FAU

DEPARTMENT OF Mathematical Sciences presents

Yan Zhang

Monday, February 17, 2020
11:00-12:00 SE 215
Refreshments provided

Functional annotation of genomic elements using deep learning techniques

Functional annotation of the genomic elements, including coding and non-coding regions, is critical for revealing biological regulations and mechanisms under disease phenotypes. Since experimental functional characterization is usually time-consuming and costly, accurate and efficient computational methods for function prediction are in high demand for generating testable hypotheses guiding large-scale experiments. In this talk, I will mainly present Graph2GO, our newly developed algorithm for annotating coding genes, in terms of predicting the functions of their protein products. Graph2GO is a graph representation learning architecture to predict protein functions by considering multiple sources of information, including protein sequence similarity, protein domains, subcellular location, and functional protein association networks. We utilize variational graph auto-encoder, a deep learning-based network representation learning algorithm, to generate informative embeddings for each protein. These embeddings are input as features for a multi-class multi-label neural network classifier to predict protein functions. Node attributes are propagated between nodes in the graph to augment the information of each protein, especially for proteins lacking enough annotations. Graph2GO achieves state-of-the-art performance on the benchmark dataset. Furthermore, I will talk about our other functional annotation tools for non-coding regions and our future research directions in deep learning applications.