Int. J. Advance Soft Compu. Appl, Vol. 9, No. 1, March 2017
ISSN 2074-8523

A Systematic Literature Review on Features of Deep Learning in Big Data Analytics

Hordri N. F.[1,2], Samar, A.[1], Yuhaniz S. S.[1,2], and Shamsuddin S. M.[2]

[1]Advanced Informatics School,
Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[2]UTM Big Data Centre, Universiti Teknologi Malaysia, Johor, Malaysia
Email: nfarhana64@live.utm.my, salireza5@live.utm.my, sophia@utm.my, mariyam@utm.my

Abstract

Context: Deep Learning (DL) is a division of machine learning techniques that based on algorithms for learning multiples level of representations. Big Data Analytics (BDA) is the process of examining large scale of data and variety of data types. Objectives: The aims of this study are to identify the existing features of DL approaches for using in BDA and identify the key features that affect the effectiveness of DL approaches. Method: A Systematic Literature Review (SLR) was carried out and reported based on the preferred reporting items for systematic reviews. 4065 papers were retrieved by manual search in four databases which are Google Scholar, Taylor & Francis, Springer Link and Science Direct. 34 primary studies were finally included. Result: From these studies, 70% were journal articles, 25% were conference papers and 5% were contributions from the studies consisted of book chapters. Five features of DL were identified and analyzed. The features are (1) hierarchical layer, (2) high-level abstraction, (3) process high volume of data, (4) universal model and (5) does not over fit the training data. Conclusion: This review delivers the evidence that DL in BDA is an active research area. The review provides researchers with some guidelines for future research on this topic. It also provides broad information on DL in BDA which could be useful for practitioners.

Keywords: *Deep Learning, Big Data Analytics, Systematic Literature Review, Features*

1 Introduction

Big Data is acknowledged as large or complex datasets that even conventional database system are insufficient to do the data processing application within a tolerable elapsed time for its user population [1]. The processing applications are analysis, capture, data curation, search, sharing, storage, transfer, visualization, querying, updating and information privacy.

Data can be classified as big data if the data consists of the three V's [2]. The first V is volume which is the data must have large number of data and it may not only refer to terabytes or petabytes but also can be measured by the number of files, records or transactions. The second V is variety where the data are in the many forms of format and can be organized in structured or unstructured way. The last V is velocity refers at which the data can be generated.

However, to extract and analyze the relevant information in large volume, varied and fast growing data is not an easy task. Analytics can be intended as intricate procedures running over large scale of data repositories as its main goal is that of mining useful knowledge kept in such repositories [3]. Therefore, there are many analytical techniques are introduced in respective to gain as much as information from unmanageable large volume and varied data. Several of these techniques are association rule learning, data mining, cluster analysis, machine learning, text analytics and crowd sourcing [4].

Machine learning techniques have been found very effective and relevant to many real world applications in bioinformatics, network security, healthcare, banking and finance, and transportations [5]. Machine learning allows computers to evolve based on empirical data. A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data [6]. For example, the U.S. Department of Homeland Security uses machine learning to identify patterns in cell phone and email traffic, as well as credit card purchases and other sources surrounding security threats [7]. They use these patterns to try to identify future threats so they can handle them before they become large problems.

In the past years, there are a few scholars who have worked in BDA using machine learning methods. In 2013, [8] have introduced a tutorial on current applications, techniques and systems with the aim of cross-fertilizing research between the database and machine learning communities. The tutorial covers current large scale applications in workflow of machine learning. This tutorial aims to inform the database community about workflows in the machine learning domain. The authors then developed a prototypical workflow of machine learning projects. The workflow consists of three phases which is example formation, modeling and deployment. They believed that the database community can provide the appropriate data management toolkits for the data scientist to operate on large scale machine learning techniques.

Later in 2014, a paper on applying big data classification for network infusion traffic was presented [9]. The author discussed the system challenges in handling big data classification using geometric representation learning techniques and the modern big data networking technologies. Also, he discussed the issues related to combining supervised learning techniques, representation-leaning techniques, machine lifelong learning techniques and big data technologies for solving network traffic classification problems.

Machine learning methods were proposed as the techniques to be used in large scale data analytics as presented by [5]. They discussed on the limitation of the traditional methods and their incremental versions for fast, scalable and accurate big data solutions. Based on their paper, they concluded that machine learning has been the most utilized tool for data analytics. The tools have been successfully used to analyze both small scale as well as large scale data using various techniques such as sampling, feature selection and distributed computations.

The historical aspects of the term "Big Data" and the associated analytics have been reviewed by [10]. They augmented 3 V's with additional attributes of big data to make it more comprehensive and relevant which then proposed BDA to include additional attributes of Business Intelligent and Statistics aspects. They have provided an overview of many popular platforms for BDA that are affordable to small and medium scale enterprises.

A powerful algorithmic framework for big data optimization, called the block successive upper-bound minimization (BSUM) was presented by [11]. The strength of the BSUM framework is its strong theoretical convergence guarantee and its flexibility. They also include discussion from viewpoint of design flexibility, computational efficiency, parallel or distributed implementation and the required communication overhead. However, they have highlighted a couple of issues such as communication delay and overhead in parallel implementation and nonlinear coupling constraints.

In summary, we could not find any related works on conducting SLR in the area of BDA with machine learning. Therefore, the purpose of this paper is to systematically review the current literature on DL approaches in BDA. Findings may assist researchers in designing double armed prognostic studies that could appropriately determine the potential of the key features of DL approaches, assess the effectiveness of DL approaches in enhancing the BDA and explore the pros and cons of DL approaches.

The remainders of this paper are divided into the following parts: the research methodology is described in Section 2, results of this review along with the discussion on the finding of this study are described in Sections 3 and 4 respectively. Finally, the conclusion of this study is expressed in Section 5.

2 Review Method

SLR has been chosen as the research method. This paper uses SLR guidelines, which is a form of secondary study that uses a well-defined methodology [12]. The SLR methodology aims to be as fair as possible by being auditable and repeatable. According to [13], the purpose of a SLR is to provide a complete of possible list of all studies that are related to certain subject area. Meanwhile, traditional reviews attempt to summarize results of a number of studies.

Based on [12], an SLR process is covered by three consecutive phases: planning, conducting and reporting. In this section, we will focus on the planning phase which involves defining the research objectives and how the review is carried out.

2.1 Review Design

This section describes the foundation of this review by defining the SLR research questions and search keywords.

2.1.1 SLR Research Questions

Defining and describing BDA with machine learning methods is a new problem since we have found the earliest related work is on 2013. Over the years, too few studies have conducted machine learning methods with BDA. Hence, this paper intended to identify the features that affect the effectiveness of DL approaches with BDA. The SLR Research Question (RQ) that we intend to answer in this paper is as follows:

“What are the features that influence DL approaches in enhancing the performance of BDA?”

In this paper, the word ‘features’ refers to the item or quality attribute that affect the effectiveness of DL in enhancing the BDA.

2.1.2 Search Process

This SLR concentrates on searching in scientific databases rather than in specific books or technical reports. An assumption was made that most of the research results in books and reports are also typically described or referenced in scientific papers. This paper has selected four databases to perform the SLR search process as follows:

1. Google Scholar (www.scholar.google.com)
2. Taylor and Francis (www.tandfonline.com)
3. Springer Link (www.springerlink.com)
4. Science Direct (www.sciencedirect.com)

The databases were chosen as they offer the most important and highest impact full-text journals and conference proceedings, covering the fields of DL and BDA in generals. The following search keywords are used to find relevant studies in paper's title, keywords and abstract;

"deep learning" OR "deep structured learning" OR "hierarchical learning" AND "big data analytics" OR "large scale data analytics"

2.2 Review Conduction

This section defines the review protocol for conducting the SLR. The SLR review protocol refers to structure and rules of conducting the review.

2.2.1 Inclusion and Exclusion Criteria

Fig. 1 shows the inclusion and exclusion criteria that have been used in this paper. Based on the Fig. 1, candidate papers which do not focus on DL and/or BDA will be excluded. This paper intention is that this SLR should concentrate on features of DL in BDA. Duplicate articles of the same study are also excluded in the SLR. The complete version of the study is included.

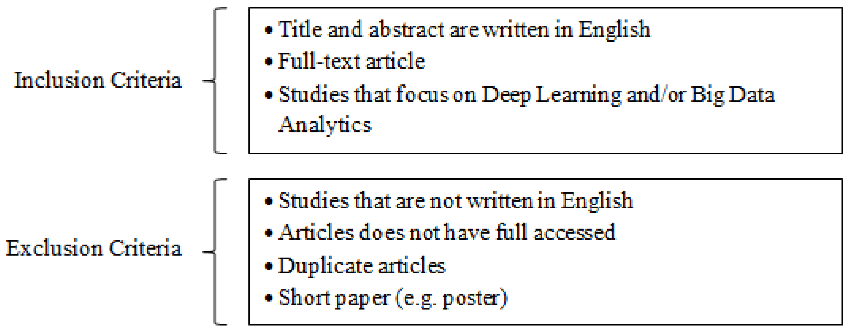

Fig. 1: Inclusion and Exclusion Criteria

2.2.2 Study Selection

The selection of studies is performed through the following processes [14]:

1. Search in databases to identify relevant studies using the search keywords
2. Exclude studies based on the exclusion criteria
3. Exclude irrelevant studied based on analysis of their titles and abstracts
4. Evaluating the selected studied based on full text read
5. Evaluation by external researcher
6. Re-evaluating the results in random
7. Obtain primary studies

2.2.3 Quality Assessment

According to the guidelines of SLR [12], four Quality Assessment (QA) questions have to be defined in order to assess the quality of the research of each proposal and to provide a quantitative comparison between them. The scoring procedures are Y (Yes = 1), P (Partly = 0.5) and N (No = 0). The quality assessment questions defined in this SLR were:

1. Was the articles referred?
   a. Yes: it either explicitly describe the features of DL approaches in BDA
   b. Partially: it only mentioned a few either
   c. No: it neither described nor mentioned features of DL approaches in BDA
2. How clearly are the work limitations documented?
   a. Yes: it clearly explained the limitation the features of DL in BDA
   b. Partially: it mentioned the limitation but did not explain why
   c. No: it did not mention the limitation
3. Were the findings credible?
   a. Yes: the study was methodologically explained so that the finding can be trust
   b. Partially: the study was methodologically explained but not in details
   c. No: the study was not methodologically explained

2.2.4 Data Extraction

Table 1 indicates the data extraction form that is employed for all selected primary studies in order to carry out an in-depth analysis.

Table 1: Data Extracted Form

No	Extracted Data	Description	Type
1	Identity of the study	Unique identity for the study	General
2	Bibliographic references	Authors, year of publication, title and source of publication	General
3	Type of study	Book, journal paper, conference paper, workshop paper	General
4	The features of Deep Learning	Description of the features of Deep Learning in Big Data Analytics	RQ
5	Findings/Contributions	Indicating findings and contributions of the study	General

2.2.5 Synthesis

Results from the analysis through SLR revealed 34 studies for further consideration. All the selected studies have been gone through but only left 20 articles that are able to answer the RQ of this SLR. Fig. 2 shows the number of studies after each defined process.

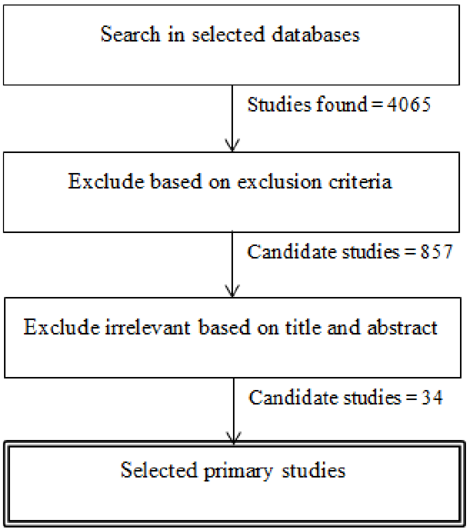

Fig. 2: Finding Primary Studies Procedure

Table 2 indicates the number of type of study which stands in selected paper for review and type journal has the higher selected study per type among conference proceeding and book chapter.

Table 2: Numbers of Selected Study per Type

Study	Count	Percentage
Journal	14	70
Conference Proceeding	5	25
Book Chapter	1	5

Table 3 cited column is obtained from Google Scholar. The journal papers by 70% of total selected papers are the most contribution on this study and book chapters and conferences proceeding stand in second and third place of contribution by 25% and 5% respectively. Table 3 shows the number of citation of selected paper. The data presented (cited column) in Table 3 only gives a rough indication of citation rates and are not meant for comparison among studies.

Table 3: Selected Papers Citation

#	Cited	#	Cited	#	Cited
S1	6	S13	5	S25	0
S2	1	S14	0	S26	0
S3	12	S15	0	S27	6
S4	1	S16	0	S28	0
S5	10	S17	0	S29	0
S6	2	S18	0	S30	1
S7	263	S19	0	S31	4
S8	77	S20	0	S32	0
S9	31	S21	0	S33	1
S10	17	S22	0	S34	0
S11	20	S23	0		
S12	0	S24	7		

Fig. 3 shows the number of primary studies by year of publication. All of these 20 articles are published from year 2014, 2015 and 2016. It shows that year 2015 has the higher selected articles compared to year 2014 and year 2016. Based on Fig. 3, the number of publication is extremely increased in year 2015 using DL approaches. Therefore, we are expected that year 2016 can reach higher than year 2015 since this SLR is written in May 2016.

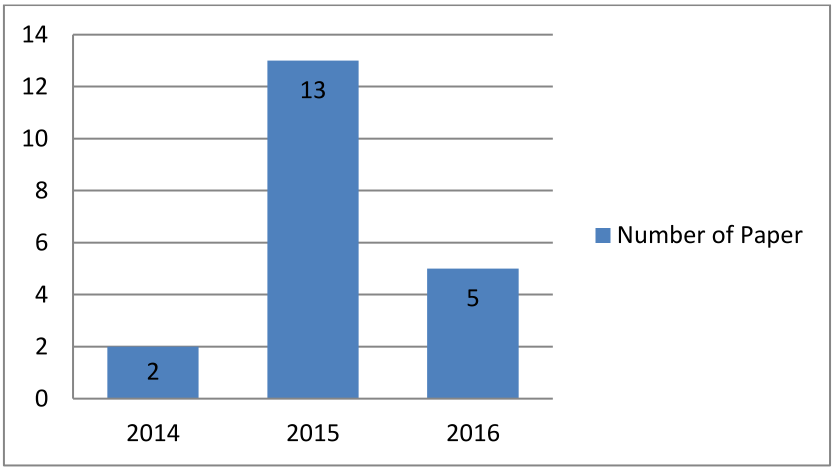

Fig. 3: Number of Papers by Year of Publication

3 Results

This section explains the findings and discussion of this review in order to answer the defined SLR research question. In addition, general discussion that obtained from selected primary studies based on defined data extract forms is expressed. As a result, this section contains finding and discussion on SLR research question.

3.1 Finding RQ

RQ: What are the features that influence DL approaches in enhancing the performance of BDA?

In order to answer this question, we analyzed the data and as a result, five features of DL are identified. The five features of DL are:

1. Hierarchical layer
2. High-level abstraction
3. Process high volume of data
4. Universal model
5. Does not over fit the training data

Table 4 depicts the number of primary studies addressing the identified features of Deep Learning. Features (2) and (3) has more than 10 papers addressed the features which is 16 papers and 11 papers respectively. Feature (1) is addressed by 9 papers. While features (4) and (5) are addressed only by one paper.

Table 4: Number of Primary Study Addressing the Identified Features

No	Features	Number of Papers	Study Identifiers
1	Hierarchical layer	9	S5, S6, S11, S12, S13, S20, S22, S28, S33
2	High-level abstraction	16	S11, S12, S13, S14, S16, S17, S18, S19, S20, S23, S27, S28, S29, S30, S31, S33
3	Process high volume of data	10	S11, S12, S13, S17, S18, S20, S23, S28, S30, S33
4	Universal model	1	S25
5	Does not over fit the training data	1	S29

3.2 Quality Assessment

Once the primary studies of the SLR had been identified, we evaluated them according to the QA questions defined in 2.2.3. The score assigned to each study for each question is shown in Table 5.

Table 5: Quality Assessment of Selected Papers

#	QA1	QA2	QA3	Total Score	% by Max S
S1	N	N	N	0	0
S2	N	N	N	0	0
S3	N	N	N	0	0

S4	N	N	N	0	0
S5	Y	P	P	2	66.67
S6	Y	P	P	2	66.67
S7	N	N	N	0	0
S8	N	N	N	0	0
S9	N	N	N	0	0
S10	N	N	N	0	0
S11	Y	Y	Y	3	100
S12	Y	Y	Y	3	100
S13	Y	Y	Y	3	100
S14	Y	P	P	2	66.67
S15	N	N	N	0	0
S16	Y	P	Y	2.5	83.33
S17	Y	P	P	2	66.67
S18	Y	P	Y	2.5	83.33
S19	Y	P	P	2	66.67
S20	Y	Y	Y	3	100
S21	N	N	N	0	0
S22	Y	P	P	2	66.67
S23	Y	P	P	2	66.67
S25	N	N	N	0	0
S25	Y	P	P	2	66.67
S26	N	N	N	0	0
S27	Y	P	P	2	66.67
S28	Y	Y	Y	3	100
S29	Y	P	P	2	66.67
S30	Y	P	P	2	66.67
S31	Y	P	P	2	66.67
S32	N	N	N	0	0
S33	Y	Y	Y	3	100
S34	N	N	N	0	0
Total	20	13	14	47	
% Total score	42.55	27.66	29.79	100	
% By max QA	100	65	70		

The row "% total score" shows the percentage of points obtained by all the selected study with regard to the total number of points obtained by all the selected studies in all the QA questions. The last row "% max QA" corresponds to the percentage of points collected by the values assigned for a given QA question over the points that would be collected if every selected study got the highest score.

The highest score with a score of 3 obtained by S11, S12, S13, S20, S28 and S33 which represents about 100% of the maximum possible. In contrast, S16 and S18 obtained a score of 2.5 and representing 83.33% of the maximum score that one primary study could get. There are 14 studies that could not get any score which mean that their title and abstract shown that it can give the answer for the research question for this SLR but after going through the full articles, there is no features of Deep Learning has been discussed. According to the Table (QA), first question are distributed over 42.55% of the total score, second question with 27.66% of total score and third question with 29.79% of total score. In view of these results, we can conclude that 12 papers that have a score obtained a minimum quality score of 66.67%.

4 Discussions

This section provides discussions about this SLR. The discussion is about the research question mentioned above in Section 2.1.2. The features of DL Learning in BDA are as follows:

4.1 Hierarchical Layer

Hierarchical layer defined as learning multiple of layers. The DL algorithm is designed to learn low-level features to extract complex high-level features for data representation through a hierarchical learning process, S12. DL has demonstrated a great promise in processing unstructured data, S5 and S28. The essential DL novelty is to design and implement models by trying to identify first lower-level categories to obtain higher-level categories, S6, S11, S12, S20 and S22. S33 added that DL also useful for complex unlabeled datasets which encapsulates machine learning algorithm for organizing the data hierarchically and exposing the most important features, characteristic and explanatory variables as high-level graph nodes. Based on S13, DL uses hierarchical representations of data for classification.

4.2 High-level Abstraction

DL is a branch of machine learning based on a set of algorithms that attempt to model high-level abstractions in data by using model architectures, with complex structures or otherwise, composed of multiple non-linear transformations, S12, S13, S14, S31 and S33. Based on S11, it not only provides complex representations of data which are suitable for artificial intelligent (AI) tasks but also makes the machines independent of human knowledge which is the ultimate goal of AI. It extracts representations directly from unsupervised data without human interference. This is extremely helpful in the case of big data because DL can flush out relationships within large amounts of data that humans would miss, due to our lower computational capacity, S17. The extraction is highly predictive and invariant features from objects with noisy inputs, varying view points, and different deformations and lighting conditions, S19 and S20. The depth of images can be extracted, S30. S18 uses DL on working with raw data without feature extraction. And thus, by analyzing the error classification rates on the Arabic handwritten characters' classification task. S28 mentioned that a feature is a measurement attribute extracted from sensory data to capture the underlying phenomenon being observed and enable more effective analytics

DL can instead reproduce complicated functions that represent higher level extractions, and replace manual domain-specific feature engineering, S29. However, with the development of the current DL techniques [15], different groups of data will be mapped to different layers to process, and the performance of a DL neural network is the current best classifier for big data, S27. According to S16, DL can learn appropriate features from the underlying textual corpus efficiently and thus surpass other state-of-the-art classifiers. However, the successful application of DL techniques is not an easy task; DL implicitly performs feature extraction through the interplay of different hidden layers, the representation of the textual input and the interactions between layers. In S23, a DL method has been compared with SVM classifier and conform that DL are faster especially in the case of raw features.

4.3 Process High Volume of Data

DL algorithm can learn relational and semantic knowledge data representations from large unsupervised raw data at high-level layers, S12. According to S18, the purpose was to take advantages of the power of these deep networks that are able to manage large dimensions input, which allows the use of raw data inputs rather than to extract a feature vector and learn complex decision border between classes. The recent achievement of DL architectures is mostly due to the ability of efficiently training such networks from very large and representative datasets at different level of abstractions (raw pixel image, edged image, image-parts etc..), S20, S30 and S33. S23 and S28 said that the "variety" aspect of MBD (Mobile Big Data) leads to multiple data modalities of multiple sensors (e.g., accelerometer samples, audio, and images). Multimodal DL [16] can learn from multiple modalities and heterogeneous input signals.

Based on S17, DL work well with large sets of data because of their peculiar quality of being a universal approximator [17], [18], which means that theoretically speaking, they can model any real function. A key concept underlying DL methods is distributed representations of the data, in which a large number of possible configurations of the abstract features of the input data are feasible, allowing for a compact representation of each sample and leading to a richer generalization, S11. According to S13, DL methods have been used in many applications, pattern recognition, computer vision, natural language processing and speech recognition. Due to exponential increase of data in these applications, DL is useful for accurate prediction from voluminous data.

4.4 Universal Model

Based on S25, universal models are defined as the machine learning models that learn the universal phenomena inductively, or adopt mathematical formulas, or physical models to characterized the universal phenomena, universal machine learning models, universal mathematical models and universal physical models. Google's artificial brain learns to identify a cat which confirmed that DL method is universal because it can recognize any cat [19].

4.5 Does Not Over Fit Training Data

In order to predict household income based on mobile communication and mobility pattern, S29 has implemented a multilayer feed forward DL architecture which while capturing the complex dependencies between different dimensions of the data, the DL algorithms do not over fit the training data as seen by their test performance.

5 Conclusions

The goal of this paper is to conduct a systematic literature review on Deep Learning in Big Data Analytics. Our aims are to investigate and identify the features of DL which influence the effectiveness of BDA. Our results reveal that there are five features of DL which need to be considered in order to enhance the BDA.

To conclude, we would like to stress an idea that has been presented throughout this study: DL in BDA is currently an active research area as shown in Fig. 3. Since BDA is becoming more mature, it is time to face the different challenges that will allow us to enhance it and make it more effectiveness.

ACKNOWLEDGEMENTS.
This work is supported by Fundamental Research Grant Scheme, FRGS PY/2016/07233 under title “Enhancing Data Analytics Algorithms using Deep Learning Approaches in Predicting Big Data Cyber-Enabled Crimes”.

Appendix: Primary Study in Review

#	Citation
S1	Hornby, G., Jennings, G., & Nulty, D. (2009). Facilitating deep learning in an information systems course through application of curriculum design principles. *Journal of Teaching in Travel & Tourism*, *9*(1-2), 124-141.

S2	Zhao, H., & Ye, X. (2013). A multidimensional OLAP engine implementation in key-value database systems. In *Advancing Big Data Benchmarks* (pp. 155-170). Springer International Publishing.
S3	Liao, G., Datta, K., & Willke, T. L. (2013). Gunther: Search-based auto-tuning of mapreduce. In *Euro-Par 2013 Parallel Processing* (pp. 406-419). Springer Berlin Heidelberg.
S4	Saletore, V. A., Krishnan, K., Viswanathan, V., & Tolentino, M. E. (2013). HcBench: Methodology, development, and full-system characterization of a customer usage representative big data/hadoop benchmark. In *Advancing Big Data Benchmarks* (pp. 73-93). Springer International Publishing.
S5	Wang, Y., Li, B., Luo, R., Chen, Y., Xu, N., & Yang, H. (2014, March). Energy efficient neural networks for big data analytics. In Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014 (pp. 1-2). IEEE.
S6	HEGER, D. (2014). Big Data Analytics—Where to go from Here. International Journal of Develkopments in Big Data and Analytics, 1(1), 42-58.
S7	Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*,*275*, 314-347.
S8	Doulkeridis, C., & Nørvåg, K. (2014). A survey of large-scale analytical query processing in MapReduce. *The VLDB Journal*, *23*(3), 355-380.
S9	Holzinger, A., & Jurisica, I. (2014). Knowledge discovery and data mining in biomedical informatics: The future is in integrative, interactive machine learning solutions. In *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics* (pp. 1-18). Springer Berlin Heidelberg.
S10	Thomas, S. A., & Jin, Y. (2014). Reconstructing biological gene regulatory networks: where optimization meets big data. *Evolutionary Intelligence*, *7*(1), 29-47.
S11	Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1), 1-21.
S12	Chen, J. C., & Liu, C. F. (2015, October). Visual-based Deep Learning for Clothing from Large Database. In Proceedings of the ASE BigData & SocialInformatics 2015 (p. 42). ACM.

S13	Kashyap, H., Ahmed, H. A., Hoque, N., Roy, S., & Bhattacharyya, D. K. (2015). Big Data Analytics in Bioinformatics: A Machine Learning Perspective. arXiv preprint arXiv:1506.05101.
S14	Adams, T., & Grimaud, J. Big Connectome Research.
S15	Zhao, Y., MacKinnon, D. J., & Gallup, S. P. (2015). Big Data and Deep Learning for Understanding DoD Data. CrossTalk, 5.
S16	Fehrer, R., & Feuerriegel, S. (2015). Improving Decision Analytics with Deep Learning: The Case of Financial Disclosures. arXiv preprint arXiv:1508.01993.
S17	McLeod, C. (2015). A Framework for Distributed Deep Learning Layer Design in Python. arXiv preprint arXiv:1510.07303.
S18	Elleuch, M., Tagougui, N., & Kherallah, M. (2015, November). Towards Unsupervised Learning for Arabic Handwritten Recognition Using Deep Architectures. In Neural Information Processing (pp. 363-372). Springer International Publishing.
S19	Aryal, J., & Dutta, R. (2015, April). Smart city and geospatiality: Hobart deeply learned. In Data Engineering Workshops (ICDEW), 2015 31st IEEE International Conference on (pp. 108-109). IEEE.
S20	Alhamali, A., Salha, N., Morcel, R., Ezzeddine, M., Hamdan, O., Akkary, H., & Hajj, H. (2015, November). FPGA-Accelerated Hadoop Cluster for Deep Learning Computations. In 2015 IEEE International Conference on Data Mining Workshop (ICDMW) (pp. 565-574). IEEE.
S21	Shamsuddin, S. M., & Hasan, S. Data Science vs Big Data@ UTM Big Data Centre.
S22	Oyedotun, O. K., & Khashman, A. Deep learning in vision-based static hand gesture recognition. Neural Computing and Applications, 1-11.
S23	Sanchez-Riera, J., Hua, K. L., Hsiao, Y. S., Lim, T., Hidayati, S. C., & Cheng, W. H. (2015). A Comparative Study of Data Fusion for RGB-D Based Visual Recognition. *Pattern Recognition Letters*.
S25	Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: a survey. *Journal of Big Data*, *2*(1), 1-32.
S25	Shen, B. Universal knowledge discovery from big data using combined dual-cycle. *International Journal of Machine Learning and Cybernetics*, 1-12.

S26	Hou, W., Guo, P., & Guo, L. (2015). Networking Big Data: Definition, Key Technologies and Challenging Issues of Transmission. In *Big Data Computing and Communications* (pp. 103-112). Springer International Publishing.
S27	Duan, L., & Xiong, Y. (2015). Big data analytics and business analytics.*Journal of Management Analytics*, *2*(1), 1-21.
S28	Alsheikh, M. A., Niyato, D., Lin, S., Tan, H. P., & Han, Z. (2016). Mobile Big Data Analytics Using Deep Learning and Apache Spark. arXiv preprint arXiv:1602.07031.
S29	Sundsøy, P., Bjelland, J., Reme, B. A., Iqbal, A. M., & Jahani, E. (2016). Deep Learning Applied to Mobile Phone Data for Individual Income Classification.
S30	He, Y., Yu, F. R., Zhao, N., Yin, H., Yao, H., & Qiu, R. C. (2016). Big data analytics in mobile cellular networks.
S31	Chen, Y., Yang, X., Zhong, B., Pan, S., Chen, D., & Zhang, H. (2016). CNNTracker: Online discriminative object tracking via deep convolutional neural network. *Applied Soft Computing*, *38*, 1088-1098.
S32	Xu, Q., Aung, K. M. M., Zhu, Y., & Yong, K. L. (2016). Building a large-scale object-based active storage platform for data analytics in the internet of things. *The Journal of Supercomputing*, 1-19.
S33	Dinov, I. D. (2016). Methodological challenges and analytic opportunities for modeling and interpreting Big Healthcare Data. *GigaScience*, *5*(1), 1.
S34	Mami, I., Bellahsene, Z., & Coletta, R. (2016). A Constraint Optimization Method for Large-Scale Distributed View Selection. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXV* (pp. 71-108). Springer Berlin Heidelberg.

References

[1] Adrian, M. Big Data. Teradata Magazine, 1(11).

[2] Russom, P. (2011). Big data analytics. TDWI Best Practices Report, Fourth Quarter, 1-35.

[3] Cuzzocrea, A., Song, I. Y., & Davis, K. C. (2011, October). Analytics over large-scale multidimensional data: the big data revolution!. In Proceedings of the ACM 14th international workshop on Data Warehousing and OLAP (pp. 101-104). ACM.

[4] Maltby, D. (2011). Big data analytics. Proceeding of Association for Information Science and Technology.

[5] Kashyap, H., Ahmed, H. A., Hoque, N., Roy, S., & Bhattacharyya, D. K. (2015). Big Data Analytics in Bioinformatics: A Machine Learning Perspective. arXiv preprint arXiv:1506.05101.

[6] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity.

[7] Miller, K. (2012). Big data analytics in biomedical research. Biomedical Computation Review, 2, 14-21.

[8] Condie, T., Mineiro, P., Polyzotis, N., & Weimer, M. (2013, June). Machine learning for big data. In Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data (pp. 939-942). ACM.

[9] Suthaharan, S. (2014). Big data classification: Problems and challenges in network intrusion prediction with machine learning. ACM SIGMETRICS Performance Evaluation Review, 41(4), 70-73.

[10]Wu, C., Buyya, R., & Ramamohanarao, K. (2016). Big Data Analytics= Machine Learning+ Cloud Computing. arXiv preprint arXiv:1601.03115.

[11]Hong, M., Razaviyayn, M., Luo, Z. Q., & Pang, J. S. (2016). A Unified Algorithmic Framework for Block-Structured Optimization Involving Big Data: With applications in machine learning and signal processing. Signal Processing Magazine, IEEE, 33(1), 57-77.

[12]Kitchenham, B., & Charters, S. Guidelines for performing systematic literature reviews in software engineering. 2007. URL http://www. dur. ac. uk/ebse/resources/Systematic-reviews-5-8. pdf.

[13]Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: a step-by-step approach.

[14] Rouhani, B. D., Mahrin, M. N. R., Nikpay, F., Ahmad, R. B., & Nikfard, P. (2015). A systematic literature review on Enterprise Architecture Implementation Methodologies. Information and Software Technology, 62, 1-20.

[15]Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep machine learning-a new frontier in artificial intelligence research [research frontier].Computational Intelligence Magazine, IEEE, 5(4), 13-18.

[16]Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal deep learning. In Proceedings of the 28th international conference on machine learning (ICML-11) (pp. 689-696).

[17]Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. Neural networks, 2(5), 359-366.

[18]Leshno, M., Lin, V. Y., Pinkus, A., & Schocken, S. (1993). Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural networks, 6(6), 861-867.

[19]Bengio, Y. (2009). Learning Deep Architectures for AI. Foundations and trends® in Machine Learning, 2(1), 1-127.