G Model
JMSY-641; No. of Pages 10
ARTICLE IN PRESS
Journal of Manufacturing Systems xxx (2018) xxx–xxx
Contents lists available at ScienceDirect
Journal of Manufacturing Systems
journal homepage: www.elsevier.com/locate/jmansys

Technical Paper

A survey of the advancing use and development of machine learning in smart manufacturing

Michael Sharp, Ronay Ak, Thomas Hedberg Jr. *

Engineering Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Stop 8260, Gaithersburg, MD 20899, USA

ARTICLE INFO

Article history:
Received 28 November 2017
Received in revised form 1 February 2018
Accepted 21 February 2018
Available online xxx

Keywords:
Machine learning
Industrie 4.0
Smart manufacturing

ABSTRACT

Machine learning (ML) (a subset of artificial intelligence that focuses on autonomous computer knowledge gain) is actively being used across many domains, such as entertainment, commerce, and increasingly in industrial settings. The wide applicability and low barriers for development of these algorithms are allowing for innovations, once thought unattainable, to be realized in an ever more digital world. As these innovations continue across industries, the manufacturing industry has also begun to gain benefits. With the current push for Smart Manufacturing and Industrie 4.0, ML for manufacturing is experiencing unprecedented levels of interest; but how much is industry actually using these highly-publicized techniques? This paper sorts through a decade of manufacturing publications to quantify the amount of effort being put towards advancing ML in manufacturing. This work identifies both prominent areas of ML use, and popular algorithms. This also allows us to highlight any gaps, or areas where ML could play a vital role. To maximize the search space utilization of this investigation, ML based Natural Language Processing (NLP) techniques were employed to rapidly sort through a vast corpus of engineering documents to identify key areas of research and application, as well as uncover documents most pertinent to this survey. The salient outcome of this research is the presentation of current focus areas and gaps in ML applications to the manufacturing industry, with particular emphasis on cross domain knowledge utilization. A full detailing of methods and findings is presented.

1. Introduction

Machine learning (ML) has seen increased usage in manufacturing over the past 20 years. Two surges in the use of ML occurred in manufacturing; the first in the 1980s, with the second currently underway. While ML saw significant attention in the 1980s, industrial adoption was not high because the methods were difficult to implement and ahead of the technology available at the time [1,2]. Many companies and researchers in industry are revisiting past work, focusing primarily on domain-specific models. We postulate there has been very little focus on cross-domain models for connecting information across the product life cycle. ML has remained “siloed” in each phase of the product life cycle: conception, design, manufacture, quality, and sustainment. With increased adoption of the Industrial Internet of Things (IIoT), Industrie 4.0, and Smart Manufacturing, even more data is being generated. Therefore, how does one effectively and efficiently take advantage of all that data?

Applications such as Total Design theory [3], Design for Six Sigma [4], and Design for Manufacturing [5,6] require knowledge of the various phases of the product life cycle. In a sampling of 35 defense-acquisition programs [7], development-cost growth averaged 57% and procurement-cost growth averaged 75%. Decision making dominated both types of cost growth. It follows that mitigating the negative effects of decisions earlier in the lifecycle could be advantageous to both the cost and the quality of a production program. Such mitigation requires knowledge of the full lifecycle and an understanding of how a decision in one phase of the lifecycle affects other phases of the lifecycle.

How does one gain such knowledge? Hedberg et al. [8] proposed three research directions to enable using knowledge earlier in the product life cycle: (1) dynamically generate knowledge bases, (2) determine minimum information requirements, and (3) data-interoperability support. ML are poised to greatly assist with the first two of these, dynamic knowledge generation and minimum information requirements. Synthesizing the work of Hedberg et al. [8] with other literature [9–13] identifies a need for automated methods to collect, transmit, analyze, and act on the most appropriate data. This sets the goal of using ML tools that can “observe” data, apply context, and generate knowledge – these tools must be cross-domain (cross-phase) observatories.

This paper provides a literature survey on the application of ML to mutli-disciplinary, cross-domain focus areas that make up the

* Corresponding author.
E-mail address: thomas.hedberg@nist.gov (T. Hedberg Jr.).

https://doi.org/10.1016/j.jmsy.2018.02.004
0278-6125/© 2018 Published by Elsevier Ltd on behalf of The Society of Manufacturing Engineers.

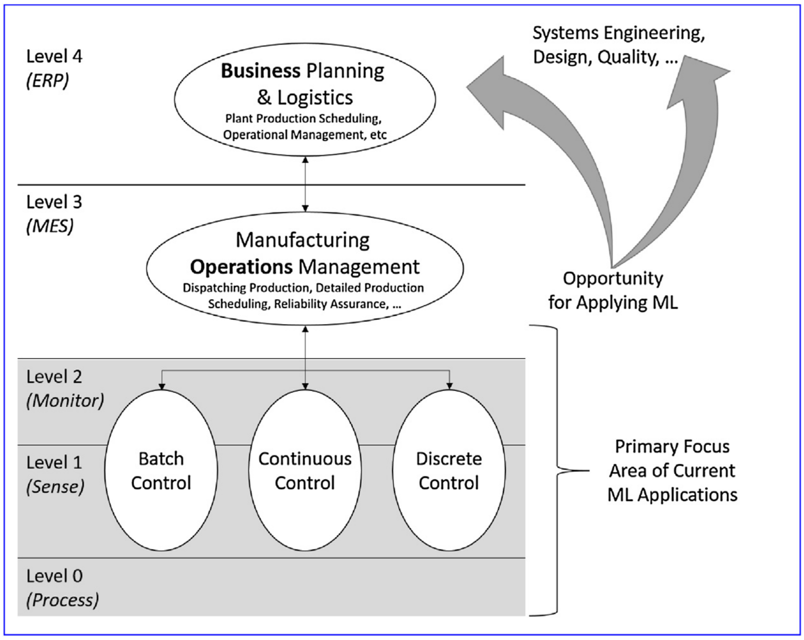

Fig. 1. Hypothesized current application areas and opportunities for applying machine learning in manufacturing and beyond (adapted from [19]).

product life cycle using manufacturing data in support of developing a life-cycle-wide "data observatory" [14]. The motivation of this work is to survey and enable the integration of previous domain-specific works, such as those described by Jennings et al. [15], Li et al. [16], Wang et al. [17], into the systems and enterprise level of the life cycle. For example, Fig. 1 presents the scope and hierarchy of the ISA-95 [18] framework and identified hypothesized current application areas and opportunities for applying ML in manufacturing and beyond. The survey was conducted with the hypothesis that most applications of machine learning are applied to low-level manufacturing problems (ISA-95 Level 0, 1, 2) and little to no application of machine learning has been applied to systems level (ISA-95 Level 3), enterprise level (ISA-95 Level 4), and other phases/domains of the product lifecycle (e.g., systems engineering, design, quality).

To accomplish extending the application of ML to cross-domain focus areas, the gaps (e.g., what questions remain) must be identified so that they may be closed through research and development. Also, to ensure successful adoption of ML solutions, the real-life applications that exist and their benefits must be determined. To accomplish this, the literature survey was conducted with two aims. (i) investigate the current state-of-the-art for ML methods; and (ii) investigate any cross-domain applications of ML in the product lifecycle.

Three survey questions were asked:

- What types of algorithms are used and with what frequency?
- Are certain applications of machine learning frequently occurring? If so, which applications and at what level of the manufacturing systems?
- Is further research needed to capture the opportunities of applying machine learning to cross-domain focus areas of the product life cycle? If so, where and what?

The scope of the survey was limited to the integration of the design, fabrication, and quality domains/functions of the product life cycle. Sustainment, or customer and product support, was considered out of scope for this survey. While sustainment is important, the initial focus was on knowledge development to support design and manufacturing decisions.

2. Motivation and background

The goal of this paper is to estimate the level of interest and actual effort being put towards the incorporation of ML technologies to the modern manufacturing industry by quantifying the presence of these concepts in the current literature. Further, this work seeks to ascertain prominent areas of the use of these technologies with both general and specific examples of applications in the literature by isolating sub populations of coordinated literature as well as targeting specific works on the subject. Last, any relevant gaps in the current level of deployment or development will be identified and presented as areas of future research.

With the availability of digital publications, it is now possible via automated techniques [20,21] to search a wider breadth and depth of literature within an area than is feasible with manual methods. Search engines and online repositories of technical documents can quickly provide a host of information based on queries of a few simple phrases. However, these searches are mostly word matching techniques and do not match the underlying concept or contextual content to a document. Most often in the English language, a collection of words or phrases presented in a particular order is required to convey a concept or idea instead of any single word. Thus, analyzing collections of words is the basis of many forms of linguistic analysis, and as related to this methodology, is what partially drove the motivation to move beyond simple key word matching as a basis for document comparison. Towards that end, key NLP techniques were identified and applied to a large corpus of technical publication abstracts in addition to simple word matching analytics. These techniques included Bag of Words/Features, and Latent Semantic Analysis (LSA) to develop a measure of 'similarity' between the documents and concepts to identify key trends [22]. This places more emphasis on not just simple word matching as with traditional searches, but core concept matching.

In this work, a large collection of digital technical abstracts is mined via ML and NLP techniques to better understand emerging trends within both industry and academia. This base corpus is created using the word matching techniques native to many online