Which Neighbours Affected House Prices in the '90s?

Hubert Baniecki, Mateusz Polakowski

DATA

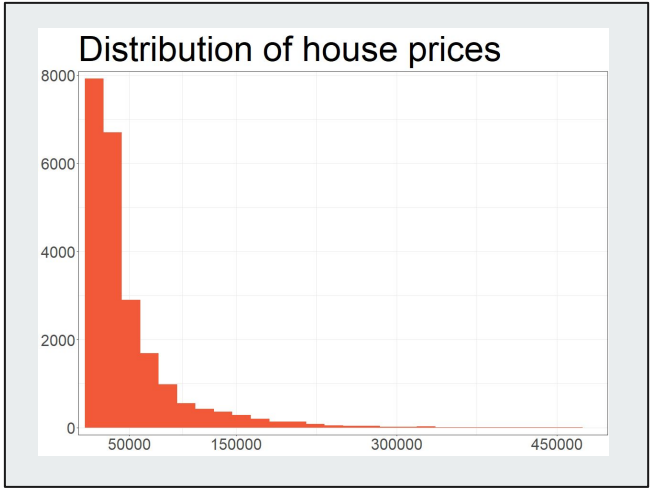
US CENSUS 1990

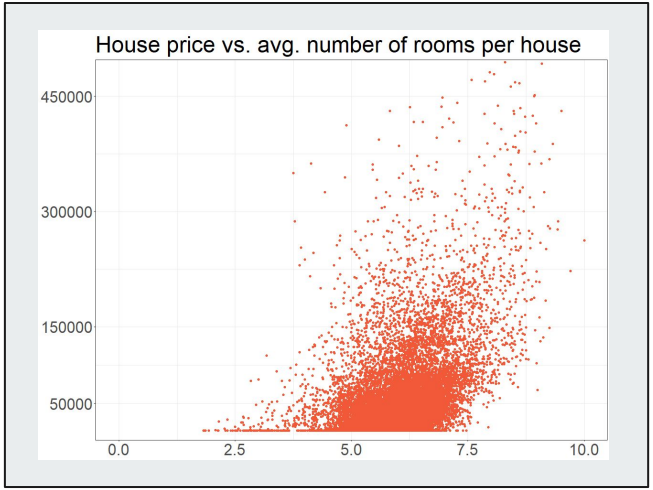
source: University of Toronto, Canada

2 datasets for comparison have been used.

Exemplary variables:

- total number of households
- percentage of people between 25-64 y.o.
- percentage of black householders
- average family size





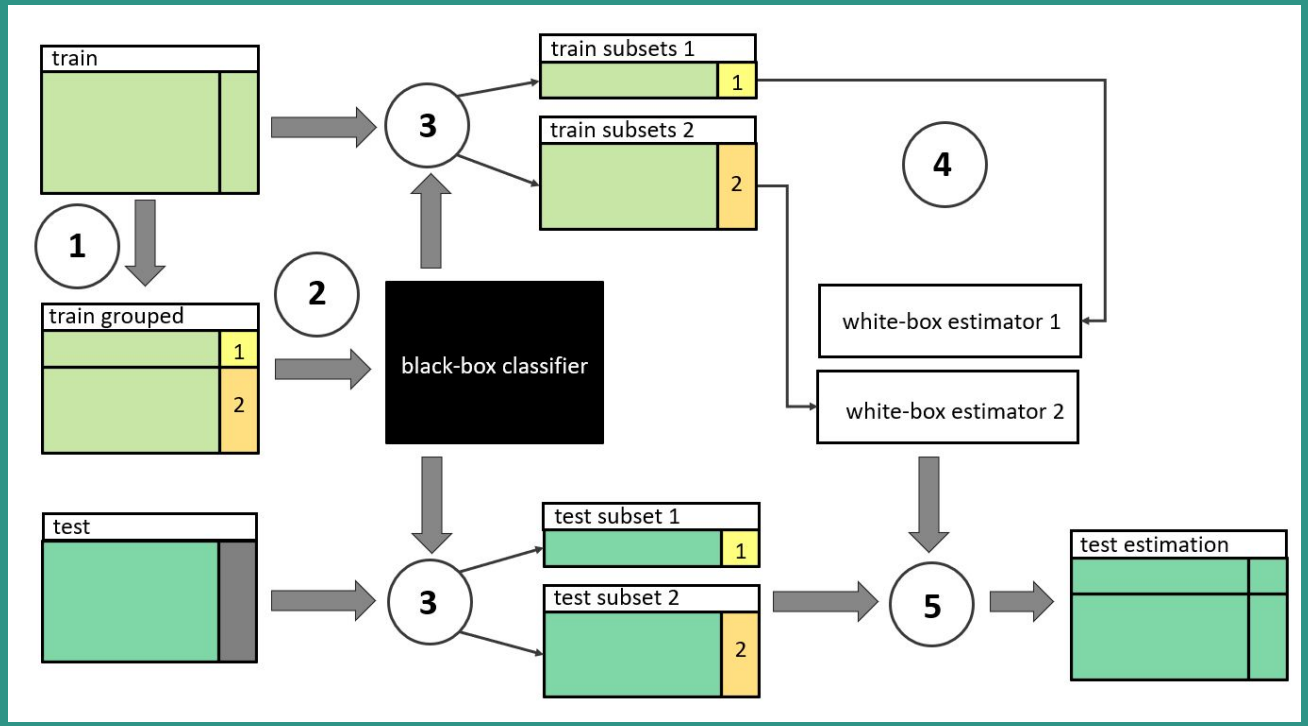
WHITE-BOX

- rpart (decision tree)
- linear model
- SAFE on ranger

BLACK-BOX

- ranger (random forest)
- xgboost

STACKED MODEL

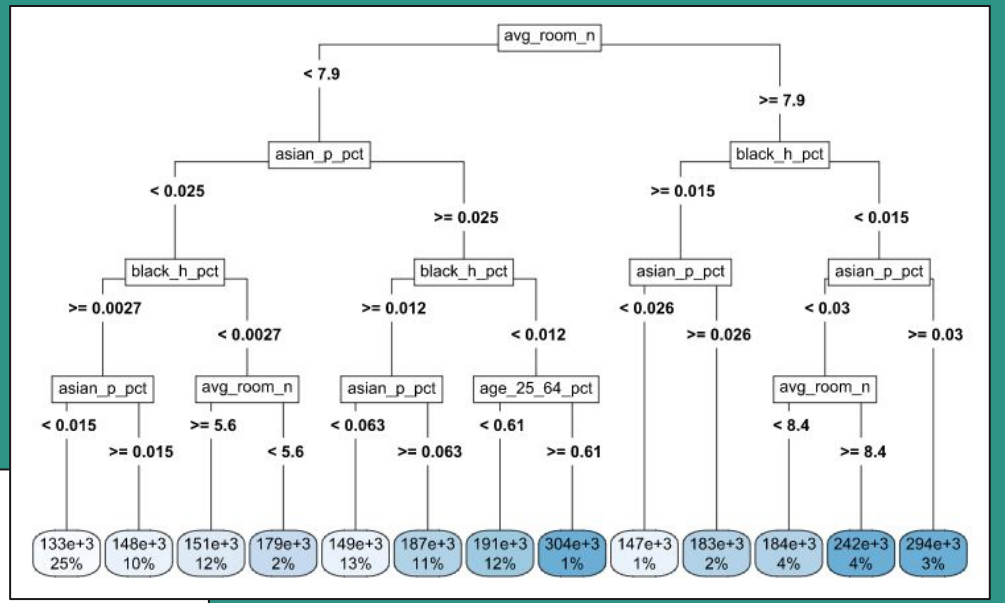


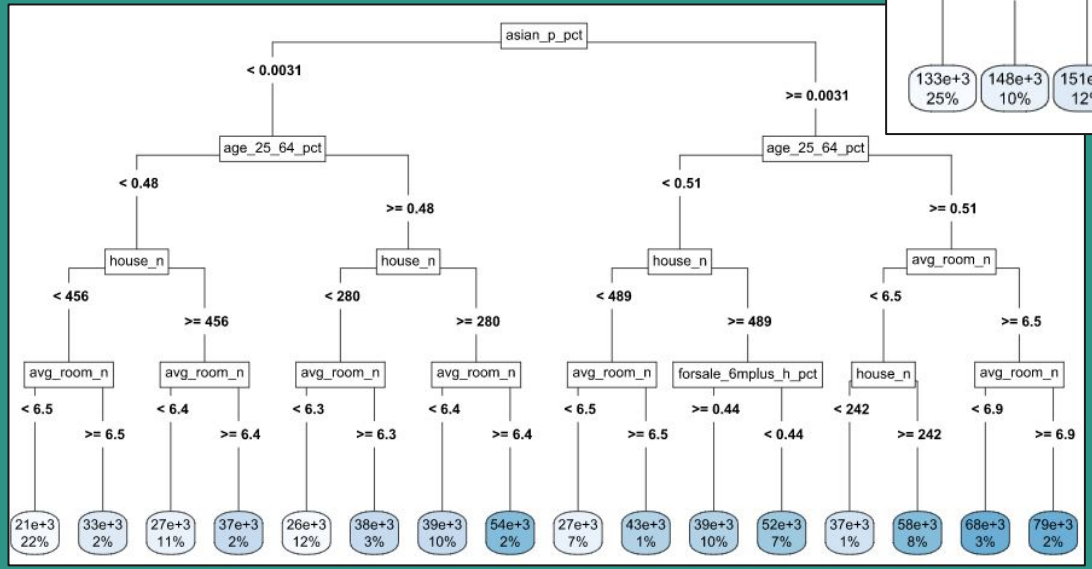
RESULTS

Table 2. Comparison of the MAE score for all of the used models on test datasets.

Model	Dataset (test)	
	house_8L	house_16H
linear model	23.1k	24.1k
SAFE on ranger	21.4k	22.6k
rpart	19.2k	22.1k
xgboost	16k	16.5k
ranger	14.8k	15.6k
stacked model	14.6k	15.3k

INTERPRETABLE MODEL





References:

- https://mini-pw.github.io/2020L-WB-Book/interpretability.html#which-neighbours-affected-house-prices-in-the-90s
- **ranger** https://www.jstatsoft.org/article/view/v077i01
- **rpart** https://github.com/bethatkinson/rpart
- **SAFE** https://arxiv.org/abs/1902.11035