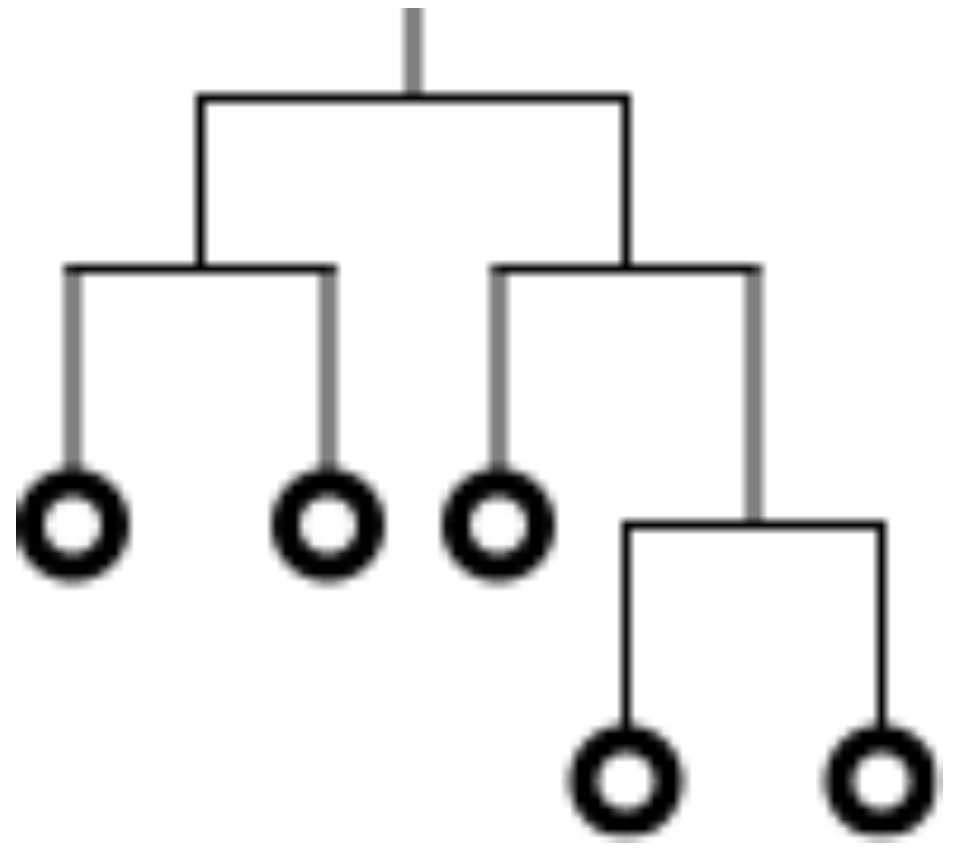
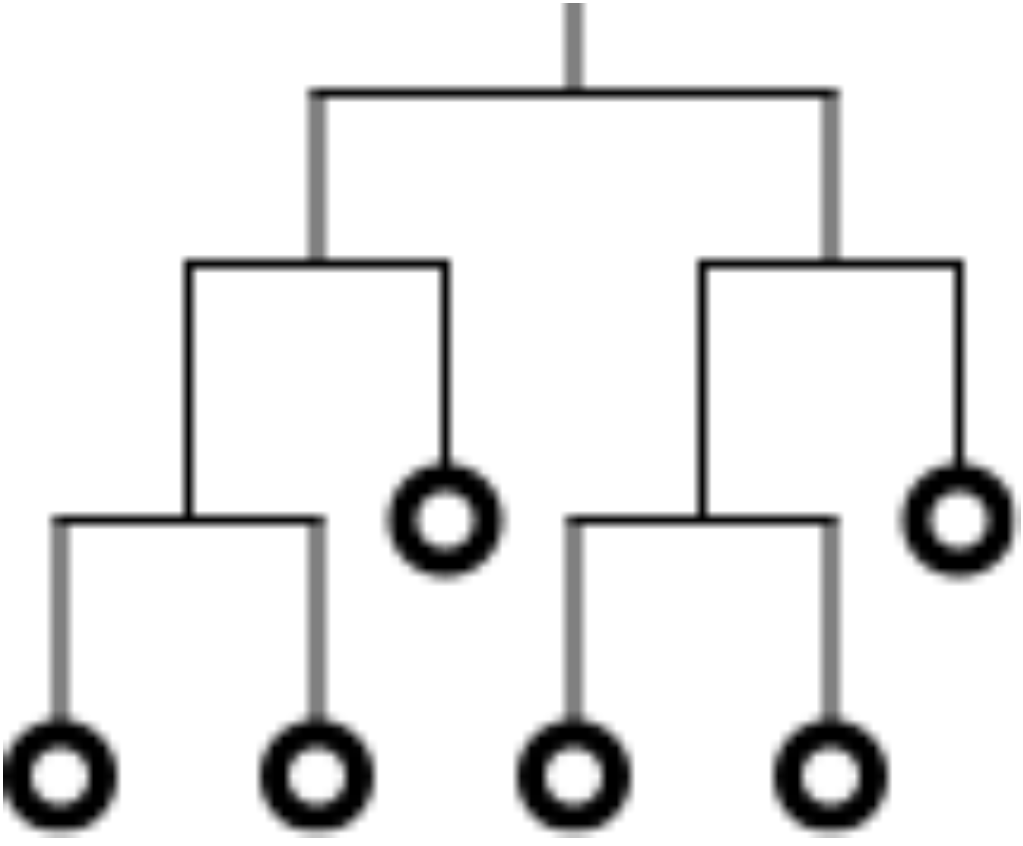
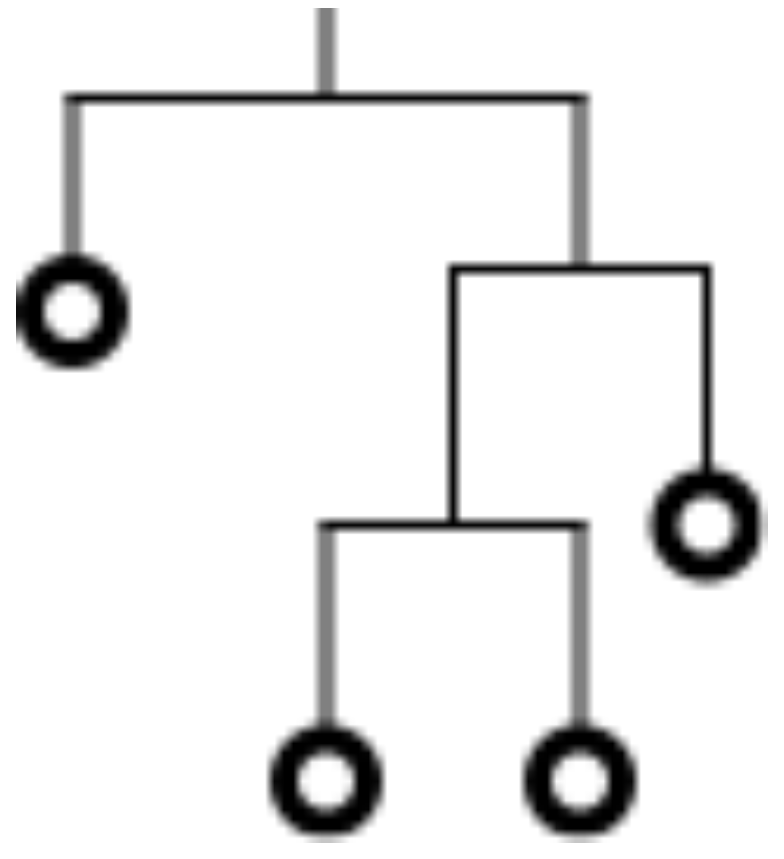
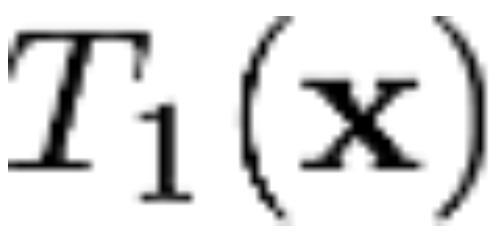
H2O.ai

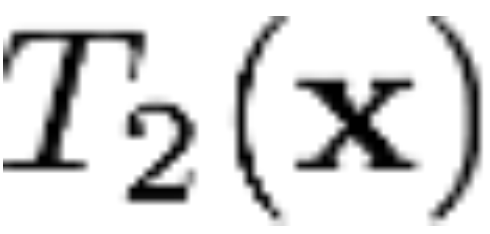
Ensemble of Trees

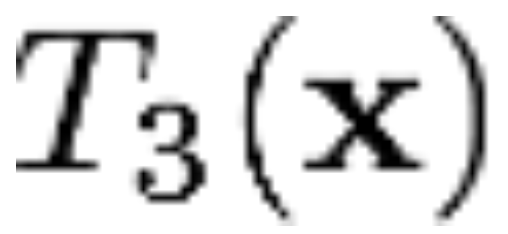


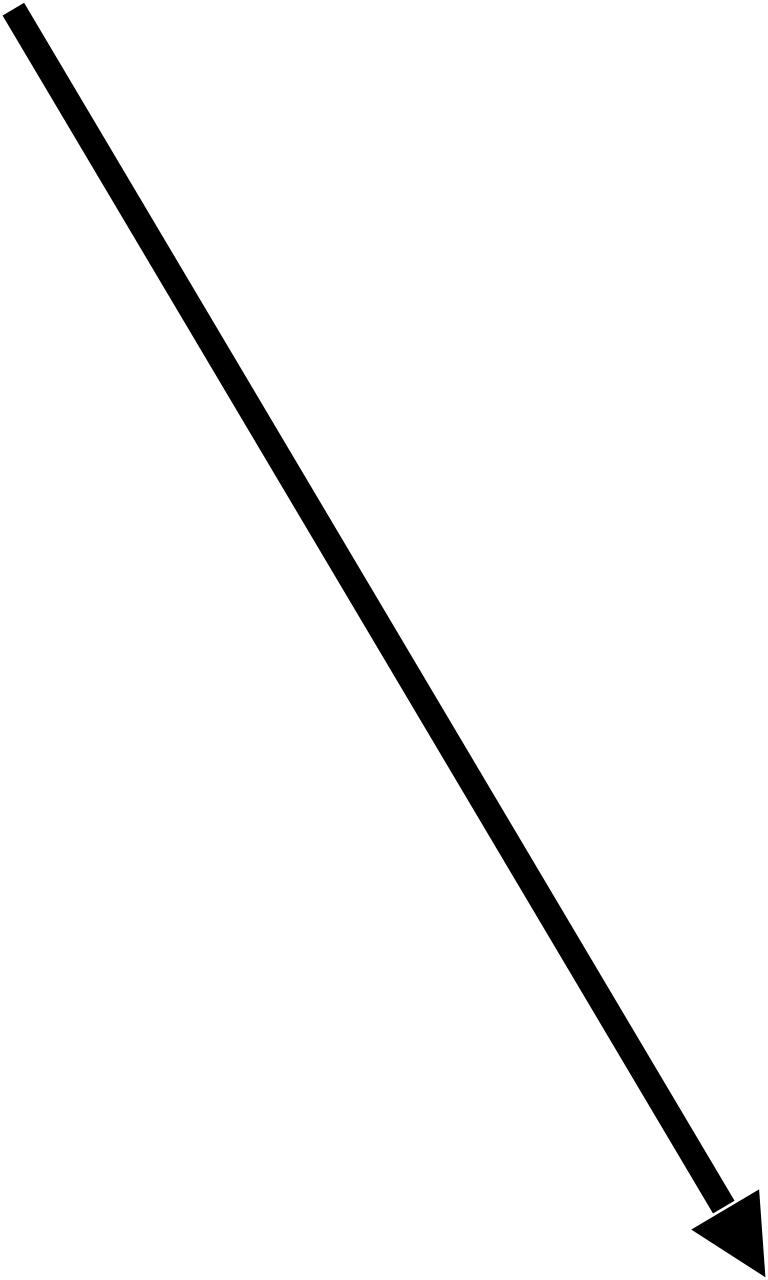


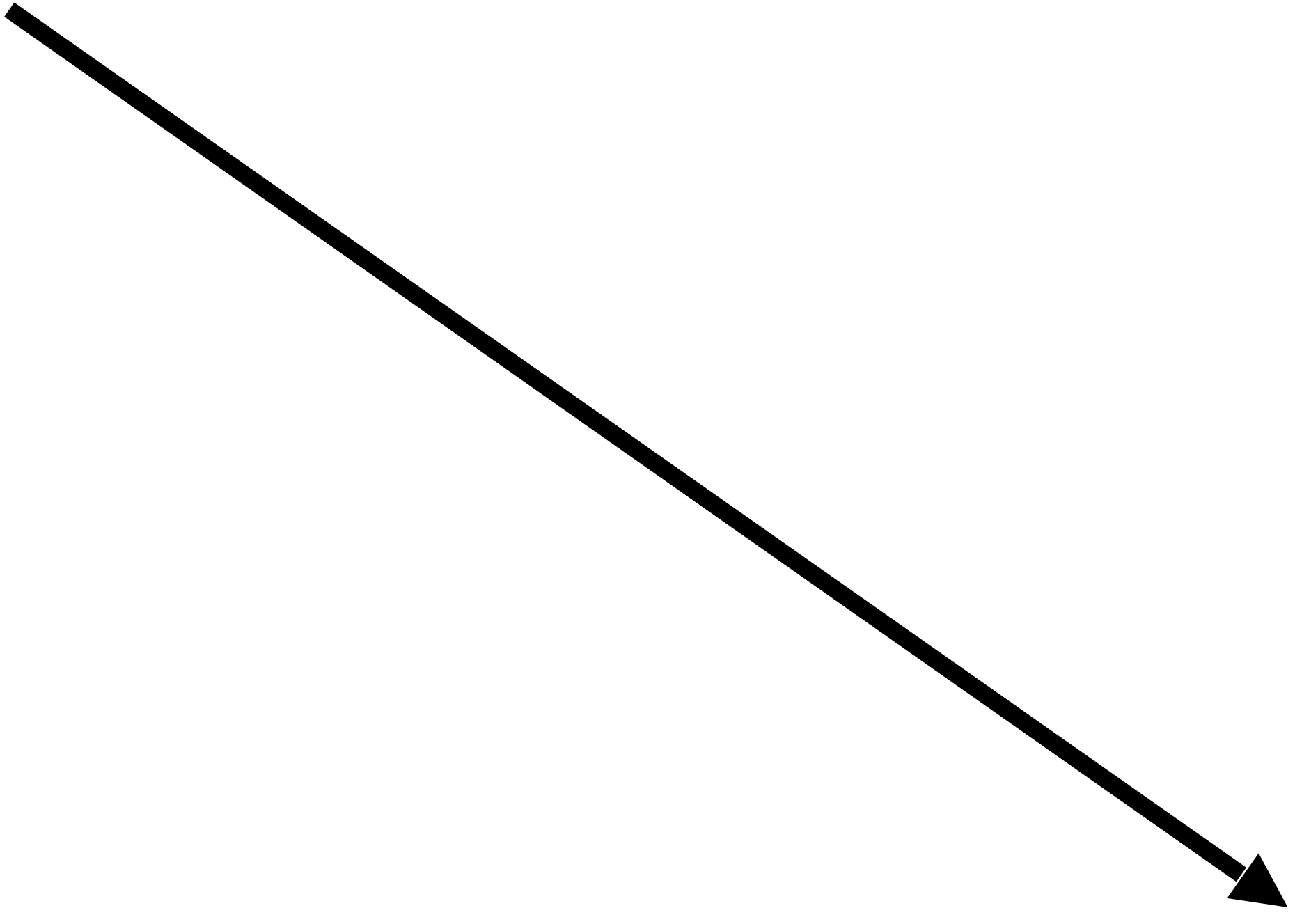


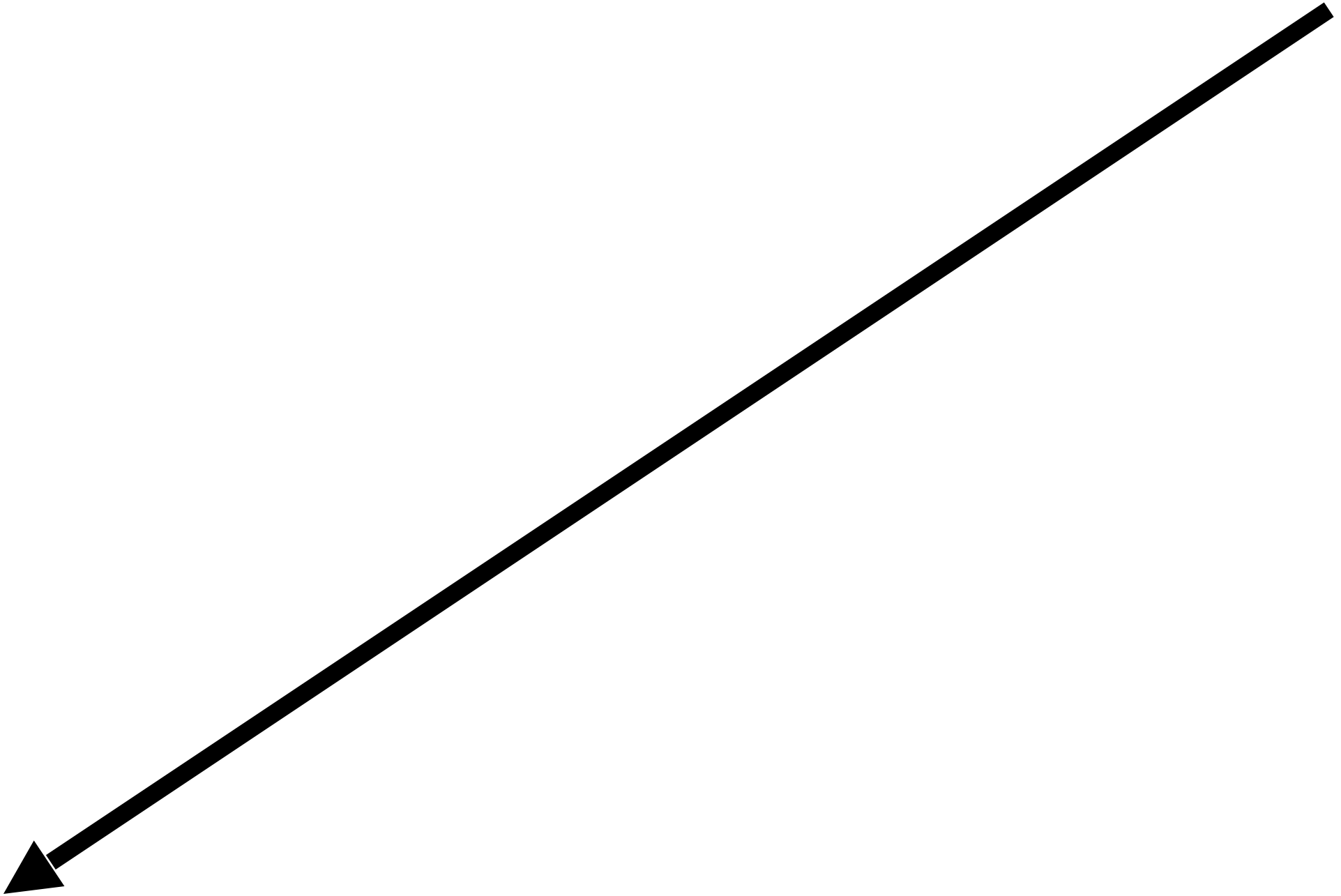


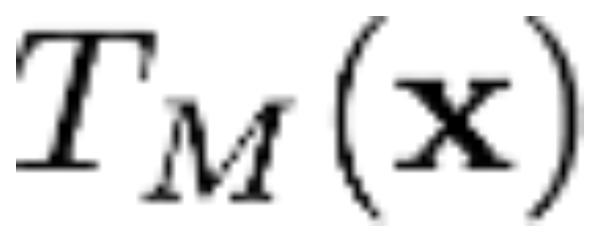












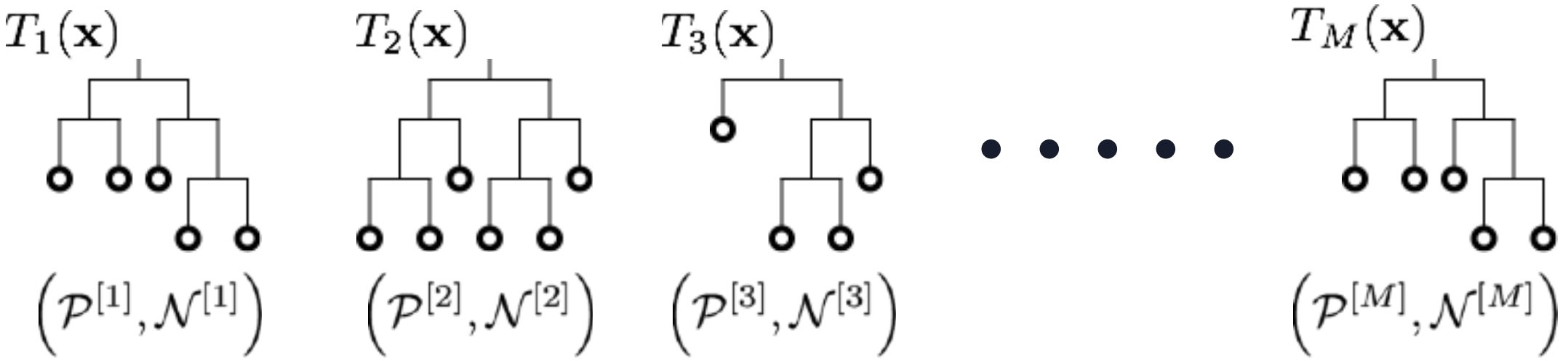
$$\hat{y} = f(\mathbf{x}; T_1, \ldots, T_M)$$

Regularization: Observation and Feature Sampling

- For each tree, select from *P* features (columns) and *N* rows

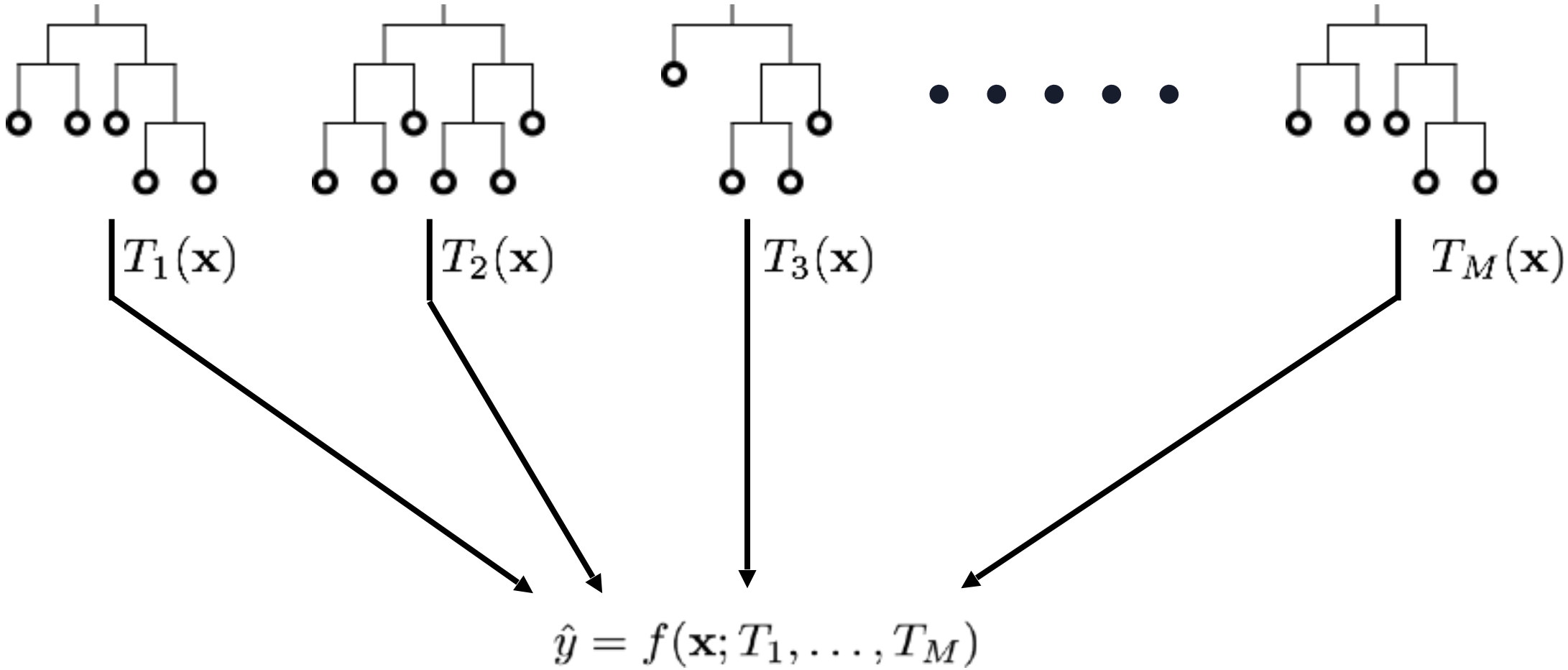
$$\mathcal{P} \in \{1, 2, \ldots, P\} \qquad \mathcal{N} \in \{1, 2, \ldots, N\}$$

$$\mathcal{P}^{[i]} \subseteq \mathcal{P},\ |\mathcal{P}^{[i]}| = P^{[i]} = \gamma_{\text{COL}} P \qquad \mathcal{N}^{[i]} \subseteq \mathcal{N},\ |\mathcal{N}^{[i]}| = N^{[i]} = \gamma_{\text{ROW}} N$$



H2O.ai

Ensemble of Trees



H2O.ai