Chapter 2
Machine Learning in Healthcare, Introduction and Real World Application Considerations

Stavros Pitoglou
Computer Solutions SA, Attica, Greece

ABSTRACT

Machine Learning, closely related to Artificial Intelligence and standing at the intersection of Computer Science and Mathematical Statistical Theory, comes in handy when the truth is hiding in a place that the human brain has no access to. Given any prediction or assessment problem, the more complicated this issue is, based on the difficulty of the human mind to understand the inherent causalities/patterns and apply conventional methods towards an acceptable solution, Machine Learning can find a fertile field of application. This article's purpose is to give a general non-technical definition of Machine Learning, provide a review of its latest implementations in the Healthcare domain and add to the ongoing discussion on this subject. It suggests the active involvement of entities beyond the already active academic community in the quest for solutions that "exploit" existing datasets and can be applied in the daily practice, embedded inside the software processes that are already in use.

INTRODUCTION

Machine Learning and Its Origins

One of the most quoted definitions of Machine Learning is:

The subfield of computer science that "gives computers the ability to learn without being explicitly programmed"… (Samuel, 1959).

That is a compact, but also a complete description of the major paradigm shift Machine Learning brings to the world of solving problems, ansoring questions and taking decisions with the use of Information

DOI: 10.4018/978-1-7998-2460-2.ch002

Technologies. It implies that we can delegate to a computer the task to make sense out of a dataset "on its own", without needing humans defining the exact course of calculations and actions, thus without us having understood the true nature of the problem at hand and the path to its solution. That way, the machine uses the data as "learning material" in order to assess and classify new or unseen data under the same context, or predict future values, eventually developing the ability to make decisions or/and define courses of action "on its own". That human-like ability is described in a definition which was given a few decades later:

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E. (Mitchell, 1997)

Taking a certain risk of oversimplification, the basic concept can be described as follows: You take a dataset that you believe (or at least hope) that contains the necessary information, a truth you cannot discover, but you value as essential in order to complete a certain task. You acknowledge, at the same time, the fact that, as the volume of data, the number of parameters that take part in the outcome and the complexity of their correlation increase, it becomes increasingly difficult (and at some point impossible) for the human mind to process, come up with a visible and intuitive hypothesis about the hidden patterns and model the acting causalities in order to provide means of accurate assessment and/or prediction. Then, you let the computer create its universe out of this data, a perception of the reality in the form of multidimensional "hyperspheres", creating vectors out of every data point, and, by the application of complex mathematic principles, calculate its way to an algorithm, that "understands" the acting causalities and "captures" the underlying patterns, thus becoming capable of being applied as "knowledge" and "experience" towards solving (or helping to solve) related problems.

Why Now?

Machine Learning is not such a new idea. In fact, already by the late 1950's, we can find a solid conceptual and theoretical background. In 1959 Dr Arthur Samuel publishes his research about "teaching" a machine to play the game of checkers, giving at the same time a definition that is still used today (Samuel, 1959). A year before that, in 1958, faces the light of publicity the mathematical description of the "Perceptron", the direct ancestor of contemporary Artificial Neural Networks (ANN's)(Rosenblatt, 1958).

Given this information, a completely valid question someone could ask is: "Why now?" If the background is more than 60 years old, then why is it that only in the last few years we experience this explosion (exponential growth) of interest in Machine Learning, evident by the unprecedented increase in the relative literature, the domination of the technology-related headlines, the undeclared but obvious race between all the big IT companies (Markoff & Lohr, 2016) and phenomena of "Rock Star" data scientists (Vincent, 2017).

The answer probably lies in the very basic principles of the Machine Learning concept, in terms of its "core ingredients". It is easily understood that due to their nature, Machine Learning techniques rely heavily (many times exclusively) on two information era "commodities": data and computational power. Data is the raw material, the "body of knowledge". It is often reported that having a significant volume of quality data can provide impressive results even with the use of moderate or un-optimized algorithms.