Supplement—Vinter et al. A role for machine learning in sex-specific prediction of successful electrical cardioversion in atrial fibrillation?

Supplement

Random forest analysis is a branch of machine learning that is based on decision or classification trees. A decision tree corresponds to a flowchart aiming to predict a characteristic or an event for an observation (patient) through a system of questions or classification criteria splitting the dataset into a number of subsets with different prediction values (Supplemental Figure 1).

Supplemental Figure 1. Example of a decision tree (hypothetical).
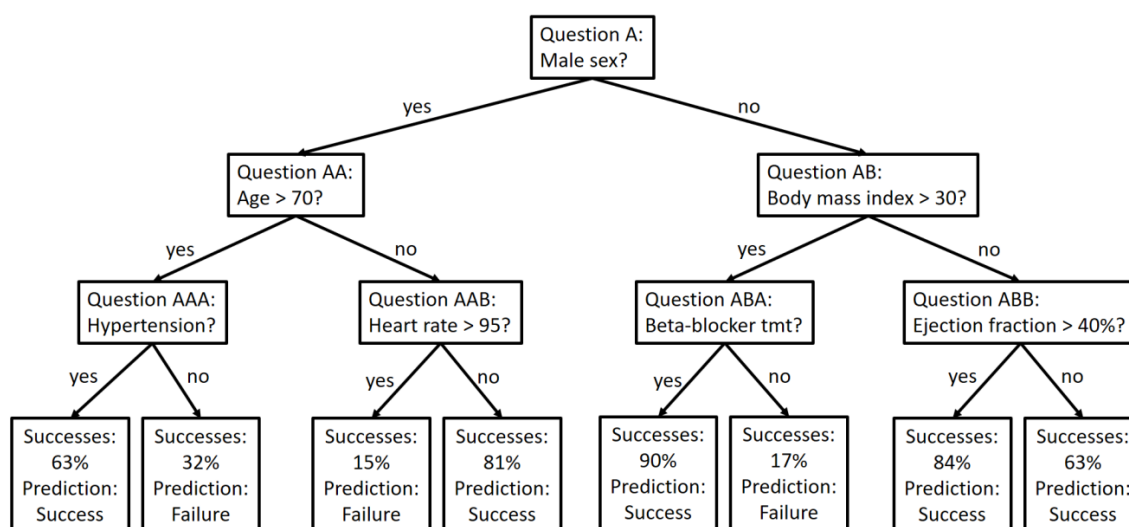
The answer to each question is allowed to determine the choice of subsequent questions, which implies that the approach has a high capability of handling complex, high order interactions between the studied characteristics/criteria. The process of assessing the most informative questions to ask in each step (i.e. the best data split) is referred to as training the tree.

In the random forest approach, a series of such decision trees are constructed on the basis of different random samples of the full dataset, and the final prediction is based on an aggregation of all these trees (Supplemental Figure 2).

Supplement—Vinter et al. A role for machine learning in sex-specific prediction of successful electrical cardioversion in atrial fibrillation?

Supplemental Figure 2. Random forest concept.