

Car Evaluation Prediction — Decision Tree with LIME-style Explanations

Auto-generated report.

1) Dataset

Rows: 1727 Columns: 7

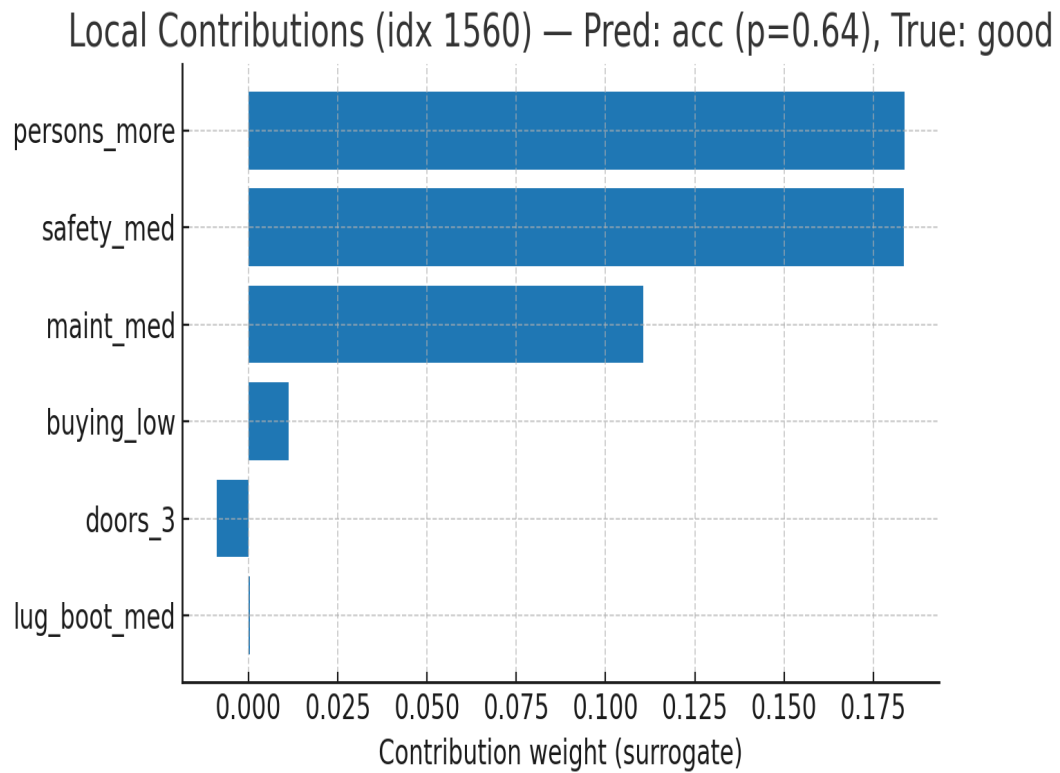
- buying: categorical
- maint: categorical
- doors: categorical
- persons: categorical
- lug_boot: categorical
- safety: categorical
- Target: class

2) Model & Training

We trained a DecisionTreeClassifier inside a scikit-learn Pipeline with a OneHotEncoder for all categorical features. Depth was capped to 5 for interpretability. The model was evaluated on a 25% holdout set with stratified sampling.

3) Evaluation

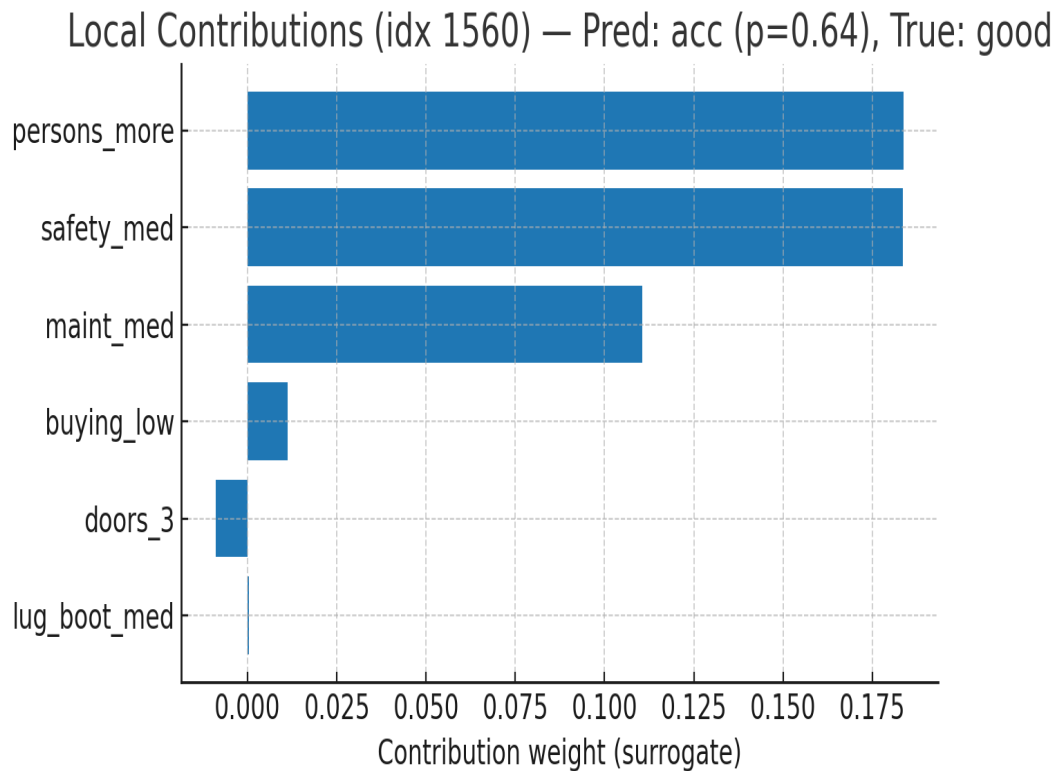
Confusion Matrix



Classification Report

4) Global Encoded Feature Importance

The tree's feature importances are computed in the one-hot encoded space. Top 15 shown below:



Encoded Feature	Importance
safety_low	0.3776
persons_2	0.2555
maint_low	0.1259
maint_vhigh	0.0810
buying_vhigh	0.0645
buying_high	0.0443
maint_med	0.0440
safety_high	0.0071
persons_4	0.0000
lug_boot_small	0.0000
lug_boot_med	0.0000
lug_boot_big	0.0000
persons_more	0.0000
doors_4	0.0000
doors_5more	0.0000

5) LIME-style Local Explanations

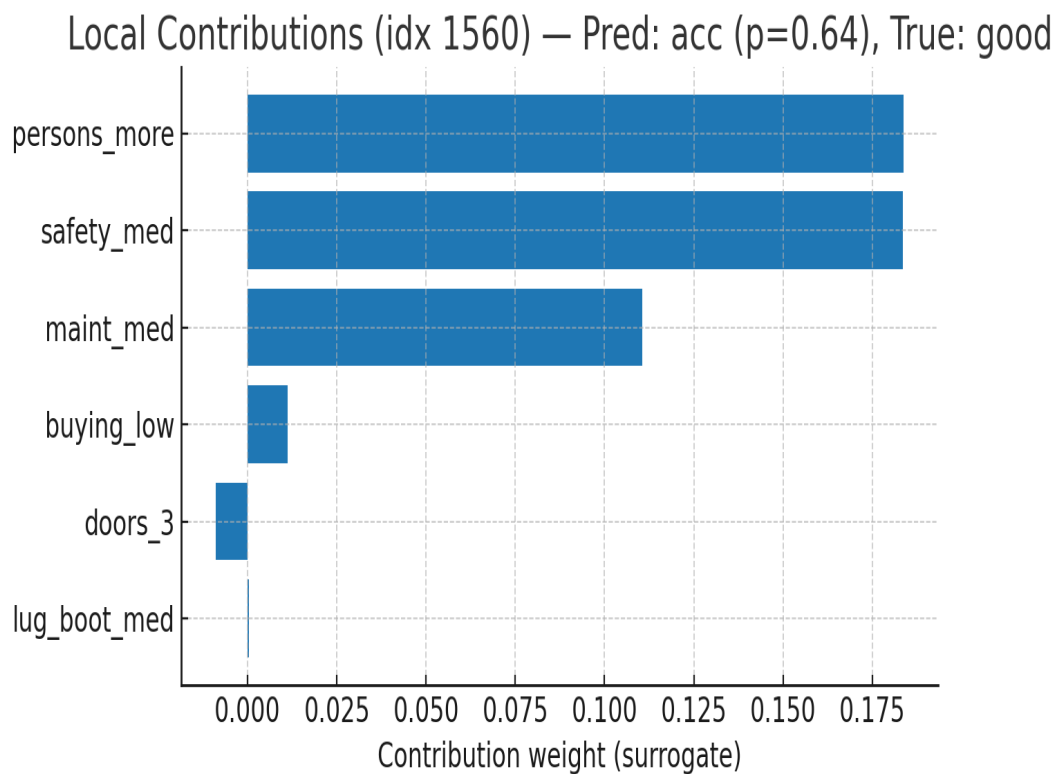
For each selected test instance, we generate perturbations by randomly changing the instance's categorical features using the training distribution, then fit a locally-weighted linear surrogate

(Ridge) to approximate the model's predicted probability near that point. The coefficients on the instance's active one-hot features form the local contributions. Positive weights push the prediction towards the predicted class; negative weights push away.

Instance 1 (index 1076)

- True label: unacc
- Predicted: unacc (p=1.000)

Feature	Value
buying	med
maint	high
doors	5more
persons	more
lug_boot	big
safety	low



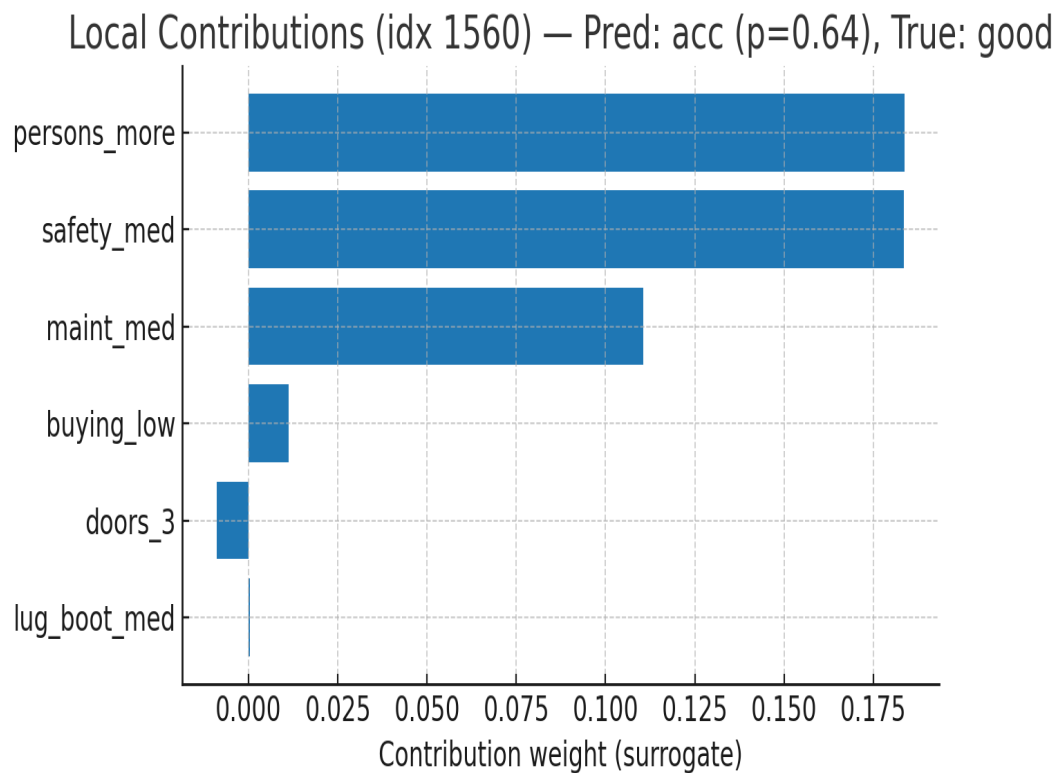
Active Encoded Feature	Weight
safety_low	0.4896
persons_more	-0.0453
buying_med	-0.0353
maint_high	0.0054
lug_boot_big	0.0024

doors_5more	-0.0012
-------------	---------

Instance 2 (index 388)

- True label: acc
- Predicted: acc (p=0.750)

Feature	Value
buying	vhigh
maint	low
doors	4
persons	4
lug_boot	small
safety	high

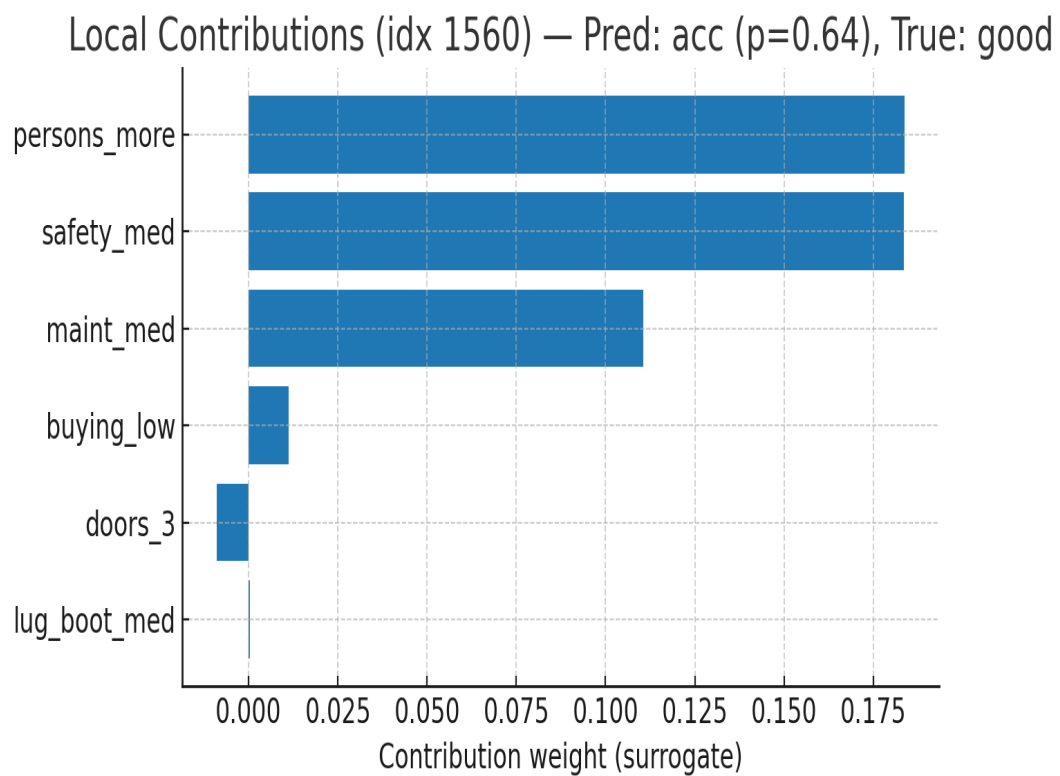


Active Encoded Feature	Weight
persons_4	0.2001
safety_high	0.1955
maint_low	0.1730
buying_vhigh	0.1639
lug_boot_small	-0.0145
doors_4	0.0021

Instance 3 (index 1560)

- True label: good
- Predicted: acc (p=0.642)

Feature	Value
buying	low
maint	med
doors	3
persons	more
lug_boot	med
safety	med



Active Encoded Feature	Weight
persons_more	0.1838
safety_med	0.1835
maint_med	0.1105
buying_low	0.0113
doors_3	-0.0090
lug_boot_med	0.0003

6) Notes

- This report uses a self-contained, LIME-inspired method to produce local explanations without external libraries.
- For publication or coursework, you may optionally compare these results with the official lime package to validate consistency.
- You can rerun with different `max_depth`, random seeds, or the number of perturbations to explore stability.