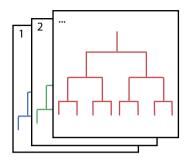
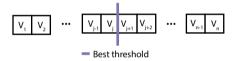
CHASMplus (Random Forest)



Decision tree split



Split criteria

$$\begin{split} i \in I &= \text{ random subset of } \sqrt{p} \text{ features} \\ \mathbf{j} &= \text{ index for values v of feature i} \\ Q &= \mathbbm{1}_{V^{(i)} \leq v_j^{(i)}} \\ p_k &= Pr(driver|Q = k) \end{split}$$

$$g(i,j) = \sum_{k=0}^{1} Pr(Q=k) * 2p_k(1-p_k)$$

$$split = \operatorname*{arg\,min}_{(i,j)} g(i,j)$$

