

**Heng Huang**

Professor

Director, Data Science Lab

University of Texas at Arlington

Large-Scale Structured Sparse Learning with Applications in Imaging Genomics

Date: Tuesday, March 22, 2016**Time: 1:00 pm - 2:00 pm****Location: NVC Room 214****Abstract:**

Data science research is accelerating the translation of biological and biomedical data to advance the detection, diagnosis, treatment, and prevention of diseases, including the recently announced BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative and Precision Medicine Initiative. Sparsity is one of the intrinsic properties of real-world data, thus the sparse learning has recently emerged as a powerful tool to obtain models of high-dimensional data with high degree of interpretability at low computational cost, and provide great opportunities to analyze the big, complex, and diverse datasets. By enforcing properly designed structured sparsity, we can integrate the specific data/feature structures and domain knowledge into the machine learning models to simplify data models and discover predictive patterns in data analytics. To address the challenging problems in current big data mining, we proposed several novel large-scale structured sparse learning models for multi-dimensional data integration, heterogeneous multi-task learning, group/graph structured data analysis, and longitudinal feature learning. We applied our new structured sparse learning models to analyze the multi-modal neuroimaging and genome-wide array data in Imaging Genomics and discover the phenotypic and genotypic biomarkers to characterize the neurodegenerative process in the progression of Alzheimer's disease and other complex brain disorders. We also utilized our new machine learning models to analyze the Electronic Medical Records (EMR) for predicting the heart failure patients' readmission and drug side effects, detect the multi-dimensional biomarkers in The Cancer Genome Atlas (TCGA) for precision medicine, predict performance and guide design of nanoparticle synthesis in Materials Genome research, and identify the brain circuitry patterns in Human Connectome.

Bio:

Dr. Heng Huang is a Professor of Computer Science and Engineering (CSE) at University of Texas at Arlington (UTA), and the director of Data Science Lab in CSE at UTA. Dr. Huang received the PhD degree in Computer Science at Dartmouth College in 2006 and then joined UTA as an assistant professor. His research areas include big data mining, machine learning, health informatics, neuroinformatics, bioinformatics, computer vision, and computational sustainability. Dr. Huang has published more than 110 papers in top-tier conferences and many papers in premium journals, such as RECOMB, ISMB, NIPS, ICML, KDD, IJCAI, AAAI, CVPR, ICCV, SIGIR, Bioinformatics, IEEE Trans. on Medical Imaging, Medical Image Analysis, TKDE, etc. He is leading one NIH newly funded \$2 million R01 project on imaging genomics based complex brain disorder study, multiple NSF funded projects on big data mining, machine learning, imaging genomics, electronic medical record data mining and privacy-preserving, computational biology, smart healthcare, cyber physical system, and also industry (e.g. Con Edison in New York City) funded projects on computational sustainability, smart metering, and smart grid.