Mesopotamian journal of Big Data
Vol. **(2022)** 2022, **pp**. 31–35
DOI: https://doi.org/10.58496/MJBD/2022/004 ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BigData

Research Article

An Overview on Deep Leaning Application of Big Data

Rofia Abada [*,1], Abdulhalim Musa Abubakar [2], Muhammad Tayyab Bilal [3]

[1] *Faculty of Architecture and Urban Planning, Salah Boubnider Constantine University, Algeria.*

[2] *Department of Chemical Engineering, Faculty of Engineering, Modibbo Adama University, Nigeria*

[3] *Computer Science Department, University of Agriculture Faisalabad Punjab, Pakistan*

ARTICLE INFO

Article History

Received 20 Apr 2022
Accepted 08 Jul 2022
Published 30 Jul 2022

Keywords

Big data

Parallel Processing
Deep Learning

Machine Learning

ABSTRACT

Big data refers to the large volumes of structured and unstructured data that are generated by businesses, organizations, and individuals on a daily basis. Deep learning is a type of machine learning that involves the use of artificial neural networks to learn patterns and relationships in data. In this paper, we discuss the applications of deep learning in the field of big data analysis. We provide an overview of deep learning and big data, and then delve into specific examples of how deep learning has been used in various domains to extract value from big data. These domains include predictive analytics, image and video analysis, natural language processing, and recommendation systems. We also discuss some of the challenges and limitations of using deep learning for big data analysis, as well as future directions for research and development in this field. Overall, deep learning has proven to be a powerful tool for extracting insights from big data, and is likely to play an increasingly important role in the field of data science.

1. INTRODUCTION

Big data[1] refers to the large volumes of structured and unstructured data that are generated by businesses, organizations, and individuals on a daily basis. With the proliferation of the internet and the widespread use of devices such as smartphones and smart devices, the amount of data being generated has increased exponentially. This data can be used to gain insights and make informed decisions, but traditional data analysis methods are often not sufficient for handling the scale and complexity of big data. Deep learning[2], on the other hand, is a type of machine learning that involves the use of artificial neural networks to learn patterns and relationships in data. Deep learning has proven to be particularly effective in tasks such as image and speech recognition, natural language processing, and predictive analytics. In this paper, we will discuss the applications of deep learning in the field of big data analysis. We will first provide an overview of deep learning and big data, and then delve into specific examples of how deep learning has been used in various domains to extract value from big data.

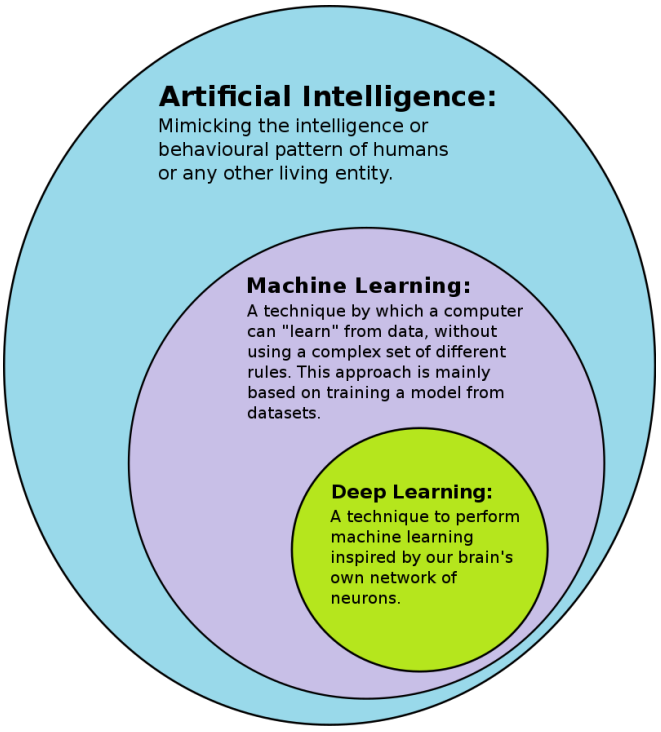

Fig. 1. Origin of deep learning.

2. DEEP LEARNING

Deep learning[3] is a subset of machine learning that involves the use of artificial neural networks, which are inspired by the structure and function of the human brain. These networks consist of layers of interconnected nodes, with each node representing a unit of computation. The nodes in the input layer receive data, and the nodes in the output layer produce the final result. Between the input and output layers are hidden layers, which contain the majority of the neural network's computational power.

Deep learning[4] algorithms are trained using a large dataset and an optimization algorithm, such as stochastic gradient descent. The algorithm adjusts the weights and biases of the network based on the input data and the corresponding output labels, in order to minimize the error between the predicted output and the true output. This process is repeated until the network reaches a satisfactory level of accuracy.

One of the key advantages of deep learning is its ability to learn features and patterns from raw data, without the need for manual feature engineering. This allows deep learning algorithms to handle complex, high-dimensional data, and to perform well on tasks such as image and speech recognition, natural language processing, and predictive analytics. There are several different types of deep learning algorithms, including:

Convolutional neural networks (CNNs)[5]: These are commonly used for tasks such as image and video analysis, and involve the use of convolutional layers to extract features from the input data.
Recurrent neural networks (RNNs): These are well-suited for tasks such as natural language processing, as they are able to process sequential data and retain information about past inputs.

Generative adversarial networks (GANs): These are used for tasks such as image generation, and involve the use of two neural networks – a generator and a discriminator – that compete with each other in order to improve the quality of the generated images.

Autoencoders: These are used for tasks such as dimensionality reduction and feature learning, and involve the use of an encoder network to compress the input data and a decoder network to reconstruct the original data from the compressed representation.
Deep belief networks (DBNs): These are a type of generative model that use a stack of restricted Boltzmann machines (RBMs) to learn a hierarchical representation of the input data.
Long short-term memory (LSTM) networks: These are a type of RNN that are able to learn long-term dependencies in sequential data, making them well-suited for tasks such as language translation and language modeling.

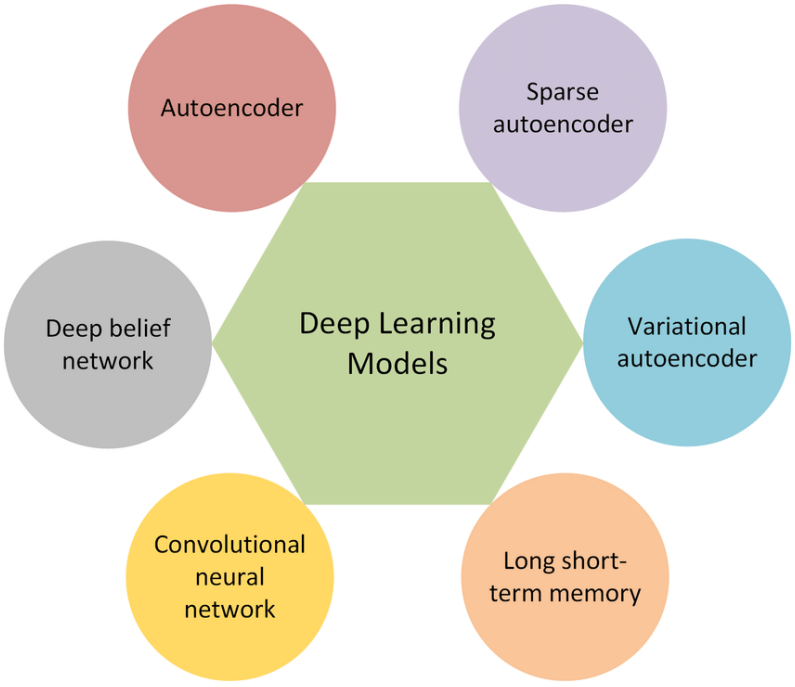

Fig. 2. Types of deep learning.

3. BIG DATA

Big data[6] refers to the large volumes of structured and unstructured data that are generated by businesses, organizations, and individuals on a daily basis. This data can come from a variety of sources, including social media, sensors, log files, and transactional databases. The volume, velocity, and variety of big data can make it challenging to store, process, and analyze using traditional data management and analysis techniques. Deep learning algorithms are trained using a large dataset and an optimization algorithm, such as stochastic gradient descent. One of the key advantages of deep learning is its ability to learn features and patterns from raw data, without the need for manual feature engineering. This allows deep learning algorithms to handle complex, high-dimensional data, and to perform well on tasks such as image and speech recognition, natural language processing, and predictive analytics.

The combination of big data and deep learning has the potential to unlock valuable insights and inform decision-making in a variety of domains. By using deep learning algorithms to analyze large datasets, businesses and organizations can identify trends, patterns, and relationships that may not be apparent from smaller datasets. However, the complexity and scale of big data often requires specialized tools and techniques for analysis, and there are challenges and limitations to the use of deep learning for big data analysis.

4. APPLICATIONS OF DEEP LEARNING IN BIG DATA ANALYSIS

There are a number of ways in which deep learning has been applied to big data analysis in various domains. Some examples include:

- Predictive analytics: Deep learning algorithms can be used to analyze large datasets in order to make predictions about future events. For example, a deep learning model might be trained on historical financial data in order to predict stock prices. These models can also be used to forecast demand for products or services, or to predict customer churn.
- Image and video analysis: Deep learning algorithms have proven to be effective at tasks such as image and video classification, object detection, and facial recognition. These algorithms can be used to analyze large datasets of images and videos in order to extract insights and identify patterns. For example, a deep learning model might be used to classify images of different types of animals, or to detect pedestrians in a video feed from a surveillance camera.
- Natural language processing: Deep learning algorithms can be used to analyze large datasets in order to make predictions about future events. For example, a deep learning model might be trained on historical financial data in order to predict stock prices. These models can also be used to forecast demand for products or services, or to predict customer churn.
- Recommendation systems: Deep learning algorithms can be used to build recommendation systems, which suggest products or services to users based on their past behavior or preferences. These systems can analyze large datasets of user interactions, such as clicks, views, and purchases, in order to make personalized recommendations.
- Fraud detection: Deep learning algorithms can be used to detect fraudulent activity in large datasets, such as financial transactions or insurance claims. By analyzing patterns in the data, these algorithms can identify anomalies and alert authorities to potential fraudulent activity.
- Healthcare: Deep learning algorithms can be used to analyze large datasets of medical records and imaging data in order to extract insights and inform decision-making in healthcare. For example, a deep learning model might be used to predict the likelihood of a patient developing a particular condition, or to identify patterns in medical images that are indicative of a particular disease.

There are also many other applications of deep learning in big data analysis, including cybersecurity, agriculture, and transportation. However, it is important to note that there are challenges and limitations to the use of deep learning for big data analysis. These include the need for large amounts of labeled data for training, the potential for bias in the data, and the complexity of designing and tuning deep learning models.

5. CONCLUSION AND FURTHER WORK

Deep learning has proven to be a powerful tool for extracting value from big data in a variety of domains. By using artificial neural networks to learn patterns and relationships in data, deep learning algorithms can handle complex, high-dimensional data and perform well on tasks such as image and speech recognition, natural language processing, and predictive analytics. While there are still challenges and limitations to the use of deep learning in big data analysis, the potential benefits make it an important area of research and development. As the volume of big data continues to grow, it is likely that deep learning will play an increasingly important role in extracting insights and informing decision-making.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors would like to express their gratitude to the Universiti Tun Hussein Onn Malaysia, the Informatics Institute for Postgraduate Studies, and the Al Salam University College for their moral support. Please accept my sincere gratitude for the useful recommendations and constructive remarks provided by the anonymous reviewers.

Funding

The authors receive no funding for this work.

References

[1] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and M. S. H. Sunny, "Application of big data and machine learning in smart grid, and associated security concerns: A review," *Ieee Access,* vol. 7, pp. 13960-13988, 2019.

[2] Y. Li, C. Huang, L. Ding, Z. Li, Y. Pan, and X. Gao, "Deep learning in bioinformatics: Introduction, application, and perspective in the big data era," *Methods,* vol. 166, pp. 4-21, 2019.

[3] F. Emmert-Streib, Z. Yang, H. Feng, S. Tripathi, and M. Dehmer, "An introductory review of deep learning for prediction models with big data," *Frontiers in Artificial Intelligence,* vol. 3, p. 4, 2020.

[4] M. A. Amanullah *et al.*, "Deep learning and big data technologies for IoT security," *Computer Communications,* vol. 151, pp. 495-517, 2020.

[5] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing," *ISPRS journal of photogrammetry and remote sensing,* vol. 173, pp. 24-49, 2021.

[6] N. L. Bragazzi, H. Dai, G. Damiani, M. Behzadifar, M. Martini, and J. Wu, "How big data and artificial intelligence can help better manage the COVID-19 pandemic," *International journal of environmental research and public health,* vol. 17, no. 9, p. 3176, 2020.