Random Forests: a description

A random forest R_f is the set of functions $T_1, ..., T_N$ where each T_j is a piece-wise function from the sample space Ω to the response space Φ. In general, Ω is defined by an n x p matrix where each column is a random variable and Φ is defined by an n x 1 vector Y.

Each tree T_j is generated on a subset of both Ω and Φ called the training set. It is then tested on a disjoint subset of Ω called the test set and the values of T_j on this test set are called the predictions.