UNIVERSITEIT GENT

FACULTY OF **ENGINEERING AND ARCHITECTURE**

Design of a diagnosis and follow-up platform for patients with chronic headaches

Kiani Lannoye & Gilles Vandewiele

Supervisors: prof. dr. ir. Sofie Van Hoecke, dr. ir. Vincent Keereman
Counsellor: ing. Olivier Janssens

Faculty of Engineering and Architecture

Headaches

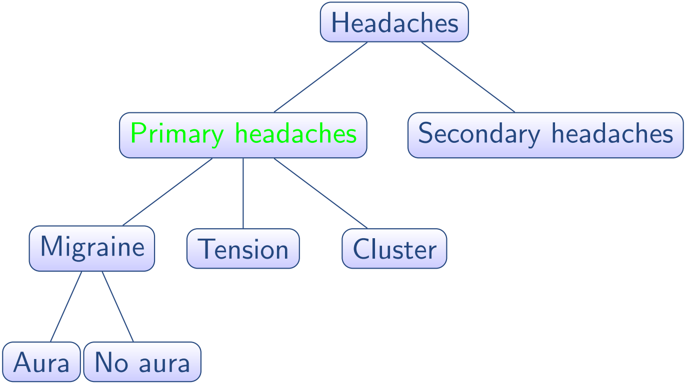
Headaches

Headaches

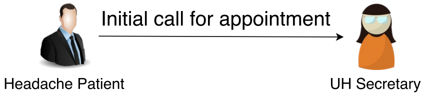
Headaches

- Headaches
  - Primary headaches
  - Secondary headaches

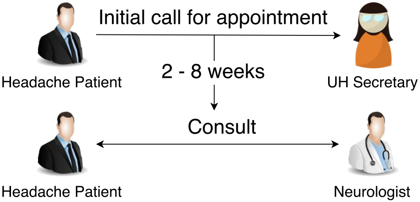
Headaches


Headaches
Primary headaches
Secondary headaches
Migraine
Tension
Cluster
Aura
No aura

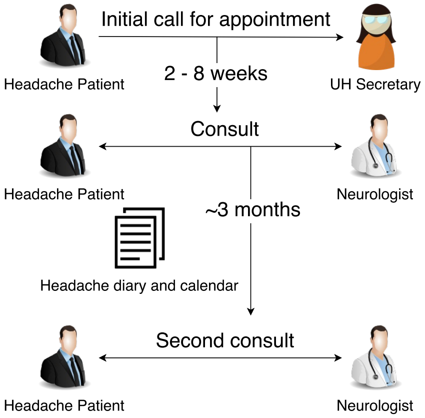
Current process at Ghent UH



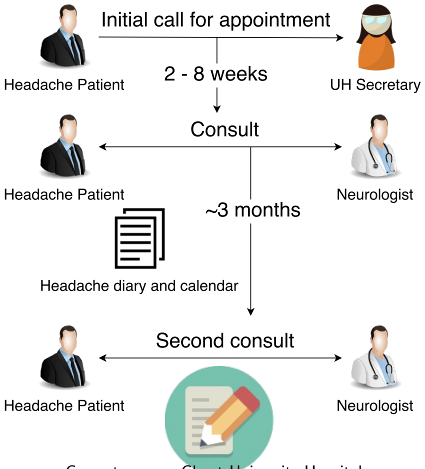
Current process at Ghent UH



Current process at Ghent UH



Current process at Ghent UH



Current process Ghent University Hospital

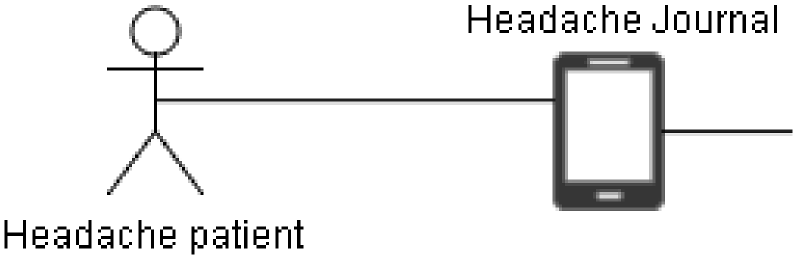
Current process at Ghent University Hospital is:

- Not digital
- cumbersome
- time consuming

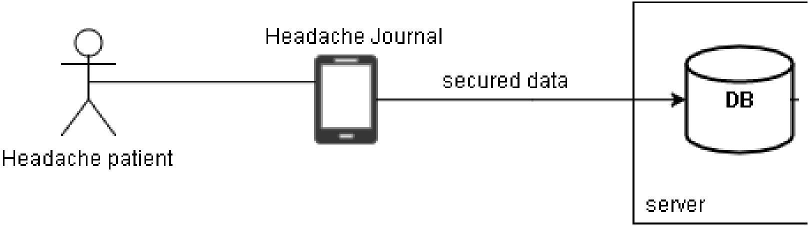
So there is need for a better (digital) alternative! This alternative has to:

- capture at least the same information as current solution
- be more efficient
- support doctors in forming a diagnosis

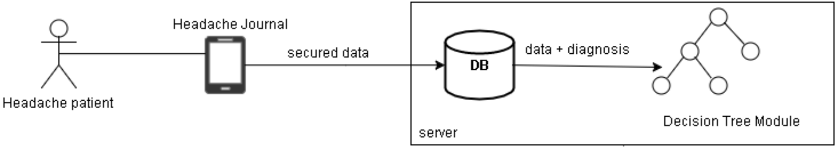
Platform requirements



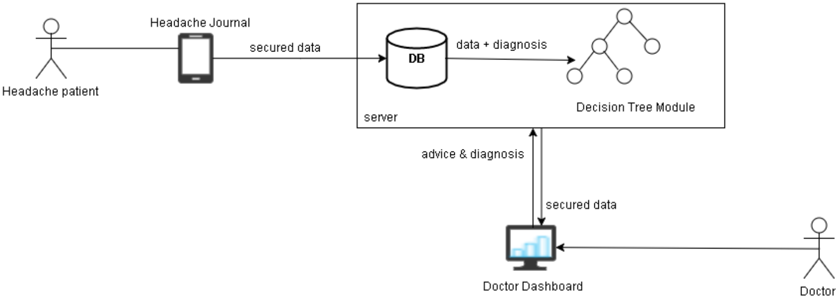
Platform requirements



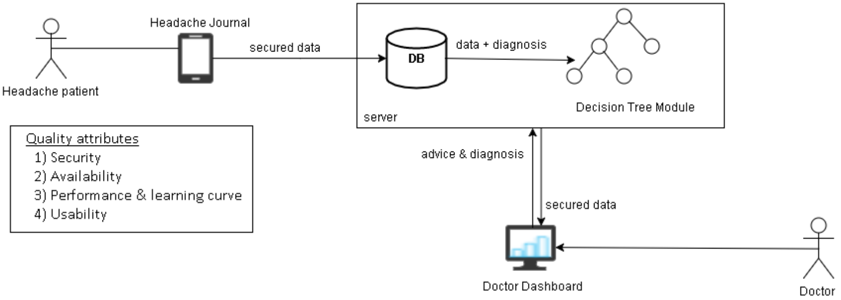
Platform requirements



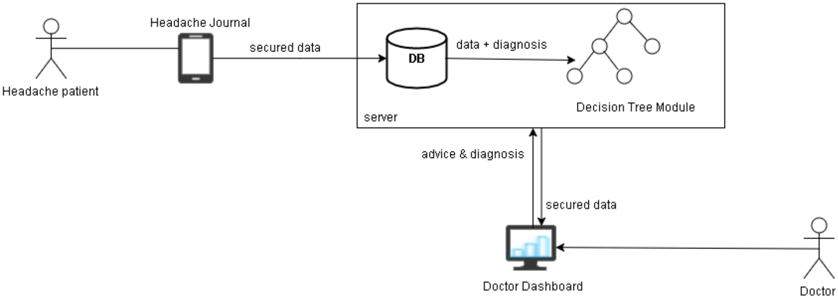
Platform requirements



Platform requirements



Mobile Application



Mobile Application

Mobile Application

Why create a new application?

Competition

- Migraine Buddy
- Headache Diary
- Pfizer headache journal

Mobile Application

Why create a new application?

Competition

- Migraine Buddy
- Headache Diary
- Pfizer headache journal

All good, but:

Mobile Application

Why create a new application?

Competition

- Migraine Buddy
- Headache Diary
- Pfizer headache journal

All good, but:

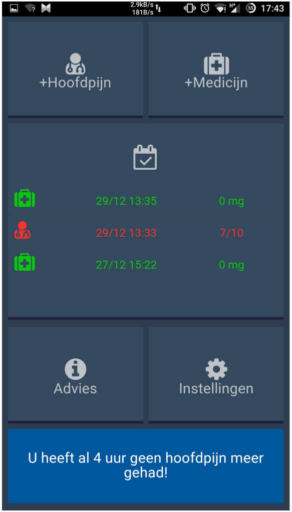
- none captures all data needed
- none offers usable data export

Cross platform vs Native

	Native	Cross-platform
+	+ Native UX + device-specific features + Better performance	+ 1 language + Write once, run everywhere + Less maintenance
-	- Multiple languages - Time consuming (development)	- Slower (lower performance) - Less device specific features - Harder to release online (Play Store/App Store)

Cross platform vs Native

	Native	Cross-platform
+	+ Native UX + device-specific features + Better performance	+ 1 language + Write once, run everywhere + Less maintenance
-	- Multiple languages - Time consuming (development)	- Slower (lower performance) - Less device specific features - Harder to release online (Play Store/App Store)

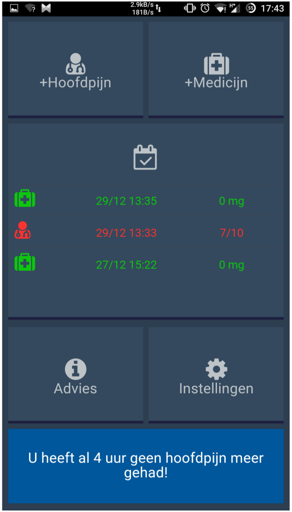

17:43
+Hoofdpijn
+Medicijn
29/12 13:35
0 mg
29/12 13:33
7/10
27/12 15:22
0 mg
Advies
Instellingen
U heeft al 4 uur geen hoofdpijn meer gehad!

HTML

CSS

Onsen UI

PhoneGap



Chronicals

Chronicals


17:44
Terug
29 dec. 17:44
?
Locatie
Toon op welke plaats(en) u pijn voelt.
Tip: u kan het model draaien
Annuleer
Overslaan
Volgende

Chronicals

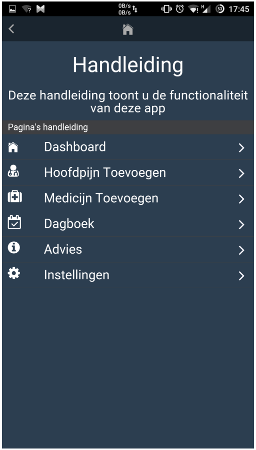

Terug
29 dec. 17:44
Locatie
Toon op welke plaats(en) u pijn voelt.
Tip: u kan het model draaien
Annuleer
Overslaan
Volgende


Terug
29 dec. 17:44
Braken
Gevoeligheid aan inspanning
Misselijkheid
Gevoeligheid aan licht
Gevoeligheid aan geluid
Gebrek aan eetlust
Hoofdpijn omschrijving
Autonoom
Aurasymptomen
Visueel aura
Sensibel aura
Motorisch aura
Spraakproblemen
Annuleer
Overslaan
Volgende

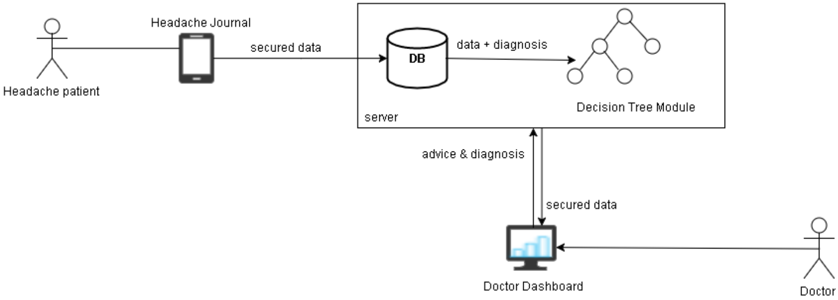
Chronicals


Terug
29 dec. 17:44
?
Locatie
Toon op welke plaats(en) u pijn voelt.
Tip: u kan het model draaien
Annuleer
Overslaan
Volgende

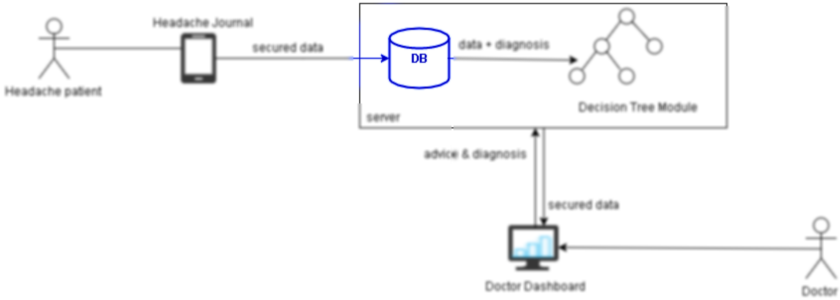

Terug
29 dec. 17:44
?
Braken
Gevoeligheid aan inspanning
Misselijkheid
Gevoeligheid aan licht
Gevoeligheid aan geluid
Gebrek aan eetlust
Hoofdpijn omschrijving
Autonoom
Aurasymptomen
Visueel aura
Sensibel aura
Motorisch aura
Spraakproblemen
Annuleer
Overslaan
Volgende


Handleiding
Deze handleiding toont u de functionaliteit van deze app
Pagina's handleiding
Dashboard
Hoofdpijn Toevoegen
Medicijn Toevoegen
Dagboek
Advies
Instellingen

Backend and data exposure



Backend and data exposure

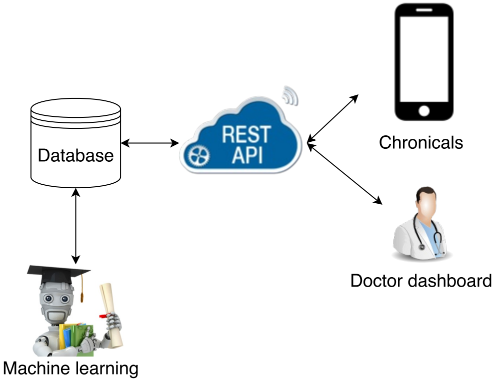


Backend and data exposure

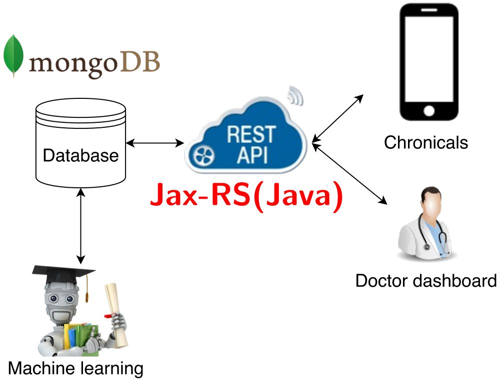
Components

- Database
- Connection to App
- Connection to Doctor Dashboard
- Connection Machine learning module

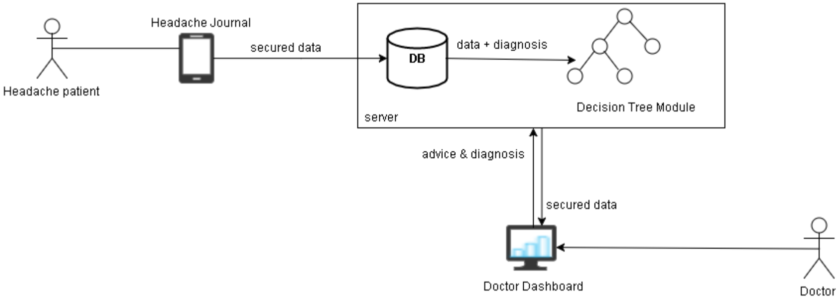
System


Database
REST
API
Chronicals
Doctor dashboard
Machine learning

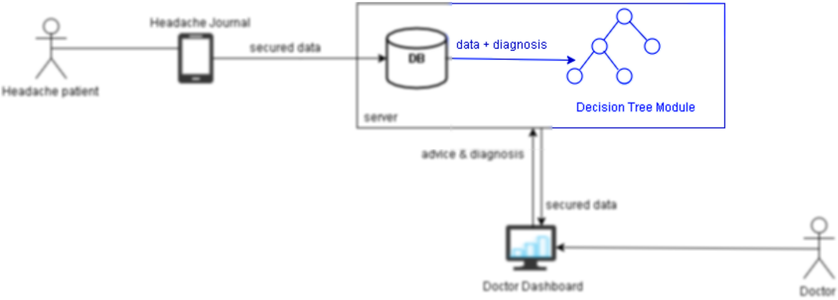
System


mongoDB
Database
REST API
Jax-RS(Java)
Chronicals
Doctor dashboard
Machine learning

Machine learning



Machine learning



1. **Introduction**

   Why white box models? - Flaws of current approaches

2. **Decision tree merging**

   Merging decision trees in a single, interpretable tree

3. **Genetic approach**

4. **Evaluation**

   Used datasets - Results - Our headache dataset

Machine Learning

Decision support ($\neq$ decision making) $\Rightarrow$ White box model

Possible models

- Decision trees
- Random Forests (Gray box)
- Bayesian networks

Machine Learning

Decision support ($\neq$ decision making) $\Rightarrow$ White box model

Possible models

- Decision trees
- Random Forests (Gray box)
- Bayesian networks

Many different DT induction algorithms

C4.5 (C5.0)

CART

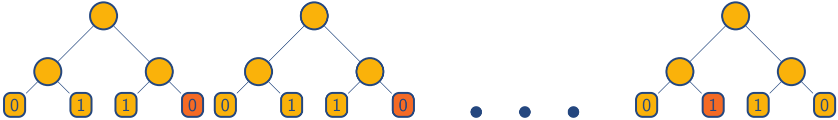
QUEST

. . .

→ **Which tree is the most beautiful?**

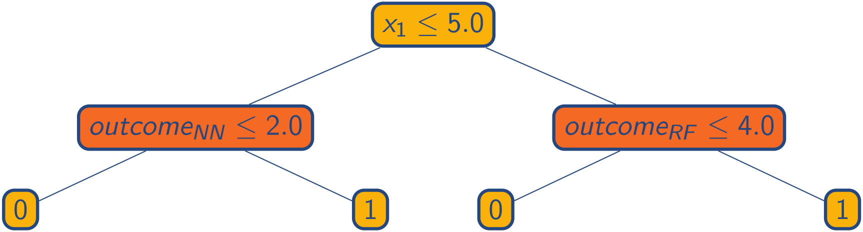
Current ensembles lack interpretability

Boosting, bagging, random forests, etc. require majority voting (classification) or mean calculation (regression) to obtain prediction

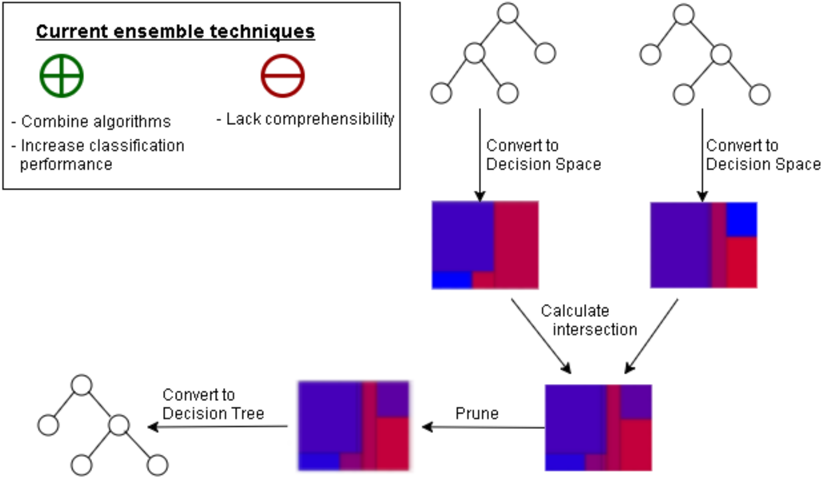


Current ensembles lack interpretability

The final decision tree obtained by **stacking** contains uninterpretable internal nodes



An ensemble technique WITH interpretability

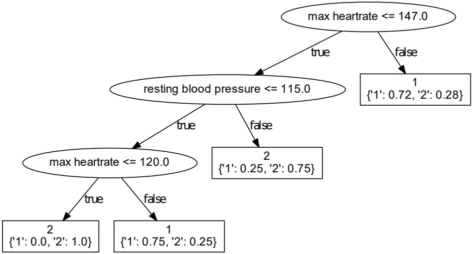


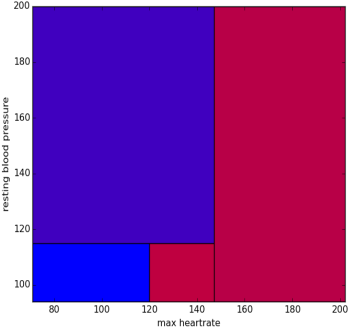
Decision tree $\rightarrow$ decision space

Converting decision trees to decision spaces

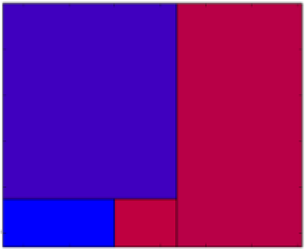
We can define a one-to-one mapping between a decision tree and a set of k-dimensional hyperplanes ($k = \#\text{features}$), called **decision space**. Each node in the decision tree corresponds to a hyperplane in the decision space.

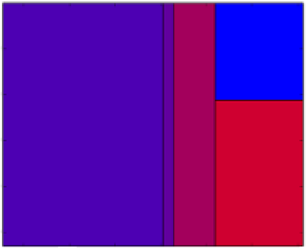
Decision tree → decision space



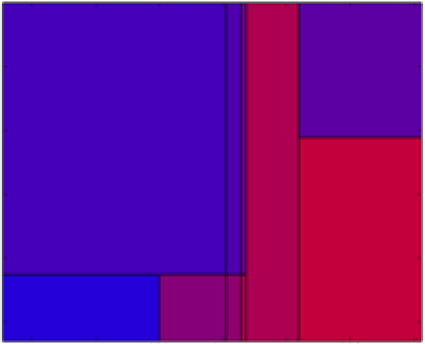


Merging decision spaces

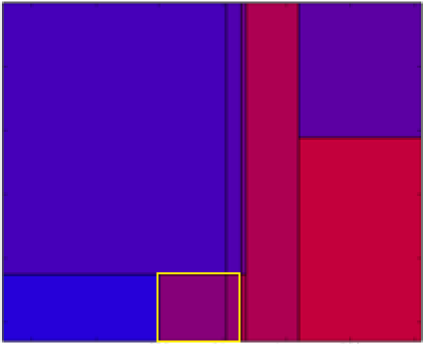




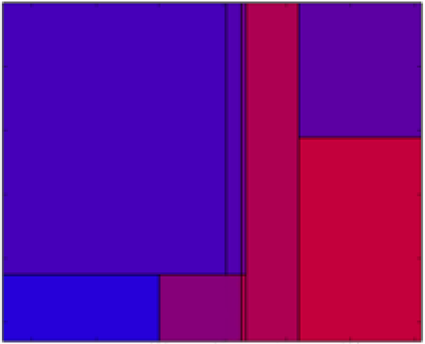
Merging decision spaces



Pruning decision spaces



Pruning decision spaces

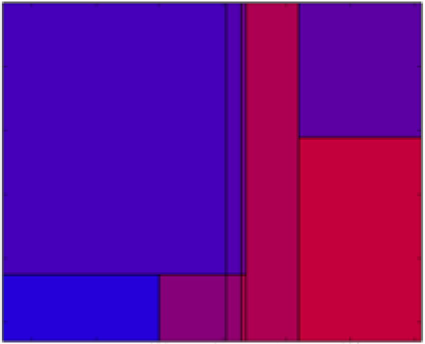


Decision space $\rightarrow$ decision tree

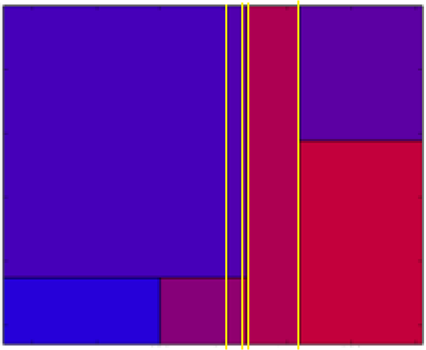
Converting decision spaces to decision trees

One-to-one mapping from decision tree to space is gone because the order is lost during conversion from DT to DS. Therefore, a **heuristic** approach must be taken, identifying **hyperplane candidates** and calculating a metric to choose the 'best' plane.

Decision space → decision tree



Decision space $\rightarrow$ decision tree



Decision space → decision tree

Finding 'best' candidate hyperplane

Apply metric function to each plane, these include:

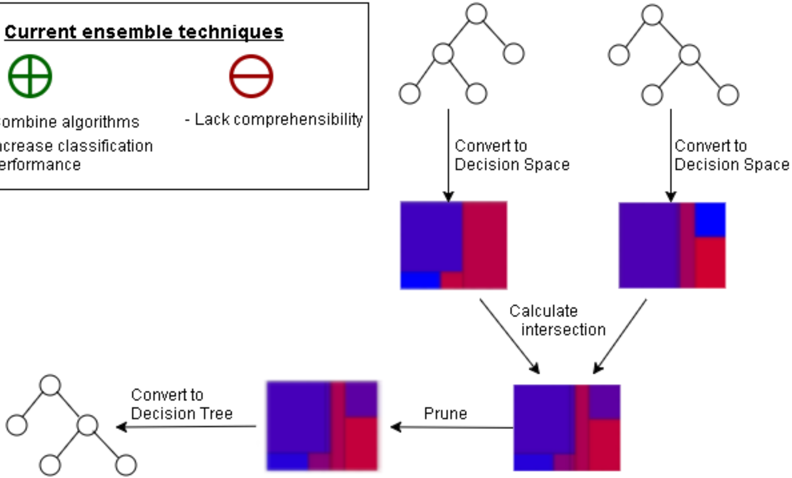
- information gain and Gini
- pick plane from most correlated feature
- pick plane that divide space in two most equal subspaces
- combination

RECAP

Current ensemble techniques

- Combine algorithms
- Increase classification performance

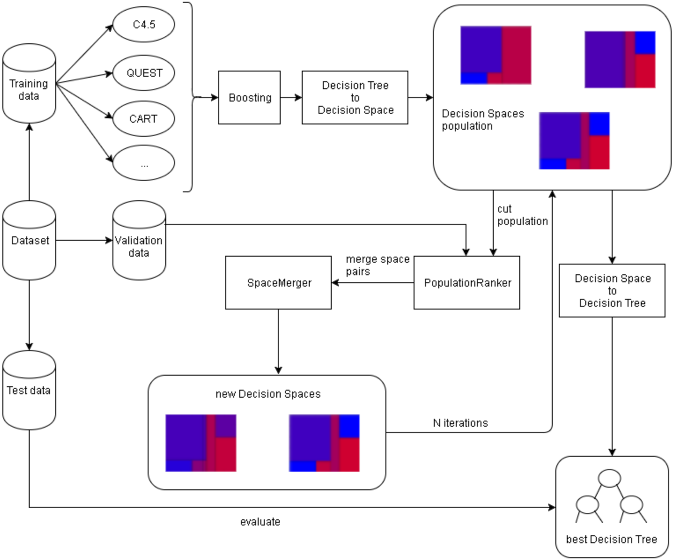
- Lack comprehensibility



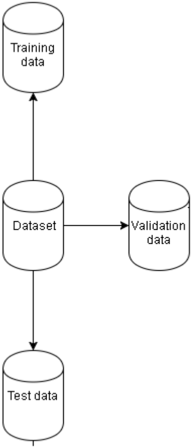
But which decision trees to merge?

Many different algorithms $\rightarrow$ Trying all combinations takes time

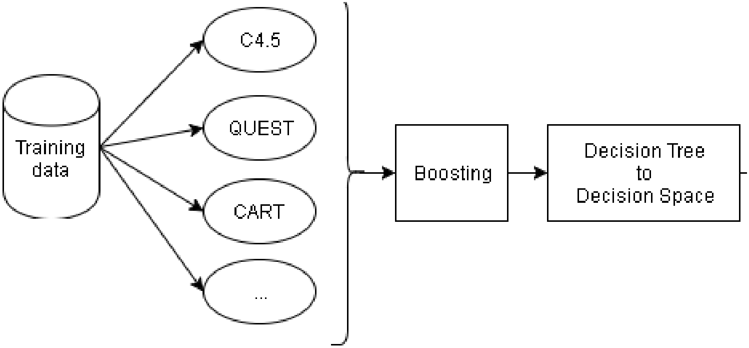
$\rightarrow$ Genetic algorithms to the rescue!


C4.5
QUEST
CART
...
Training data
Boosting
Decision Tree to Decision Space
Decision Spaces population
Dataset
Validation data
cut population
merge space pairs
SpaceMerger
PopulationRanker
Decision Space to Decision Tree
Test data
new Decision Spaces
N iterations
evaluate
best Decision Tree

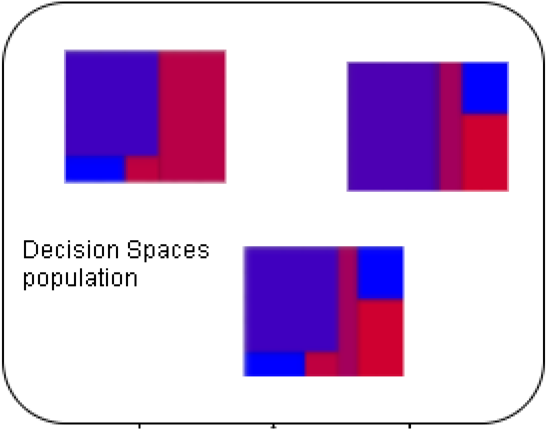
Splitting the data



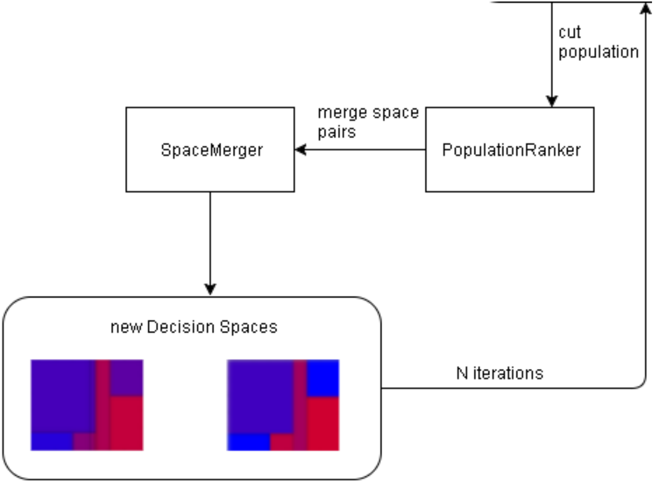
Generate different decision trees



Generate different decision trees



Genetic merging

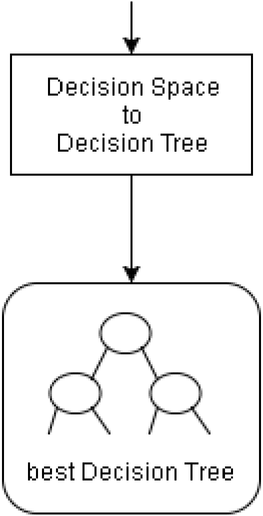


PopulationRanker

Fitness function

A high accuracy is the most important property of a decision tree, followed by its' size ($\rightarrow$ comprehensibility). Genetic algorithms are well suited for **multi-objective optimization**.

Final iteration

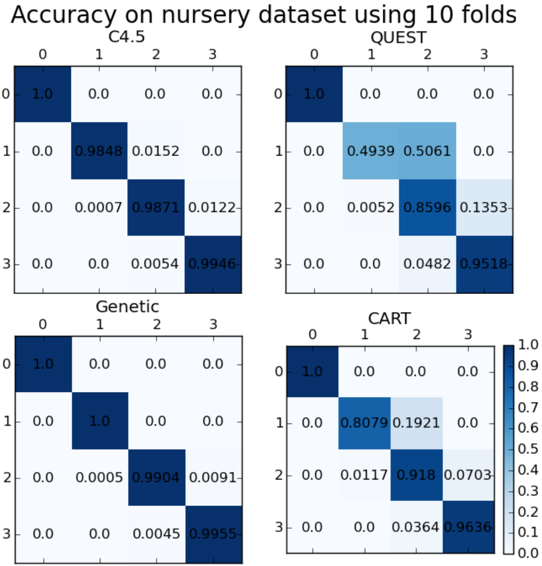


Evaluating our algorithm

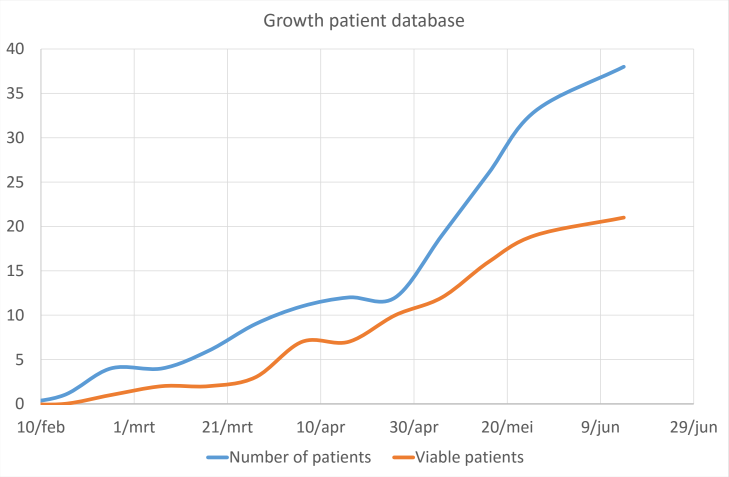
5 datasets from UCI
optimal parameters, feature selection when needed and k-fold CV

Name	#Samples	#Disc	#Cont	#Class	Imbalance rate
Heart	270	7	6	2	0.058
Car	1728	6	0	4	0.225
Iris	150	0	4	3	0
Shuttle	14500	0	9	7	0.18308
Nursery	12960	8	0	5	0.1498

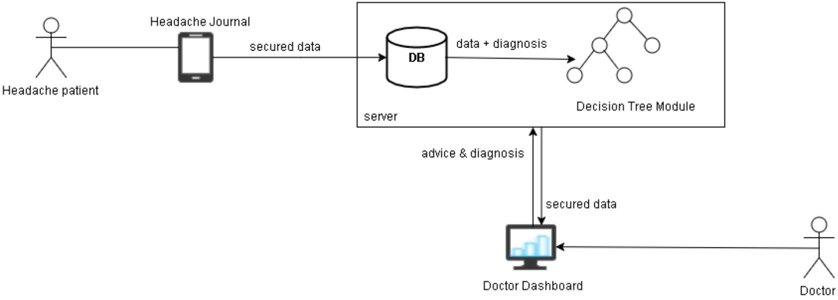
Dataset	Folds	C4.5	CART	QUEST	Genetic
Heart disease	5	**0.8067**	0.7844	0.7844	**0.8067**
	10	**0.8104**	0.7732	0.7881	0.7993
Iris	3	0.9533	0.9467	0.9467	**0.96**
	5	0.9467	0.9333	0.9467	**0.9533**
Cars	3	**0.9722**	0.9693	0.9229	0.9693
	5	0.9711	0.9682	0.9241	**0.9786**
	10	0.9756	0.9751	0.9265	**0.9803**
Shuttle	3	0.9987	0.9983	0.9964	**0.9988**
	5	0.9986	0.9981	0.9962	**0.9988**
	10	0.9990	0.9987	0.9941	**0.9992**
Nursery	3	0.9890	0.9431	0.9147	**0.9914**
	5	0.9918	0.9498	0.9251	**0.9958**
	10	0.9937	0.9568	0.9259	**0.9954**


Accuracy on nursery dataset using 10 folds
C4.5
0 1 2 3
0 1.0 0.0 0.0 0.0
1 0.0 0.9848 0.0152 0.0
2 0.0 0.0007 0.9871 0.0122
3 0.0 0.0 0.0054 0.9946
QUEST
0 1 2 3
0 1.0 0.0 0.0 0.0
1 0.0 0.4939 0.5061 0.0
2 0.0 0.0052 0.8596 0.1353
3 0.0 0.0 0.0482 0.9518
Genetic
0 1 2 3
0 1.0 0.0 0.0 0.0
1 0.0 1.0 0.0 0.0
2 0.0 0.0005 0.9904 0.0091
3 0.0 0.0 0.0045 0.9955
CART
0 1 2 3
0 1.0 0.0 0.0 0.0
1 0.0 0.8079 0.1921 0.0
2 0.0 0.0117 0.918 0.0703
3 0.0 0.0 0.0364 0.9636
1.0 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0

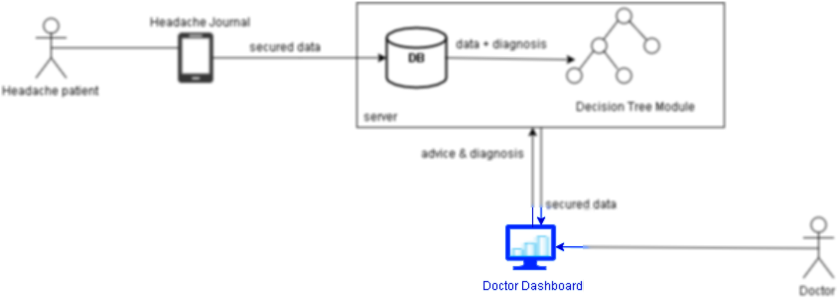
Headache dataset



Doctor dashboard



Doctor dashboard



Doctor Dashboard

- Web application in order for the doctors to access the data exposed by our REST API
- Preferably in the form of visualizations, which allow to process a lot of data in a small amount of time
- Developed by Maarten Vanden Berghe

Conclusion

The current process in the UH of Ghent can be completely digitized:

- Collect information using a mobile application
  → More efficient than paper calendars
- Present the data through a web application
  → Visualizations allow to process a lot of information quickly

This leads to an increased efficiency and reduced frequency of consults, resulting in lower health care costs.

Conclusion

A new ensemble technique was developed and tested on very varying datasets:

- increases classification performance
- preserves excellent interpretability in contrast to current ones
  $\rightarrow$ Can easily be visualized

When a higher number of headache patients are registered in our system, the resulting decision tree can be used to support physicians in forming a diagnosis

Future work

- Develop native applications for iOS and Android to enhance look-&-feel
- Re-evaluate our machine learning models on a larger headache dataset
- Implement more induction algorithms and ensemble techniques to create a more diverse initial population
- Experiment with other selection techniques and fitness functions
- Optimize the heuristic approach to convert decision spaces to decision trees

UNIVERSITEIT GENT

FACULTY OF **ENGINEERING AND ARCHITECTURE**

Thank you for your attention!

Intro

Platform requirements

Mobile application

Backend and data exposure

Machine learning

Doctor dashboard

Conclusion & future work