



Special Issue on Healthcare Knowledge Discovery and Management

Feichen Shen¹ · Vipin Chaudhary² · Yan Liu³ · Che Ngufor¹ · Hongfang Liu¹

Published online: 28 May 2019
© Springer Nature Switzerland AG 2019

1 Knowledge Discovery and Management in Healthcare Domain

Healthcare informatics research is a scientific endeavor that investigates the use of informatics and data science approaches for addressing healthcare needs in order to optimize both health service and patient care outcomes. A large number of healthcare data are collected and stored in various formats, including electronic health records (EHRs), clinical databases, biomedical knowledge bases/ontology, and medical images. With the wide adoption as well as the increasing volume of the aforementioned data, how to incorporate insights from both data-driven and knowledge-driven learning strategies to perform dynamic and intelligent learning from big heterogeneous healthcare data becomes an important research topic.

2 The Special Issue

This special issue provides a platform for novel findings from both data-driven and knowledge-driven perspectives and the latest investigation in leveraging informatics and data science techniques for healthcare knowledge discovery and management. In this special issue, a strict scientific review process was adopted, and the final five studies were selected from a total of 16 submissions.

The study done by Yeruva, Junaid, and Lee describes a solution to enhance the decision-making process for healthy eating and obesity prevention by combining contextual word embeddings and topic modeling approaches. The proposed framework ContWEB leveraged Word2Vec and Natural Language Processing techniques based on

✉ Feichen Shen
shen.feichen@mayo.edu

¹ Mayo Clinic, Rochester, MN, USA

² University at Buffalo, Buffalo, NY, USA

³ University of Southern California, Los Angeles, CA, USA

big data environment has shown its capability to explore the relationships between healthy dieting and obesity from Crowds' and Experts' perspectives.

Chandrababu and Bastola proposed an integrated approach to recognize potential protective effects of culinary herbs against chronic diseases. From a knowledge-driven perspective, they developed a graph-based database, HerbMicrobeDataBase (HMDB), to integrate knowledge from key biological entities associated in maintaining gut health and provide efficient storage/retrieval and graphical presentation of botanical, biochemical, and pharmacological data for culinary herbs and the human microbiome. They demonstrated the utility of the HMDB in understanding the molecular mechanism of metabolite production as well as their therapeutic or toxicological effects on gut microbes.

Moon, Liu, Chen et al. investigated on an interesting research topic on how to use natural language processing techniques to capture clinical concept from outside medical records and clinical notes to assist transferred patients at the point of care. Specifically, they conducted a comprehensive analysis on large amount of outside medical records, initial consultant notes, and follow-up consultant notes in Cardiovascular Medicine at Mayo Clinic. This study represented an initial step towards automated information extraction from outside medical records originating from diverse healthcare providers.

Due to the large availability and ambiguity of temporal information in clinical text, machines have difficulties in understanding and processing temporal information from electronic medical records automatically. Under this drive, Moharasan and Ho presented an approach to extract temporal information from clinical narratives. They divided tasks into three sub-tasks: temporal expression recognition, temporal events extraction, and temporal relation classification. To complete the first subtask, they proposed a novel feature set along with the adoption of the HeideITime system. For the second subtask, they developed a novel semi-supervised framework to extract temporal events from clinical narratives by extending the training data. For the third subtask, they adopted the Naïve Bayes classifier to generate candidate pairs.

Wyk, Williams, MacMillan et al. proposed a data-driven-based approach to conduct a cost-benefit analysis of automated physiological data acquisition system. In order to showcase the immediate benefit, they focused on the early detection of sepsis using random forest, recurrent neural network, and multilayer perceptron. This study showed potential utility to improve precision medicine in healthcare institutions.

The guest editors would like to appreciate all the authors, reviewers, editors, and the publisher for their contributions to this special issue.

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