

Deep Learning Security Systems

C. Rajeev, K. Kumar Swamy

Abstract— Big Data Analytics and Deep Learning are not supposed to be two entirely different concepts. Big Data means extremely huge large data sets that can be analyzed to find patterns, trends. One technique that can be used for data analysis so that able to help us find abstract patterns in Big Data is Deep Learning. If we apply Deep Learning to Big Data, we can find unknown and useful patterns that were impossible so far. With the help of Deep Learning, AI is getting smart. There is a hypothesis in this regard, the more data, the more abstract knowledge. So a handy survey of Big Data, Deep Learning and its application in Big Data is necessary.

Index terms : Deep Learning, Bigdata, Artificial Intelligence.

I INTRODUCTION

If we want to have a glance at the data generation history from 1960, we can see this trend in overall: 1960-1990, relational databases; 1990-2000, OLAP technology; 2000-2010, column based data storages and cloud computing; and 2010- 2016, Big Data applications. These days, Knowledge plays a key role to get success. Many companies need more abstract knowledge. This need can be satisfied by a combination of two major domains: Big Data and Deep Learning. Each device can generate data. This situation will become worse if each device can be connected to other devices to use their information.

In other words, with the emergence of Internet of Things, we are facing with huge amount of data that needed to be stored and managed, one Example of Big Data. in brief, with the advances in digital devices such as digital sensors, large amounts of data have been generated at a fast speed that resulted in an area named Big Data. Big Data is not only about producing data from sensors; It can be provided by humans, texts, images and so on. Big Data has a great impact on technologies and computing. In other words, we have more data these days that current methods cannot deal with these data. In simple word, the term of Big Data means collecting, processing and presenting the results of huge amounts of data that come at high speed in a variety of formats. Traditional Machine Learning tools have shortcoming when they face with Big Data and want to solve Big Data area problems.

Big Data and deep learning are two important words in data science now days. Big Data Analytics and Deep Learning are two high-focus of data science. The large volumes of data collected by organizations are utilized for various purposes such as for solving problems in marketing, technology, medical science, national intelligence, fraud detection etc. Traditional data processing systems are not

adequate to handle, analyze and process as the collected data are unlabelled, uncategorized and very complex.

Deep learning is appropriate for exploiting large volumes of data and for analyzing raw data from multiple sources and in different styles. A key benefit of Deep Learning is the analysis and learning of massive amounts of unsupervised data, making it a valuable tool for Big Data Analytics where raw data is largely.

II. REVIEW OF LITERATURE

2.1 BIGDATA:

While the term “big data” is relatively new, the act of gathering and storing large amounts of information for eventual analysis is ages old. The concept gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the three Vs:

A) Volume. Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data. In the past, storing it would've been a problem – but new technologies (such as Hadoop) have eased the burden.

B) Velocity. Data streams in at an unprecedented speed and must be dealt with in a timely manner. RFID tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time.

C) Variety. Data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

At SAS, we consider two additional dimensions when it comes to big data:

D) Variability. In addition to the increasing velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Is something trending in social media? Daily, seasonal and event-triggered peak data loads can be challenging to manage. Even more so with unstructured data.

E) Complexity. Today's data comes from multiple sources, which makes it difficult to link, match, cleanse and transform data across systems. However, it's necessary to connect and correlate relationships, hierarchies and multiple data linkages or your data can quickly spiral out of control.

Why Is Big Data Important?

The importance of big data doesn't revolve around how much data you have, but what you do with it. You can take data from any source and analyze it to find answers that enable 1) cost reductions, 2) time reductions, 3) new product development and optimized offerings, and 4) smart decision making.

Deep learning is an aspect of artificial intelligence (AI) that is concerned with emulating the learning approach that human beings use to gain certain types of knowledge. At its simplest, deep learning can be

Revised Manuscript Received on 14 August, 2019.

C. Rajeev, Assistant Professor, Department of Computer Science and Engineering,

K. Kumar Swamy, Malla Reddy Engineering College for Women, Maisammaguda, Hyderabad

thought of as a way to automate predictive analytics. Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.

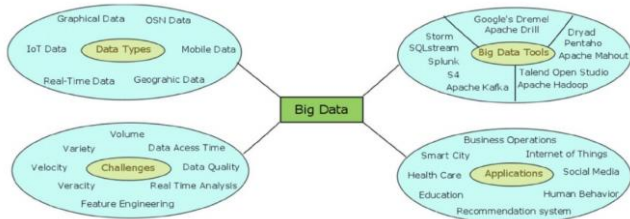


Fig 2.1: Big data literature Classification

2.2 DEEP LEARNING:

2.2.1 Deep Learning is Large Neural Networks

Andrew Ng from Coursera and Chief Scientist at Baidu Research formally founded Google Brain that eventually resulted in the productization of deep learning technologies across a large number of Google services. In the 2013 talk titled “Deep Learning, Self-Taught Learning and Unsupervised Feature Learning” he described the idea of deep learning as: Using brain simulations, hope to:

- Make learning algorithms much better and easier to use.
- Make revolutionary advances in machine learning and AI.

He also commented on the important point that it is all about scale: “for most flavors of the old generations of learning algorithms ... performance will plateau. ... deep learning ... is the first class of algorithms ... that is scalable. ... Performance just keeps getting better as you feed them more data”.

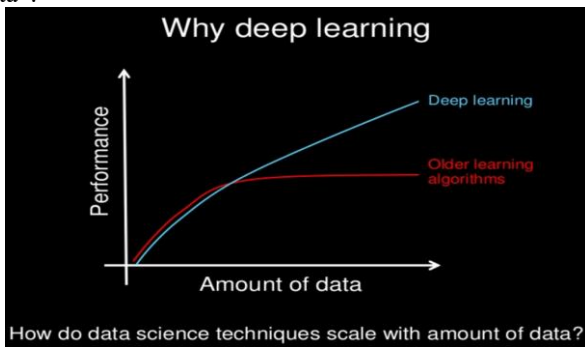


Fig 2.2 : how do data science techniques scale with amount of data.

2.2.2 Deep Learning is Hierarchical Feature Learning:

In addition to scalability, another often cited benefit of deep learning models is their ability to perform automatic feature extraction from raw data, also called feature learning. Yoshua Bengio is another leader in deep learning although began with a strong interest in the automatic feature learning that large neural networks are capable of achieving. He describes deep learning in terms of the algorithms ability to discover and learn good representations using feature learning.

“Deep learning algorithms seek to exploit the unknown structure in the input distribution in order to discover good representations, often at multiple levels, with higher-level learned features defined in terms of lower-level features”.

III. THE DEEP LEARNING PROCESS

In deep Learning applications, a data scientist or other analyst:

1. understand the problems and whether deep learning is good fit.
2. Identifies relevant data sets and prepares them for analysis.
3. Choose the type of deep learning algorithm to use.
4. Trains algorithm on large amount of labeled data.
5. Tests the model’s performance against unlabeled data.

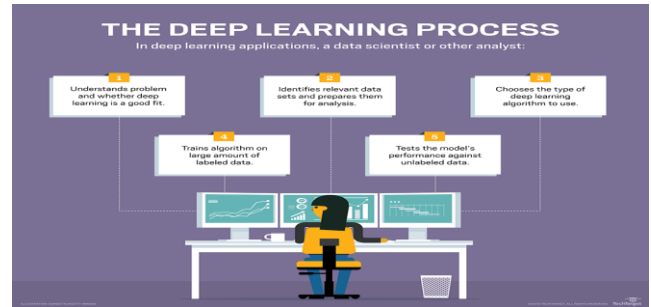


Fig 2.3: Deep Learning process.

IV. APPLICATIONS OF DEEP LEARNING IN BIG DATA

Deep Learning algorithms have great potential for research into the automated extraction of complex data representations. Deep Learning algorithms can develop a layered, and hierarchical architecture of learning and representing data. Deep Learning in Big Data Analytics has become a high-focus of data science.

Big Data has now become important as several organizations are collecting massive amounts of domain-specific information that can be used to solve problems related to national intelligence, cyber security, fraud detection, marketing, and medical informatics. Deep Learning Big Data allows extraction of high-level, complex abstractions as data representations through a hierarchical learning process. A key benefit of Deep Learning is Big Data analysis that it can learn from massive amounts of unsupervised data. This makes it a valuable tool for Big Data Analytics where huge amounts of raw data are uncategorized.

4.1. Semantic Indexing:

Information retrieval is a key task that is associated with Big Data Analytics. Efficient storage and retrieval of information is a growing problem in Big Data analysis, as data in large-scale quantities such as text, image, video, and audio is being collected and made available across various domains. Thus, strategies and solutions that were previously used for information storage and retrieval are challenged by massive volumes of data. Semantic indexing proves to be an efficient technique as it facilitates knowledge discovery and comprehension, thereby making search engines work more quickly and efficiently.

4.2 Conducting Discriminative Tasks:

While performing discriminative tasks in Big Data Analytics, Learning algorithms allow users to extract complicated nonlinear features



from the raw data. It also facilitates the use of linear models to perform discriminative tasks using extracted features as input. This approach has two advantages: Firstly, by extracting features with Deep Learning adds nonlinearity to the data analysis, thereby associating discriminative tasks closely to AI, and secondly applying linear analytical models on extracted features is more efficient computationally. These two benefits are important for Big Data because it allows practitioners in accomplishing complicated tasks related to Artificial Intelligence like object recognition in images, image comprehension, etc

4.3. Semantic Image and Video Tagging:

Deep Learning techniques help in semantic tagging. Deep Learning mechanisms can facilitate segmentation and annotation of complex image scenes. Deep Learning can also be used for action scene recognition as well as video data tagging. It uses an independent variant analysis to learn invariant spatiotemporal features from video data. This approach helps in extracting useful features for performing discriminative tasks on image and video data. Deep Learning has been successful in achieving remarkable results in extracting useful features. However, there is still considerable work that remains to be done for further exploration that includes determination of appropriate objectives in learning good data representations and performing other complex tasks in Big Data Analytics.

4.4. Healthcare:

This exciting domain covers a whole amount of applications like diagnostics including non-invasive diagnostics, monitoring certain parameters through apps using a variety of sensors, predicting health and probable ailments and prescribing personalized medicine on the basis of Biobank data.

4.5. Predicting Demographics and Election Results:

50 million Google Street View images were used to explore what a Deep Learning network can do with them. Among the impressive results, the computer learned to localize and recognize cars. It detected over 22 million cars including their make, model, body type, and year. Going further from there, the model was actually able to predict the demographics of each area by the car makeup of that area. Among the many insights that were generated based on the demographics deduced from the car makeup, here's a cool example: "If the number of sedans encountered during a 15-minute drive through a city is higher than the number of pickup trucks, the city is likely to vote for a Democrat during the next Presidential election (88% probability); otherwise, it is likely to vote Republican (82% probability).

4.6. Finance:

Today, stocks are frequently traded via electronic exchanges. Traders are continuously submitting, cancelling, and executing buy and sell orders in the exchange's limit order book. Since market events are often reported at the nanosecond granularity, the limit order book data generated over time runs into terabytes and petabytes. Deep learning algorithms can model key quantities, such as the probability

distribution of future price movements, given the current state of supply and demand in the market.

4.7 Manufacturing:

Smart manufacturing refers to using advanced data analytics to complement physical science for improving system performance and decision making. With the widespread deployment of sensors and IoT (Internet of Things), there is a growing need for handling big manufacturing data characterized by high volume, high velocity, and high variety. Deep learning provides advanced analytics tools for processing and analyzing big manufacturing data.

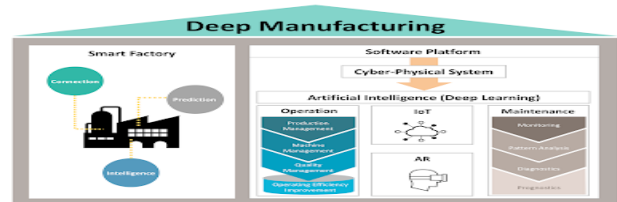


Fig 4.7: Deep Manufacturing

V. CONCLUSION

Big Data presents significant challenges to deep learning. Deep learning technique can be used for data analysis so that able to help us find abstract patterns in Big Data. If we apply Deep Learning to Big Data, we will find unknown and useful patterns that were impossible. Deep Learning has an advantage of potentially providing a solution to address the data analysis and learning problems found in massive volumes of input data. So, Deep Learning over Big Data is an emerging paradigm to mine value from the massive amount of gathered data. So, Now-a-days deep learning in Big data plays a major role .

FUTURE ENHANCEMENT

In future work, we will see how Machine Learning as a service has help to manage big data better in data analytics and its applications. So, Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

REFERENCES

1. Yann LeCun, Yoshua Bengio and Geoffrey Hinton, "Deep learning", Nature, Vol 521, May 2015.
2. Xue-Wen Chen and Xiaotong Lin, "Big Data Deep Learning: Challenges and Perspectives", IEEE, Volume 2, May 2014.
3. Shui Yu, "Big Privacy: Challenges and Opportunities of Privacy Study in the Age of Big Data", IEEE, Volume 4, June 2016.
4. Daniele Ravi, Charence Wong, Fani Deligianni, Melissa Berthelot, Javier Andreu-Perez, Benny Lo, and Guang-Zhong Yang. "Deep Learning for Health Informatics", IEEE Journal of Biomedical and Health Informatics, Vol. 21, No. 1, January 2017.
5. Sri Satish Ambati, "Deep learning: A brief guide for practical problem solvers", New Tech Forum, 2015.
6. Jaseena K U and Julie M David, " Issues, Challenges, and Solutions : Big Data Mining", NeTCoM, CSIT, 2014, pp. 131– 140
7. B M Wilamowski, Bo Wu, and Janusz Korniak, "Big data and Deep Learning", , 20th Jubilee IEEE International Conference on Intelligent Engineering Systems, June 30-July 2, 2016.

8. X. Chen, and X. Lin, "Big Data Deep Learning: Challenges and Perspectives". In Access, IEEE, vol.2, pp.514,525,doi:10.1109/ACCESS.2014.2325029, 2014.
9. L. Deng, "A tutorial survey of architectures, algorithms, and applications for Deep Learning", In APSIPA Transactions on Signal and Information Processing, 2013.
10. H. Larochelle, Y. Bengio, J. Louradour, and P. Lamblin, "Exploring strategies for training deep neural networks," J. Mach. Learn. Res., vol. 10, pp. 1–40, Jan. 2009.
11. K. Kavukcuoglu, M. A. Ranzato, R. Fergus, and Y. LeCun, "Learning invariant features through topographic filter maps," in Proc. Int. Conf. CVPR, 2009, pp. 1605–1612.
12. V. Vanhoucke, A. Senior, and M. Mao, "Improving the speed of neural networks on CPUs," in Proc. Deep Learn. Unsupervised Feature Learn. Workshop, 2011.
13. Fact Sheet: Brain Initiative, Office of the Press Secretary, The White House, Washington, DC, USA, 2013.
14. D. Laney, The Importance of 'Big Data': A Definition. Stamford, CT, USA: Gartner, 2012.
15. Torralba, R. Fergus, and W. Freeman, "80 million tiny images: A large data set for nonparametric object and scene recognition," IEEE Trans. Softw. Eng., vol. 30, no. 11, pp. 1958–1970, Nov. 2008.
16. J. Wang and X. Shen, "Large margin semi-supervised learning," J. Mach. Learn. Res., vol. 8, no. 8, pp. 1867–1891, 2007.
17. J. Weston, F. Ratle, and R. Collobert, "Deep learning via semi-supervised embedding," in Proc. 25th Int. Conf. Mach. Learn., Helsinki, Finland, 2008.
18. K. Sinha and M. Belkin, "Semi-supervised learning using sparse eigenfunction bases," in Proc. Adv. NIPS, 2009, pp. 1687–1695.
19. R. Fergus, Y. Weiss, and A. Torralba, "Semi-supervised learning in gigantic image collections," in Proc. Adv. NIPS, 2009, pp. 522–530.
20. J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Ng, "Multimodal deep learning," in Proc. 28th Int. Conf. Mach. Learn., Bellevue, WA, USA, 2011
21. N. Nithiyandam, K. Venkatesh, M. Rajesh, Transfer The Levels Of The Monitored Carbon, Nitrogen Gases From The Industries, International Journal of Recent Technology and Engineering, Volume-7 Issue-6S3 April, 2019.
22. Sivanesh Kumar, A., Brittoraj, S., Rajesh, M., Implementation of RFID with internet of things, Journal of Recent Technology and Engineering, Volume-7 Issue-6S3 April, 2019.
23. Rajesh, M., Sairam, R., Big data and health care system using mlearningJournal of Recent Technology and Engineering, Volume-7 Issue-6S3 April, 2019.
24. Rajesh, M., and J. M. Gnanasekar. "Path Observation Based Physical Routing Protocol for Wireless Ad Hoc Networks." Wireless Personal Communications 97.1 (2017): 1267-1289.