Now I’d like to share some experiment and conclusion.

First we compared tree in tree with cart on some datasets. Cart is the normal decision tree.

In this experiment, we ensure that tnt and cart has the same amount of internal nodes. We can see that in general, tree in tree outperforms cart, but the execution time of tnt is much slower than cart. This is because the time complexity of tnt is big and also the implementation is used for demonstration, not focused on efficiency.

And we use dot language to draw some images of tnt. We find some interesting things. First, the redundant node. From the image, we find this kind of nodes. It’s obviously redundant, so maybe we can make some judges and merge them. Also, when the number of node increasing, some parts of tnt is similar to neural network. We can see that the two layers of nodes make pairwise connection. We guess tree in tree might be the Intermediate between decision tree and neural network.

Now it comes to conclusion. The core idea of this paper is that every node in the decision tree is also a decision tree. Therefore, it forms a graph. This idea can be easily used in other decision tree frameworks. And it outperforms cart. However, the limitation is that tnt is really time-consuming.