Complex & Intelligent Systems (2022) 8:705–707
https://doi.org/10.1007/s40747-021-00539-y

EDITORIAL

Special issue on interpretation of deep learning: prediction, representation, quantification and visualization

Nian Zhang[1] · Zhaojie Ju[2] · Chenguang Yang[3] · Dingguo Zhang[4] · Jinguo Liu[5]

Published online: 17 September 2021

While Big Data offers the great potential for revolutionizing all aspects of our society, harvesting of valuable knowledge from Big Data is an extremely challenging task. The large scale and rapidly growing information hidden in the unprecedented volumes of non-traditional data requires the development of decision-making algorithms. Deep learning is currently an extremely active research area in machine learning and pattern recognition society. In contrast to the conventional classification methods, deep learning models can learn a hierarchy of features by building high-level features from low-level ones, thereby automating the process of feature construction for the problem. The deep learning approach exploits many nonlinear processing layers to develop representations of data at increasing levels of abstraction. It has demonstrated best-in-class performance in a range of applications, including image classification, and been successfully applied in industry products that take advantage of the large volume of digital data. Companies like Google, Apple, and Facebook, who collect and analyze massive amounts of data on a daily basis, have been aggressively pushing forward deep learning techniques.

While deep learning has achieved unprecedented prediction capabilities, it is often criticized as a black box because of lacking interpretability, which is very important in real-world applications such as healthcare and cybersecurity. For example, healthcare professionals would appropriately trust and effectively manage prediction results only if they can understand why and how a patient is diagnosed with pre-diabetes. There has been an explosion of interest recently in the related research directions, such as (a) analyzing the information bottleneck for efficient learning, (b) inferring and regularizing the network structure for stable and robust prediction, and (c) interpreting the learned representations and generated decisions. This special issue will focus on the interpretability of deep learning from representation, modeling and prediction, as well as the deployment of interpretability in various applications.

Qin et al. propose a projection-based continuous-time algorithm to solve distributed optimization problems with equality and inequality constraints over multi-agent systems. By exact penalty method, the distributed optimization problem is reformulated to a new one without inequality constraints and consensus constraints. It is proved that from any initial points, the states of continuous-time algorithm will enter the equality constraint set of the transformed distributed optimization problem in fixed time. States of continuous-time algorithm are proved to be convergent to an optimal solution of the original distributed optimization problem. Compared with existed models and approaches, the continuous-time algorithm has the advantages of owning fewer state variables. Besides, the states of continuous-time algorithm can find an optimal solution of distributed optimization problem under mild assumptions.

Chen et al. propose a novel progressive genetic algorithm named Progressively Searching Tensor Ring Network Search (PSTRN), which has the ability to find optimal rank precisely and efficiently. Through the evolutionary phase and progressive phase, PSTRN can converge to the interest region quickly and harvest good performance.

✉ Nian Zhang
nzhang@udc.edu

Zhaojie Ju
zhaojie.ju@port.ac.uk

Chenguang Yang
charlie.yang@uwe.ac.uk

Dingguo Zhang
d.zhang@bath.ac.uk

Jinguo Liu
liujinguo@sia.cn

[1] University of District of Columbia, Washington, DC, USA

[2] University of Portsmouth, Portsmouth, UK

[3] University of West England, Bristol, UK

[4] University of Bath, Bath, UK

[5] Shenyang Institute of Automation, Shenyang, China

Experimental results show that PSTRN can significantly reduce the complexity of seeking rank, compared with the enumerating method. Furthermore, the proposed method is validated on public benchmarks like MNIST, CIFAR10/100 and HMDB51, achieving the state-of-the-art performance.

Zeng et al. present the memristor crossbar architectures for implementation of deep neural networks, which include architectures for fully connected layer, convolutional operation, and average pooling operation. Memristor-based multilayer neural network (MbMNN) and convolutional neural network (MbCNN) are built to evaluate the performance of these memristor crossbar architectures. The networks are in-situ trained by two kinds of weight update schemes, which are the fixed-voltage update and the approximately linear update, and simulation results show that the networks trained by the weight update schemes result in satisfying performances. The robustness of MbMNN and MbCNN to conductance variations of memristors is also analyzed. The memristor-based DNNs constructed by presented memristor crossbars perform satisfactorily in pattern recognition tasks and have certain robustness to imperfections of hardware.

Le et al. investigate the enduring question of how to recover pruned models. Knowledge distillation is used as a perfect tool for transferring knowledge, increasing accuracy, and compressing models. An improved integrated framework of pruning combining knowledge distillation strategy is proposed. Experiments have been performed on different image classification CNNs with three knowledge distillation methods. The results show that knowledge from the original network of various forms helped the pruned network recover with higher accuracy. It is also observed that different knowledge distillation methods suited different architecture of CNNs.

Cheng et al. propose a new experimental paradigm and only use fNIRS signals to complete the classification task of six subjects in order to improve the accuracy of classification. The experiment is carried out in a non-laboratory environment, and movements of motion imagination are properly designed. When the subjects are imagining the motions, they are subvocalizing the movements to prevent distraction. Next, the signals are classified by nine classification methods, and the different characteristics and classification methods are compared. The results show that under this new experimental paradigm, the classification accuracy of 89.12% and 88.47% can be achieved by using support vector machine and random forest methods, and the paradigm is effective. Finally, by selecting the five channels with the largest variance after empirical mode decomposition of the original signal, similar classification effect can be achieved.

Yu et al. propose a botnets detection system based on the FP-growth (Frequent Pattern Tree) frequent item mining algorithm to improve the detectability of botnet activities. The detection system is composed of three parts: packet collection processing, rule mining, and statistical analysis of rules. Its characteristic feature is the rule-based classification of different botnet behaviors in a fast and unsupervised fashion. The effectiveness of the approach is validated in a scenario with 11 Peer-to-Peer host PCs, 42,063 Non-Peer-to-Peer host PCs, and 17 host PCs with three different botnet activities (Storm, Waledac and Zeus). The recognition accuracy of the proposed architecture is shown to be above 94%. The proposed method is shown to improve the results reported in literature.

Wang et al. propose a linear parameterization method in order to fully utilize the system information. Firstly, by applying the Lagrange's mean-value theorem, the linear parameterization method is applied to transform the nonlinear system into a linear one with time-varying bounded uncertain terms. Secondly, a new generalized convex combination lemma is proposed to dispose the relationship of bounded uncertainties with respect to their boundaries. Thirdly, sufficient conditions are established to ensure the FTS by resorting to Lyapunov Krasovskii theory, convex combination technique, Jensen's inequality, linear matrix inequality, etc. Finally, the simulation verifications indicate the validity of the theoretical results.

Dai et al. propose a robust semi-supervised non-negative matrix factorization method for binary subspace learning (RSNMF) to handle Salt and Pepper noise and Contiguous Occlusion problems. In order to achieve better clustering performance on the data that have been contaminated by outliers and noise, a weighted constraint on the noise matrix is proposed. In addition, the manifold learning is imposed into non-negative matrix factorization. Moreover, the discrete hashing learning method is utilized to constrain the learned subspace, which can achieve a binary subspace from the original data. Experimental results validate the robustness and effectiveness of RSNMF in binary subspace learning and image clustering on the face dataset corrupted by Salt and Pepper noise and Contiguous Occlusion.

Acknowledgements The guest editors wish to thank Dr. Yaochu Jin (IEEE Fellow, Editor-in-Chief of the Complex and Intelligent Systems) for providing the opportunity to edit this special issue. We would also like to thank Ms. Mangayarkarasi Lakshmiram, the Editorial Assistant, and the referees who have thoroughly evaluated the papers and the editorial staff for their support.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.