Computer Science
Seminar

Federated Tensor Factorization for Collaborative Health Data Analytics

Jing Ma

Emory University

Abstract: Modern healthcare systems are collecting a huge volume of healthcare data from a large number of individuals with various medical procedures, medications, diagnosis, lab tests and so on. Tensor factorization has been demonstrated as an efficient approach for computational phenotyping, where massive electronic health records (EHRs) are converted to concise and meaningful clinical concepts. However, the EHR data is also fragmented and is always distributed among independent medical institutions, and they are prohibited from being shared and exchanged. Recently, federated learning offers a paradigm for collaborative learning among different entities, which seemingly provides an ideal potential to further enhance the tensor factorization-based collaborative phenotyping to handle sensitive personal health data. This poses challenges to preserving the privacy of the exchanged intermediary results in order to protect the sensitive patient information. Meanwhile, efforts still need to be made to overcome the limitations of the federated tensor factorization, including the restrictions to the classic tensor factorization, high communication cost and reduced accuracy. Furthermore, it is essential to develop the decentralization techniques for federated tensor factorization to deal with the vulnerability of the central server to malfunction and external attacks. To deal with these challenging problems, we propose 1) a privacy-preserving collaborative tensor factorization method for computational phenotyping which is able to deal with heterogeneous data with rigorous privacy guarantee and achieves less communication cost and comparable accuracy; 2) a communication efficient federated generalized tensor factorization, which is flexible enough to choose from a variate of losses to best suit different types of data in practice; 3) a communication efficient decentralized generalized tensor factorization method which enables the absence of the central server and further reduces the communication cost.

Tuesday, June 22, 2021, 2:00 pm

https://emory.zoom.us/j/8187241545

Computer Science
Emory University