

## 2.6 Method of pruning.

As it was already mentioned, the elementary method of tree construction described above, can give a very “optimistic” decision. In this method, the estimation of the tree quality will be carried out on the same training sample on which the decision tree is formed.

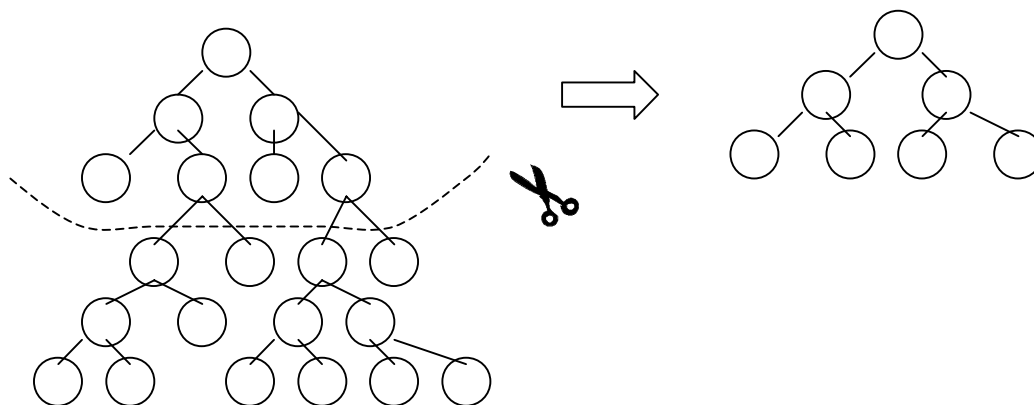
For more objective estimation and elimination of random laws, it is necessary to use a sample which did not participate in tree construction.

In a pruning method, the training sample is divided in two parts. The first part is used for tree construction by a method of consecutive branching. Parameters of a stop are set in such way to provide the maximal possible accuracy of the received decision. The number of leaves of the tree can be very large.

The second part of the sample serves for pruning (“simplification”) of the received tree. For this purpose, the following steps are carried out.

- 1) All internal nodes of the tree are considered one by one.
- 2) Operation of pruning for considered node is carried out.
- 3) Using the second part of the sample, we estimate the prediction error for the truncated variant of the tree.

The variant with minimal error is a result (figure 13).



**Fig. 13**

The described method gives more objective estimation of quality, however, if the initial tree is far from optimum, then the truncated variant will not be ideal also.