



# Estimating energy performance of buildings

Machine learning for energy label prediction

[rein@geodan.nl](mailto:rein@geodan.nl), a co-production with

[s.hettinga@vu.nl](mailto:s.hettinga@vu.nl) [h.j.scholten@vu.nl](mailto:h.j.scholten@vu.nl) [maarten.kriekaert@geodan.nl](mailto:maarten.kriekaert@geodan.nl)



Knowledge Representation  
and Reasoning

kadaster



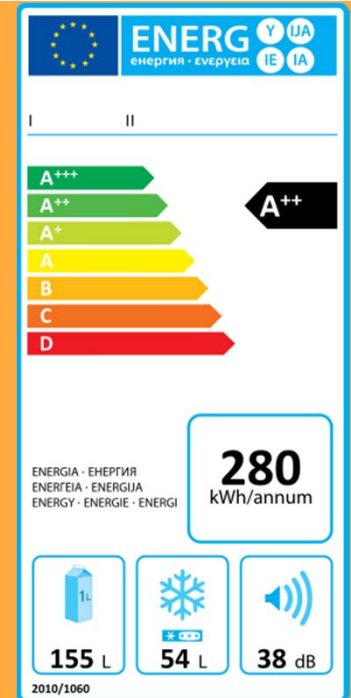
Geodan

SPIN

LAB

SPATIAL INFORMATION LABORATORY

# Energy labels: for appliances *and* buildings



# What do we know about buildings in the Netherlands?

## Why use machine learning?

Netherlands: TABULA prediction accuracy 21 %

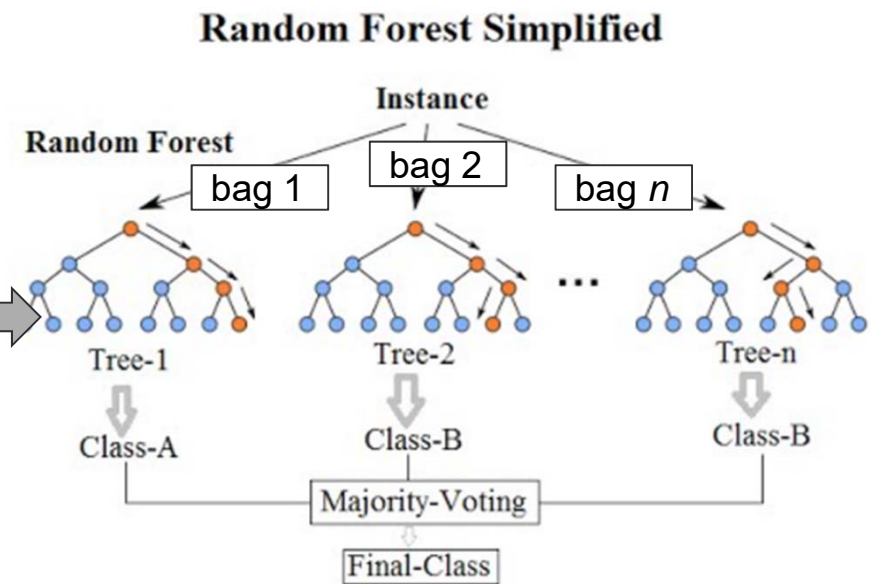
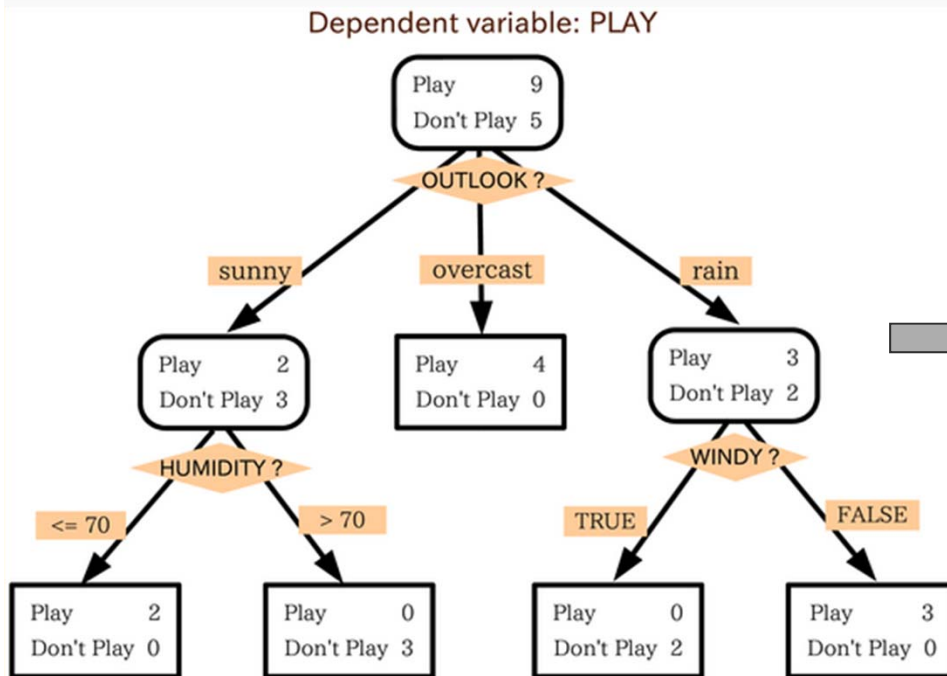
## Variables (mostly from [PICO](#) data set)

- Year of construction
- Surface area
- Habitable surface area
- Exposed perimeter
- Acquisition price
- Tax valuation
- Gas consumption 2017
- Purpose
- Building type

## Model types

- [Classification](#):
  - [Decision tree](#)
  - [Random forest](#)
  - [Gradient boosting](#)
- [Regression](#):
  - [Regression tree](#)
  - [Random forest](#)
  - [Gradient boosting](#)

# From simple decision tree to random forest



source: <https://community.tibco.com/wiki/random-forest-template-tibco-spotfirer-wiki-12345>

# Results

Model type	Model	Training subset	Estimators	Maximum depth	Accuracy	Training time
<i>Baseline accuracy</i>	<i>Data set base accuracy</i>	Unknown?			~ 86 %	
	Majority class ('C')				29 %	
Baseline hand-tuned	TABULA typological				21 %	
	TABULA specified				16 %	
Single estimator	Decision tree	100%	1	40	68 %	0h30m
	Regression tree	100%	1	48	67 %	1h
Ensemble	<b>Random forest classifier</b>	<b>100%</b>	<b>128</b>	<b>32</b>	<b>71 %</b>	<b>3h30m</b>
	Random forest regressor	100%	64	48	67 %	20h
	Gradient boosting classifier	25%	32	32	64 %	18h
	Gradient boosting regressor	100%	64	24	67 %	9h30m

# Outro

Me:

[rein@geodan.nl](mailto:rein@geodan.nl)

[rvr900@vu.nl](mailto:rvr900@vu.nl)

This presentation:

<http://reinvantveer.github.io/presentations.html>

Further links:

<https://ec.europa.eu/jrc/communities/community/digitranscope>

The Digitranscope

project

<http://webtool.building-typology.eu>

TABULA webtool

<https://doi.org/10.1016/j.enpol.2018.02.015>

Article by Sanne Hettinga on energy planning

<https://youtu.be/WWRYecd1LuI>

The 'prequel' talk by Sanne Hettinga on energy labels