Book Review

Reviewed by Zhongxian Wang

Email: wangz@mail.montclair.edu

Artificial Intelligence for Big Data: Complete Guide to Automating Big Data Solutions Using Artificial Intelligence Techniques
by: Anand Deshpande and Manish Kumar
Published 2018
by Packt Publishing Ltd., 35 Livery St., Birmingham, B3 2PB, UK, 384pp
ISBN: 978-1-78847-217-3

The onset of Big Data along with significant advances in computing enabled machine learning to take off. Like humans, machines build their internal models based on observations and environmental stimuli in order to form a decision and complete thought (Young, 2017). Artificial Intelligence for Big Data explores the processes and functions of analysing data in an ever-changing world that seeks to evolve. The human brain is the basis of comparison for the creation of machines. Learning machines have become an innovative project that continues to evolve with the accumulation of information and technology (Kelleher and Tiernew, 2018). Furthermore, Caldarelli and Chessa (2016) integrate ideas from data science and complex networks to create a toolkit for tackling big data challenges.

In Chapter 1, authors merge human intelligence and machine intelligence and the connection between them to explain our process of learning. The results pyramid expresses how our decisions impact our quality of life and how this theory does not cause us to achieve favourable or non-favourable results when the same actions are taken. Machines have evolved from being considered dumb to highly intelligent due to continuous innovation. Overall, this introduction provides background on the history of machines and the impact that they have on big data.

Chapter 2 dives into ontology which is a set of components and categories in a subject area or domain, showing their properties and the relationships between them. This concept is paired with the human brain to explain how information is stored and processed. The generic characteristics of intelligent agents are a key component of the process ontologies use to build them. This chapters explains how ontology is a catalyst for its adaptation as it is viewed by the world.

The concept of learning from big data is explained in Chapter 3 through algorithms. These algorithms are very widely used to understand the broad categories of machine learning. The Spark programming model aids our understanding of how activity is processed and coordinated. The driver assigns processes that the spark executor is responsible for and stores data in in-memory data structures called RDD's which then report the code execution back to the driver. The Spark MLlib is a library that stores machine learning algorithms and utilities that are designed to make machine learning

easier. The overview of this topic explains the deeper understanding of the process of machine learning.

Chapter 4 involves the concept of the evolution of neural networks in big data. This is the most important concept because it involves the realisation of intelligent machines. These machines cannot be developed without the evolution of artificial neural networks. These networks are modelled much like the human biological brain to process information with the same efficiency. The building blocks of these artificial neural networks is also a major topic of this chapter as it used to produce outcomes for new datasets through training and generalisation.

A further understanding of Artificial Neural Networks is clarified in Chapter 5. As the human brain has many neural networks, ANNs are modelled the same way. The concept of deep learning has evolved and has become the most popular algorithm used to solve complex problems. The practical approaches for building data preparation pipelines were also looked at to show the application of regularisation. An important concept that is brought up in this chapter is the processing power that is needed for the neural networks that are computationally heavy. With the ability to have more processing power, more data can be added.

Chapter 6 deals more with machine learning and developing the understanding of Natural Learning Processing (NLP). NLP is an area of algorithms that is focused on processing data that is unstructured. Ontology was a topic that reappeared in this chapter to take a look at NLP and how intelligent machines can function at the human level to understand large amounts of information and knowledge. Through this understanding, the proper actions and decisions will take place and be made in accordance with the comprehension that is reached.

In Chapter 7, the fundamental theory of fuzzy logic enlightens the importance of building intelligent machines that are continuously evolving and that are capable of processing data from various sources. The concept of real world interaction through the form of interface was also approached to display how machines need to have that same ability as human beings do. The goal is to create machines that have similar abilities of a human being which would classify them as intelligent.

Chapter 8 describes genetic algorithms that allow artificial intelligence systems to be inspired by the natural process of evolution. Fundamental algorithms are introduced to show how to efficiently use genes to mine data. Through this process, genetic algorithms aim to implement the same concept of inheritance, variation in partner selection, and attributes of the offspring on a level that is conducive to machine learning. Through the exploration of genetic algorithms, a quick attribute search with Weka was reached to provide an easy-to-use and rich user interface.

Chapter 9 develops an understanding of how the intelligent behaviour of living creature and their concepts can create intelligent machines of the future. With the combination of swarm intelligence and data mining techniques, a better understanding of big data analytics problems is reached. The effectiveness of these algorithms is based on their ability to solve big data analytics problems in the real-world.

In Chapter 10, machine learning is described as a topic that cannot be explained without reinforcement learning. The key component of reinforcement learning is the intelligent agent which is designed to learn the right behaviour based on the reward that it receives in its given environment. This process is critical to designing machines that connect capabilities of a human being and processes of intelligent machines. In this

chapter the fundamentals of reinforcement are explained as well as mathematical theory to reach a deeper understanding of their impact on the evolution of intelligent machines.

Chapter 11 focuses on cyber security and the role it plays in accordance with big data. The security of critical applications is threatened and is required to be secure to process data properly. Batch processing and real-time processing are the two fundamental types of big data in regard to streaming data sources. The most important type of processing when dealing with cyber security threats is real-time stream-based processing. This is taken into consideration when building systems because machine learning and artificial intelligence are threatened by cyber security attacks. This chapter highlights the importance of more research needing to be done in the field of cyber security to protect data assets.

Cognitive computing is the main topic of Chapter 12 and it illuminates how cognitive intelligence aims to be closely related to human intelligence. The goal of this chapter is to achieve strong artificial intelligence by using the five senses. The mind is mentioned as the sixth sense in relation to cognitive abilities that are to be implemented into the process of developing artificial intelligence. The history of cognitive systems is viewed to provide a background of the systems and picture for the future of the evolution that can be achieved.

This book is aimed to provide useful information for data scientists, members of technical staff in IT products and service companies, technical project managers, architects, business analysts, and anyone who deals with data assets. I think that this book effectively explains various concepts regarding artificial intelligence and its analysis of big data throughout the twelve chapters. Overall it was very informative on the topic of machine learning and its processes and aided in the understanding of the role the human brain plays in the explanation.

You will not regret reading this book.

References

Caldarelli, G. and Chessa, A. (2016) *Data Science & Complex Networks*, Oxford University Press. Great Clarendon St., Oxford, OX26DP, UK.

Kelleher, J.D. and Tiernew, B. (2018) *Data Science*, MIT Press, US One Rogers Street Cambridge, MA 02142-1209, USA.

Young, M. (2017) 'Machines learning astronomy: the new era of artificial intelligence & big data is changing how we do astronomy', *Sky & Telescope*, Vol. 134, No. 6, pp.20–27.