Machine learning on big data for future computing

Young-Sik Jeong[1] · **Houcine Hassan**[2] · **Arun Kumar Sangaiah**[3]

1 Introduction

Recently, the integration of big data and machine learning technologies has emerged as an exciting new paradigm for industry and life. These technologies are fast-growing areas of research that cover the fields of Information and Communication Technologies (ICT) for future computing (FC). It becomes a preferable solution for solving complex problems in different fields, such as artificial intelligence, large-scale analysis systems, natural language processing, pattern recognition, video analysis, and telematics. For instance, AlphaGo, developed by Google DeepMind, changed the current paradigm of machine learning on big data after its steady win in the Google DeepMind Challenge. Machine learning is the art and science of designing computer algorithms from the dataset, and AlphaGo learns and improves performance by relying on massive data associated with huge games played by humans to improve artificial intelligence. Likewise, there are many issues to realize and provide effective services and applications with enormous attentions, and efforts have been focused on machine learning on big data for FC. Contributions to theoretical research presenting advanced technologies and concepts, analysis and reporting of experiences, and experiences related to the implementation and application of theories, as well as tutorials on emerging trends are required for these fields of research. As such, the purpose of this special issue is to provide an overview of the state-of-the-art technologies and solution guidelines for these areas of research.

This special issue covers pure research and applications within novel scopes related to machine learning on big data for future computing. In addition, it deals

✉ Young-Sik Jeong
   ysjeong@dongguk.edu

   Houcine Hassan
   husein@disca.upv.es

   Arun Kumar Sangaiah
   sarunkumar@vit.ac.in

1   Dongguk University-Seoul, Seoul 04620, Republic of Korea

2   Polytechnic University of Valencia, Valencia 46022, Spain

3   VIT University, Vellore, Tamil Nadu 632014, India

with hardware/software technologies, new frameworks and architectures, efficient data processing, specific mathematical models, and theories of future computing.

2 Machine learning on big data for future computing

In this special issue, 'Machine Learning on Big Data for Future Computing' is presented in the twenty-four accepted articles [1–21]. All accepted articles are categorized into four different dimensions: computational optimization, privacy and security, optimized data processing and sharing, and communication network.

2.1 Computational optimization

In the computational optimization dimension, six articles were accepted. Peng et al. [1] designed a graphics processing unit (GPU) parallel algorithm, in particular, to speed up the process of weighting the characteristics of classification based on data gravitation. They tackled the problem of reducing the computational complexity of the classification model based on data gravitation using GPU. Hur et al. [2] proposed an entropy-based pruning technique to determine the threshold by considering the average amount of output weight information while training to lighten the computational complexity of neural networks.

To address the problems of steering road feel, and optimization of the steering and sensitivity of the automobile, a novel vehicle handling stability evaluation method is proposed by Chen et al. [3]. They analyzed the influence of the important parameters on the steering system using the method of absolute sensitivity.

In the field of distance education of higher education institutions, e-learning has improved the traditional teaching method and provided a better way to share educational resources for human beings. Another research conducted by Panwar et al. [4] studies emotional states and musical cultures in relation to radio songs broadcast in four different cities and in four different regions. They proposed a music perception model for machine learning to identify the emotional content of a given audio stream and to study the emotional effects of music.

2.2 Privacy and security

In the privacy and security dimension, seven articles were accepted. To solve security issues related to wireless devices, Lin et al. [5] proposed a radio frequency fingerprint technique to identify authenticated devices. They included the combination of the dimension reduction method and the classifier to meet the needs of the application. A semantic approach to making a machine-readable large-scale attack graph to improve its readability for a large-scale attack graph is presented by Lee et al. [6]. Kaur et al. [7] proposed a multi-party privacy preserving classification scheme in the cloud for arbitrarily distributed data among multiple parties so that parts of different regions can collaborate effectively.

For vehicular ad-hoc networks, an efficient certificateless aggregate signature scheme is presented by Kumar et al. [8]. They claimed that the signature scheme is tamper proof against adaptively chosen message attacks. They also demonstrated the preservation of conditional confidentiality in the message generated by a vehicle. Another scheme is proposed for continuous leakage and tampering resilient of CCA secure public-key encryption schemes by Hu et al. [9].

Using the blockchain technique, Huh et al. [10] proposed an integrated automatic login platform based on mobile fingerprint recognition. Due to the limitations of existing digital door locks easily opening with a stolen PIN code or electric shock, they have focused on enabling users to perform their authentication in a more secure and convenient way.

2.3 Optimized data processing and sharing

Five articles were accepted in the dimension of optimized data processing and sharing. Recently, to expand the transfer radius between stations and allow passengers to choose their destinations more flexibly, sharing plat formless bikes is gaining popularity around the world. Ai et al. [11] developed a solution to the recursive fuzzy linear regression model with transfer tolerance to explore the functional relationships between passengers. Experimental results indicate that foot-transferring passengers will be more distance sensitive, while users of shared bikes without a dock are more concerned about the time cost of finding the available bike and the delay of the traffic lights.

Precise segmentation of the gallbladder and gallstone from the ultrasound scan image is a difficult task in medical image processing. To solve this problem, Muneeswaran et al. [12] presented a bio-inspired algorithm based on the spider web construction process for the image segmentation process. They focused on the detection of gallbladder shape and the segmentation of gallstones and polyps within the gallbladder using a computerized algorithm. Further research is being conducted to provide a personalized health-focused referral system in the IoT-based healthcare system to reduce travel time and costs of the travelers [13]. The personalized health-focused referral system takes into account climatic attributes based on the user's personal choice and nutritional value.

In the area of autonomous driving, the data computing process is an element for detecting and tracking objects in close motion avoiding collisions and navigation in case of heavy traffic and pedestrian zones. Zhang et al. [14] proposed a multi-object tracking method based on light detection and ranging data to track an object, monitor surrounding objects, and predict moving states of objects in real time. Another research is conducted to create a realistic three-dimensional visualization with which a remote driver can intuitively control a remote unmanned robot using the massive sensor data characteristics [15]. They proposed a three-dimensional object reconstruction method using real-time fusion and multi-sensor data analysis.

2.4 Communication network

Finally, six articles were accepted in the communication network dimension. Recently, spectrum requirements for data transmission have increased rapidly due to the tremendous growth in the adaptive capacity of smart wireless devices and services. To solve this problem, Fu et al. [16] proposed two parallel cooperative heuristic-based spectrum-sensing strategies with time-based exhaustion and time heuristic detection approaches. Bedi et al. [17] presented an efficient and secure multi-cloud storage approach by leveraging the fog computing technique to address the limitation of resource-constrained mobile devices. Another research is conducted to reduce the cost of communication by providing a method for detecting detection errors and recovering the inter-pulse interval using thresholds. There is a method that reduces the cost of communication compared to previous approaches, not only energy efficient but also key generating [18]. The efficiency of the recovery method is also demonstrated. Diwakaran et al. [19] proposed a data-aware energy conservation prediction model for wireless sensor network to reduce the redundant data transmission. To avoid an increase in the memory usage of the algorithm when the matching algorithm is applied due to a large number of transfer edges, Sun et al. [20] proposed an improved method of the concatenation of transfer edges using a range of characters, with several consecutive characters represented by character intervals. Lastly, Kim et al. [21] proposed a framework to enhance the security capacity of small- and medium-sized enterprises when using big data to build security capability. The framework includes an Android application implemented with Java Android to conduct an empirical analysis of the contents of diagnostics elements of major security capability.

3 Conclusions

Special thanks go to Journal of supercomputing's Editor-in-Chief, as well as to all editorial teams for their invaluable support throughout the preparation and publication of this special issue. In addition, we thank the external reviewers for their invaluable help in reviewing the papers.

References

1. Peng L, Zhang H, Hassan H, Chen Y, Yang B (2018) Accelerating data gravitation-based classification using GPU. J Supercomput. https://doi.org/10.1007/s11227-018-2253-5
2. Hur C, Kang S (2018) Entropy-based pruning method for convolutional neural networks. J Supercomput. https://doi.org/10.1007/s11227-018-2684-z
3. Chen H, Chen X, Chen C, Peng Z, Gong Y, Chen G (2017) Multidisciplinary design optimization for vehicle handling stability of steering-by-wire system. J Supercomput. https://doi.org/10.1007/s11227-017-2208-2

4. Panwar S, Rad P, Choo KKR, Roopaei M (2018) Are you emotional or depressed? Learning about your emotional state from your music using machine learning. J Supercomput. https://doi.org/10.1007/s11227-018-2499-y

5. Lin Y, Zhu X, Zheng Z, Dou Z, Zhou R (2017) The individual identification method of wireless device based on dimensionality reduction and machine learning. J Supercomput. https://doi.org/10.1007/s11227-017-2216-2

6. Lee J, Moon D, Kim I, Lee Y (2018) A semantic approach to improving machine readability of a large-scale attack graph. J Supercomput. https://doi.org/10.1007/s11227-018-2394-6

7. Kaur H, Kumar N, Batra S (2018) ClaMPP: a cloud-based multi-party privacy preserving classification scheme for distributed applications. J Supercomput. https://doi.org/10.1007/s11227-018-2691-0

8. Kumar P, Kumari S, Sharma V, Li X, Sangaiah AK, Islam SH (2018) Secure CLS and CL-AS schemes designed for VANETs. J Supercomput. https://doi.org/10.1007/s11227-018-2312-y

9. Hu C, Yang R, Liu P, Li T, Kong F (2018) A countermeasure against cryptographic key leakage in cloud: public-key encryption with continuous leakage and tampering resilience. J Supercomput. https://doi.org/10.1007/s11227-018-2534-z

10. Huh JH, Seo K (2018) Blockchain-based mobile fingerprint verification and automatic log-in platform for future computing. J Supercomput. https://doi.org/10.1007/s11227-018-2496-1

11. Ai Y, Li Z, Gan M (2018) A solution to measure traveler's transfer tolerance for walking mode and dockless bike-sharing mode. J Supercomput. https://doi.org/10.1007/s11227-017-2211-7

12. Muneeswaran V, Rajasekaran MP (2018) Automatic segmentation of gallbladder using bio-inspired algorithm based on a spider web construction model. J Supercomput. https://doi.org/10.1007/s11227-017-2230-4

13. Subramaniyaswamy V, Manogaran G, Logesh R, Vijayakumar V, Chilamkurti N, Malathi D, Senthilselvan N (2018) An ontology-driven personalized food recommendation in IoT-based healthcare system. J Supercomput. https://doi.org/10.1007/s11227-018-2331-8

14. Zhang W, Cho S, Chae J, Sung Y, Cho K (2018) Object tracking method based on data computing. J Supercomput. https://doi.org/10.1007/s11227-018-2535-y

15. Cho S, Cho K (2019) Real-time 3D reconstruction method using massive multi-sensor data analysis and fusion. J Supercomput. https://doi.org/10.1007/s11227-019-02747-3

16. Fu S, Zhang G, Fujii T (2018) A heuristic method-based parallel cooperative spectrum sensing in heterogeneous network. J Supercomput. https://doi.org/10.1007/s11227-018-2250-8

17. Bedi RK, Singh J, Gupta SK (2018) MWC: an efficient and secure multi-cloud storage approach to leverage augmentation of multi-cloud storage services on mobile devices using fog computing. J Supercomput. https://doi.org/10.1007/s11227-018-2304-y

18. Kim J, Cho K, Kim YK, Lim KS, Shin SU (2018) Study on peak misdetection recovery of key exchange protocol using heartbeat. J Supercomput. https://doi.org/10.1007/s11227-018-2616-y

19. Diwakaran S, Perumal B, Devi KV (2018) A common prediction model based data collection for energy efficient wireless. J Supercomput. https://doi.org/10.1007/s11227-018-2437-z

20. Sun R, Shi L, Yin C, Wang J (2018) An improved method in deep packet inspection based on regular expression. J Supercomput. https://doi.org/10.1007/s11227-018-2517-0

21. Kim HK, So WH, Je SM (2019) A big data framework for network security of small and medium enterprises for future computing. J Supercomput. https://doi.org/10.1007/s11227-019-02815-8