KIMON FOUNTOULAKIS, Cheriton School of Computer Science, University of Waterloo
Graph Attention Retrospective

Graph-based learning is a rapidly growing sub-field of machine learning with applications in social networks and bioinformatics. One of the most popular models is graph attention networks. They were introduced to allow a node to aggregate features of neighbour nodes in a non-uniform way, in contrast to simple graph convolution which does not distinguish the neighbours of a node. In this presentation I will discuss multiple results on the performance of graph attention for the problem of node classification for a contextual stochastic block model. The node features are obtained from a mixture of Gaussians and the edges from a stochastic block model. I will show that in an "easy" regime, where the distance between the means of the Gaussians is large enough, graph attention is able to distinguish inter-class from intra-class edges. Thus it maintains the weights of important edges and significantly reduces the weights of unimportant edges. Consequently, I will show that this implies perfect node classification. In the "hard" regime, I will show that every attention mechanism fails to distinguish intra-class from inter-class edges. In addition, I will show that graph attention convolution cannot (almost) perfectly classify the nodes even if intra-class edges could be separated from inter-class edges. Beyond perfect node classification, I will discuss a positive result on graph attention's robustness against structural noise in the graph. In particular, the robustness result implies that graph attention can be strictly better than both the simple graph convolution and the best linear classifier of node features.