

Deep Learning for Cancer Prognosis Prediction

Andy Te-Cheng Hsu

Cancer is one of the world's leading mortality causes. Accurate prognosis predictions in the early stages are essential for doctors to provide appropriate treatments. Genetic and clinical data include rich information, and several cancer biomarkers have been identified for predicting cancer prognosis. Recently, powerful deep learning models are established to learn from large scale data. However, fundamental challenges when adopting deep learning models for biological data lie in certain aspects such as curse of dimensionality, small sample size, and missing values. Therefore, robust biomarker selection algorithms along with powerful deep learning model designs for cancer prognosis prediction are of crucial importance. The talk walks through our recently published works on biomarker selection and cancer prognosis prediction. In the first part of the talk, fundamental challenges and deep learning basics are reviewed. Several gene biomarker selection strategies and deep learning prediction models combining heterogeneous patient data are also discussed. The second part of the talk focuses on issues resulting from small sample size and missing labels. Advanced generative semi-supervised learning can either generate synthetic data to help model training or utilize large censored unlabeled patient data. With appropriately designed dimension reduction approaches and deep learning models, we can extract meaningful biological insights from high dimensional data such that accurate risk predictions can be made.