BLU SPHINX

Artificial Intelligence: Myths and Realities

Thomas Dillig, PhD Aurelie M.H. Beaumel

The buzz around Artificial Intelligence (AI) can't seem to stop these days. But increased interest comes with increased uncertainty. How will AI affect our lives and societies? How can it be used to stay ahead, in business and other domains? What can it do?

To cut through the marketing hype and answer these questions, one must start by understanding what AI is. The field of AI is not new. It emerged in the 1950s as a multi-disciplinary endeavor to try to represent and replicate human intelligence using computers. It involved not only computer science (the 'Artificial' in AI), but also cognitive science, linguistics, psychology, and neuroscience (the 'Intelligence' in AI). In the decades that followed, AI went through cycles of booms and busts ('AI winters'), due to (too) high expectations about AI's potential that inevitably led to disappointments when the potential did not materialize.

Today, AI is used as an umbrella term to refer to all sorts of computational methods that allow machines to perform specific tasks automatically. However, 99% of research and efforts fueling the current AI boom are concentrated in a sub-field known as Machine Learning (ML). ML started to blossom in the late 1990s thanks to the increased availability of digital data, better micro-processors, and new algorithms that allowed computers to process vast amounts of data quickly to identify patterns.

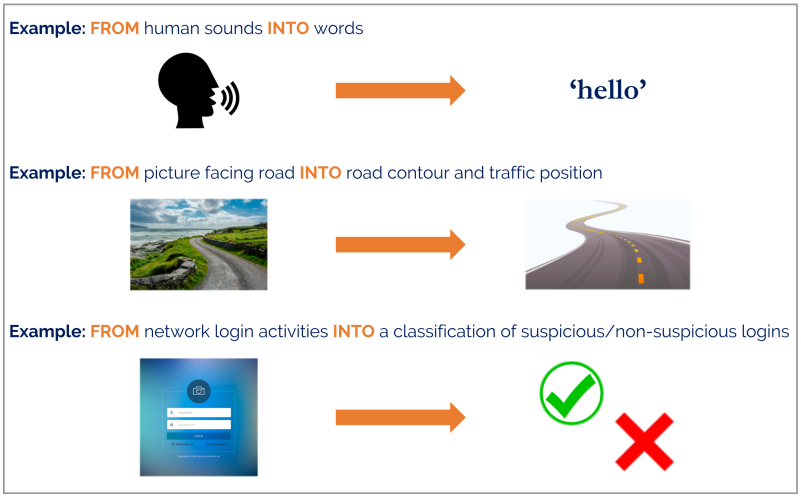

Fig. 1: *Machine Learning transforms detailed, low-level, disorganized data into more abstract concepts.*

What is Machine Learning?

The goal of Machine Learning (ML) is to transform detailed, low-level, disorganized data into more abstract concepts (Fig. 1).

Most ML algorithms follow three main steps (Fig. 2). They turn input data into numbers called features and then "learn" a function that maps inputs into outputs, using training data that has been annotated with the correct input-output mapping, often by humans. There are many different techniques for this learning, such as (Deep) Neural Networks, Decision Trees, SVMs, etc. Once a ML algorithm has learned the function with the training data, it can be applied to new input data to generate output data.

Potential and Risks

Some of the limitations and risks of ML are related to its probabilistic nature. There is no guarantee that the results of a ML algorithm will be error-free. Since ML learns from training data, the quality of ML results are directly related to the quality of the training data (which can be costly to obtain). Moreover, ML is vulnerable to manipulations of its training data. Training data could be hacked using a cyber attack to alter the results of a ML algorithm, without anyone noticing.

It is important to remember that ML functions like a 'black

Machine Learning is a very powerful technique to perform specific tasks more efficiently. It can be a game changer for specific and well-defined applications. Its successful utilization requires not only deep expertise in AI/ML algorithms, but also strategic vision, knowledge about business processes, and understanding of human skills.

box'. It is very hard or impossible to explain why a ML algorithm produces a specific result (and not another). This has serious implications for domains where unexpected ML behavior may lead to accidents and endanger human lives (e.g. autonomous vehicles). The use of ML raises ethical, social, and legal questions that must be addressed as early as possible, and are often overlooked in the current AI buzz.

For all its prowess in different areas (game playing, face recognition, natural-language processing, etc) ML remains in the realm of specific AI. It is not general AI, and researchers currently have no idea how to model general intelligence. ML performs no reasoning, deduction, or induction. A ML algorithm cannot expand its knowledge beyond the specific rules and data it has been given; it cannot generalize on its own.

Risks and limitations do not negate the potential of AI/ML, but they highlight the importance of understanding AI/ML well to use it properly. ML is a very powerful technique to perform specific tasks more efficiently. It can be a game changer for specific and well-defined applications. Its successful utilization requires not only deep expertise in AI/ML algorithms, but also strategic vision, knowledge about business processes, and understanding of human skills. Very seldom are ready-made AI products appropriate. AI/ML needs to be customized. The most successful users of AI/ML understand the technical limitations of AI. They carefully choose the specific domain(s) in which to apply AI and have a clear view on how it will improve their business performance. They develop a comprehensive vision and architecture where AI is well integrated with other IT systems, business processes, and human users.

Using AI requires thinking. Simply applying AI on a messy problem without understanding it is a guarantee for an expensive failure. Humans possess a capacity for thinking, reasoning, judgment, and common sense that AI does not have. The purpose of AI is not to replace humans, but to be a tool humans can use to become better at their jobs.

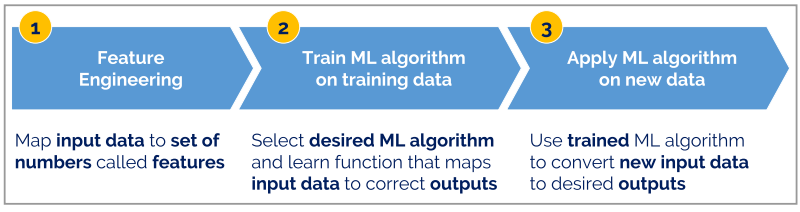

Fig. 2: *The three fundamental steps of any Machine Learning algorithm.*

Authors

Thomas Dillig *is a digital engineer and co-founder of BLUSPHINX. He is an expert in computer science, AI, cyber security, software, and project management.*

Aurelie Mei-Hoa Beaumel *is a digital architect and co-founder of BLUSPHINX. She is an expert in digital strategy, AI, data insights, and resilience.*

BLUSPHINX® is an independent firm of digital architects and engineers. We are experts in getting the most value out of digital technology, while minimizing risks, within constraints such as money, time, or resources. We work with business owners, leaders, and homeowners.

For more information, visit our website: `www.blusphinx.com`

You can contact the authors at `hello@blusphinx.com`.