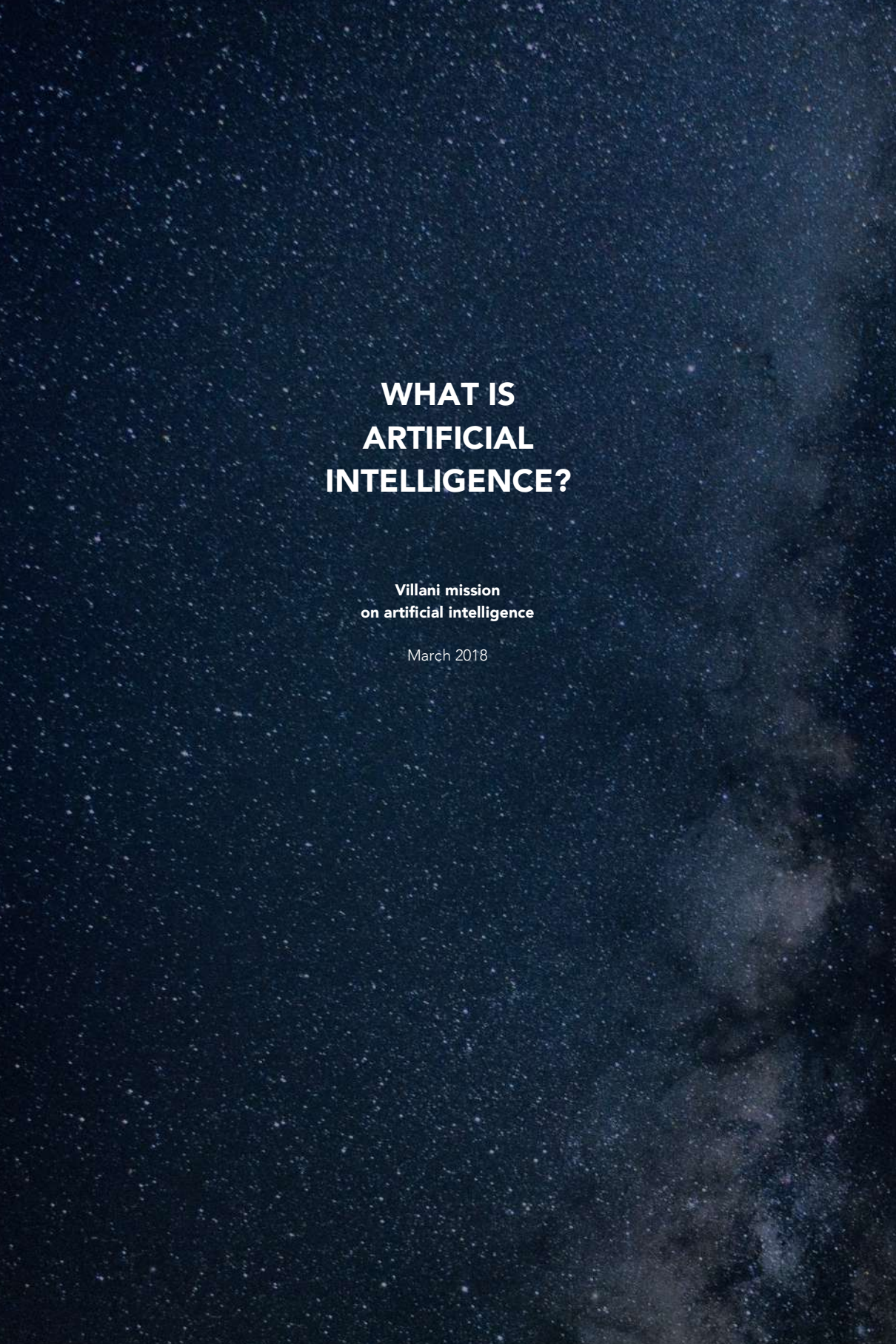


WHAT IS ARTIFICIAL INTELLIGENCE?

Villani mission on artificial intelligence

March 2018

Defining artificial intelligence (AI) is no easy matter. The field is so broad that it cannot be limited to a specific area of research. AI is more of an ambition: it seeks to understand how human cognition works by creating cognitive processes that emulate those of human beings.

AI is at the crossroad of multiple fields: computer science, mathematics (logics, optimisation, analysis, probabilities, linear algebra…), cognitive science… These core scientific disciplines also need to be mixed with the specific knowledge of the fields they are applied to, and each algorithm in AI is supported by a mix of techniques: semantic analysis, symbolic computing, machine learning, exploratory analysis, deep learning and neural networks…

Defining AI

Imitate human capabilities

Marvin Lee Minsky, who is considered as one of the founding fathers of AI, defines it as follows: "the science of making machines do things that would require intelligence if done by men. It requires high-level mental processes such as: perceptual learning, memory and critical thinking."

In other words, artificial intelligence is the science of building computer programs that aim to perform tasks that would require some intelligence if they were done by human beings. Therefore, no human activity seems to be out of reach: moving from one place to another, learning, reasoning, socializing, creativity etc. Nevertheless, we are still far from creating a machine that would be able to match or outperform human capabilities in all fields.

Grounded on logic

Some researchers such as John McCarthy, also one great figure in the AI field, argue that rational logic is a more relevant "standard" of intelligence measurement than human capabilities. This approach to AI takes advantage of the tools offered by mathematical logic to formalize the complex tasks to be accomplished by artificial intelligence machines. The main issue this approach has to deal with is the formalization of the tasks. For example, we can ask ourselves which apprehensible rules of reasoning allow us to distinguish almost immediately a picture of a cat from that of a dog among the hundreds of possible races? Moreover, how to take into account the fact that the image could be blurred? Besides, in our daily decisions, uncertainty is constantly present and managed by other rules than logic: habit, emotions and intuitions play an essential role in these mechanisms.

What is the goal of AI?

One could wonder why we even bother to build smart machines. The primary goal would be that they make our lives easier by taking care of complex and repetitive daily tasks: find our best itineraries to reach our friend's house, select the most relevant content according to our tastes, translate texts… That is what experts refer to when they talk about "weak AI".

These programs are only able to perform one single and delimited task, hence the "weak" qualification. Algorithms building our favorite playlists, finding the most relevant search results or identifying our friends on social media are not able to do anything more than this. Yet, this does not make these algorithms any less useful, but it could lead to some artificial stupidity: if one AI technique is

trained to distinguish cats versus dogs, if provided a human face picture it will still answer either dog or cat without flinching.

To the other end sits what is called "strong AI", able to match and even outperform human capabilities in all fields. We are still very far from building such a machine, and any unconvinced reader just has to engage in a short discussion with the most advanced chatbot to be fully convinced. The fact that the prospect of a "strong AI" machine is highly unlikely anytime soon does not prevent people, experts and regular citizens alike, from passionately debating its destructive or positive impact. It is no accident that the sci-fi culture, which has been central to the development of AI, is full of this kind of argument. Further, previous technological advances also raised similar fears, and it is not surprising that we observe the same phenomenon with AI. Nevertheless, policymakers and regulators are already urged to consider the fundamental questions brought about by the development of AI; especially considering diverse projections regarding its potential dramatic effect on employment. That's a tricky matter for regulators; little accustomed to such speculative topics.

How to build AI?

Symbolic AI: grounded on logic

Symbolic AI refers to a class of approach to AI that aims to reproduce the reasoning mechanisms of humans. These mechanisms, put in logical form, are produced by a priori modeling of knowledge and rules of logic to perform the task at hand. One of the distinguished illustrations of symbolic AI technique is the expert systems algorithm, capable of producing reasoning from known facts and rules. For instance, diagnostic rules defined by a physician are transcribed into code to automatically produce diagnoses.

As an a priori approach, these techniques require to split into fixed classes the objects processed by the algorithms according to some predefined features. As such, it is very challenging to capture the complexity of the real world and establish fixed categories.

Machine learning

Alongside symbolic AI, machine learning techniques were developed in order to model cognitive process directly from previous experience.

Every machine learning technique has two steps: the first one is the learning phase that uses the input data (e.g. pictures of cats and dog for the classification task) to find the parameters that best fit the task at hand. The second step takes the learned parameters as input and performs the task accordingly. This is the inference phase.

Among machine learning approach, we can identify a few broad classes of problems.

Supervised learning

In supervised learning, one learns a set of rules in order to perform a specific task according to a given set of examples for which we are given the expected result. For example, we could be given a million of pictures of cats and dogs identified as such. The learning phase progressively changes the parameters of the model in order to make the error as small as possible on the known (and unknown) examples. In the inference phase, the algorithm performs the task on unknown data—e.g. it classifies dogs and cats on previously unseen pictures.

Unsupervised learning

In unsupervised learning, we only get unlabeled data with no additional knowledge. In the learning phase, the goal is to find the underlying structure of the data such as categories. The inference phase is the same as the supervised one. The unsupervised approach is especially useful to find behavioral profiles from the recorded activity without being given beforehand any prior knowledge about these profiles.

These approaches are at the core of current research projects since their success would enable AI techniques to rely less on the availability of labeled data, which is very expensive to produce.

Reinforcement learning

Reinforcement learning is different from supervised and unsupervised learning because it does not rely on any pre-existing set of examples as input. Instead, leaning happens directly through interactions with the environment: the machine performs an action, the state of the environment changes according to this action, and the machine gets a reward depending on the result.

The reward enables the machine to discover the best actions to perform according to the result it aims to get. A recent and famous example of reinforcement learning is AlphaGo: here, the action performed is a move on the board, the environment is the board with the stones on it, and the reward is the outcome of a game, either win or loss.

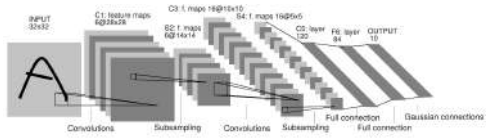
Deep learning

Deep learning is a subfield of machine learning that has come to a new era since 2006. These models are built from basics components called "neurons" that are organized into successive layers. These neurons are linked by connections whose weight is adjusted in the learning phase. Every neuron maps its input to its output with a transfer function which is schematically inspired by the brain, but it is well known that this is a crude analogy.

After the learning phase, the neural network is theoretically able to split the input data into a hierarchy of features representing multiple abstraction layers. For example in the case of face recognition, the first layer identifies the elementary patterns such as lines, edges, corners, the next ones find large patterns (lips, forehead, eyes…), and the description is progressively refined (size of the nose…) until the description is complete.

Neural networks are the go-to approach in signal processing (sound, images, language, video etc) since they are able to extract complex spatial and temporal structures.

Two kinds of models are especially well known: convolutional networks (especially in computer vision) and recurrent networks for time series and language processing applications, in which the notion of layer is a little fuzzier.



Applications

Let's finish with some concrete examples.

Image and video recognition

Interpreting an image—recognizing a person or an object and its surrounding environment—is a relatively easy task for a human being. Every day, our brain process effortlessly complex visual information: a family picture, a car, a landscape. However, it is a very challenging task for a computer.

Yet the stakes are high because the development of autonomous cars (for the perception of its surrounding environment), the automating labeling of images, the improvement of identification systems, the detection of pathologies from medical imaging devices are all dependent on advances in image and video recognition.

How do social media platforms recognize faces on pictures?

They use convolutional neural networks. According to the same principle of supervised learning for the classification of the pictures of dogs and cats described above, the system learns to distinguish the faces of user's friends for whom he has labeled images. When a photo is uploaded to the platform, the system only has to categorize the new faces present by matching them with the labelled faces from its database.

Translation

In order to build a translation software, we start by building a large database of texts translated by human translators that serve as models. These translations often come from books, official documents produced by international organizations (United Nations, European Commission) and authoritative websites. We talk about millions of texts…

Then, a neural network goes process the texts to find statistical patterns. From these patterns, it builds equivalences between sequences of words, which allow it to translate an unknown text.

However, translation softwares have fewer documents translated for some language pairs. For instance, there are more documents translated from French into Spanish than from Danish to Romanian. That's why the quality of translations varies a lot from one language to another.

Content recommendation

To recommend content to their users, online marketplace or streaming platforms, use AI systems which operate according to a different approach from those of the learning techniques exposed above. Here, input data are composed of all the past choices made by users. This dataset is used to create fake user profiles and product categories, here "fake" means that they represent average user behaviours. Because of the wide variety of users and products, it is impossible to infer consumption behaviours categories directly from their consumption choices. Each user can then be analyzed from these fake categories, which makes it possible to compute a "fake proximity" between users more reliably than simply counting the products they ordered. Afterwards, the system can recommend to a specific user a set of products likely to match his/her preferences because they have been selected according to similar fake users' preferences.