Pruning Neural Networks with Classification Activation Paths

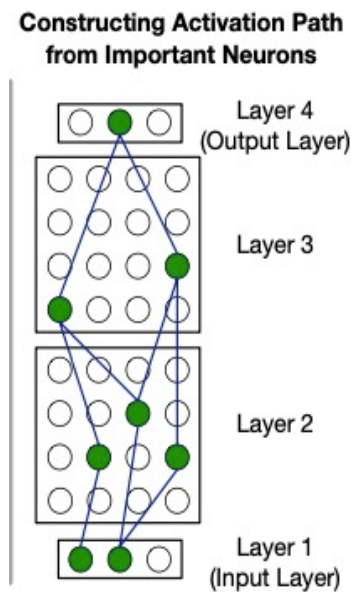**Constructing Activation Path from Important Neurons**

Figure1: The activation path constructed from important neurons.

In recent years, deep neural networks (DNNs) have achieved remarkable breakthroughs. However, there are a huge number of multiply-accumulate operations in DNNs, which restricts their applications in resource-constrained platforms, e.g., mobile phones. To reduce the computation complexity of neural networks, various pruning methods have been developed to reduce the size of models while minimizing loss in accuracy or performance. In the existing neural network pruning methods, weights whose absolute values are small are removed from a trained model. However, these methods do not consider the relationship between different weights.

In this master thesis, a novel pruning method considering classification activation paths will be explored. The basic idea is shown in Figure 1. When a neural network recognizes an image of a dog, this image triggers a special activation path in the neural network. Such activation paths for various classes will be used to prune neural networks in this master thesis.

If you are interested in this topic for master thesis, please contact:
Dr.-Ing. Li Zhang (grace-li.zhang@tum.de) with your CV and bachelor and master transcripts.