Chapter 56
Investigation on Deep Learning Approach for Big Data: Applications and Challenges

Dharmendra Singh Rajput
VIT University, India

T. Sunil Kumar Reddy
Sri Venkateswara College of Engineering and Technology, India

Dasari Naga Raju
Sri Venkateswara College of Engineering and Technology, India

ABSTRACT

In recent years, big data analytics is the major research area where the researchers are focused. Complex structures are trained at each level to simplify the data abstractions. Deep learning algorithms are one of the promising researches for automation of complex data extraction from large data sets. Deep learning mechanisms produce better results in machine learning, such as computer vision, improved classification modelling, probabilistic models of data samples, and invariant data sets. The challenges handled by the big data are fast information retrieval, semantic indexing, extracting complex patterns, and data tagging. Some investigations are concentrated on integration of deep learning approaches with big data analytics which pose some severe challenges like scalability, high dimensionality, data streaming, and distributed computing. Finally, the chapter concludes by posing some questions to develop the future work in semantic indexing, active learning, semi-supervised learning, domain adaptation modelling, data sampling, and data abstractions.

DOI: 10.4018/978-1-7998-0414-7.ch056

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

INTRODUCTION

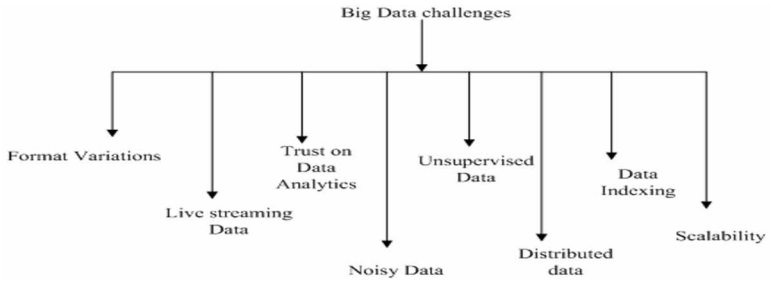
In the recent years, machine learning concepts made major impact on different fields. The machine learning is the concept of defining the input data and generalizes the patterns for the data which are used for the future purpose. The good data representation leads to the improvement in the performance of the machine learning concepts and poor representation of data causes the reduction of performance of any advanced machine learners. Therefore, the present research is concentrated on developing the data representations and exploiting concrete features from the raw data (Domingos, P.2012).

Deep learning approach is one of the feature engineering methods applied for the complex data sets to retrieve the abstract features. This type of algorithms follows the hierarchical and layered structures for representing the data, where the data is represented in low level and high level abstractions. The hierarchical structure in the deep learning approach is inspired by the data perception process of the human brain (Dalal, N, &Triggs, B.2005 and Lowe DG 1999). Deep learning algorithms are more advantages in dealing with huge volumes of unsupervised data and it follows the greedy procedure for data representations. Research studies proved that data representation using feature extractions will help in improving the machine learning outputs. For instance, invariant data representations (Goodfellow et al., 2009), probabilistic models (Salakhutdinov, R & Hinton GE, 2009) and improved classification models (Larochelle, H, et al., 2009). Deep learning made major positive impact on different machine learning approaches such as computer vision (Krizhevsky A, et al., 2012; Hinton GE,et al., 2006 and Bengio, Y., et al., 2007), speech recognition (Dahl et al., 2012; Mohamed et al., 2012; Seide et al ., 2011; Hinton et al., 2012 & Dahl et al., 2010) and NLP (Socher et al., 2011; Mikolov et al., 2011 and Bordes et al., 2012).

Big data is the recent buzz word in the data science field. It creates solution to the problems generated by large volumes of unsupervised application data with respect to the specific domain. Recent advancements in the field of data storage and computational resources have contributed lot more to the development of big data analytics (Tiwari & Thakur, in press). Major competitors like, Google, amazon, yahoo and Microsoft are managing larger proportions of data (i.e., exabytes).The users of some social media companies like Facebook, Instagram, Twitter and YouTube are posting huge volumes of data in their daily activities. Different leading companies developed their analytics platform to monitor, analyse and simulate the data for future business needs.

Data mining and data extraction are the basic operations performed on the big data for data prediction and decision making (Tiwari et al., 2010). Moreover, the data mining in big data pose many challenges which are represented in Figure 1.

Figure 1. Challenges of Big Data in data mining



12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/investigation-on-deep-learning-approach-for-big-data/237918

Related Content

Simultaneous Evolution of Network Architectures and Connection Weights in Artificial Neural Networks
Ruhul A. Sarkerand Hussein A. Abbass (2006). *Artificial Neural Networks in Finance and Manufacturing (pp. 28-42).*
www.irma-international.org/chapter/simultaneous-evolution-network-architectures-connection/5347

Higher Order Neural Networks with Bayesian Confidence Measure for the Prediction of the EUR/USD Exchange Rate
Adam Knowles, Abir Hussain, Wael El Deredy, Paulo G.J. Lisboaand Christian L. Dunis (2009). *Artificial Higher Order Neural Networks for Economics and Business (pp. 48-59).*
www.irma-international.org/chapter/higher-order-neural-networks-bayesian/5276

Neural Network Based Automated System for Diagnosis of Cervical Cancer
Seema Singh, V. Tejaswini, Rishya P. Murthyand Amit Mutgi (2020). *Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications (pp. 1422-1436).*
www.irma-international.org/chapter/neural-network-based-automated-system-for-diagnosis-of-cervical-cancer/237943

Efficient Imbalanced Multimedia Concept Retrieval by Deep Learning on Spark Clusters
Yilin Yan, Min Chen, Saad Sadiqand Mei-Ling Shyu (2020). *Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications (pp. 274-294).*
www.irma-international.org/chapter/efficient-imbalanced-multimedia-concept-retrieval-by-deep-learning-on-spark-clusters/237877

Deep Learning Techniques for Prediction, Detection, and Segmentation of Brain Tumors
Prisilla Jayanthiand Muralikrishna Iyyanki (2020). *Deep Neural Networks for Multimodal Imaging and Biomedical Applications (pp. 118-154).*
www.irma-international.org/chapter/deep-learning-techniques-for-prediction-detection-and-segmentation-of-brain-tumors/259491