Application of Tree-based Machine Learning Models for Energy Performance Certificate Predictions of Dwellings in the Netherlands

Magan, I.H. (Isaï)

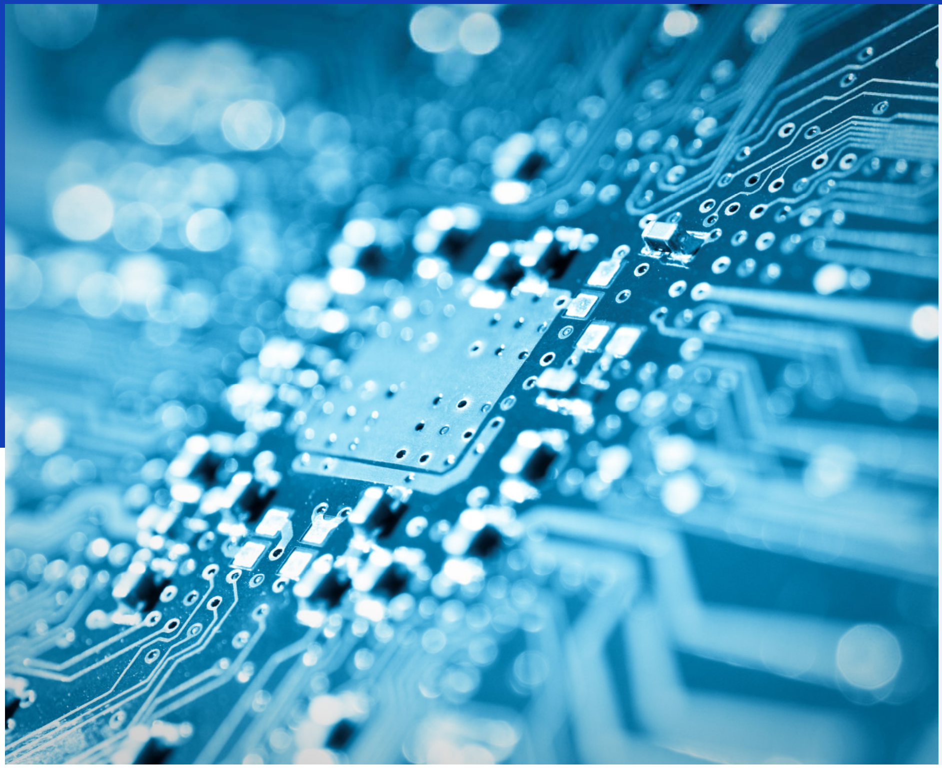
Tejada Arango, D.A. (Diego)

Zwamborn, A. (Arjan)

ENERGY EVALUATION

TNO innovation for life

Agenda



1. Introduction
2. Methodology
3. Results
4. Discussion
5. Conclusion
6. Questions

Introduction

- EU ambitions for increasing the energy performance of dwellings
  - EU Energy Performance of Buildings Directive (EU/2024/1275)
    - Fully decarbonize building stock by 2050
    - Reduce average primary energy consumption of dwellings with at least 16% by 2030 (RESCoop EU, 2024)

Figure 1: EU Directive (DR Deutsche Recycling Service GmbH, 2024)

- Energy Performance Certificate
  - Used as a metric to assess the energy efficiency of a household
  - Often represented in classes (A, B, C...)
    - Underlying classification expressed in some consumption metric such as $kWh/m^2/year$

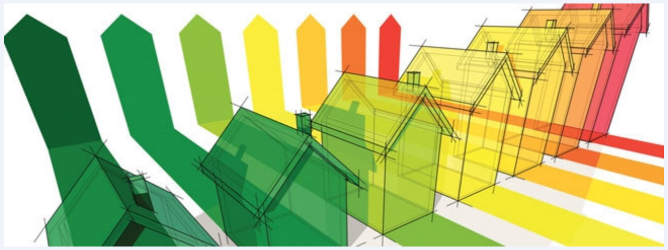


Figure 2: Dwelling classes (EPG, 2025)

Introduction

- Default methodology to assess energy performance of dwellings in the Netherlands – NTA 8800
  - Mandatory when a dwelling is constructed or sold
  - Not easily reproducible
  - Time consuming
  - Has not assigned energy labels to all dwellings yet (~40% missing as of 1st of January 2024)
  - Widely used by municipalities, architects, construction companies

- The aim of this study was to develop an alternative methodology
  - Tree-based machine learning
  - Assign labels and interpret the assignment

Figure 3: NTA 8800 (Duresta, 2019)

Methodology

- Tree-based machine learning model
  - Predicts using decision trees
  - Trees split data into smaller groups to make decisions
  - These decisions relate to classification

  - Random forest
    - Trees are developed in parallel
    - Combine many trees and average their predictions
  - Boosted trees (XGBoost)
    - Trees are developed sequentially
    - Each tree attempts to improve the performance of the former
    - Final tree is taken for prediction purposes

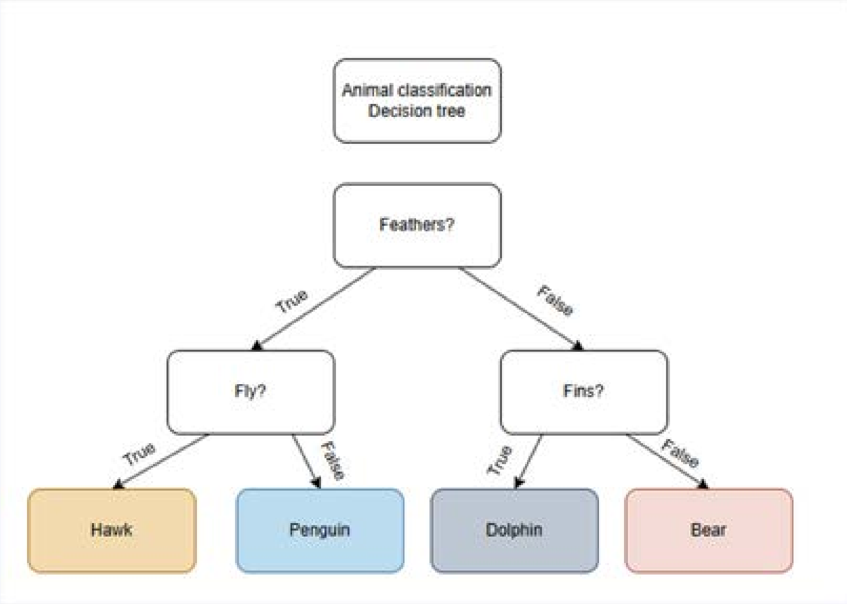


Figure 4: Example of a decision tree for animal multiclass classification

Methodology

- The machine learning models are trained on specific datasets that contain information on already labelled dwellings
  - Training set (~70%)
  - Stratified sampling

- These datasets contain the following dwelling-specific information:
  - Building year
  - Surface area
  - Compactness*
  - Ownership type
  - Ventilation type
  - Heating type
  - Surface and insulation quality of individual components
    - Roof, windows, doors, floors, walls, panels

Table 1: Example of dwelling-specific information in table format used for training the models

Parameter	Dwelling #1	Dwelling #2	Dwelling #3
Surface area	69	137	92
Area of heat loss	88.991	377.218	420.422
Compactness	1.29	2.753	4.57
Surface area of the walls	18.902	157.955	174.910
Surface area of the floor	0	87.15	67.813
Surface area of the roof	50.274	109.185	156.786
Surface area of the windows	17.116	16.699	15.787
Surface area of the doors	2.699	6.229	4.112
Surface area of the panels	0	0	1.012
Building year	2002	1900	1935
Building type	5	2	1
Ventilation type	Mechanical	Natural	Natural
Water heating type	HR	Other	HR
General heating type	HR	Other	Heat pump
Ownership type	2	0	0
Insulation quality of the walls	2	0	0
Insulation quality of the floors	Missing	1	0
Insulation quality of the roof	2	0	0
Insulation quality of the panels	Missing	Missing	0
Insulation quality of the windows	2	3	2
Insulation quality of the doors	1	1	1
Insulation quality of the sealing material	0	0	0

ENERGY EVALUATION

*Measure of the ratio between area of potential heat loss and surface area

TNO innovation for life

Results

- Random forest model performs slightly better than boosted trees
    - Assessment of models is done through the *accuracy* metric (how often is the prediction correct?)
- Average accuracy of the model is around 80%
    - Randomly assigning labels would result in an accuracy of ~1/7 (14.3%)
- When inaccurate, model tends to predict a more efficient label

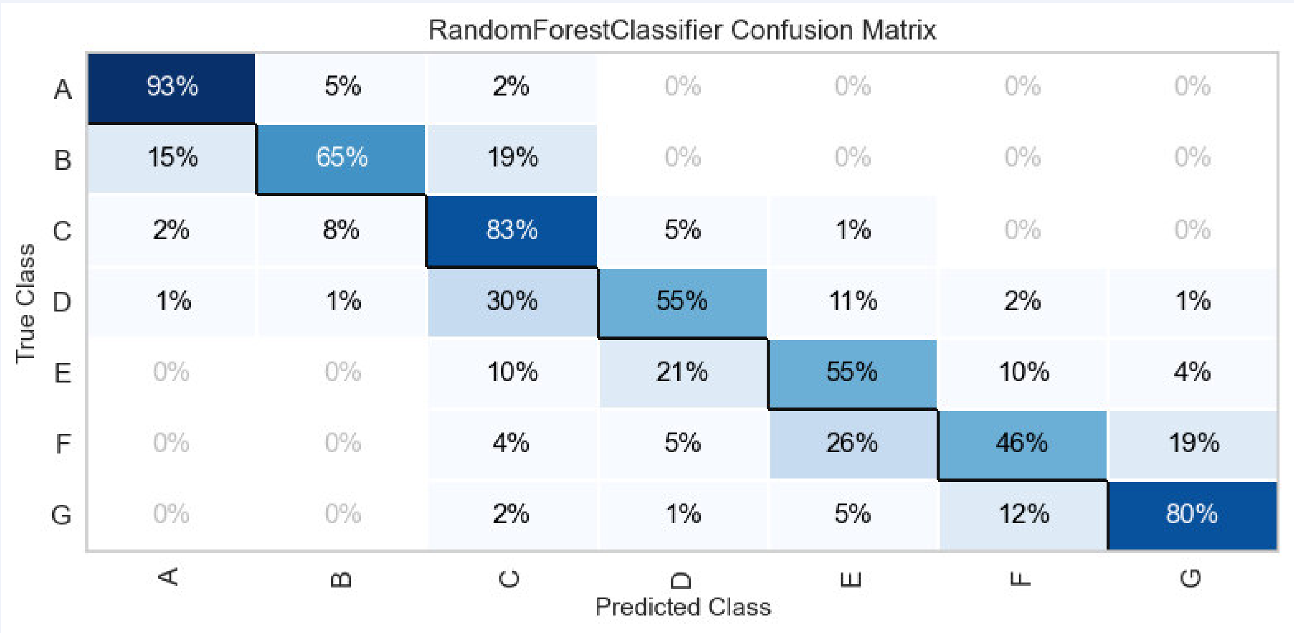


Figure 5: Confusion matrix result for the random forest classification model depicting predicted class against true class

Results

- Interpretability – influence of different individual building characteristics on EPC assignment
  - SHAP value indicates magnitude of impact (SHapley Additive exPlanations)
  - Results available per EPC assignment class – see below for 'label G' (worst performing dwellings)

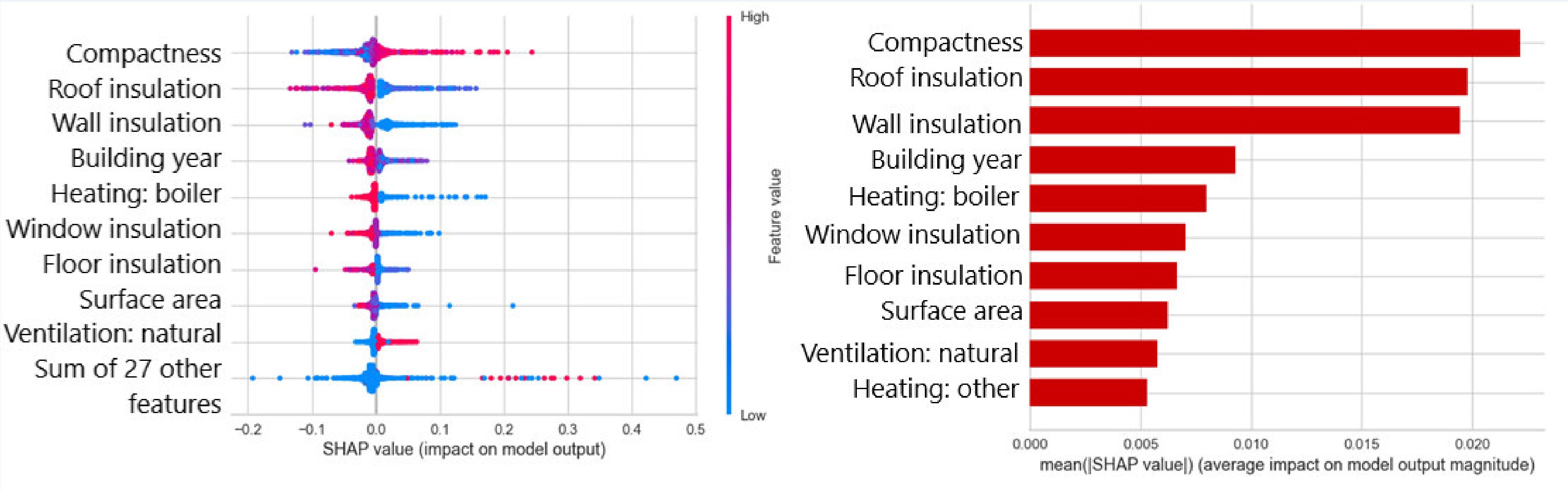


Figure 6: Interpretability plot of the impact of dwelling characteristics on the EPC assignment of the worst performing dwellings

Discussion

- Aim was to establish an alternative methodology
  - Machine learning model
  - However, subject to data quality
    - Skewed dataset
    - Limited number of parameters (including non-actionable such as building year)
    - Some impactful parameters excluded (solar PV, ventilation)

- Interpretability analysis
  - Provides insights on impact of specific characteristics on assignment
  - Can differentiate between different EPC classes (e.g. best- or worst-performing dwellings)
  - However, what are follow-up steps or research topics necessary for policy makers?
    - Non-actionable characteristics provide little to no available follow-up steps
    - Requires follow-up work to analyse improvement strategies of dwelling stock

Conclusion

- Alternative for the standard NTA 8800 procedure
  - Random forest classification model
  - Interpretability functionality
  - Not intended as a substitute, rather as a complimentary tool
  - Generalized enough to be applicable outside of the Netherlands

- We find that for the worst performing dwellings in the Netherlands are mostly impacted by:
  - The quality of the insulation of the roof and facades
  - The type of space heating
- Ongoing effort to train similar models on additional dwelling-specific criteria
  - Instead of multi-class classification, another option is to estimate energy efficiency directly (regression)
- Follow-up work necessary to translate into effective measures for improving energy efficiency metrics in dwellings

Questions

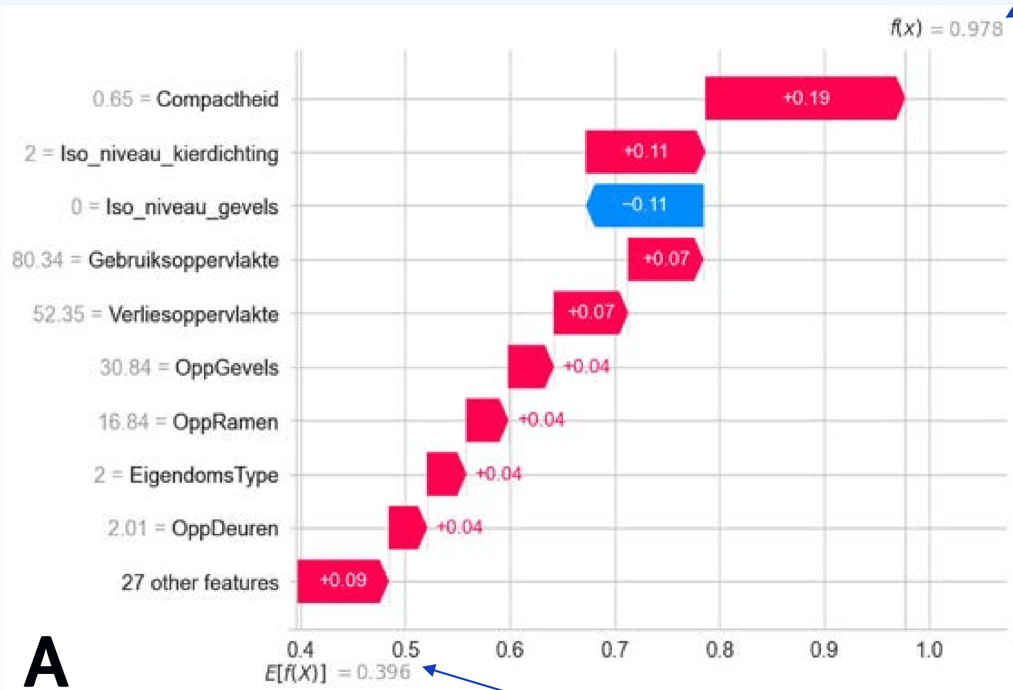
References

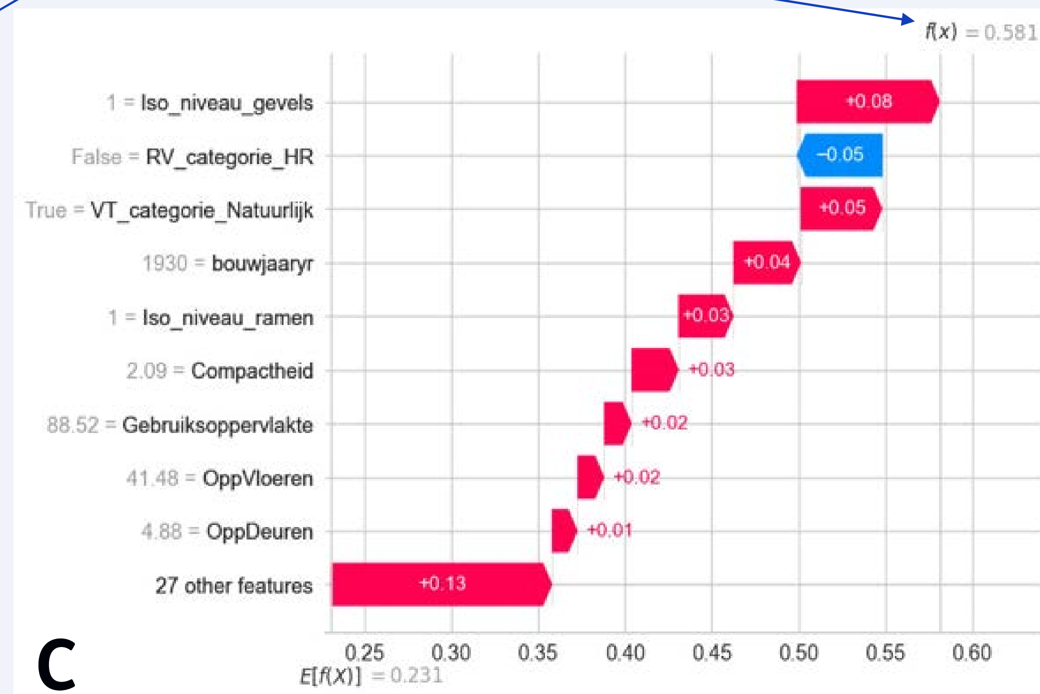
- DR Deutsche Recycling Service GmbH. (2024). *EU Directive for EPR – Packaging Act, WEEE & Bat*. Deutsche Recycling. https://deutsche-recycling.com/blog/eu-directive-for-epr-packaging-act-weee-battg/
- Duresta. (2019). *Welke invloed heeft de Nta8800 op het wws?*. Duresta. https://www.duresta.nl/2019/11/nta8800/
- EPG. (2025). *Herziening Bijlage AA (2025) bij NTA 8800:2024 'Energieprestatie van gebouwen' gepubliceerd*. Energieprestatie voor Gebouwen. https://www.gebouwenergieprestatie.nl/herziening-bijlage-aa-2025-bij-nta-88002024-energieprestatie-van-gebouwen-gepubliceerd/
- REScoop EU (2024). *Directive (EU) 2024/1275 (recast European Performance of Buildings Directive)*. RESCoop EU. https://www.rescoop.eu/uploads/rescoop/downloads/EPBD-Policy-Briefing-2nd-Generation-of-Energy-Communities-Legislation.pdf

Appendix 1 – Interpretability of the Model

- Example of two dwellings with energy labels A and C.

Probability of the energy label

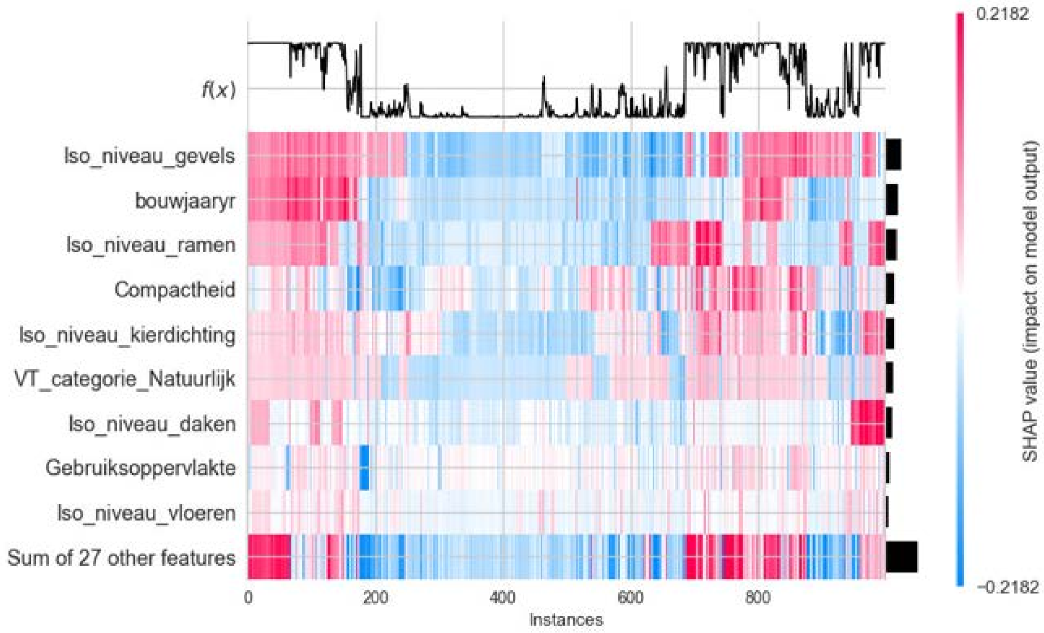




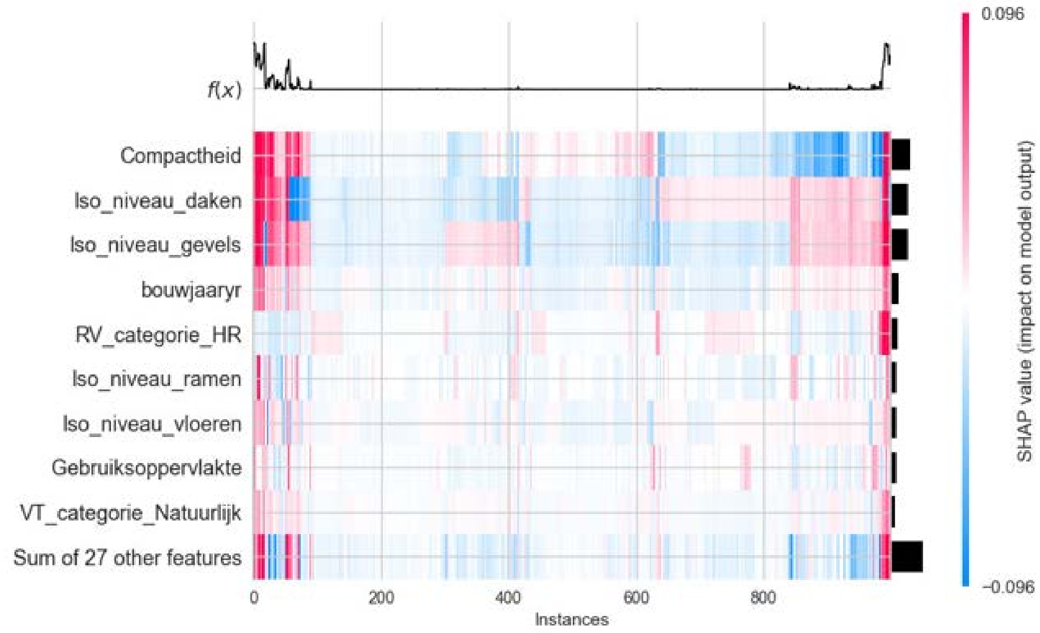
Expected probability of the energy label

Appendix 2 - Interpretability of Label A and G in 1000 dwellings

A



G



Appendix 3 – example decision tree