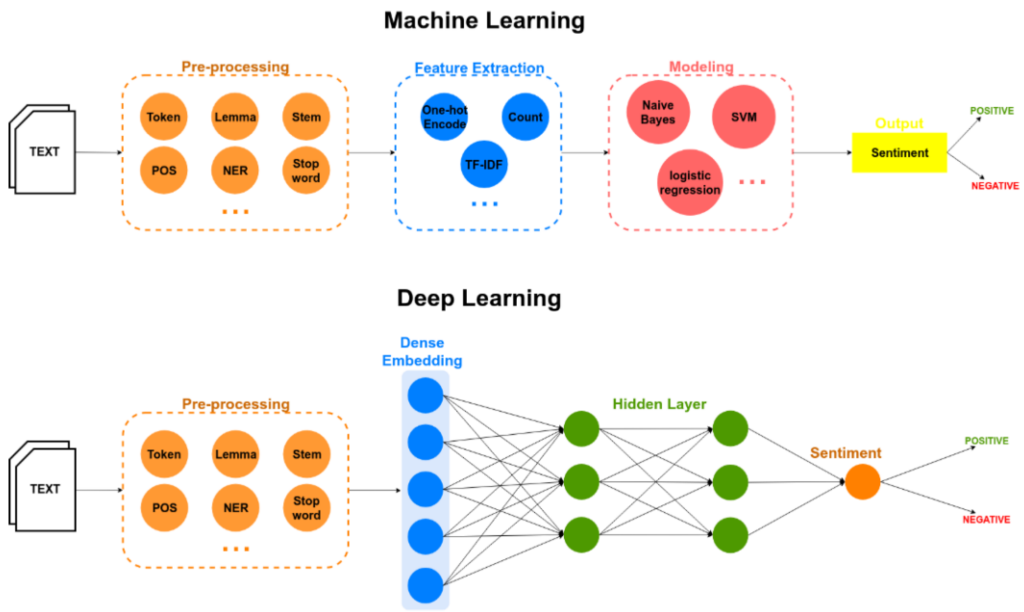
Named entity recognition

- Named entity recognition (NER) refers to the methods for identifying and classifying important nouns into categories such as organizations, people, and times.

- It is fundamental for sentiment analysis: recognize the brand, the time, the item etc…

Sentiment Analysis

- Sentiment analysis is a method of understanding how the speaker feels about a particular object or subject.
- There are several methods for understanding sentiment, including using machine learning, statistics, knowledge-based methods, or a hybrid of these;
- While knowledge-based methods rely more heavily on determining sentiment in definite terms, machine learning methods allow for more flexibility in meaning.



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Computer Vision in Fashion

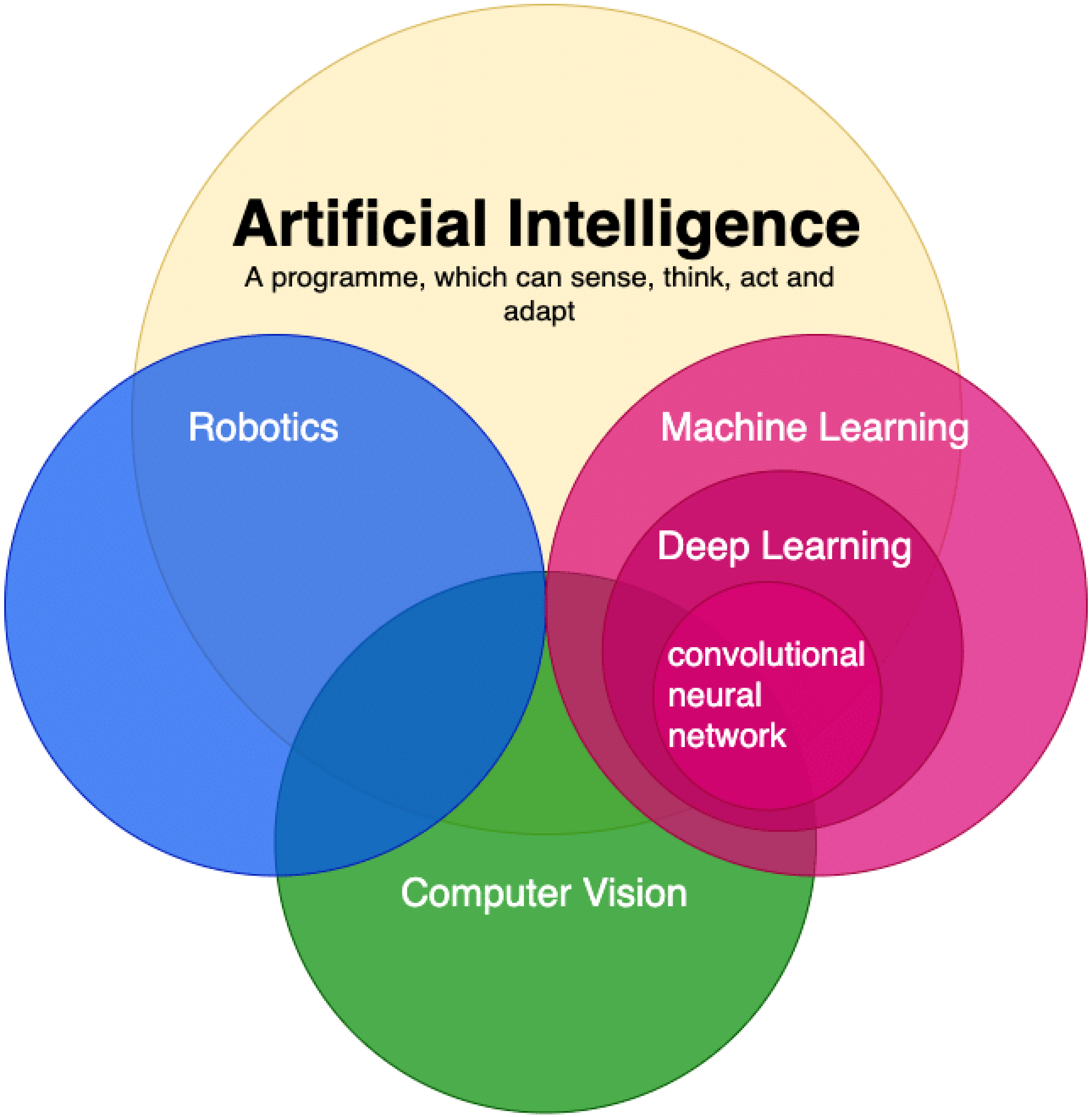
Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

Computer Vision

- Computer vision (CV) is used to process and analyze images and videos;
- CV automates tasks we might associate with the human visual system and more;
- Although computer vision is a field of its own, artificial intelligence has played a major role in recent progress;
- Computer vision is frequently used in fashion applications because the fashion industry is so visual;
- **In the fashion industry, computer vision is being used in technologies such as garment generation, trend analysis/forecasting, visual search, smart mirrors but also in virtual reality, and augmented reality.**


Artificial Intelligence
A programme, which can sense, think, act and adapt
Robotics
Machine Learning
Deep Learning
convolutional neural network
Computer Vision

- Despite such definition, nowadays computer vision heavily relies on machine and deep learning;

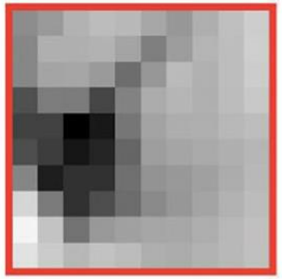
- Thanks to DL and ML, we reached state of the art performance in different tasks: autonomous driving, cancer detection, virtual try-on.

- We will see the types of computer vision fashion tasks that deep learning allowed to advance;

- Before that, we must understand in detail how “computer vision works”.

How computer vision is implemented

- In order to understand an image, computers must translate an image into numbers.
- The basic unit of a digital image is a pixel, a set of numeric values that correspond to a color. For example, in an 8-bit black-and-white image, the number 0 would display as black, 255 as white, and a value in between like 200 as a light - gray.
- Machines use the value from that pixel and surrounding pixels to apply algorithms to the image. Mathematical calculations based on patterns in these values are what machines use to determine the content of an image.

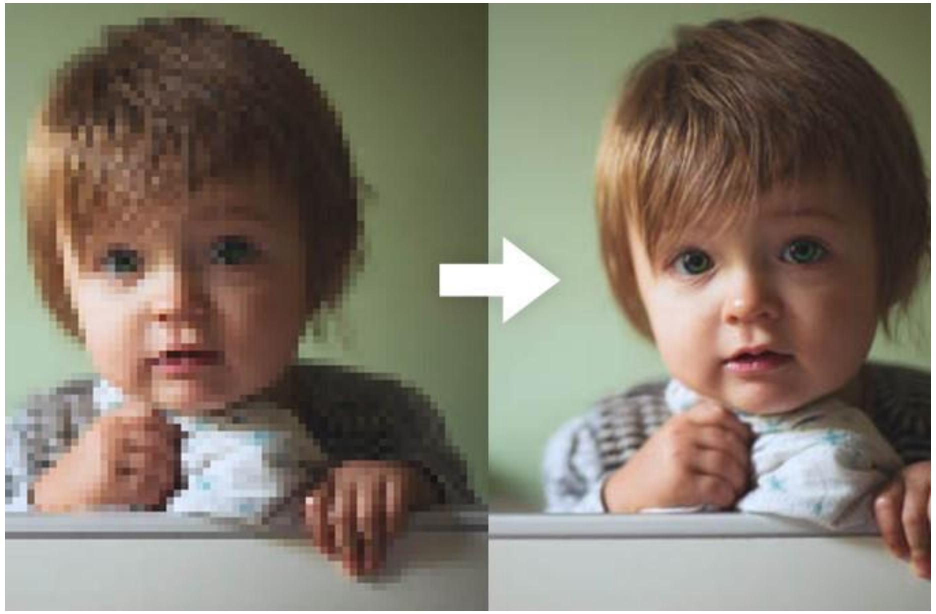


135	175	178	164	126	158	152	191	205	206
164	166	166	105	109	151	176	189	204	206
164	166	166	155	109	171	186	188	200	204
80	123	123	130	123	163	173	175	183	189
65	65	0	37	160	171	169	170	175	182
61	122	29	37	111	150	161	162	175	181
134	49	77	77	138	153	158	164	177	188
195	114	77	108	140	138	164	177	183	192
244	203	144	153	166	174	171	183	184	190
226	203	183	194	188	174	171	178	184	190

Image processing

- Image processing refers to a broad set of techniques used to manipulate an image;
- These techniques are used in a wide variety of applications by photographers and engineers alike;
- The main difference between standard image processing and image processing in computer vision lays primarily in the goal;
- In computer vision applications, the goal is generally to enhance an image to make it more readable by a machine.

Super resolution



Denoising

Image transformation

- By using the numeric values that represent a digital image, machines are able to apply mathematics to images in image transformations.

- An example of image transformation consist in adding the value of 20 to each pixel on the grid. In grayscale images, lower pixel values closer to 0 are darker, and higher values closer to 255 are lighter. The result is that the entire image is lightened.

Image filtering

- Filtering is usually used to enhance an image, extract information, or detect patterns before sending it through a machine learning model;

- By applying filters to images, machines may have an easier time with more-complicated tasks down the road; Hugely applied in Photoshop but also in programming languages such as Python (automatize your task);

- **In this example a Gaussian Blur is applied;** used to reduce noise in images. **(N.B., pay attention, book definition is wrong)**

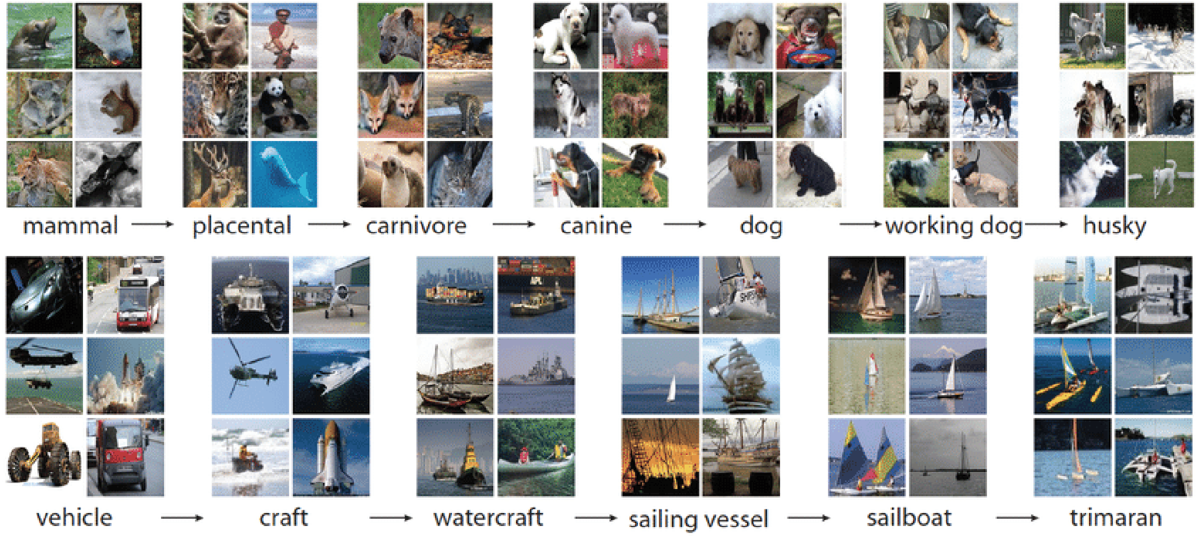
Feature extraction

- Do you remember the dataset features described earlier? What is a feature in an image?
- In computer vision, a feature is vaguely described as an interesting part of an image and more precisely as a higher-level description of raw data.
- Feature extraction refers to processes used to find features. Features include edges, corners, interest points, blobs, and ridges;
- Finding features allows a machine to also define a local segment in an image where an object or area of interest might be found.
- Without a repeatable and reliable feature-detection process as a base, most of the state-of-the-art performance we cited before, wouldn't be reached;
- Nowadays, **feature extraction is made by means of Deep Learning;**

How to use image features?

- All the modern computer vision tasks strongly relies on feature extraction to perform;

- Image classification is one of those It refers to the process of classifying images into one of many predetermined possible categories.



- One of the major limitations of image classification is its reliance on a dataset of labelled images for training (supervised learning);

- Modern algorithms are trying to be unsupervised;

Why images are so important in fashion?

- It’s hard to imagine an industry that relies on images more than the fashion industry.
- Almost every process, from manufacturing to marketing, revolves around images.
- When might you want to know what is in an image?
- In retail, the ability for customers to search for a particular garment style on a web site gives them access to the products they are searching for;
- Even better, the ability to discover products from styled images gives customers the ability to browse inspiration and to access the product (which item to buy to achieve the desired look).
- For this reason, Computer Vision is one of the central AI branches adopted in fashion and this is also the reason why we will introduce Deep Learning, the most important subfield of artificial intelligence talking about images.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Computer Vision and the Deep Learning revolution

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

Neural Networks

- Neural networks is a type of machine learning model;

- They were originally modelled after our understanding of the behaviour of neurons in the human brain: in the brain, a single neuron takes in input, processes it, and sends output;

- Neuroscience has moved away from this idea. We now know brains don't work like this, and the statistics behind neural networks in machine learning have been developed independently of neuroscience;

- Neural networks are typically created with layers that compute information in parallel. They're composed of interconnected nodes.

- Knowledge in these systems is represented by the patterns that are taken on by nodes passing information to each other.

https://www.youtube.com/watch?v=aircAruvnKk

- Usually, the composition of a neural network includes three basic parts:

    - ***Input layer***: Contains input data;

    - ***Hidden layer(s)***: Contains the synapse architecture;

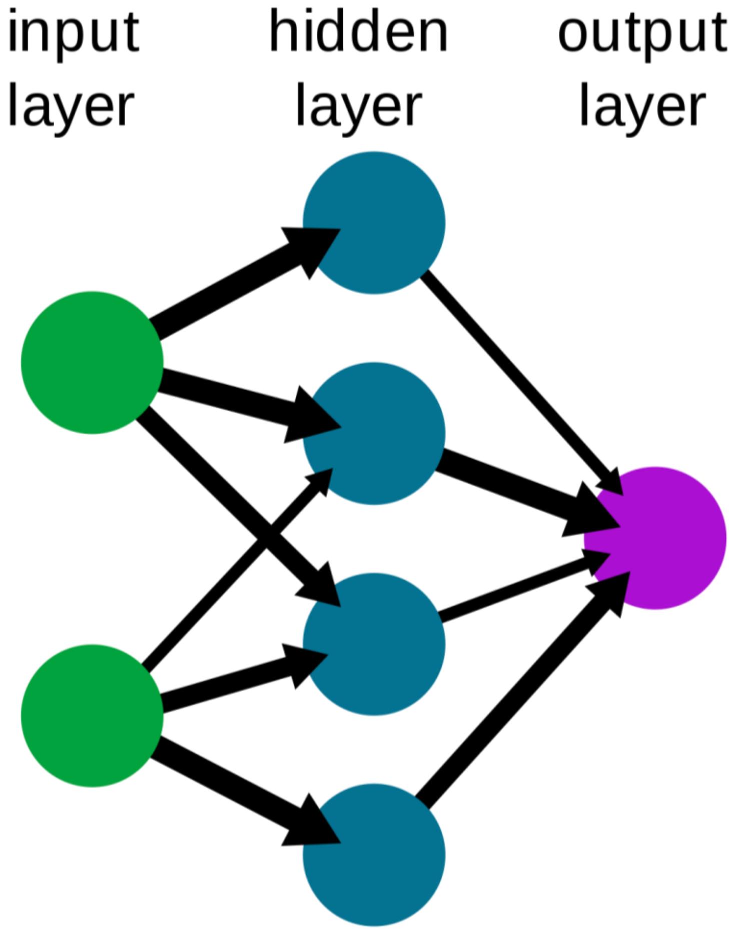
    - ***Output layer***: Provides the outputs;

- Within this framework, neural networks can present themselves with different facets;

- In their implementation, training is also an important part of the process and involves sending data through the neural network.

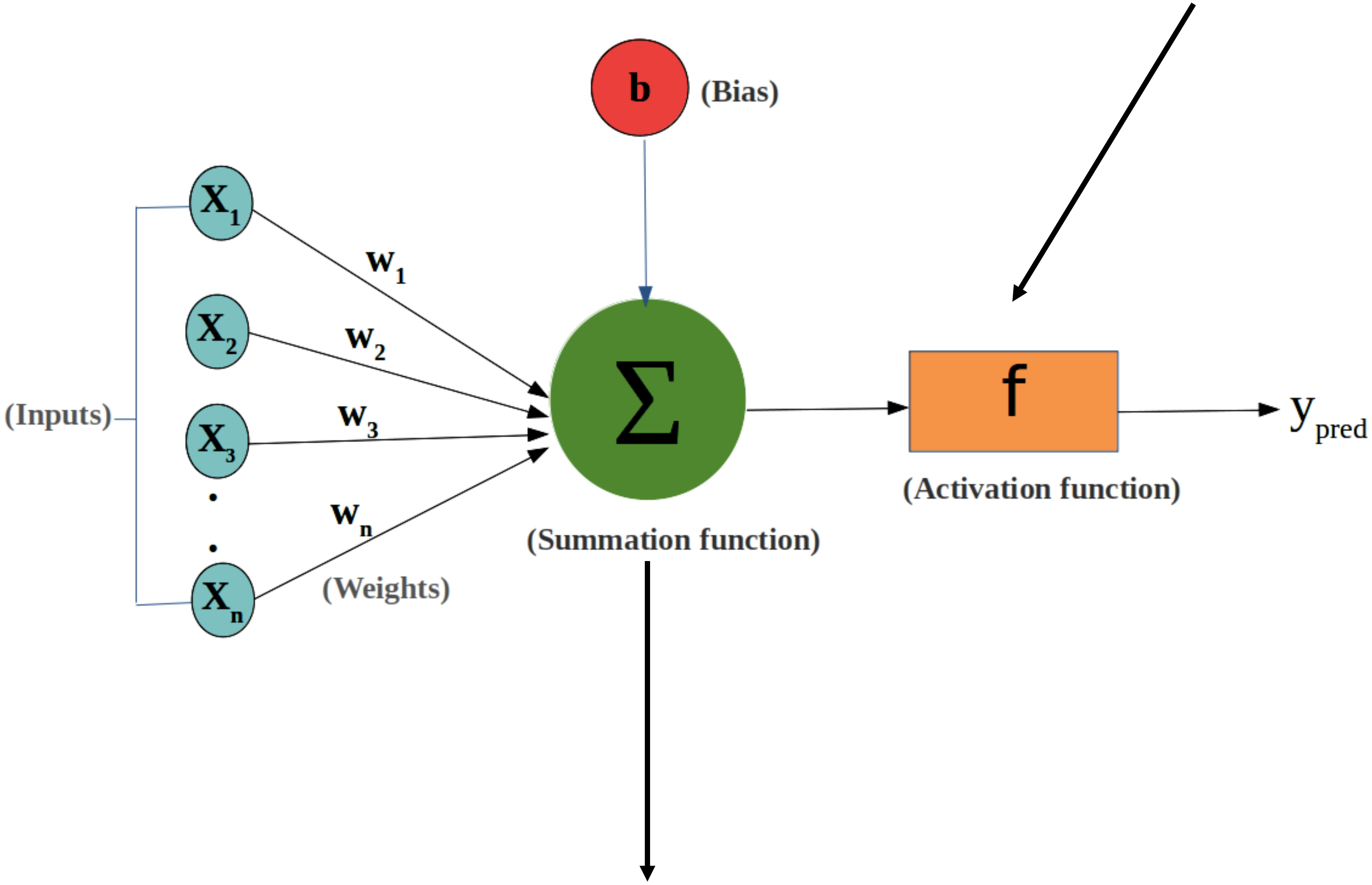
- In many instances, the network's effectiveness is reliant on high-quality and high-quantity data.

A simple feed-forward neural network



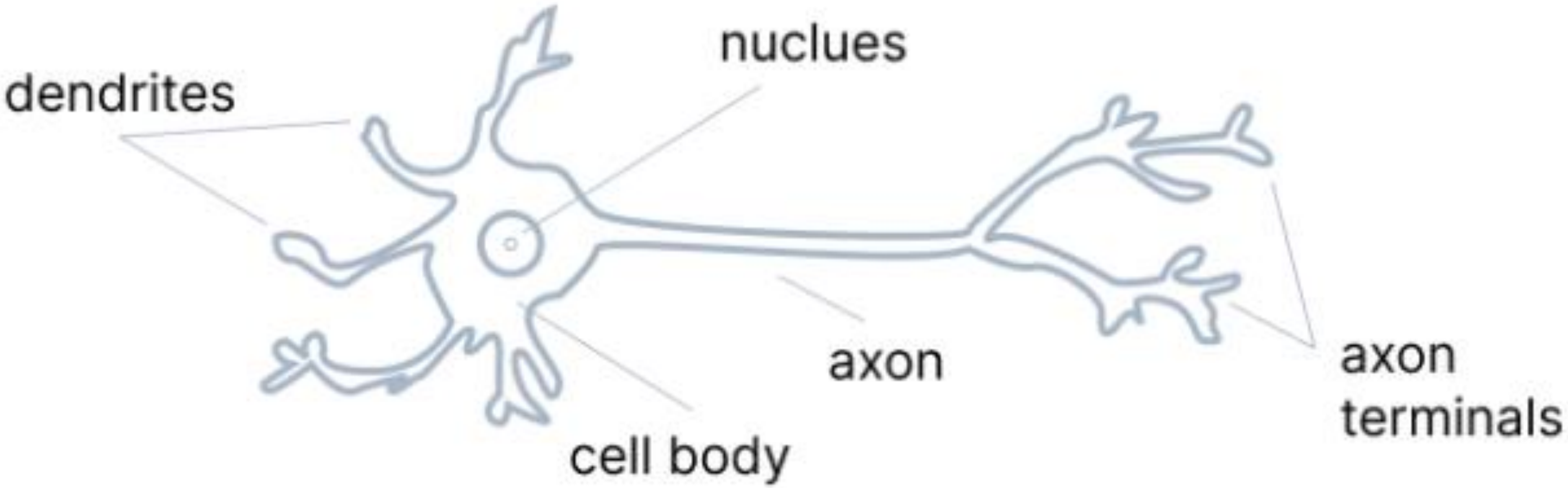
Let’s start with a single neuron

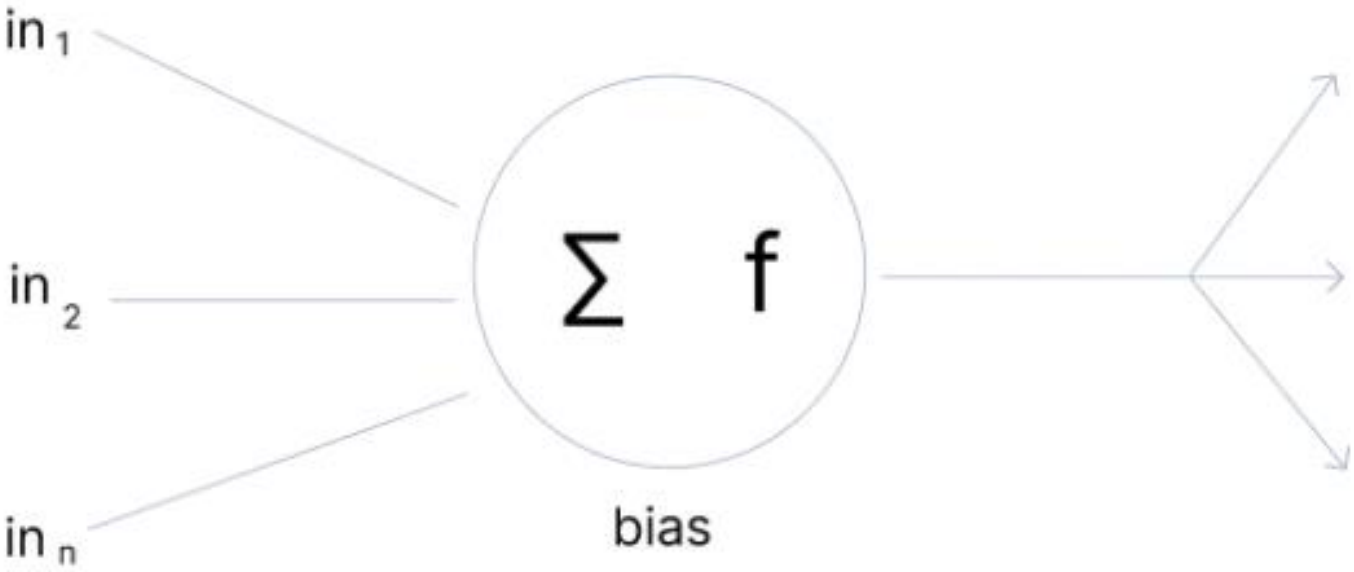
An Activation Function **decides whether a neuron should be activated or not and** also its value.

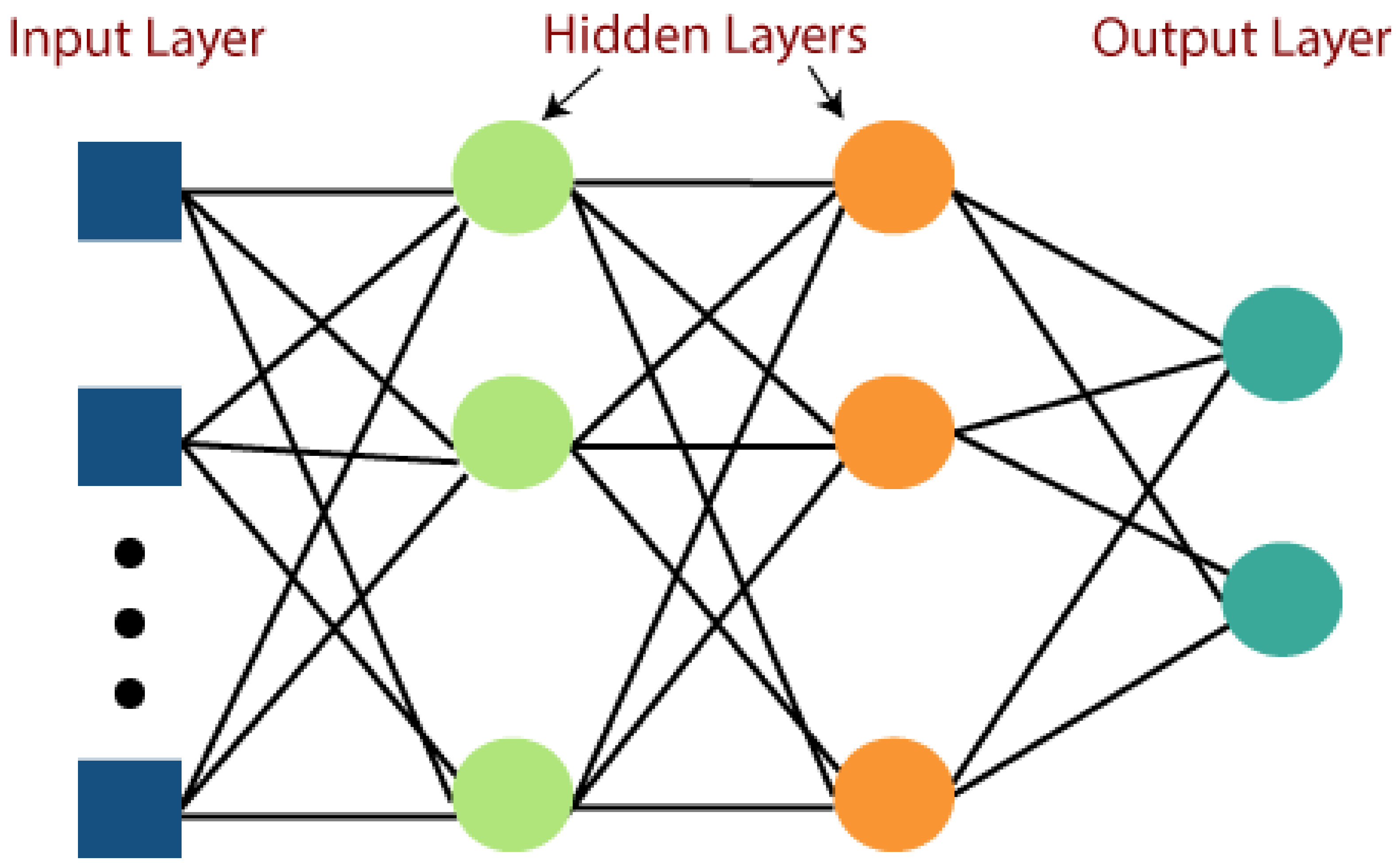


((x1*w1) + (x2*w2) … + (xn*wn))

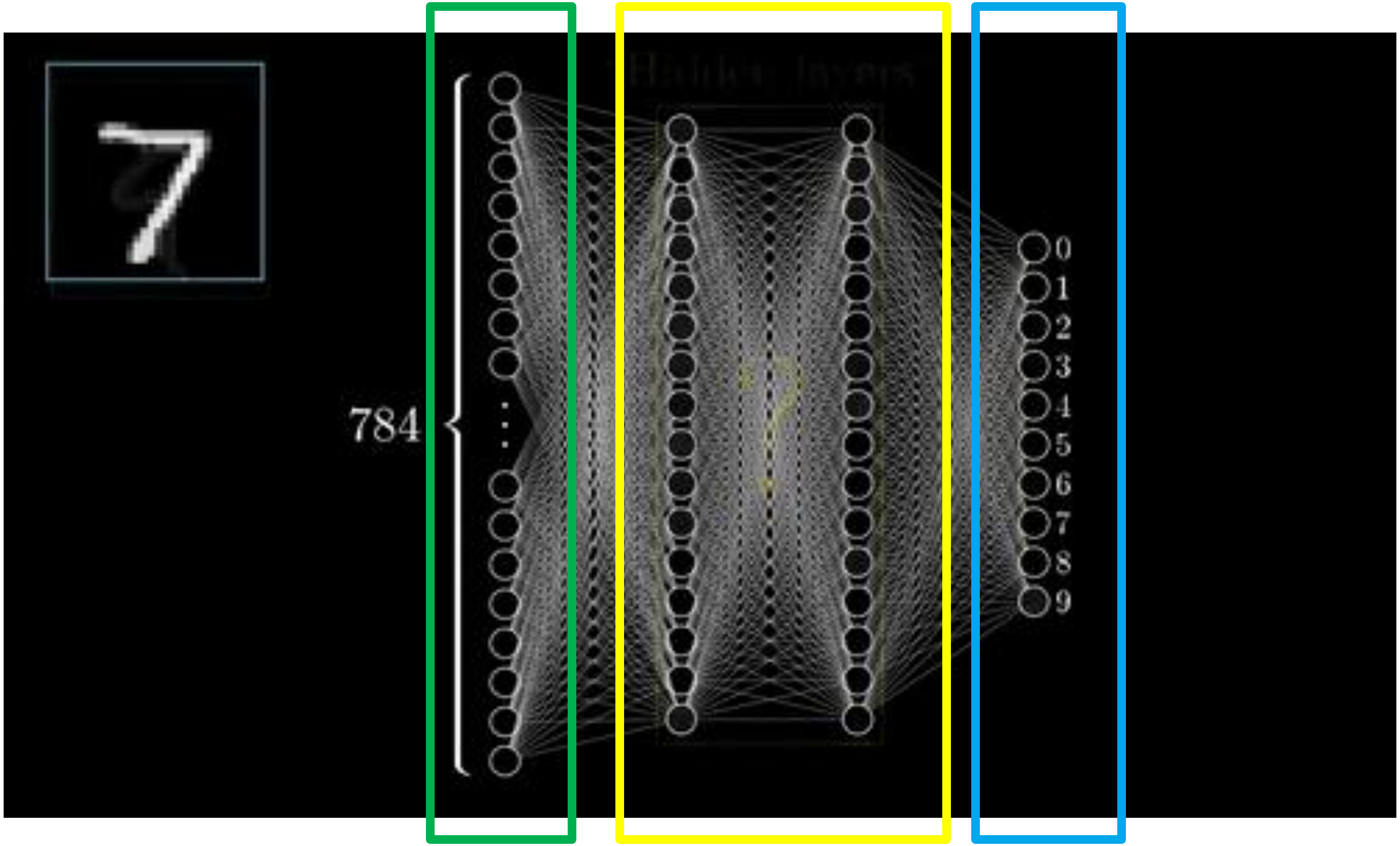
- The weights and the bias values are learned while **training the network**!


dendrites
nuclues
axon
axon
terminals
cell body


in 1
in 2
in n
Σ f
bias


Input Layer
Hidden Layers
Output Layer

Input layer Hidden Layers Output layer



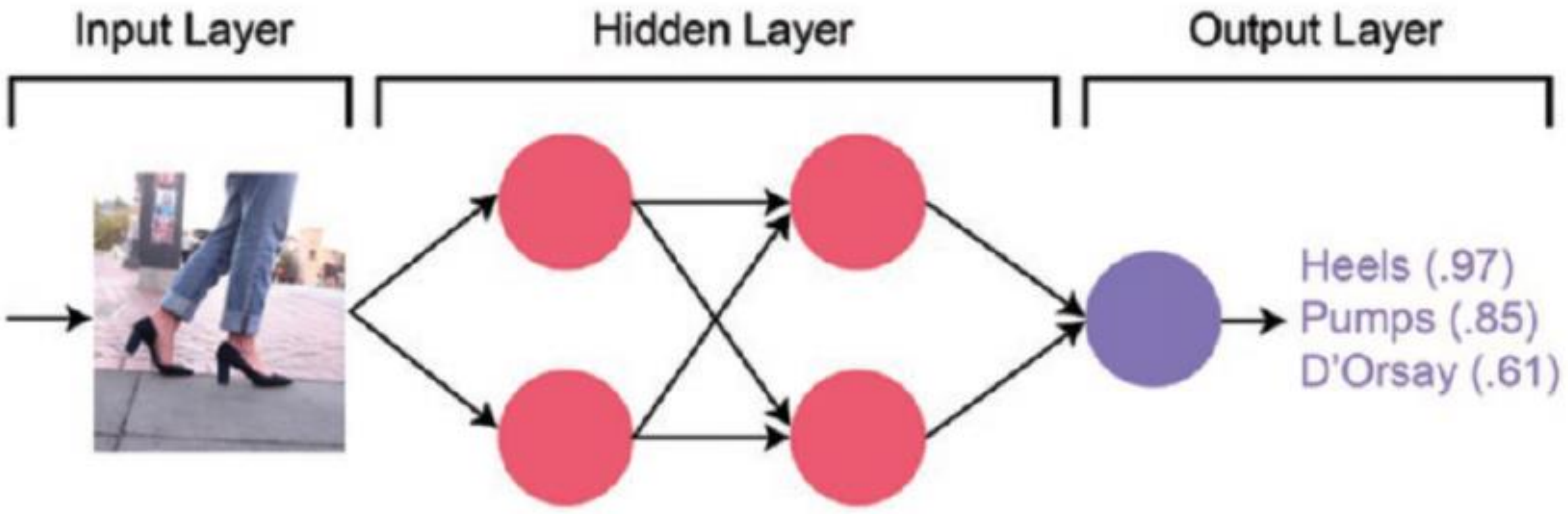
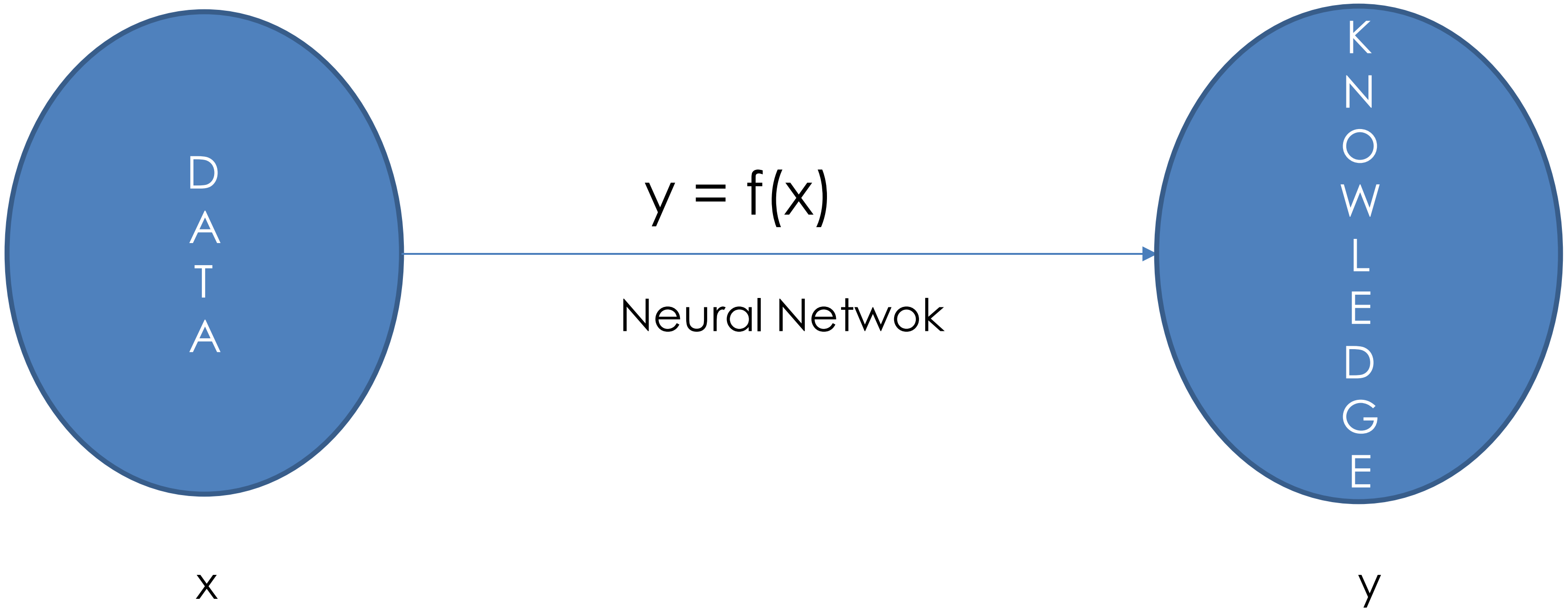


Figure 4-4. *The flow of information, input, and output in a simplified illustration of a neural network. There isn't usually a single input node, but rather a collection of images or other types of data.*

Machine learning algorithms are parametric mapping functions

- Machine learning is about exploring and developing mathematical models for learning from data;

- In practice, each machine learning algorithm models an **optimal** (parametric) **mapping function** between the data domain and something we want to learn;



How neural networks learn?

- We will not discuss about the gradient descent and in general optimization algorithms, because they require a strong mathematical foundation (derivative, partial derivatives, space functions);

- However, you must know that all the models based on neural network, optimize their weights exploiting gradient descent, to minimize a loss function;

- In practice, a loss function provides a mathematical value that inform us of how much our model is wrong while learning (e.g., if the model propose a cat on a dog picture is badly wrong)!

- In **neural network**, the gradient descent is used to minimize this error along with **backpropagation**, which propagates the error to all the neurons in the network, from the output layer to the input one;

- The **learning rate** is the factor that control the force of the changes in weights;

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

What is deep learning?

Lorenzo Stacchio

PhD student in Computer Science

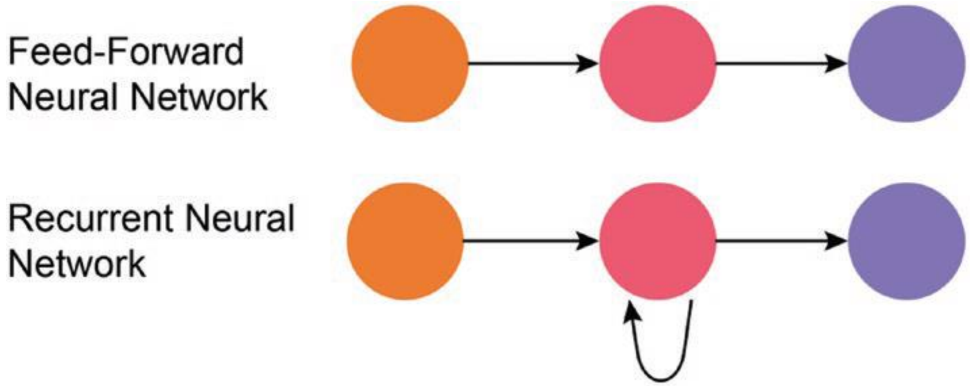
Department for Life Quality studies

Deep learning and universal approximation

- Deep learning is sub-field of machine learning methods based on neural networks;
- Deep-learning architectures such as **recurrent neural networks and convolutional neural networks** have been applied to fields including **computer vision, speech recognition, natural language processing, machine translation** where they have produced results comparable to and in some cases surpassing human expert performance;
- The adjective "deep" in deep learning refers to the use of multiple layers in the network;
- A network with one hidden layer is a universal function approximator, if we use a **nonpolynomial activation** function with one hidden layer of unbounded width can.
- This means that, no matter what f(x) is, there is a network that can approximately approach the result and do the job! This result holds for any number of inputs and outputs.

Recurrent neural network

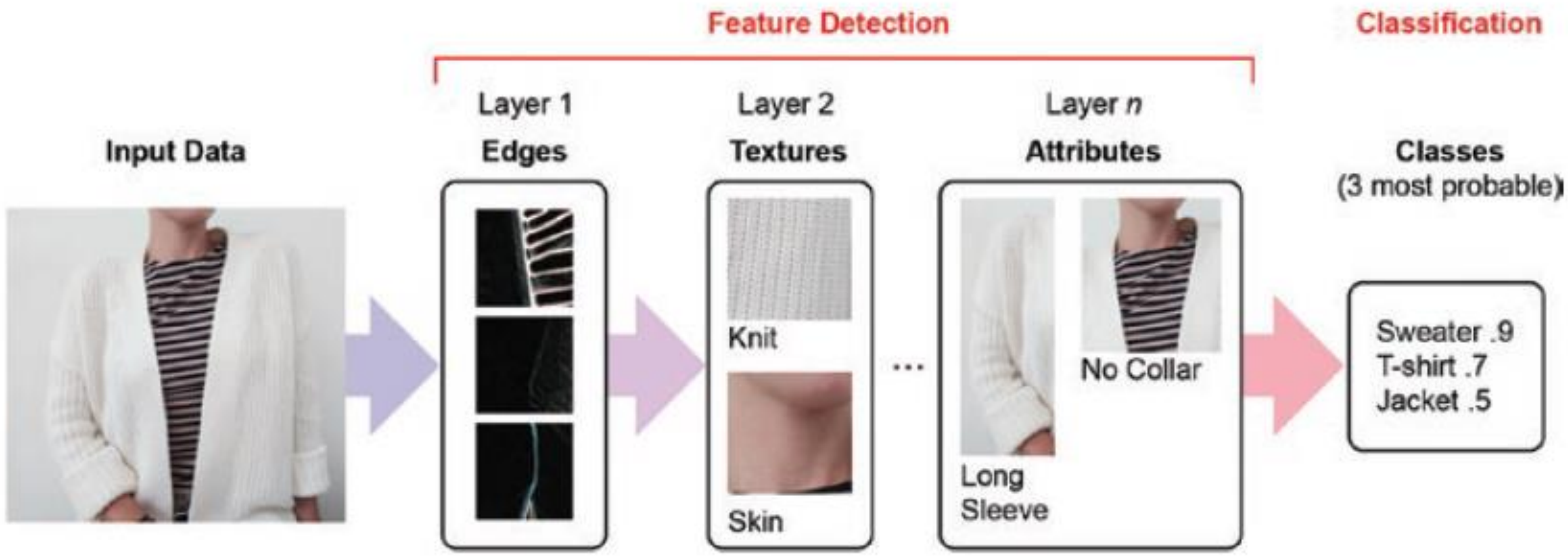
- Recurrent neural networks (RNNs) are particularly useful when it comes to sequential data.
- Arranging data in order is important for applications like natural language processing and speech recognition.
- Recurrent networks can have multiple hidden layers.
- While feed-forward neural networks can also have multiple layers, they allow signal to travel in only
- one direction, from input to output.
- This Figure shows a simple recurrent network cycling through the hidden layer, where the data is processed multiple times using the same function and parameters.



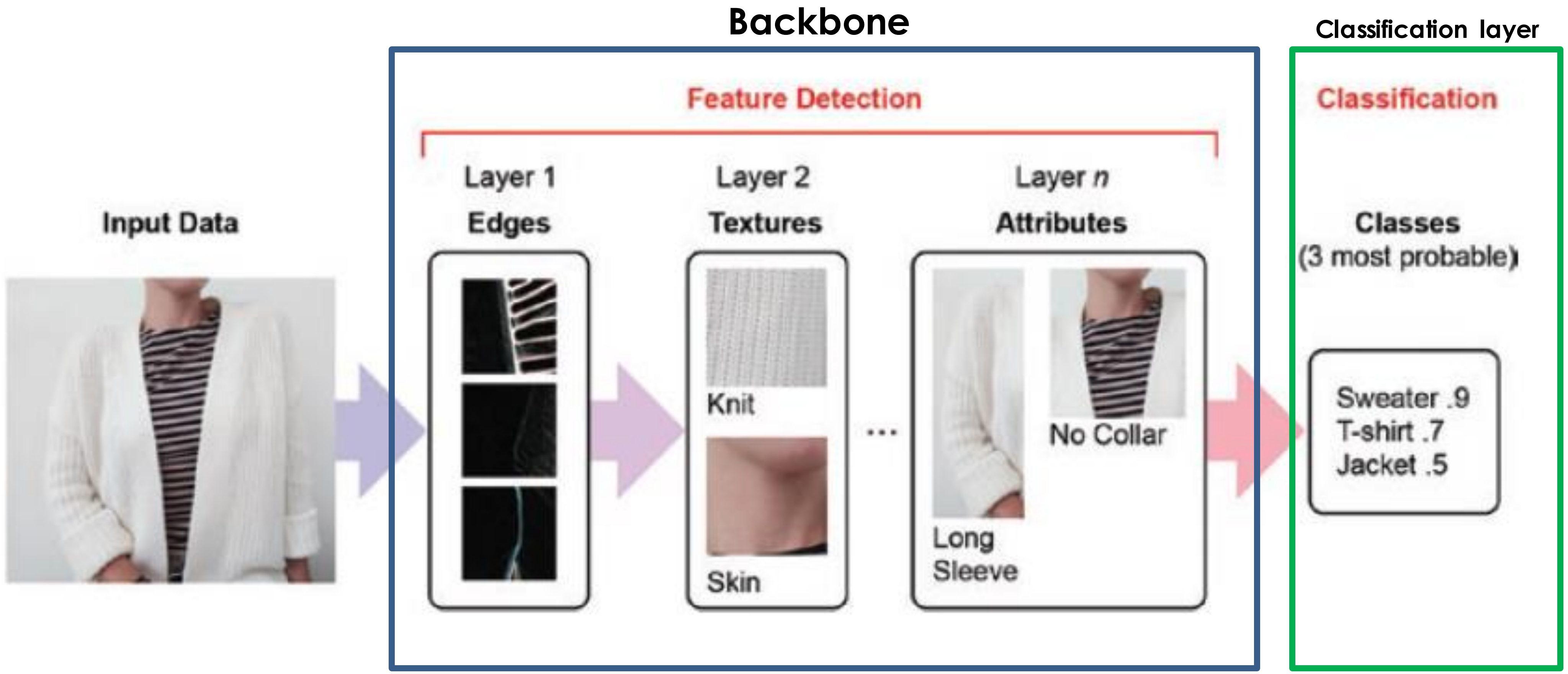
ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

Convolutional Neural Networks

- Convolutional neural networks (CNNs) are better suited for working with images.
- In CNNs, the neural network will find features in a large dataset and use those to determine what is in the image.
- The design of CNNs was inspired by the visual cortex system specifically for image-based tasks.
- It can be difficult to understand what is going on in the hidden layers of any neural network, but in a CNN, there are two major parts: **feature extraction and classification**.
- Features of an image are being detected, narrowing down what is contained in that image.
- An image with a sweater might be classified by characteristics such as knit structure, the presence of skin, sleeves, and collar;



- A CNN can have tens or hundreds of hidden layers.
- Each layer can detect different features within an image, increasing complexity with each layer, as in the image shown.



- In our language, the feature detection or extraction part is called **Backbone**;
- The classification is made by a final layer of neurons, named as **fully connected layer**;

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

What can we do with Deep Learning for fashion images?

Lorenzo Stacchio

PhD student in Computer Science

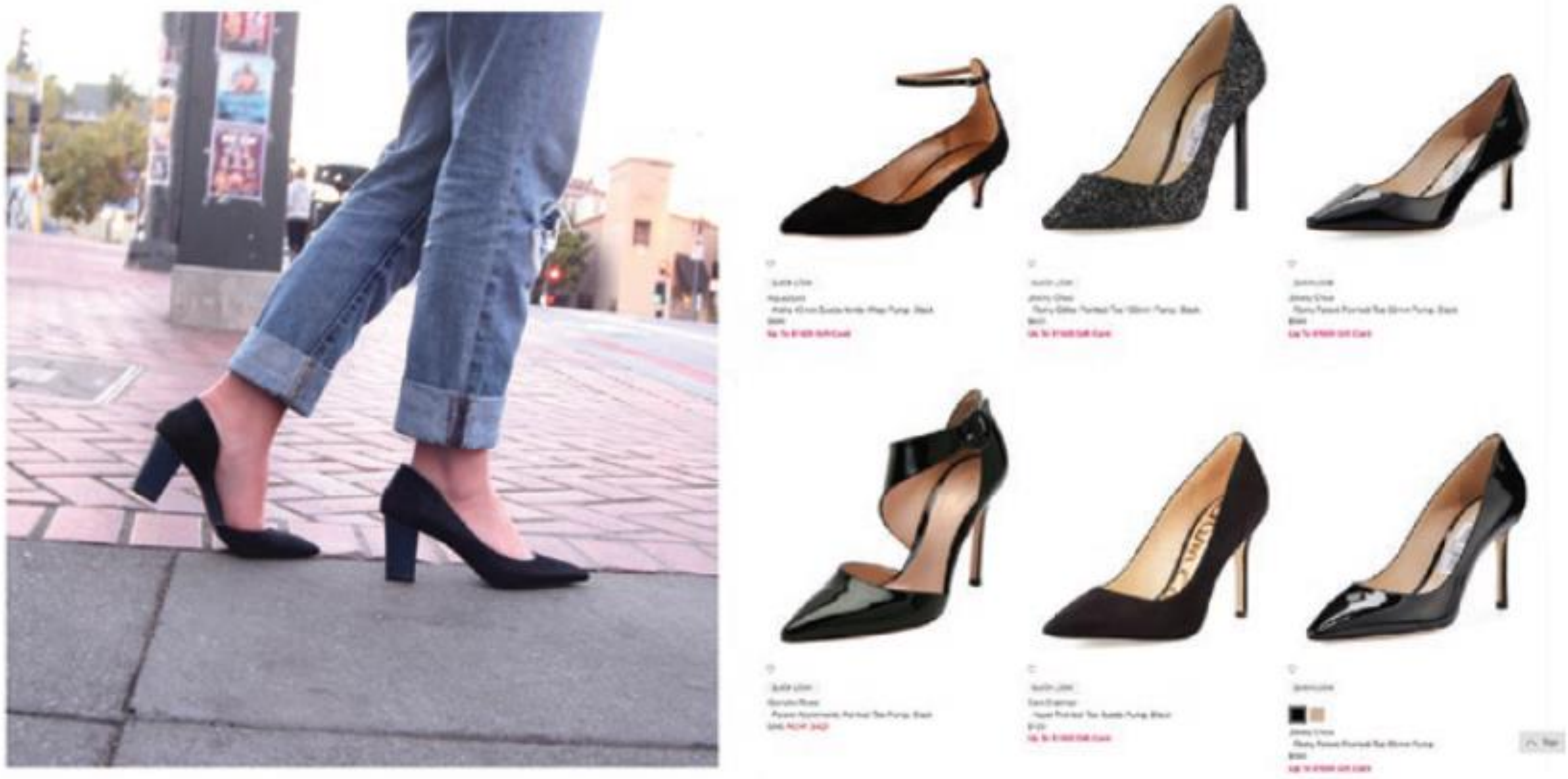
Department for Life Quality studies

Some Computer Vision task for fashion [2]

Semantic segmentation [2]

Clothing attributes recognition [2]

Visual search



Virtual Try on [9]

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Fashion trend analysis and forecasting

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

Predictive analysis

- Predictive analytics uses a variety of methods that use historical information to predict events that will happen in the future. These methods range in complexity and include data mining, basic statistics, and machine learning;

- **Recommender systems** are part of predictive analytics. They seek to understand user or customer behaviour and recommend products or services that the user is likely to like or purchase. They play a critical role for discovering products in e-commerce, but also in areas like music and video streaming;

- **Demand forecasting** is used to optimize supply-chain planning. By predicting demand for products, the fashion industry can reduce overproduction, thereby cutting costs and reducing waste;

- **Trend forecasting** is a novel sub/field, and is used to predict future trends by analysing social media. By predicting trends, the fashion industry can reduce overproduction and focus on what really matters;

Predictive analysis

- Predictive analytics uses a variety of methods that use historical information to predict events that will happen in the future. These methods range in complexity and include data mining, basic statistics, and machine learning;

- **Recommender systems** are part of predictive analytics. They seek to understand user or customer behaviour and recommend products or services that the user is likely to like or purchase. They play a critical role for discovering products in e-commerce, but also in areas like music and video streaming;

- **Demand forecasting** is used to optimize supply-chain planning. By predicting demand for products, the fashion industry can reduce overproduction, thereby cutting costs and reducing waste;

- **Trend forecasting** is a novel sub/field, and is used to predict future trends by analysing social media. By predicting trends, the fashion industry can reduce overproduction and focus on what really matters;

Trend forecasting

- Trend forecasting has always been an elusive industry, with large brands paying hefty fees for consultancies to give them advice about the future.
- This forecasting toggles between art and science.
- A trend forecasting consultancy that started as an art collective making commentary about the corporate world, to researchers at Cornell University, who have taken to social media data to study fashion's anthropology around the world [14];
- Today, social media platforms provide a place where information is aggregated, leaving a treasure trove for data mining;
- This data can be used to unlock the scientific piece of the puzzle, at least, when it comes to understanding and forecasting the cultural trends that drive the fashion industry.

- The traditional methods of trend forecasting, including relying on sales data and trend reports, are often how brands gain an understanding of customer desires and behaviors.

- In today's age of digital personalization and social media, consumer demands are changing their purchasing behavior.

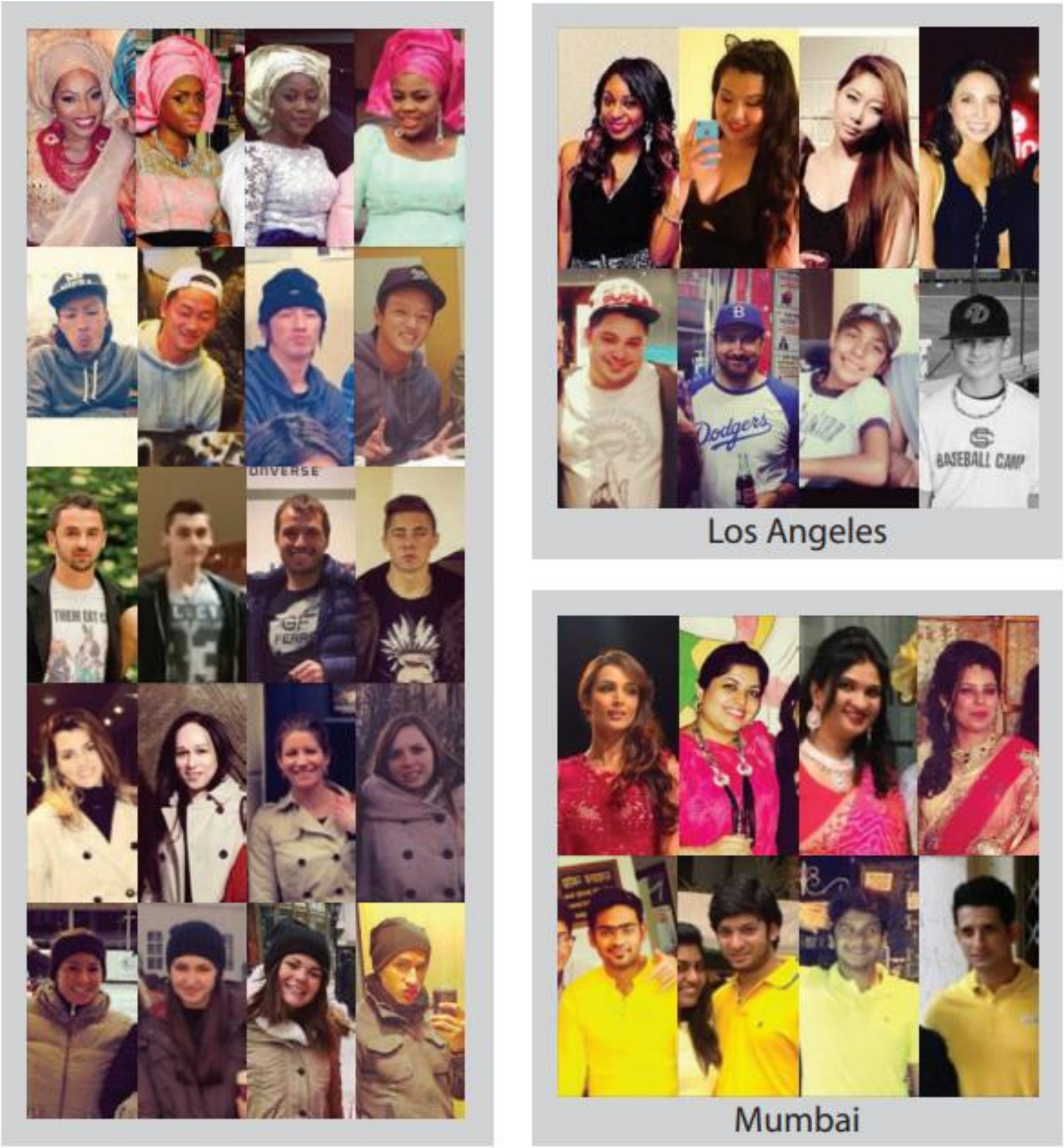
- The forecasted trends end up in the hands of dozens, if not hundreds, of designers in the industry. Color trends are propagated to textile designers and manufacturers, and the predictions can drive the supply more than the other way around;

- In the meantime, consumers might already be expressing their perspectives relating to these trends in a public forum such as Twitter, Instagram, or Facebook. They might already be expressing boredom with a color that designers have chosen to be dominant in next season's palette.

- In 2017, researchers at Cornell University released a paper titled "StreetStyle: Exploring Worldwide Clothing Styles from Millions of Photos."

- The main hypothesis was that by referring to millions of images uploaded by users on social media, they could index local and global trends in what the users were wearing.

- Their approach took an anthropological angle of dissecting differences in garments through time, place, and style.



(b) Global clusters

(c) Representative clusters per-city

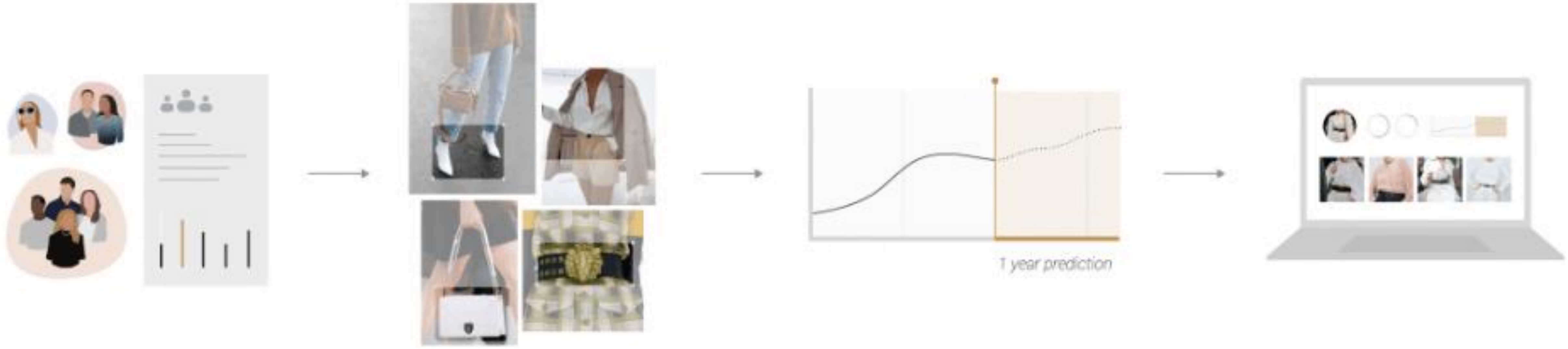
Euritech case study [18]

- Heuritech's methodology defines audience panels on social media and applies its image recognition technology to social media images to assess more than 2,000 components, from shapes and attributes to fabrics, prints, and colors.

- Heuritech's machine learning algorithms can predict trends up to one year in advance, which brands can find in our platform to make decisions on assortment, marketing, merchandising, and more.

- Heuritech was founded in 2013 by PhDs in Machine Learning who believed that images could tell powerful stories if one knew how to read them.

- They developed a proprietary image recognition technology to analyze fashion images on social media. In today's world, where millions and millions of pictures are shared each day on Instagram and other platforms, it's a powerful tool to analyze within seconds what's happening out there.



Step 1
Define representative panels

Step 2
Applying our computer vision technology to millions of social media images stemming from these panels

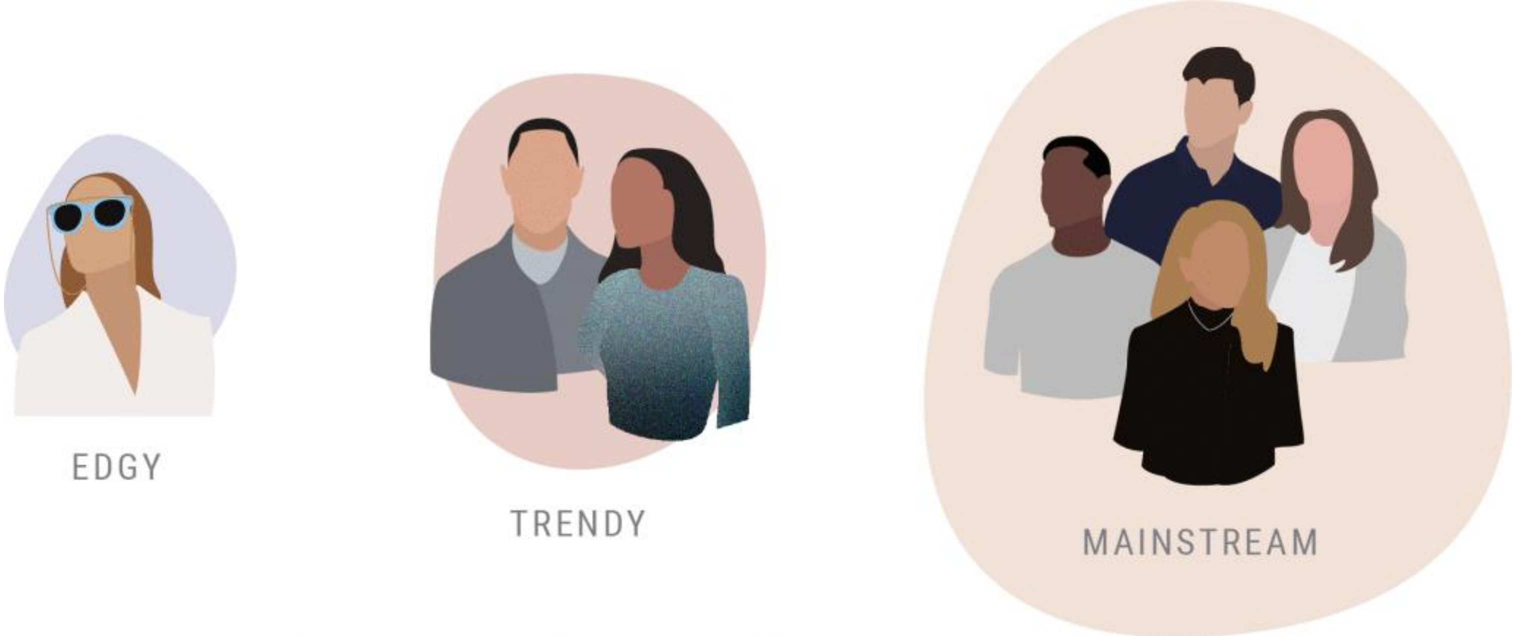
Step 3
Use our machine learning forecasting algorithms to predict trends up to one year in advance

Step 4
Insert Heuritech's data into our market intelligence platform

Heuritech's AI-based trend forecasting methodology

Step 1: Defining audiences' panels on social media to have a holistic view

- 100 million pictures are shared each day on social media.
- One can easily get lost in the amount of data, especially fashion pictures: indeed, according to Instagram, #fashion is the 4th most used hashtag on Instagram.
- Furthermore, the question to ask here is what, and whom, to analyze – indeed, we want to understand the underlying influence groups on Instagram to interpret the data and analyze it with purpose.
- Whom to analyze: We want to be representative of the whole population to be statistically reliable.
- Criteria include:
  - Accounts demonstrating an interest in fashion;
  - Many accounts that keep on posting regularly to have a holistic view;
  - Representative of the different style segments of the fashion market: edgy, trendy, and mainstream accounts



- **Edgy accounts**: People with bold and distinctive style. They represent the smallest segment. Still, their content is very niche, and they tend to post often. They cover all the main fields of the fashion industry, whether they are professionals (stylists, journalists, influencers) or industry authorities (luxury, sports, fast fashion, high street, beauty);

- **Trendy accounts**: Fashionable people looking for the latest styles and who tend to help spread trends across the market. They are more substantial in number than edgy accounts but remain much more qualified and rarer than mainstream accounts. They also tend to post often;

- **Mainstream accounts**: People looking for safe clothing choices and who will follow a trend rather than shape it. They represent the largest panel. Their content is more casual and occasional.

Step 2: Applying our computer vision technology to millions of social media pictures

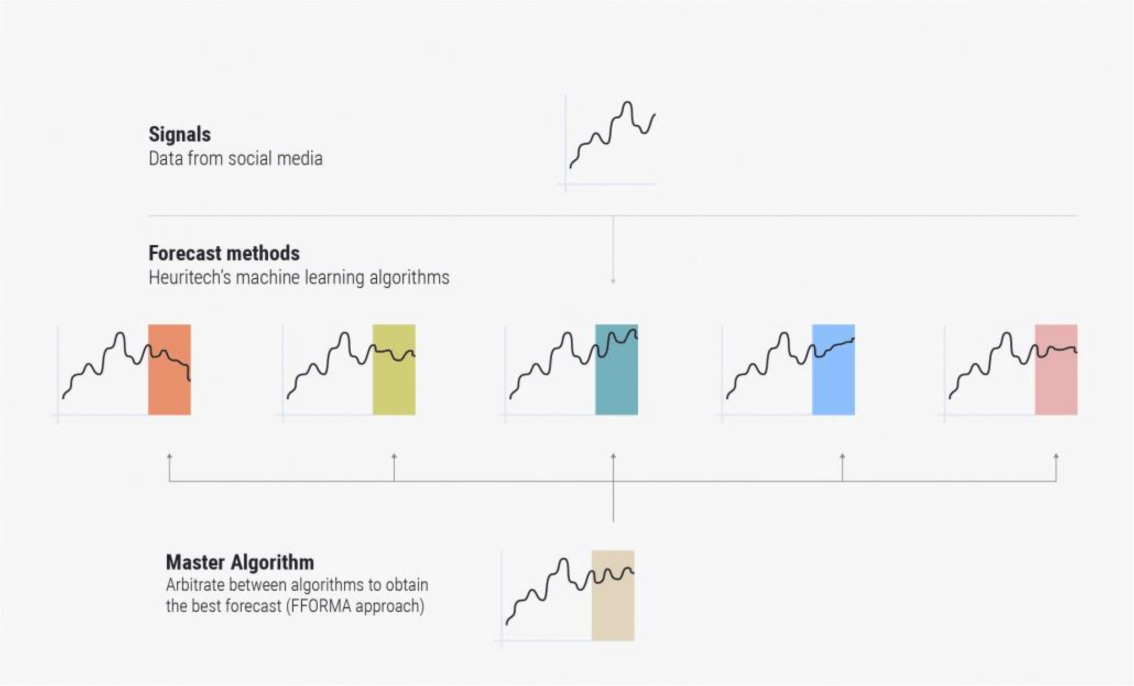
- In any given picture, Heuritech's proprietary computer vision technology tailored to fashion can detect and categorize more than 2,000 components, from shapes and attributes to fabrics, prints, and colors;

- Combining these components allows us to precisely identify trends for different panels;

- For example, they can recognize boots that are beige, pointed, and have a kitten heel;

- Thanks to the millions of images analyzed each day, we are able to assess the volume of the beige, pointed, kitten heel boot trend;

Step 3: Use our machine learning algorithms to predict fashion trends up to one year in advance



- It goes without saying that the main interest when focusing on data would be to know how these trends will behave;

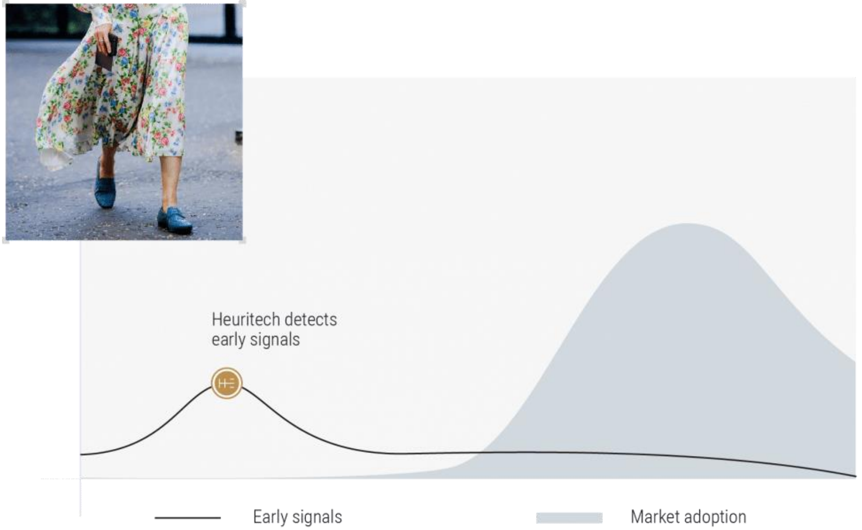
- Predicting the future is notoriously hard, and a tricky endeavor. At Heuritech, a forecaster toolkit is adopted: a series of forecasting methods and a master algorithm, whose purpose is to arbitrate between these methods to automatically obtain the best consensus

- Most of the forecast methods stem from a classical statistics approach, relying only on the past dynamics of a trend to predict its future;

- While this approach provides good overall results and is state-of-the-art in many applications, it ignores the general context of a trend and may fail to predict what is most interesting: sudden, emerging, niche trends which will become the new, mainstream fashion trends.

- To be able to forecast these seemingly random evolutions, we developed an in-house deep-learning approach that feeds with what is defined as **early signals**;

- Those blips of activity among edgy influencers or sub-cultures segments of the trend have the potential to tilt the way we perceive and wear fashion items;



Euritech’s clients

VOGUE

“Heuritech has turned social media into an artificial intelligence tool that predicts the development of trends”

BOF

“The world’s first artificial intelligence (AI) service that can predict fashion trends”

Forbes

“Visual recognition specialist Heuritech is working with Louis Vuitton and Dior to provide predictive analysis on trends.”

adidas

havaianas

PRADA

LV

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Virtual style assistants

Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

Virtual style assistants (VSA)

- The virtual style assistant is useful when it comes to fashion sales;
- Bringing personal stylists to the retail and e-commerce environment can improve a brand's ability to match consumers with desired products and guide context-based decision-making;
- It can also be brought into the home of the consumer, making better use of existing wardrobes.
- An AI stylist can help consumers discover apparel items that meet a wide variety of expectations: flatter their figure, work well as an outfit, align with current trends and values, and provide a personalized experience.
- To build a VSA we need three building blocks: **Personal AI stylist, Virtual assistants and Voice interfaces**;

Personal stylist

- Personal stylists help individuals to look their best by curating clothing, outfits, makeup, and other aspects of personal style. Hiring a personal stylist is not cheap;
- Several solutions on the market use technology to address a desire for personal styling. Platforms that act as social networks between stylists and consumers bridge this gap.
- However, because they continue to rely on services by humans, these services will not scale at the rate that an AI-style assistant can. It is because of its scalability and potential to capture value in this field that the AI stylist is a compelling use case;
- AI stylists just aren't very good compared to humans right now. Style encompasses many things machines still don't understand. Indeed, Stylist curate aesthetics for their clients, often interpreting between the lines as the clients describe their personal style;
- Nevertheless, there are many interesting works that embrace this topic [11, 12, 13].

Virtual assistants and voice interfaces

- Virtual assistants provide the basis for the virtual style assistant and have become increasingly prevalent in consumer electronics.
- These assistants generically refer to a software agent that provides services to individuals.
- Today, this is often carried out through voice command prompts. By brand name, virtual assistants refer to Apple's Siri, Google's Google Home and Google Assistant, Amazon's Alexa, and other similar AI-based assistants.
- The first step of the process for these systems is interpreting human input. For today's virtual assistants, that usually means converting speech to text.
- Then, those text is computationally managed to execute commands or run other softwares;
- **Think of Alexa's skills.**

Features of the VSA

- The virtual style assistant is unique compared to other virtual assistants in that it emphasizes the use of images more than any other use case.

- Images are critical to giving style advice.

- The design of a virtual style assistant requires a few key components:

    - the ability to take a photo of oneself;
    - the ability to store those photos in the application;
    - It also requires underlying technologies such as computer vision capabilities for **image recognition and visual search, recommendation engines, analytics, and access to fashion products**;

Existing examples

- Images collected by a virtual style assistant can be used in several ways. In this and other existing virtual style assistants, images are used to catalog a wardrobe, make product recommendations, and provide insights into the user's style preferences;
- Amazon's Echo Look **was** the most public example of a virtual-style assistant built on top of computer vision;
- Various startups have begun to develop technologies for this purpose;
- For example, Lookastic, which makes recommendations based on your response to their "What's in your closet?" interface;
- Defining personal style is difficult, however, Fashion is one of their most popular types of content;
- We will reach good results in no time.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

The role of Extended reality in fashion

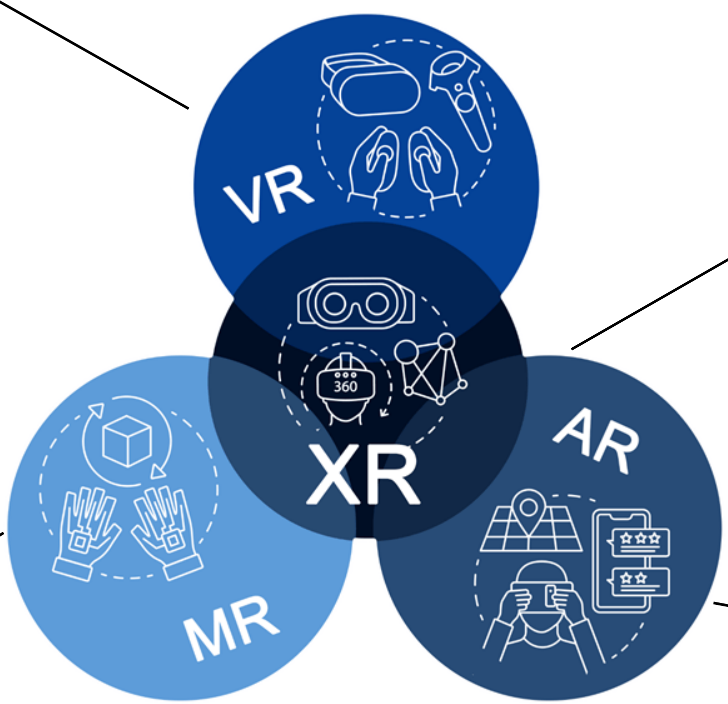
Lorenzo Stacchio

PhD student in Computer Science

Department for Life Quality studies

VR creates a fully computer-generated 360° environment. Users are totally immersed in the virtual world and interact with it, while the real world remains suspended. With a head-mounted display, users visually perceive the virtual world, while the real one remains completely excluded and not visible.

XR: **through immersive technology, users find themselves in a virtual world or interact with an augmented world. Virtual contents are therefore perceived as extremely close to reality. The degree of immersion is closely linked to the possibility of interaction with the digital environment.**



MR: is the union of elements of AR and VR. This form of extended reality combines our real world with the virtual environment and thus giving rise to a new environment. The user interacts simultaneously with the real and virtual environment. Physical objects in the real world affect digital elements.

AR: combines virtual objects and information in the real world, thus expanding the latter. Virtual elements are for example images, texts and animations. In augmented reality, the real environment remains at the center of perception. For users, interaction with virtual objects is only limited or even impossible.

VR

Replacing
your environment
with digital content

AR

Enhancing
your environment
with digital content

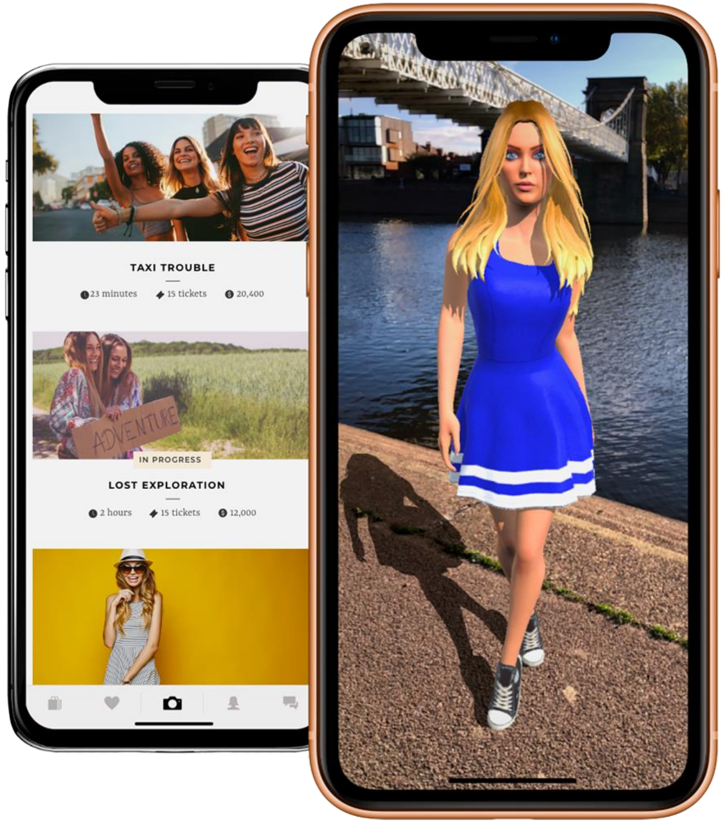
MR

Why talk about XR?

- Extended reality or “XR” is just one of the many disruptive tools changing the way that we innovate and interact;
- Alongside machine learning as one of the exciting new opportunities of a digital future;
- For instance, with extended reality solutions, teams could work together on building new products and prototypes in a virtual environment, without wasting resources; Using AI technology, it could be possible to determine which solutions really work using information from previous tests and data;
- Another examples is information fruition for decision making AI helps predicting future scenario and Extended reality helps to visualize and manipulate variables in that scenario;
- **What from a consumer perspective?**

Augmented reality for fashion

Smart mirrors



3D-model visualization

Virtual Reality

Virtual Fashion shows

Virtual Fashion shows with 360 video

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Virtual Reality: X-commerce and Voice assistants [10]

- The widespread of eXtended Reality (XR) technologies is supported by a rapid increase of their performances along with a progressive decrease of their costs, making them more and more attractive for a large class of consumers.;

- This may foster new opportunities for e-commerce strategies, giving birth to an XR-based commerce (x-commerce) ecosystem, that could give a more brick-and-mortar store-like experience.

- One interesting and consolidated one amounts to the interactions among customers and shop assistants inside fashion stores;

- For this reason, we designed and implemented an XR-based shopping experience, where vocal dialogues with an Amazon Alexa virtual assistant are supported, to experiment with a more natural and familiar contact with the store environment;

- **We discover that, even if the proposed implementation with Amazon Alexa as a voice assistant has imposed some design limitations, the presence of a voice assistant may improve the perceived immersion.**

Sources

- [1] Zou, X., & Wong, W. (2021). fAshIon after fashion: A Report of AI in Fashion. arXiv preprint arXiv:2105.03050.
- [2] Cheng, W. H., Song, S., Chen, C. Y., Hidayati, S. C., & Liu, J. (2021). Fashion meets computer vision: A survey. ACM Computing Surveys (CSUR), 54(4), 1-41.
- [3] Russell, S., & Norvig, P. (2002). Artificial intelligence: a modern approach.
- [4] Campbell, M., Hoane Jr, A. J., & Hsu, F. H. (2002). Deep blue. Artificial intelligence, 134(1-2), 57-83.
- [5] Sagiroglu, Seref, and Duygu Sinanc. "Big data: A review." 2013 international conference on collaboration technologies and systems (CTS). IEEE, 2013.
- [6] Weizenbaum, J. (1983). Eliza—a computer program for the study of natural language communication between man and machine. Communications of the ACM, 26(1), 23-28.
- [7] Yuan, Y., & Lam, W. (2021). Sentiment Analysis of Fashion Related Posts in Social Media. arXiv preprint arXiv:2111.07815.
- [8] Rasool, A., Tao, R., Marjan, K., & Naveed, T. (2019, March). Twitter sentiment analysis: a case study for apparel brands. In Journal of Physics: Conference Series (Vol. 1176, No. 2, p. 022015). IOP Publishing.

- [9] Lewis, K. M., Varadharajan, S., & Kemelmacher-Shlizerman, I. (2021). Tryongan: Body-aware try-on via layered interpolation. ACM Transactions on Graphics (TOG), 40(4), 1-10.

- [10] Morotti, E., Stacchio, L., Donatiello, L., Roccetti, M., Tarabelli, J., & Marfia, G. (2021). Exploiting fashion x-commerce through the empowerment of voice in the fashion virtual reality arena. Virtual Reality, 1-14.

- [11] Hsiao, W. L., & Grauman, K. (2017). Learning the latent" look": Unsupervised discovery of a style-coherent embedding from fashion images. In Proceedings of the IEEE International Conference on Computer Vision (pp. 4203-4212).

- [12] Di, W., Wah, C., Bhardwaj, A., Piramuthu, R., & Sundaresan, N. (2013). Style finder: Fine-grained clothing style detection and retrieval. In Proceedings of the IEEE Conference on computer vision and pattern recognition workshops (pp. 8-13).

- [13] Lin, Y., Ren, P., Chen, Z., Ren, Z., Ma, J., & De Rijke, M. (2019). Explainable outfit recommendation with joint outfit matching and comment generation. IEEE Transactions on Knowledge and Data Engineering, 32(8), 1502-1516.

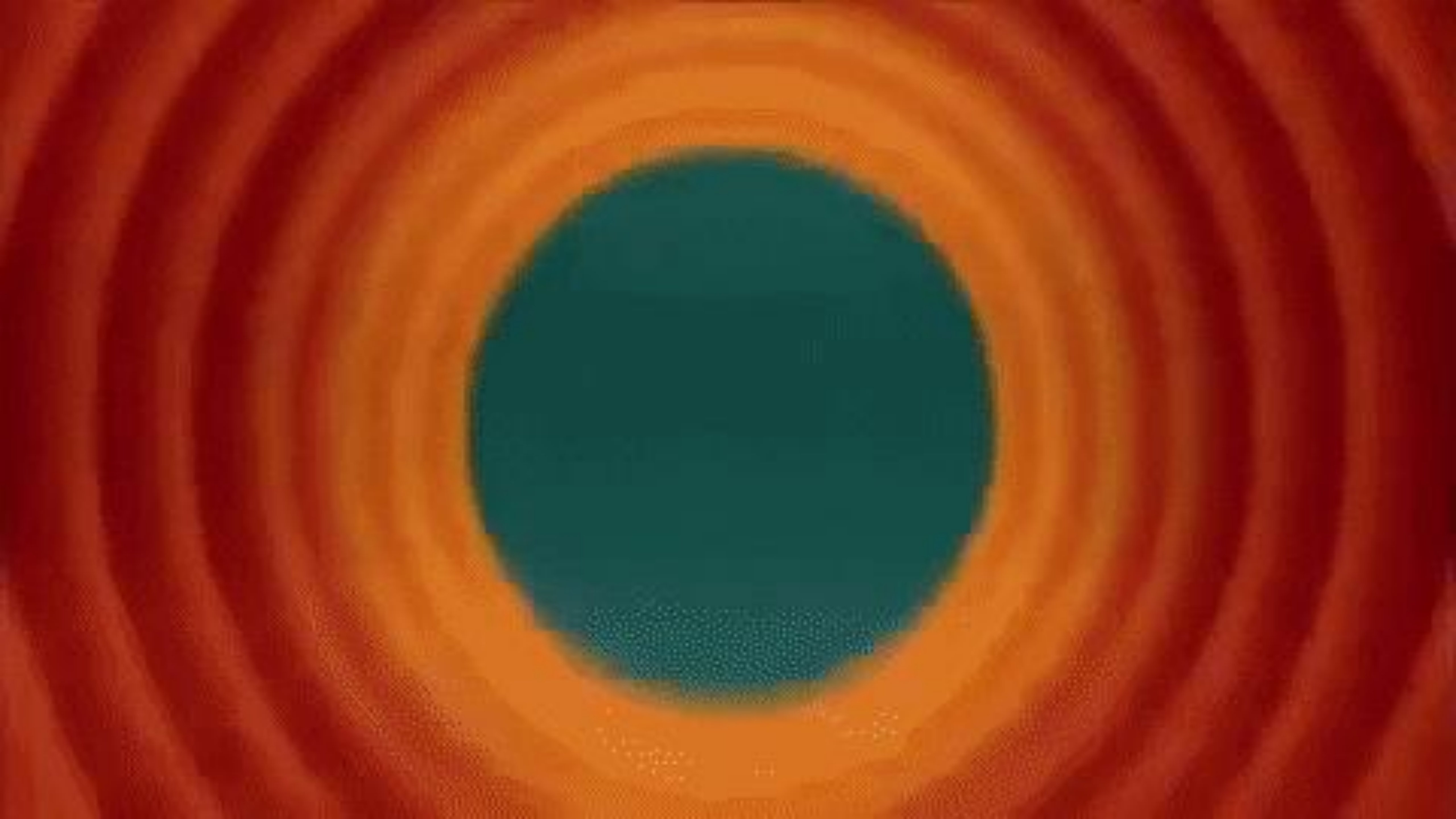
- [14] Matzen, K., Bala, K., & Snavely, N. (2017). Streetstyle: Exploring world-wide clothing styles from millions of photos. arXiv preprint arXiv:1706.01869.

- [15] Grauman, K. (2020, January). Computer vision for fashion: From individual recommendations to world-wide trends. In Proceedings of the 13th International Conference on Web Search and Data Mining (pp. 3-3);

- [16] Al-Halah, Z., Stiefelhagen, R., & Grauman, K. (2017). Fashion forward: Forecasting visual style in fashion. In Proceedings of the IEEE international conference on computer vision (pp. 388-397).

- [17] Lee, D. J. L., Han, J., Chambourova, D., & Kumar, R. (2017). Identifying fashion accounts in social networks. KDDW on ML meets fashion.

- [18] Euritech, How Heuritech forecasts fashion trends thanks to AI, https://www.heuritech.com/articles/fashion-solutions/how-heuritech-forecasts-fashion-trends-thanks-to-artificial-intelligence/.



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Lorenzo Stacchio

Department for Life Quality studies

lorenzo.stacchio2@unibo.it