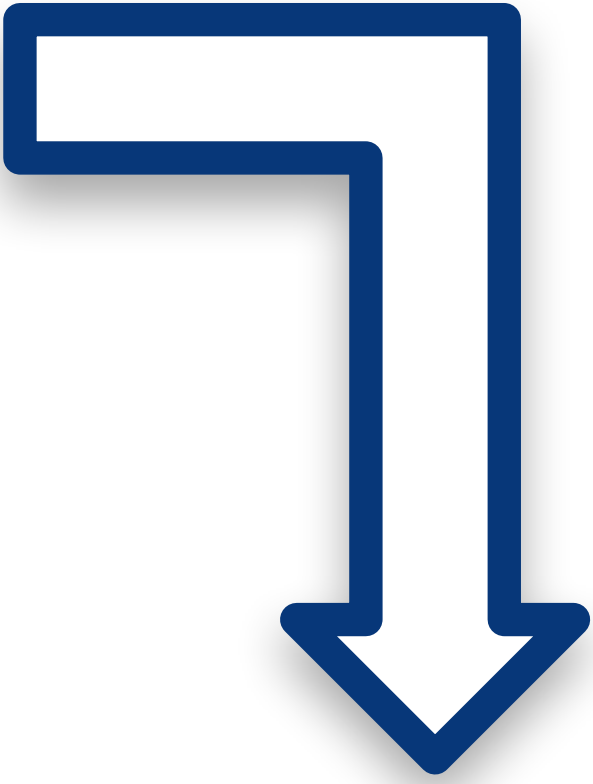
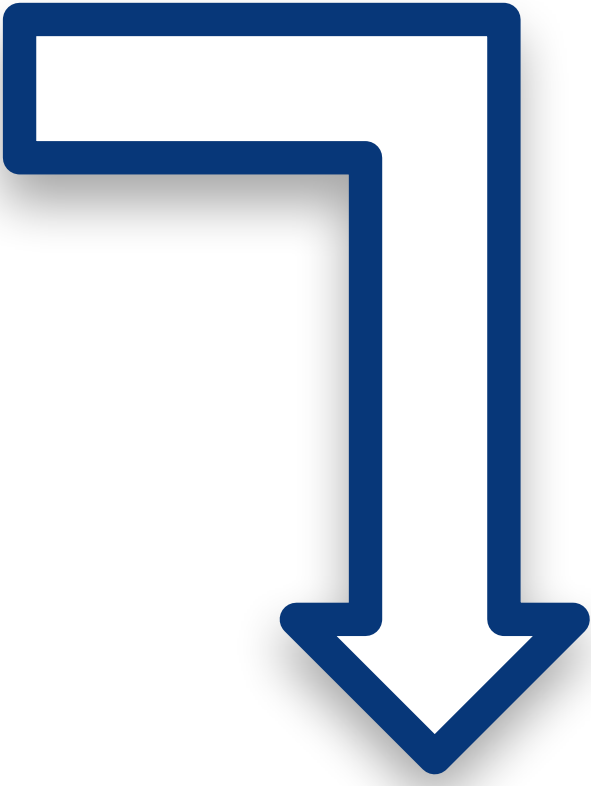
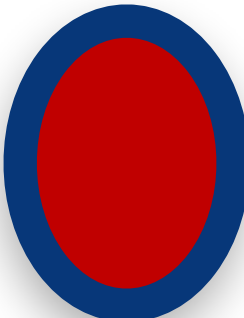
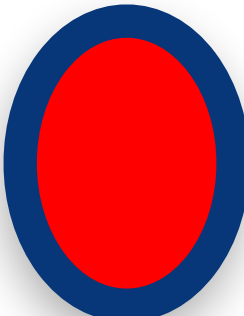

H2O.ai

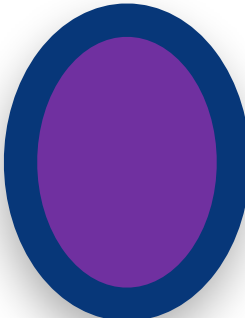


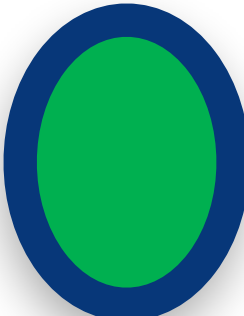


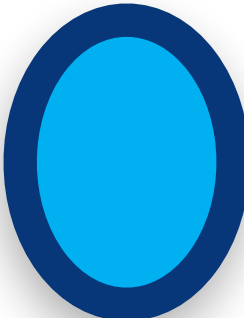
RF Tree Variance: Bagging

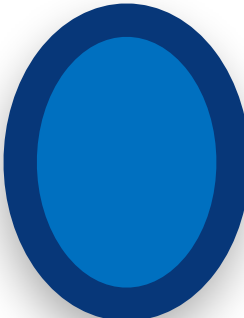


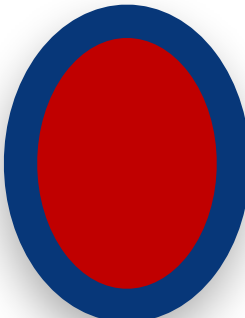


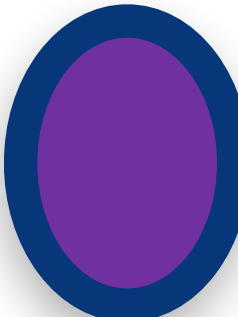


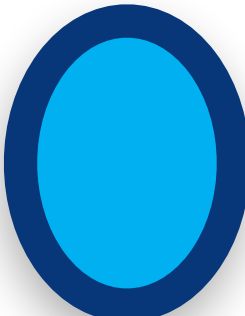


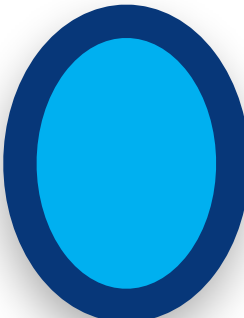


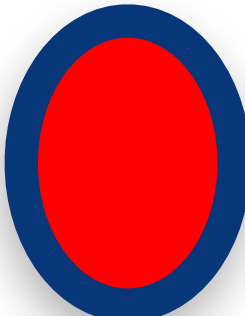


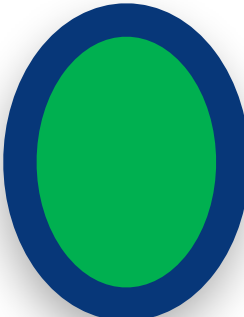


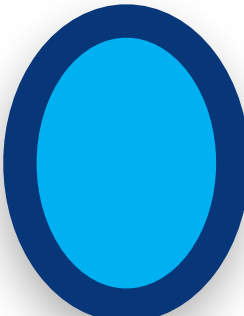


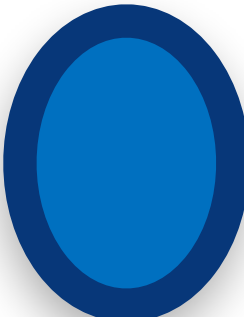


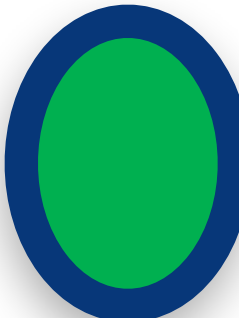


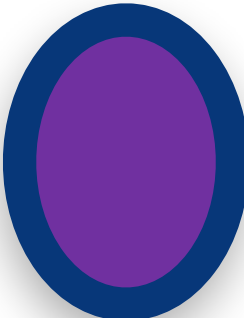


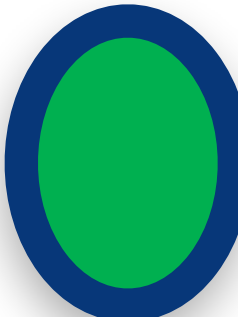


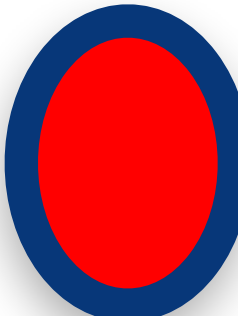


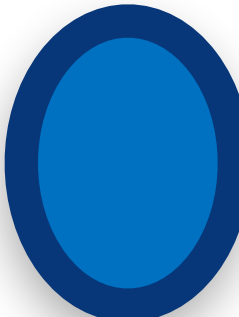


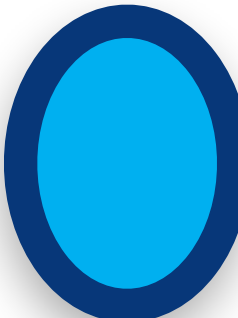


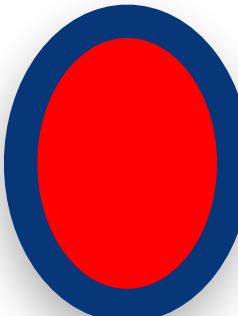


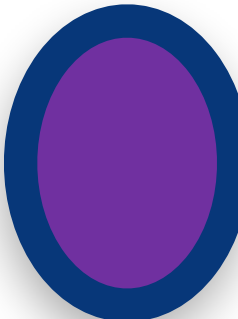


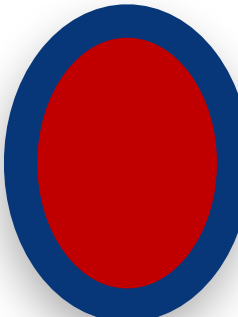


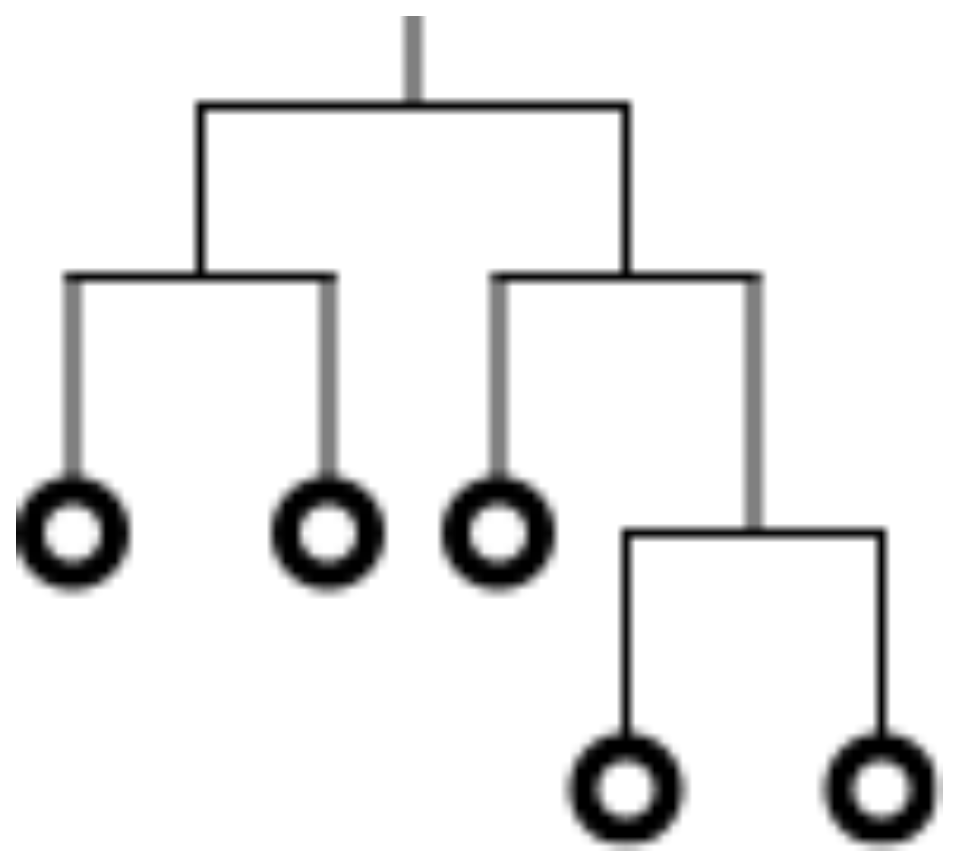


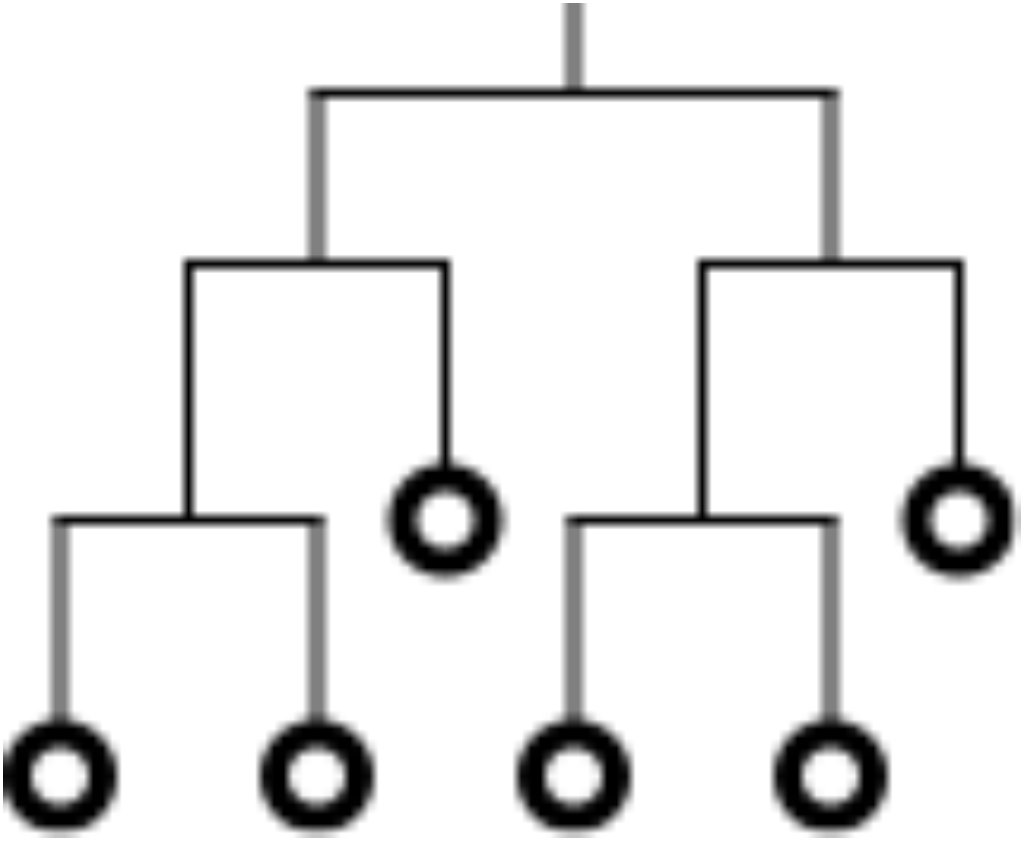


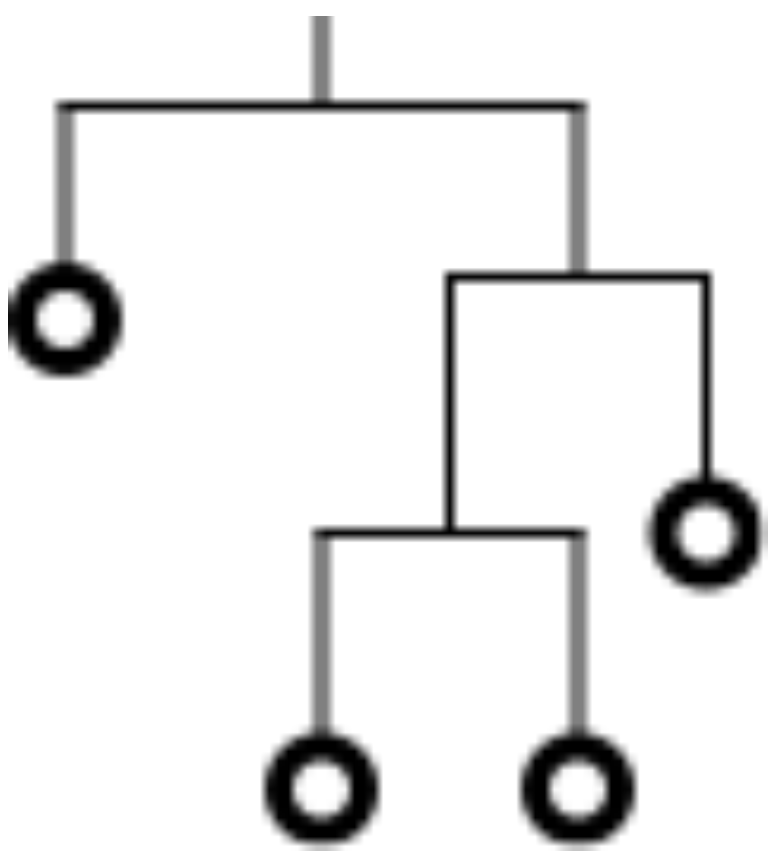












1) Bootstrap Aggregating ('bagging')

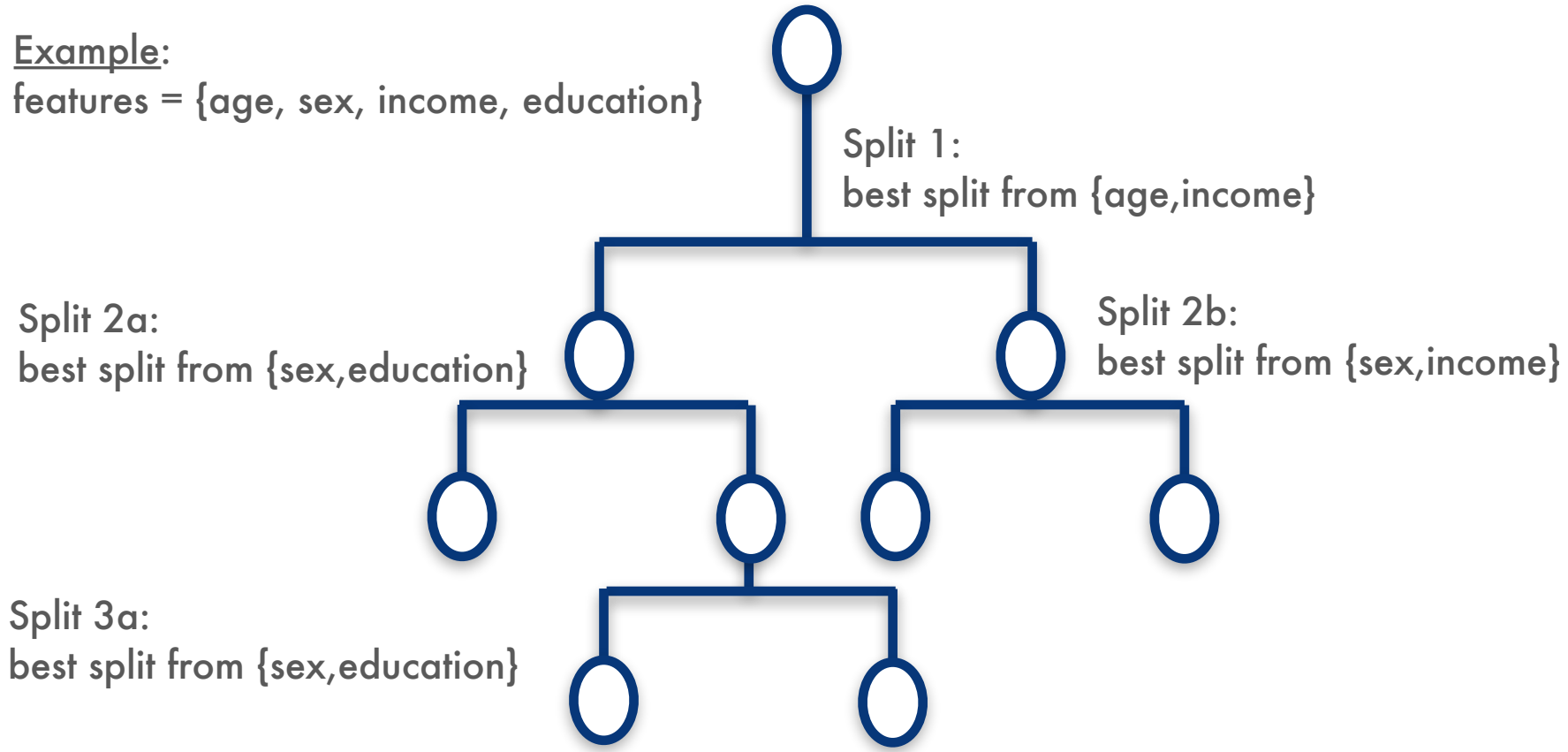
Sample n, with replacement

$$f_M(\mathbf{x}) = \sum_{i=1}^{M} T_i(\mathbf{x})$$

RF Tree Variance: Splits

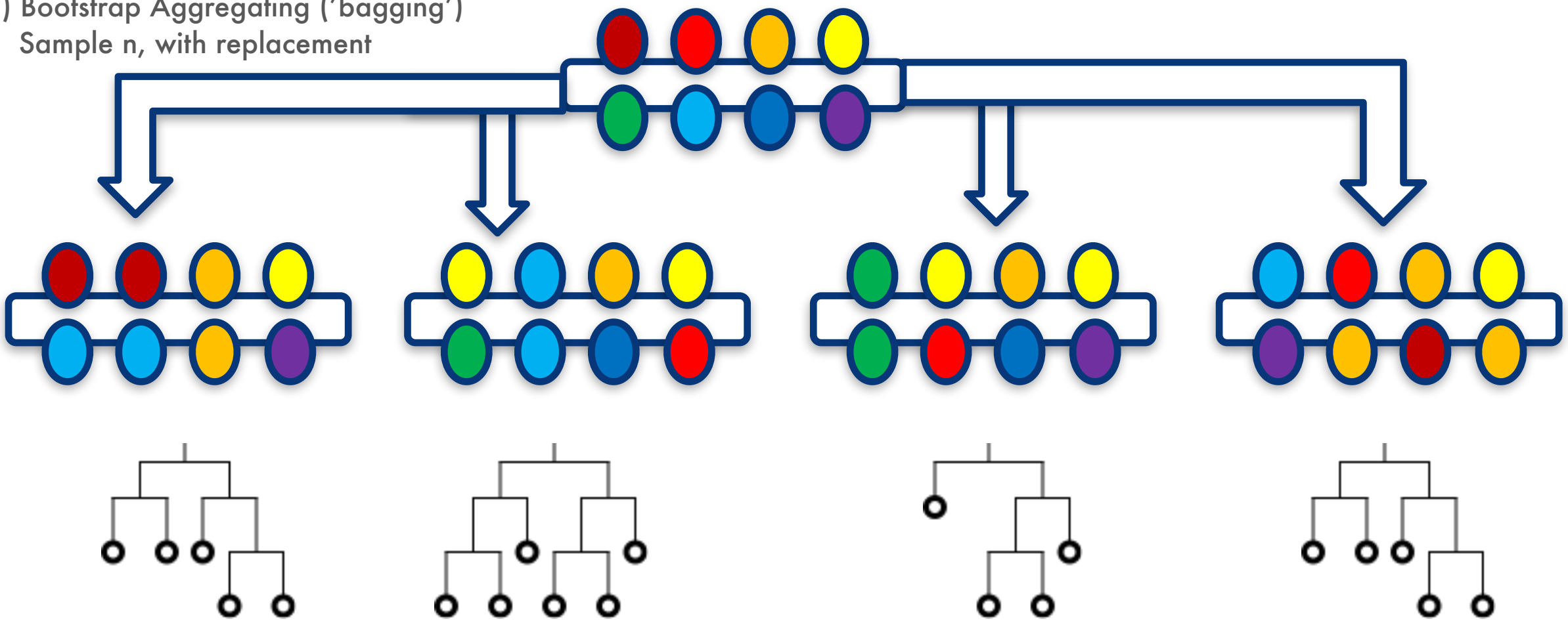
2) Do not consider all of the features for each split

- default is often $\sqrt{n}$
- also speeds up computational time



RF Tree Variance: Bagging

1) Bootstrap Aggregating ('bagging')
   Sample n, with replacement



$$f_M(\mathbf{x}) = \sum_{i=1}^{M} T_i(\mathbf{x})$$