## **Original model**

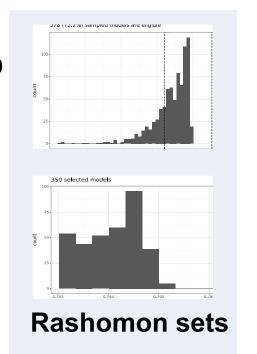
(Fairness-unaware counterpart model)



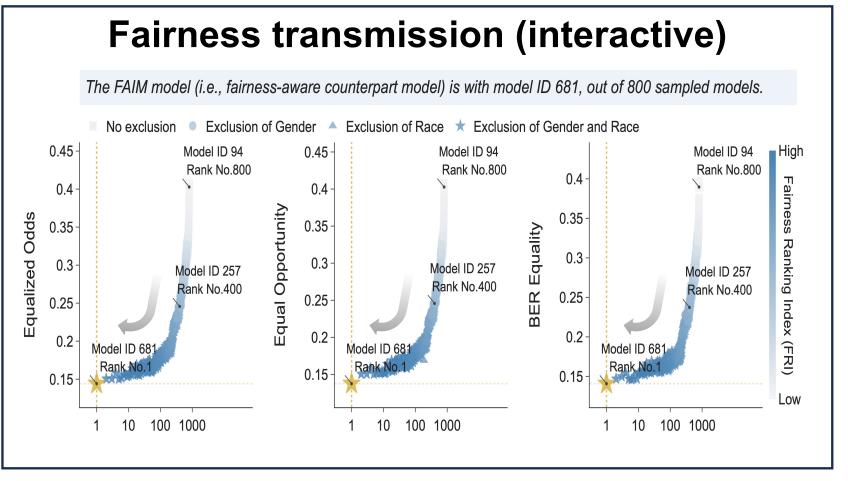
## **Nearly-optimal model generation**

Toggling sensitive variables  $\{s_i\}_{i=0}^k$ 

- No exclusion
- Individual exclusion of  $s_0, ..., s_k$
- Pair exclusion of  $\{s_0, s_1\}$ ,  $\{s_0, s_2\}$ , ...,  $\{s_{k-1}, s_k\}$
- Exclusion of  $\{s_i\}_{i=0}^k$

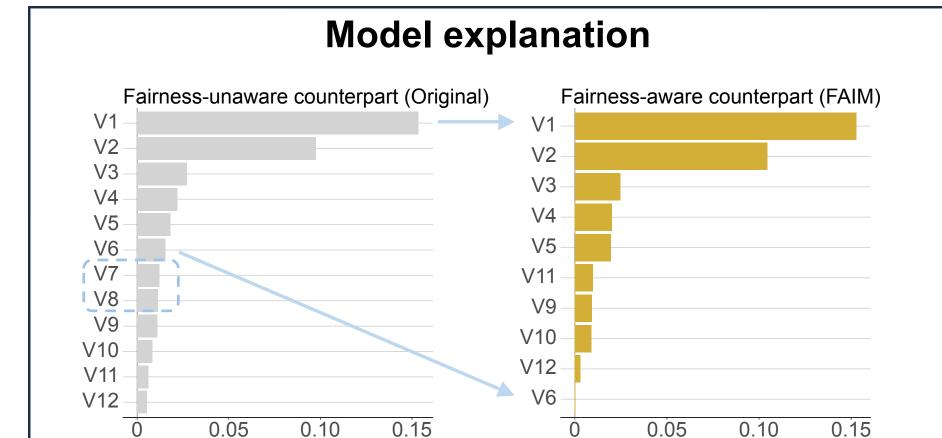








FAIM output model
(Fairness-aware counterpart model)



 Decisions of inclusion/exclusion of sensitive variables (e.g., V7 and V8)

Feature importance

Changes in variable importance (e.g., V6)

Feature importance

Consistency in variable importance (e.g., V1)