Methods xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Guest Editor's Introduction

Interpretable machine learning in bioinformatics

Machine learning has emerged as an indispensable tool impacting a wide range of areas, including (but not limited to) computer vision, natural language processing, business analytics, bioinformatics, and medical fields. Machine learning learns complex patterns or rules from big data automatically, mainly for data representation and prediction problems, instead of explicitly defining them based on prior domain knowledge or observation. Although an exponentially increasing number of machine learning techniques have been applied to solve challenging related problems, the difficulty of model interpretation is a barrier that continues to cause hesitation in utilizing machine learning in certain areas. Furthermore, the recent advent of deep learning, which is considered as a black box, makes the difficulty more challenging, despite its outstanding predictive performance in many applications.

The ultimate goal of interpretable machine learning is to obtain useful domain knowledge from a machine learning model, rather than only to optimize predictive performance. Interpretable machine learning is a slightly different concept from explainable machine learning which focuses on explaining how machine learning works mathematically. Thus, interpretable machine learning can be widely used as an advanced analytics tool. Interpretable machine learning has mainly been implemented by the following approaches: (1) embedding a domain bias into a model by explicitly modeling phenomena of interest, or specific patterns; (2) incorporating prior knowledge or databases; and (3) integrating multi-modal data or databases into a model.

This special issue of *Interpretable Machine Learning in Bioinformatics* introduces several novel, interpretable machine learning models in various fields of bioinformatics, including medical image analysis, drug-drug interaction prediction, clinical outcome prediction, association studies, and electronic health record data analysis.

The advent of digitizing whole-slide images has highlighted the advancement of deep learning-based automatic tissue examination. While most deep learning methods have suggested improving predictive performance in digital pathology, identifying unknown morphological patterns to understand clinical outcomes has only recently been given attention. Kosaraju et al. [1] developed a novel deep learning model for histopathology (named Deep-Hipo) that takes multi-scale patches of the same size, in both high and low magnification levels simultaneously, not only to improve predictive performance, but also to connect morphological patterns to domain knowledge. The method was assessed with the publicly available medical imaging data, The Cancer Genome Atlas (TCGA), Stomach Adenocarcinoma (TCGA-STAD), and TCGA Colon Adenocarcinoma (TCGA-COAD), and the experimental results were clinically verified by a pathologist.

Histopathological image retrieval methods can be useful to search similar whole slide images or cases for reference, when a pathologist analyzes ambiguous slide images. Yang et al. [2] developed a deep metric learning-based histopathological image retrieval method to learn an embedding function that can map original whole slide images into a predefined metric space, where similar images are close to each other. The authors claim that the proposed method achieves high recall in the datasets.

Mozaffari et al. [3] proposed two novel deep neural network approaches (named BowNet and wBowNet) that extract tongue contours from ultrasound images in real time. Dilated convolutional neural networks learned local and global contexts from time-series ultra-image data to make sharper segmentation results. The authors stated that the accurate and automatic tongue contour extraction could be applied to study healthy or impaired speech production. Additionally, Google recently proposed an advanced version of the proposed BowNet method for semantic segmentation application.

Drug-drug interaction (DDI) prediction can optimize synergistic reactions of two or more drugs to treat diseases, while minimizing their antagonistic reactions. Zhang et al. [4] proposed a drug representation learning method based on multi-modal deep autoencoders (named DDI-MDAE) to predict drug-drug interactions. DDI-MDAE learned DDIs via unified drug representations by integrating the multiple drug feature networks of deep auto-encoders. Then, a random forest classifier was adopted in the positive-unlabeled learning settings to predict the DDIs.

Chen et al. [5] developed a Graph Convolutional Network with Bond-award Message Propagation (GCN-BMP) to predict DDIs by alleviating inductive bias from noisy data. The proposed model is based on graph neural networks, where molecule graph attention modules provide the insight of domain knowledge.

Accurate molecule toxicity prediction makes drug development efficient at early stages by excluding drug candidates that may not be effective in clinical trials. Peng et al. [6] developed a novel molecule representation method and a deep learning-based framework (named TOP) for toxicity prediction, coupled with an recurrent neural network based on a bi-directional gated recurrent unit (BiGRU) and two fully connected neural networks. TOP employed structural features and physiochemical properties simultaneously for toxicity prediction, for better model interpretation. A domain-specific data augmentation method was also proposed to enrich the training dataset for training deep neural networks effectively. The authors claimd that incorporating physiochemical properties into the model helped improve the predictive performance toxicity.

Drug property prediction and *de novo* drug design play important roles in precision medicine, pharma-co-genetics, and pharma-co-informatics. Jo et al. [7] applied a Message Passing neural network (MPNN)-based attention network that learns both local and global features from irregularly formed molecular sequence data, whereas most conventional machine learning methods are based on graph representations. Molecular sequence data were considered as a chemical string, and analyzed by the MPNN-based algorithms. The message passing

https://doi.org/10.1016/j.ymeth.2020.05.024

algorithm learns chemical properties.

Understanding the roles of long non-coding RNAs (lncRNAs) is essential to unveiling various biological processes, including cell growth and development, gene expression regulation, alternative splicing, and nuclear organization. Zeng et al. [8] developed a novel hybrid computational framework (named SDLDA) for lncRNA-disease association prediction by integrating singular value decomposition matrix factorization and deep learning techniques to extract the linear and nonlinear features of lncRNAs and diseases, respectively. The proposed framework demonstrated more outstanding AUC, AUPR, and accuracy than various benchmark methods.

Identifying disease-related RNAs plays an important role in understanding complex human diseases at the molecular level for disease diagnosis, therapy, and prognosis. Zhang et al. [9] proposed a novel Graph Attention Adversarial Network (GAAN) to predict disease-RNA associations. GAAN learns the latent distribution of the node representation for disease-RNA networks, incorporating biological prior knowledge in the adversarial process.

Integration of current biological knowledge makes a model interpretable to characterize biological mechanisms. Jo et al. [10] proposed a novel computational framework (named PL-Subtype) that integrates biological pathways and gene expression profile data to characterize the biological mechanisms of breast cancer subtypes. Pathways were modeled by Pathway logic, which is a logic-based formal system for modeling biological pathways, including post-transitional modifications. PL-Subtype identifies signaling paths from receptors to transcript factors activated in a specific subtype, and discovers potential associations between pathways. The authors analyzed TCGA BRCA data, where they found basal-specific mechanisms composed of signaling and TF networks in breast cancer.

Electronic Health Records (EHR) data often consists of heterogeneous cohorts with different subpopulation-specific risk profiles. The challenge is to identify patient subgroups with health risks, and to discover their risk factors within each subpopulation. Shou et al. [11] developed two supervised mixture models: a Supervised Gaussian Mixture model (SGMM) and a Supervised Bernoulli Mixture model (SBMM). Then they learned logistics regression predictors for each subpopulation simultaneously, which can be highly interpretable. The proposed models identified patient subpopulation groups with different health risks, and pinpointed their risk factors by analyzing the high cost drivers of Medicaid expenditures for inpatient stays.

Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) is widely-adopted, comprehensive, and evolving clinical reference terminology for interoperability between electronic health records. Identifying inconsistencies in SNOMED CT can provide an efficient solution for quality assurance. Agrawal et al. [12] proposed a machine learning quality assurance technique that identifies inconsistent attributes, which require manual auditing. The proposed method was used to conduct intensive text mining analyses to compute the similarity of concepts. The contextual based machine learning makes the analyses interpretable to help reduce inconsistencies in SNOMED CT.

Interpretable machine learning has been applied not only to improve predictive performance, but also to understand complex biomedical systems. We hope this special issue, including a set of representative works in interpretable machine learning, provides readers references for further research in related areas.

Finally, the guest editors would like to thank all of the authors for their contributions and high-quality work in this special issue. We would also like to thank all anonymous reviewers for their great efforts and constructive feedback, as well as the journal's editorial members for their editorial work.

References

[1] Sai Chandra Kosaraju, Jie Hao, Hyun Min Koh, Mingon Kang, Deep-Hipo: multi-scale receptive field deep learning for histopathological image analysis, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.012.

[2] Pengshuai Yang, Yupeng Zhai, Lin Li, Hairong Lv, Jigang Wang, Chengzhan Zhu, Rui Jiang, A deep metric learning approach for histopathological image retrieval, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.015.

[3] M. Hamed Mozaffari, Won-Sook Lee, Encoder-decoder CNN models for automatic tracking of tongue contours in real-time ultrasound data, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.011.

[4] Yang Zhang, Yang Qiu, Yuxin Cui, Shichao Liu, Wen Zhang, Predicting drug-drug interactions using multi-modal deep auto-encoders based network embedding and positive-unlabeled learning, Methods, http://dx.doi.org/10.1016/j.ymeth.2020.05.007. (in press).

[5] Xin Chen, Xien Liu, Ji Wu, GCN-BMP: Investigating graph representation learning for DDI prediction task, Methods, http://dx.doi.org/10.1016/j.ymeth.2020.05.014. (in press).

[6] Yuzhong Peng, Ziqiao Zhang, Qizhi Jiang, Jihong Guan, Shuigeng Zhou, TOP: a deep mixture representation learning method for boosting molecular toxicity prediction, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.013.

[7] Jeonghee Jo, Bumju Kwak, Hyun-Soo Choi, Sungroh Yoon, The message passing neural networks for chemical property prediction on SMILES, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.009.

[8] Min Zeng, Lu. Chengqian, Fuhao Zhang, Yiming Li, Wu. Fang-Xiang, Yaohang Li, Min Li, SDLDA: lncRNA-disease association prediction based on singular value decomposition and deep learning, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.002.

[9] Jinli Zhang, Zongli Jiang, Xiaohua Hu, Bo Song, A novel graph attention adversarial network for predicting disease-related associations, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.010.

[10] Kyuri Jo, Beatriz Santos-Buitrago, Minsu Kim, Sungmin Rhee, Carolyn Talcott, Sun Kim, Logic-based analysis of gene expression data predicts association between TNF, TGFB1 and EGF pathways in basal-like breast cancer, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.008.

[11] Xiao Shou, Georgios Mavroudeas, Malik Magdon-Ismail, Jose Figueroa, Jason N. Kuruzovich, Kristin P. Bennett, Supervised mixture of experts models for population health, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.016.

[12] Ankur Agrawal, Kashifuddin Qazi, Detecting modeling inconsistencies in SNOMED CT using a machine learning technique, Methods (2020), https://doi.org/10.1016/j.ymeth.2020.05.019.

Young-Rae Cho
Computer and Telecommunication Engineering Division, Yonsei University, Wonju, South Korea
E-mail address: youngcho@yonsei.ac.kr.

Mingon Kang
Department of Computer Science, University of Nevada, Las Vegas, NV, USA
E-mail address: mingon.kang@unlv.edu.