Shapley values and PCA reconstruction error model
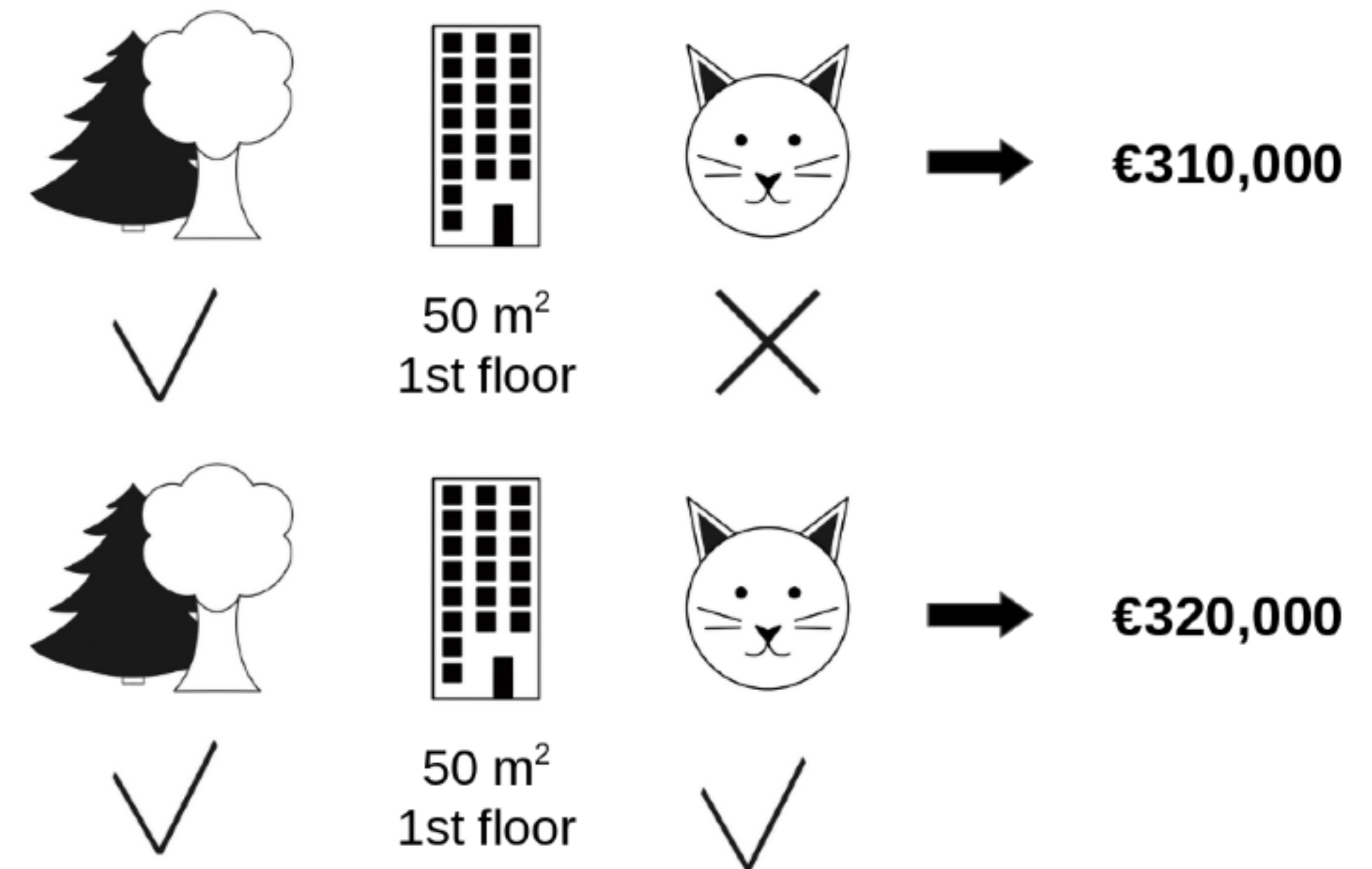
Shapley value: **Average marginal contribution of a feature value across all possible coalitions.**

- Model is tested with iterations of unique coalitions.
- The predicted outcome is computed with and without the feature value of interest for each coalition.
- The mean of the marginal contributions from all iterations is calculated as the Shapley value.

The Shapley value is the average contribution of a feature value to the prediction in different coalitions. The larger the Shapley value, the more it affects the model prediction.

€310,000

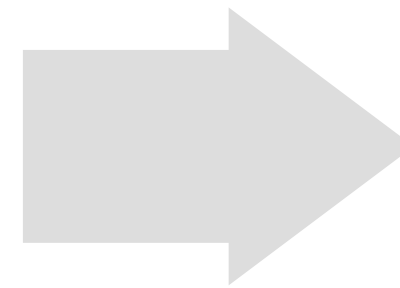50 m² 1st floor

€320,000

50 m² 1st floor

Reconstruction error model:
By considering Shapley values for reconstruction error, we can identify features that lead to higher reconstruction errors in PCA.
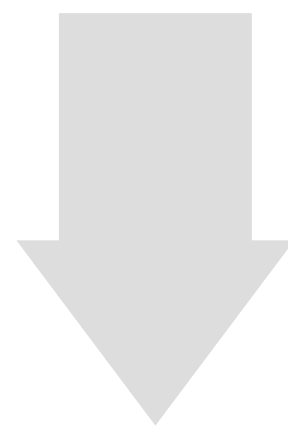
SHAP

4

Background | Goals | **Methods** | Results | Summary
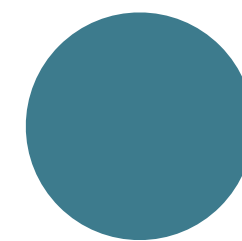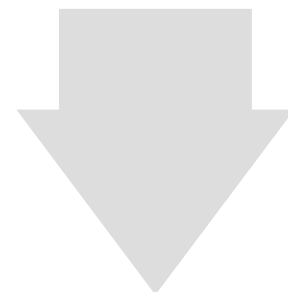
Design of study

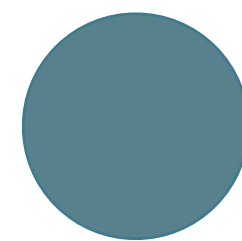25 clinical features → **Baseline PCA**

Feature selection using Shapley value reconstruction error model.
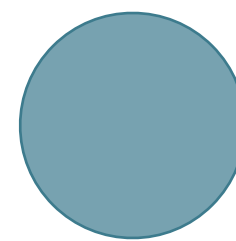Identifying features causing the supra threshold reconstruction error (x)
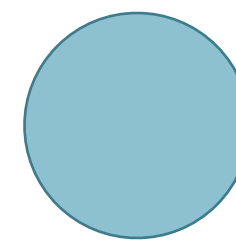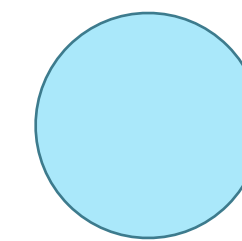
25 - x clinical features

PC1   PC2   PC3   PC4   …   PCn