Data Science for Industry 4.0

Indraneel Dabhade
https://orcid.org/0000-0002-2988-890X
O Automation, India

INTRODUCTION

Minsky defines *artificial intelligence* as "the science of making machines do things that would require intelligence if done by [people]" (Minsky, 1969). "...any software with as much as an 'if statement' can be considered a form of narrow Artificial Intelligence" (Yampolskiy & Spellchecker, 2016). The learning systems modeled are a result of our understanding of learning theories. The most prominent of these systems are impressed with an error reduction objective. This chapter tries to highlight the effects of these intelligent systems in commercial applications. The sections included in this article are learning systems, big computing, semantics, supply chain, manufacturing, corporation, information security, quantum computing, and autonomous systems. The chapter also sheds light on the safety, ethics, and limitations of applying advanced learning techniques to self-governed systems.

LEARNING MACHINES

Thinking is a direct product of consciousness levels. Dehaene et al. (2017) distinguish conscious computation into two crucial dimensions. They are C1: global availability of relevant information and C2: a cognitive system that self-monitors. The current state of machine learning algorithms encompasses both these dimensions. The most complex learning challenges lie in modeling systems replicative of human emotions. Fear is one such emotion, difficult to capture holistically with static rules. Machines capable of learning fear display better autonomous behavior (Hutson, 2019). The safest and most reliable systems are those that showcase zero errors.

One of the earliest contributors to experimental machine learning, Arthur Samuel, while devising procedures for a machine to play checkers, introduced the concepts of 'rote learning' and 'learning by generalization' (Samuel, 1959). Donald Michie expounded on this idea with the 'Parable of Self-Improvement,' which states '...rarely occurring problems will gravitate to the bottom and frequently occurring problems to the top' (Michie, 1968). Michie asserts:

- that the apparatus of evaluation associated with any function shall comprise a "rule part" (computational procedure) and a "rote part" (lookup table);
- that evaluation in the computer shall on each given occasion proceed whether by rule, or by rote, or by a blend of the two, solely as dictated by the expediency of the moment;
- that the rule versus rote decisions shall be handled by the machine behind the scenes; and
- that various kinds of interaction be permitted to occur between the rule part and the rote part.

DOI: 10.4018/978-1-7998-9220-5.ch004

Today, machine learning algorithms have developed to handle higher statistical complexity. Non-monotonicity (i.e., rejection and reform of prior decisions) and the ability to work on high-dimensional data have added power to the learning models.

BIG COMPUTING

Jim Gray's seminal work on data processing in the '90s laid the foundation for high-speed access to data residing on nodes. In his 2007 lecture to the Computer Science and Technology Board of the National Research Council, he introduced the fourth research paradigm of 'Data-Intensive Science' to the existing three (positivism/postpositivism; interpretivism/constructivism; and critical theory), thus defining 'eScience' as the synthesis of technology and science. 'GrayWulf' set an example for other applications, including CERN's Large Hadron Collider (LHC), Virtual Observatory (VO) for astronomical data, and the National Center for Biotechnology Information (NCBI) in genomics (Szalay A. S., 2009).

Users, capitalizing on the interconnected nature of global information networks, push to scale computational processes and storage resources. Redundancy, fault tolerance, security, and speed are features that add trust and confidence to this adoption. A hurdle in running complex algorithms, including those of nonlinear optimizations, is a time-based constraint. Problems, including getting stuck at a local minimum of a solution landscape, handle suboptimal solutions. Improved cloud computing infrastructure reduces the computation expense for finding better solutions to everyday problems. To measure big data, Laney (2001) introduced the 3Vs.

Volume: Volume of the data.
Velocity: Data transfer speed.
Variety: Type of data handled.

In addition, are

Veracity (added by IBM): Quality of the data.
Variability (added by SAS): Change in data loads.
Value (added by Oracle): Value offered by the data.

While industry standards may exist for data communication, there is a need for a solid standard to exchange the science and information collected from the data. Wirth & Hipp (2000) opened the efforts to establish such a standard by introducing the <u>CR</u>oss <u>I</u>ndustry <u>S</u>tandard <u>P</u>rocessing for <u>D</u>ata <u>M</u>ining (CRISP-DM).

Industry wise, a tradeoff between real-time and batch processing has introduced several big data platforms, including Pentaho (suitable for batch processing) and SQLstream s-Server (suitable for real-time processing) (Philip Chen & Zhang, 2014). They inducted seven principles for designing big data systems.

Principle 1: Good architectures and frameworks are necessary and on the top priority.
Principle 2: Support a variety of analytical methods.
Principle 3: No size fits all.
Principle 4: Bring the analysis to data.
Principle 5: Processing must be distributable for in-memory computation.

Related Content