Guest Editorial: Data Analytics for Public Health Care

GROWING and intensive attention for the use of data analytics in healthcare due to its proven benefits to gain insights on the macro and micro depths to support decision-making at the patient, clinical, bench side and business levels. Health data analytics contributed to the improvements to patient care, faster and more accurate diagnoses, preventive measures, more personalized treatment, narrow the gap between bedside and bench side and provide more informed decision-making. However, the recent pandemic of COVID-19 draw the world attention to need to extend the power of data analytics beyond the legacy data channels including Electronic Health Records (EHRs) and Patient Portals. Data analytics can assist population health management in improving patient outcomes, enhancing care management, and address social determinants of health. As the population health analytics continues to grow in the healthcare space, healthcare systems must be able to gather data from multiple sources, apply variety of data analytics (e.g., Descriptive Analytics, Predictive Analytics Prescriptive Analytics), and manage the provide insights for the population. The data analytics systems need to be prepared to collect data or mine data, examine current and historical data, evaluate raw data, build predictive models and automate reports as well as be able to integrate heterogeneous data and process them in a uniform way. This type of need have forced the healthcare industry to rewire the entire ecosystem to stay apace with the new phase of data analytics. Solving this challenge will let the pharma and healthcare institutions not to waste time just wrangling through massive data but rather to accelerate research and deliver actionable insights. Digital transformation in healthcare cannot accelerate without effective data analytics that can reach the public and beyond the walls of healthcare centers. Data analytics in healthcare including for public health data can help streamline patient data, encourage forward-thinking treatments, reduce costs, and, most importantly, save lives. It is the way forward. This special issue provided a snapshot of data analytics in healthcare with applications spanning from novel techniques, innovative methods, and useful use-cases demonstrating significant advancement in health and wellbeing. The response from the international scientific community was very positive, and the amount of submitted papers well indicates the large appeal of the topics involved. Although all those papers presented very high quality and interesting applications, only seven papers met the highly demanding criteria for publications followed by the Editorial Board, being selected for final inclusion in the special issue. These selected papers showed a wide variety of cutting-edge topics, which covered a large part of the issues mentioned in the Call for Papers.

The first paper, by Chen *et al.* [A1], exploited the differences between detecting objects in natural scene images and detecting cervical lesions in colposcopic images, and recognized that the CLD task in colposcopic images demands discriminative features to help lesion regions become more distinguishable. To achieve this, they proposed a discriminative CLD network, CervixNet, which incorporates two new modules, global class activation (GCA) and local bin excitation (LBE), in an end-to-end fashion. Specifically, they were among the first to effectively develop both global and local feature enhancing mechanisms, and our approach not only captures discriminative features for cervical lesions, but also learns local excitation to further emphasize the lesions.

Zhu *et al.* [A2] are the first to successfully predict the causal sequence of death using neural machine translation frameworks to support the timely, accurate, and complete death reporting. They also evaluate the model performance using three different accuracy scores, achieving 81.68% accuracy in generating the individual codes in the output sequence. Furthermore, they visualized the attention scores to interpret the causal relationship of diagnosis codes from the discharge records. Specifically, they identify the death-related conditions from available symptoms by mapping all diagnosis codes in the input sequence against all causes of death codes in the output sequence. Lastly, they demonstrate a FHIR-based mobile app to retrieve, modify, and upload cause of death data to improve clinical integration.

The third paper by Beheshti *et al.* [A3] aimed to comprehensively evaluate various regression models for estimating Brain Age not only on CH individuals but also in clinical population. We assessed 22 different regression models on a dataset comprising CH individuals as a training set. We then quantified each regression model on independent test sets composed of CH individuals, MCI subjects, and AD patients. Our comprehensive evaluation suggests that the type of regression algorithm affects downstream comparisons between groups, and caution should be taken to select the regression model in clinical settings.

Gómez-Expósito *et al.* [A4] addressed the problem of monitoring and tracking the evolution of a viral epidemic, such as Covid-19, through the application of signal processing techniques to the time series of data reported by governments and health agencies. Three main contributions can be pointed out: 1) the exclusive use of time-varying geometric ratios of daily data to track the disease, rather than the customary virus reproductive

Date of current version April 13, 2022.
Digital Object Identifier 10.1109/JBHI.2022.3159347

number ($R0$); 2) the development of a simple algebraic model relating the geometric ratio of infectious people, (n), with those of positives, reported and dead; 3) the application of a nonlinear KF, along with a smoothing technique, to estimate the evolution of (n). By properly fitting the estimated values of (n) to a decreasing exponential, an accurate prediction of the epidemic peak can be made, as early as two weeks before the peak actually takes place.

The paper by Tanveer *et al.* [A5] presented a novel ensemble model (DTE) for the classification of Alzheimer's disease. The DTE utilizes a combination of deep learning, transfer learning and ensemble learning. DTE exploits the diversity of individual models with randomly chosen hyper parameter with low generalization error to produce more accurate and robust results. For the large ADNI baseline dataset, the DTE achieved a maximum classification accuracy of 99:09% for NC vs AD and 98:71% for MCI vs AD classification tasks. For the small dataset chosen from ADNI, the DTE achieved a maximum classification accuracy of 85% for NC vs AD.

Subramaniyaswamy *et al.* [A6] in the sixth paper introduced a new method using digital colposcopy to diagnose Cervical cance. The method is called the Faster SmallObject Detection Neural Networks (FSOD-GAN) to automatically detects the cervical spot using Faster Region-Based Convolutional Neural Network (FR-CNN) and performs the hierarchical multiclass classification of three types of cervical cancer lesions. Experimentation was done with colposcopy data collected from available open sources consisting of 1993 patients with three cervical categories, and the proposed approach shows 99% accuracy in diagnosing the stages of cervical cancer.

Finally, the paper by Feng *et al.* [A7] proposed a heterogeneous ensemble learning method to predict the survival of patients with neuroblastoma. To construct a high-quality base learner pool, a novel feature selection method, HFS, is proposed to obtain the optimal feature subset of each base learner, and the obtained optimal feature subset is then used to guide the construction of the base learners as a priori knowledge. Furthermore, the disadvantages of the base learner SVM are improved to further improve the integration effect. During the integration process, an effective integration mechanism called WAUCE is proposed for the integration of the five heterogeneous base learners, and the posteriori probability obtained for each class is used to guide the post-hoc interpretation analysis.

Before we end this Editorial, we do wish to thank all the reviewers and the contributing authors, for their original ideas and solutions, and for choosing our Special Issue. Last but by no way least, we wish to wholeheartedly express our most sincere gratitude to everybody at the Journal of Biomedical and Health Informatics for their effective support in the management of this Special Issue.

SABAH MOHAMMED, *Guest Editor*
Lakehead University, Canada

TAI-HOON KIM, *Guest Editor*
Beijing Jiaotong University, China

RUAY-SHIUNG CHANG, *Guest Editor*
President, National Taipei University of Business Taipei, Taiwan

CARLOS RAMOS, *Guest Editor*
ISEP, Polytechnic of Porto

APPENDIX: RELATED ARTICLES

[A1] T. Chen *et al.*, "Discriminative cervical lesion detection in colposcopic images with global class activation and local bin excitation," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, Apr. 2022, doi: 10.1109/JBHI.2021.3100367.

[A2] Y. Zhu and M. Li, "Public health informatics: Proposing causal sequence of death using neural machine translation," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, Apr. 2022, doi: 10.1109/JBHI.2022.3163013.

[A3] I. Beheshti, M. A. Ganaie, V. Paliwal, A. Rastogi, I. Razzak, and M. Tanveer, "Predicting brain age using machine learning algorithms: A comprehensive evaluation," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, Apr. 2022, doi: 10.1109/JBHI.2021.3083187.

[A4] A. Gómez-Expósito, J. A. Rosendo-Macías, and M. A. González-Cagigal, "Monitoring and tracking the evolution of a viral epidemic through nonlinear Kalman filtering: Application to the Covid-19 case," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, Apr. 2022, doi: 10.1109/JBHI.2021.3063106.

[A5] M. Tanveer, A. H. Rashid, M. A. Ganaie, M. Reza, I. Razzak, and K.-L. Hua, "Classification of Alzheimer's disease using ensemble of deep neural networks trained through transfer learning," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, Apr. 2022, doi: 10.1109/JBHI.2021.3083274.

[A6] V. Subramaniyaswamy, R. Elakkiya, V. Vijayakumar, and A. Mahanti, "Cervical cancer diagnostics healthcare system using hybrid object detection adversarial networks," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, Apr. 2022, doi: 10.1109/JBHI.2021.3094311.

[A7] Y. Feng, X. Wang, and J. Zhang, "A heterogeneous ensemble learning method for neuroblastoma survival prediction," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, Apr. 2022, doi: 10.1109/JBHI.2021.3073056.