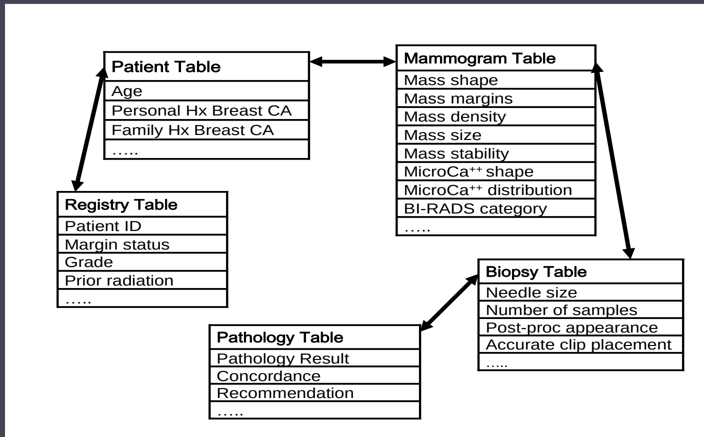


Multi-Relational Learning with applications to Health

Inês Dutra
Departamento de Ciência de Computadores
Faculdade de Ciências, Universidade do Porto
CRACS INESC TEC LA

Agenda

- Motivation
- Data in Health x Healthy Data ☺
- Case studies:
  - Breast cancer
  - Heart Diseases in children
- Parallelisation
- Wrap up

Motivation

```
is_malignant(A) :-
   'BIRADS_category'(A,b5),
   'MassPAO'(A,present),
   'MassesDensity'(A,high),
   'HO_BreastCA'(A,hxDCorLC),
    in_same_mammogram(A,B),
   'Calc_Pleomorphic'(B,notPresent),
   'Calc_Punctate'(B,notPresent).
```

Motivation

```
is_malignant(A) :-
   'BIRADS_category'(A,b5),
   'MassPAO'(A,present),
   'MassesDensity'(A,high),
   'HO_BreastCA'(A,hxDCorLC),
    in_same_mammogram(A,B),
   'Calc_Pleomorphic'(B,notPresent),
   'Calc_Punctate'(B,notPresent).
```

42 malignant and 11 benign findings

(435 total malignant + 65,000 benign)

Data (in Health)

- Unstructured x Structured
- Different sources of information
- Data from multiple tables
- Different formats

Data → Information → $\boxed{\text{Knowledge}}$ → Understanding → $\boxed{\text{Wisdom}}$

Some data science principles

- (Medical) systems are complex
- Data is dirty: deal with it!
- SvOT = LoL!
- Data munging, taming and wrestling > 70% time
- Simplification. Reduction. Distillation.
- Curiosity. Empiricism. Skepticism.

(Somehano, DataScience London)

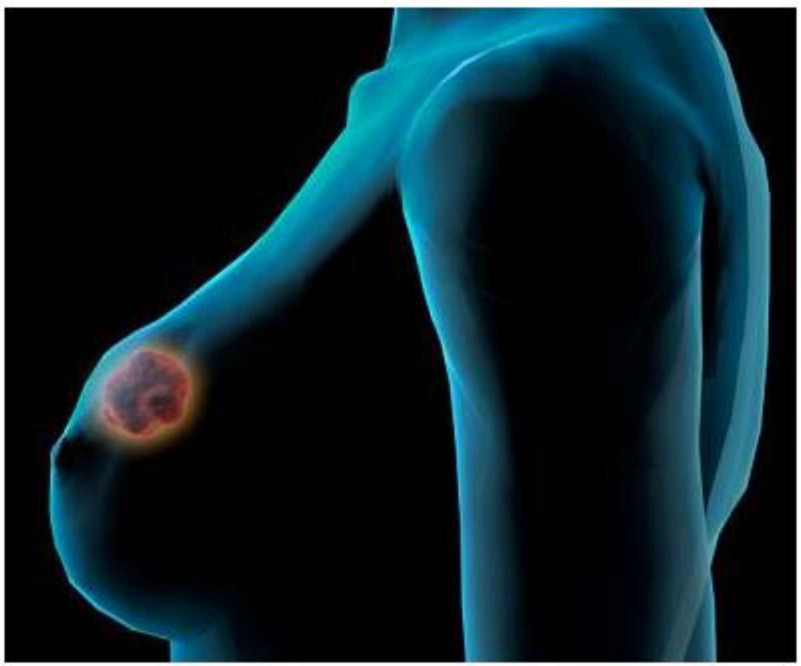
Learning from data is tricky

Statistics	Vs.	Machine learning
Supervised	Vs.	Unsupervised
Induction	Vs.	Deduction
Correlation	Vs.	Causation
Sampling	&	Confidence intervals
Probability	&	Distribution
Deviation	&	Variance
Causation	&	Prediction

Case Studies

- Breast Cancer Diagnosis
- Cardiac pathology detection

Breast Cancer: General Scenario



- USA:
  - 1 woman dies of breast cancer every 13 minutes
  - In 2013:
    - 232.340 invasive cancers
    - 39.620 (≈ 17%) expected to die

  Source: *U. S. Breast Cancer Statistics* – accessed January 2014

- Portugal:
  - Per year:
    - 4500 new cases
    - 1500 deaths (33%)

  Source: *Liga Portuguesa Contra o Cancro* – accessed January 2014

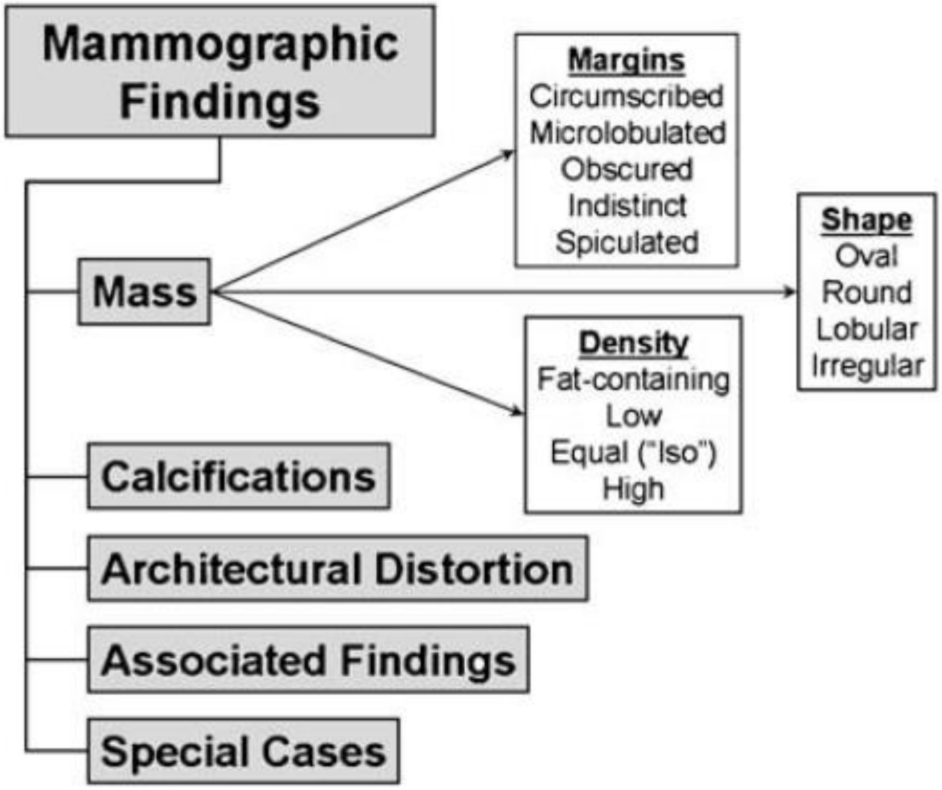
BI-RADS© Lexicon

Descriptors annotated by a specialist in mammography (X-Ray of the Breast)

- Breast screening

Cheapest screening method to detect tumours at an early stage

- Image annotated according to the BI-RADS lexicon



BI-RADS© categories

- Degree of malignancy of a finding:
  - B1, B2, B3: indicative of a **benign** finding
  - B4, B5: indicative of high-risk **malignancy**
- Medical performance:

PPV = 0.67 (**B5**)
PPV = 0.06 (**B4**)
PPV = 0.09 (B3)

- in G. Kennedy, et al., "**Predictive value of BI-RADS classification for breast imaging in women under age 50**", in Breast Cancer Res Treat, 2011.

What if something suspicious is found?

- Patients need to go through a **biopsy**
  - Invasive treatment
- Not all biopsies are conclusive!
- General practice: send every patient to excision
  - **Less than 15%** of patients that go through excision have, in fact, a malignant finding

☹

Goals

- To improve:
  - classification of malignant and benign tumors
    - MammoClass
    - ILP Rules
    - ExpertBayes
  - classification of multiple diseases
    - Not yet approached
  - BI-RADS© categorization
    - ongoing

Goals

- To be able to:
  - deal with multiple tables
  - correlate multiple exams of the same patient
    - T-SPPA
  - reduce the number of patients that go to excision when results of biopsies are inconclusive
  - extract BI-RADS© features from texts written in Portuguese
  - integrate physician advice to the automatic learning process
    - ExpertBayes

Other goals

- To study the value of information added by features obtained using image processing algorithms
- To produce interactive tools that can be used by specialists

Related Projects

- ABLe: Advice-Based Learning applied to Health Care
  - FCT-funded
- Integrating Machine Learning and Physician Expertise for Breast Cancer Diagnosis
  - NLM-funded (National Library of Medicine)
- QREN (Project 38667) with NLPC
  - Interactive tool for data analysis and discovery

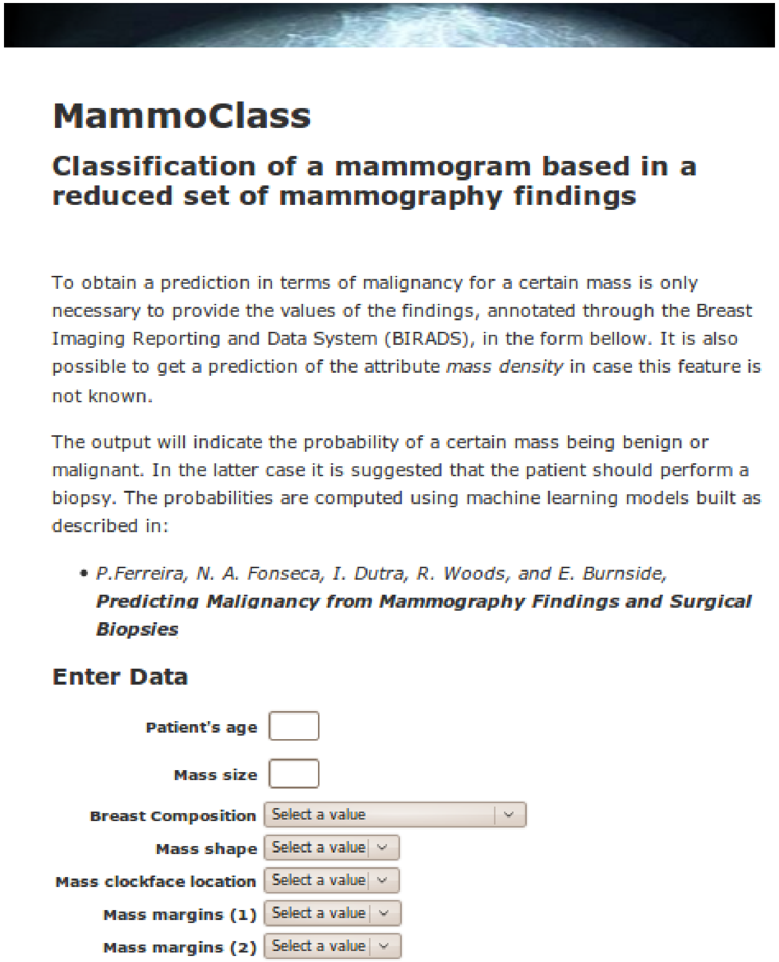
CRACS INESCTEC

Tools

- WEKA (Black-box classifiers)
  - Propositional classification algorithms
  - Feature selection
  - Filtering
- Aleph (White-box classifiers)
  - Inductive Logic Programming (ILP)
  - Multi-relational learning
  - First order-logic
  - Interpretable results
- YAP Prolog
  - Efficient implemention of Prolog
  - Several optimizations to ILP
- ExpertBayes (graphical model)
  - In-house implementation of Bayes networks refinements

MammoClass

- Online application freely available at:
  - http://cracs.fc.up.pt/mammoclass/



MammoClass

Classification of a mammogram based in a reduced set of mammography findings

To obtain a prediction in terms of malignancy for a certain mass is only necessary to provide the values of the findings, annotated through the Breast Imaging Reporting and Data System (BIRADS), in the form bellow. It is also possible to get a prediction of the attribute *mass density* in case this feature is not known.

The output will indicate the probability of a certain mass being benign or malignant. In the latter case it is suggested that the patient should perform a biopsy. The probabilities are computed using machine learning models built as described in:

- *P.Ferreira, N. A. Fonseca, I. Dutra, R. Woods, and E. Burnside,* ***Predicting Malignancy from Mammography Findings and Surgical Biopsies***

Enter Data

Patient's age

Mass size

Breast Composition Select a value

Mass shape Select a value

Mass clockface location Select a value

Mass margins (1) Select a value

Mass margins (2) Select a value

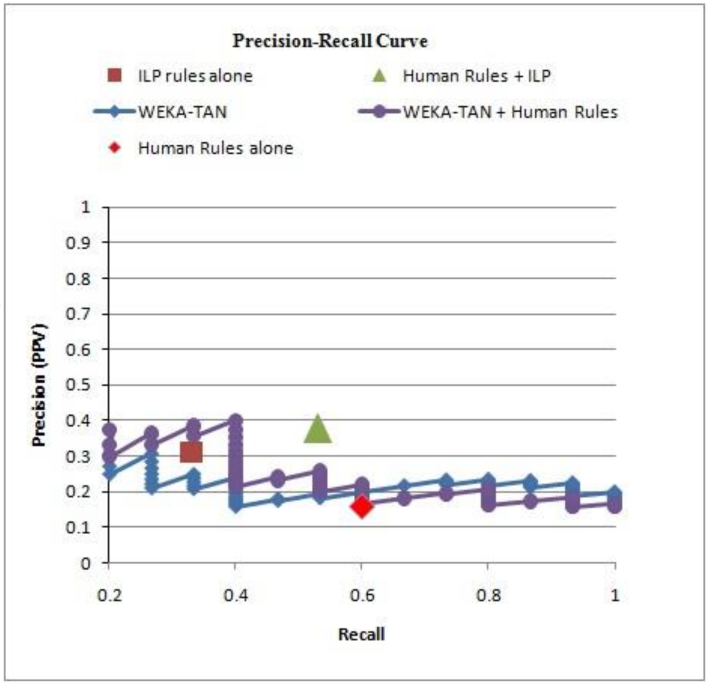
First-order rules (ILP)

(1) Probability of upgrade increases for stereotactic biopsy if the finding is microcalcifications and biopsy results reveals atypical ductal hyperplasia (ADH)

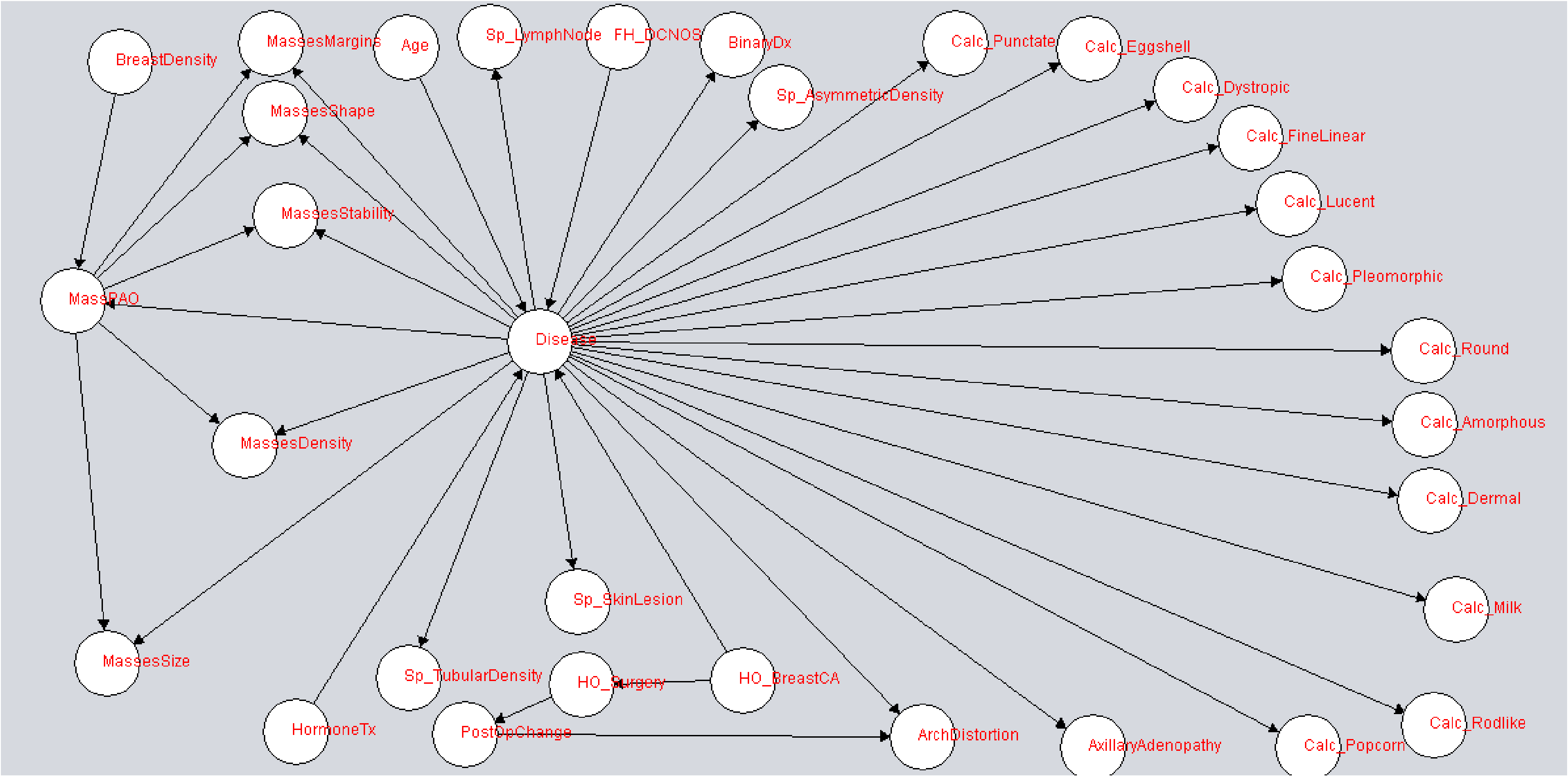
upgrade(A) IF
  pathDx(A,atypical_ductal_hyperplasia),
  calcifications(A,'a,f') and
  biopsyProcedure(A, stereo)

(2) Probability of Upgrade increases for ultrasound (US) biopsy if the finding is a mass and biopsy results reveals benign breast tissue

upgrade(A) IF
  concordance(A,d),
  biopsyProcedure(A, US_core) and
  pathDx(A,benign_breast_tissue)



ExpertBayes



CRACS INESCTEC

People

Universidade do Porto	University of Wisconsin Madison	Radboud University Nijmegen
Inês Dutra	Elizabeth Burnside	Marina V. Velikova
Vítor Santos Costa	Mai Elezaby	
Nuno A. Fonseca	Heather Neuman	
Pedro Ferreira	Molly Klein	
Ezilda Almeida	Roberta Strigel	
Tiago T. V. Vinhoza	Jude Shavlik	
Filipe Cunha	Houssam Nassif	
Tiago Silva	Ryan Woods	
Ricardo Cruz-Correia	David Page	
Inês C. Moreira	Finn Kuusisto	
Eliana Sousa	Yirong Wu	
Liliana Dores	Jingwei Li	
Joana Côrte-Real		
Gustavo Augusto		
Ricardo Sousa		

We acknowledge the *Comissão de Ética da Universidade do Porto* and the University of Wisconsin Institutional Review Board

DigiScope

Heart Diseases in Children

- **6 million**
  - children **worldwide** suffer from heart disease [1]
- **500**
  - cardiac surgeries in children per year in **Portugal** [2]
- **8-10 out of 1000**
  - babies are born with a congenital heart disease in **Portugal**, **Brazil** and **USA** [2,3,4]

Sources:

1) *European Society of Cardiology* – June 2013
2) *Apifarma, Portuguese Association of the Pharmaceutical Industry* – June 2013
3) *Revista Brasileira de Cirurgia Cardiovascular*– June 2013
4) *Lucile Packard Children's Hospital at Stanford*– June 2013

DigiScope

State of the Art

- [1] D. Aha and D. Kibler, "**Instance-based prediction of heart-disease presence with the Cleveland database**", tech. rep., University of California, Mar. 1988.
  - **Accuracy: 75.7%**

- [2] S. M. Kamruzzaman, A. R. Hasan, A. B. Siddiquee, and M. E. H. Mazumder, "**Medical diagnosis using neural network**", in 3rd International Conference on Electrical & Computer Engineering (ICECE), pp. 28–30, Dec. 2004.
  - **Accuracy: 87.5%**

- [3] B. O'Hora, J. Perera, and A. Brabazon, "**Designing radial basis function networks for classification using differential evolution**", inProc. International Joint Conference on Neural Networks (IJCNN), pp. 2932 –2937, 2006.
  - **Accuracy: 84%**

- [4] J. Wu, J. Roy, and W. F. Stewart, "**Prediction modeling using EHR data: Challenges, strategies, and a comparison of machine learning approaches**", Medical Care, vol. 48, pp. 106–113, Jun. 2010.
  - **detection of heart failure** more than **6 months before** the actual date of clinical **diagnosis**
  - **AUC: 0.77**

DigiScope

Goals

- Study relations between demographic and physiological features
- Integrate clinical features with features extracted from the heart sounds
    - ongoing
- Build classifiers that, in an automatic way, distinguish between normal and pathological cases
    - ...to support decision during screening
- Create a database for cardiology

DigiScope

Collaboration

Data

- Recife, Pernambuco – Brazil
- Collected between **October 2003 and September 2009**
- [**2-19**] year old **children**
  - Average age: 8.6 (year)

DigiScope

Methodology

17k

1st phase

7603

2nd phase

7199 instances

404 instances removed
from phase 1 to phase 2

Dataset

Preprocessing tasks

Munging, Taming, Wrestling ☺

data cleaning

data transformation

data normalization

2507 (34.8%)
pathological

4692 (65.2%)
normal (-)

DigiScope

Methodology

Dataset

Preprocessing tasks

7199 instances

33 attributes

removal of irrelevant features*

17 attributes

* patient ID, name of the physician, health insurance information, etc.

DigiScope

Methodology

Attribute
Height (cm)
Weight (kg)
Sex
Age Range
Body Mass Index Percentile
Systolic Blood Pressure (SBP)
Diastolic Blood Pressure (DBP)
Result-SBP-DBP
Murmur
Second Heart Sound (S2)
Pulses
Heart Rate (bpm)
Current Disease History 1 (CDH 1)
Current Disease History 2 (CDH 2)
Primary Reason
Secondary Reason
Pathology (class)

Dataset

17 attributes

Note:

Some of the attributes are **annotations provided by a cardiologist**, not features extracted from the sound wave files

DigiScope

Results: Feature Importance

All 7199 cases

Feature	Score
Murmur	0.61
Secondary Reason	0.10
Weight	0.09
Primary Reason	0.05
HR	0.05

DigiScope

Classification – Results Predicting abnormalities

7199

Metrics	Nested c.v.	internal test
CCI (%)	93.31	**93.32**
Sensitivity	0.85	**0.85**
Specificity	0.98	**0.98**
AUC	0.93	**0.93**

Best algorithm in all folds: **NaiveBayes**

169[5]

Metrics	Nested c.v.	internal test
CCI (%)	91.56	90.53
Sensitivity	0.72	0.70
Specificity	0.98	0.97
AUC	0.85	0.83

- [5] P. Ferreira et al., "**Detecting cardiac pathologies from annotated auscultations**", in Proc. International Symposium on Computer-Based Medical Systems (CBMS), 2012.

Related project

- DigiScope
- Tools:
  - WEKA
  - Aleph
  - YAP Prolog
  - R
  - And yes…Excel!

People

Universidade do Porto	Real Hospital Português	Universidade Federal de Minas Gerais
Inês Dutra	Sandra Mattos	Zilma Reis
Miguel Coimbra	Felipe Mourato	
Daniel Pereira	Thiago Tavares	
Pedro Ferreira	Renata Grigorio	
Ana Isabel Dias		
Tiago T. V. Vinhoza		
Ana Castro		
Ricardo Sousa		
Pedro Gomes		
Sergio Faria		
Fabio Hedayioglu		
Tiago Silva (past)		
Alípio Jorge		
André Fernandes		

We acknowledge the *Comissão de Ética da Universidade do Porto* and the Real Hospital Português

Parallelisation

- Map-reduce construct for Prolog
  - Implemented in YAP Prolog
  - User level library
  - Useful to program many applications
  - Shared memory (thread-based)
  - Distributed memory (MPI)
  - Hybrid (ongoing)
  - Dynamic data scheduler

Some Results: speedups

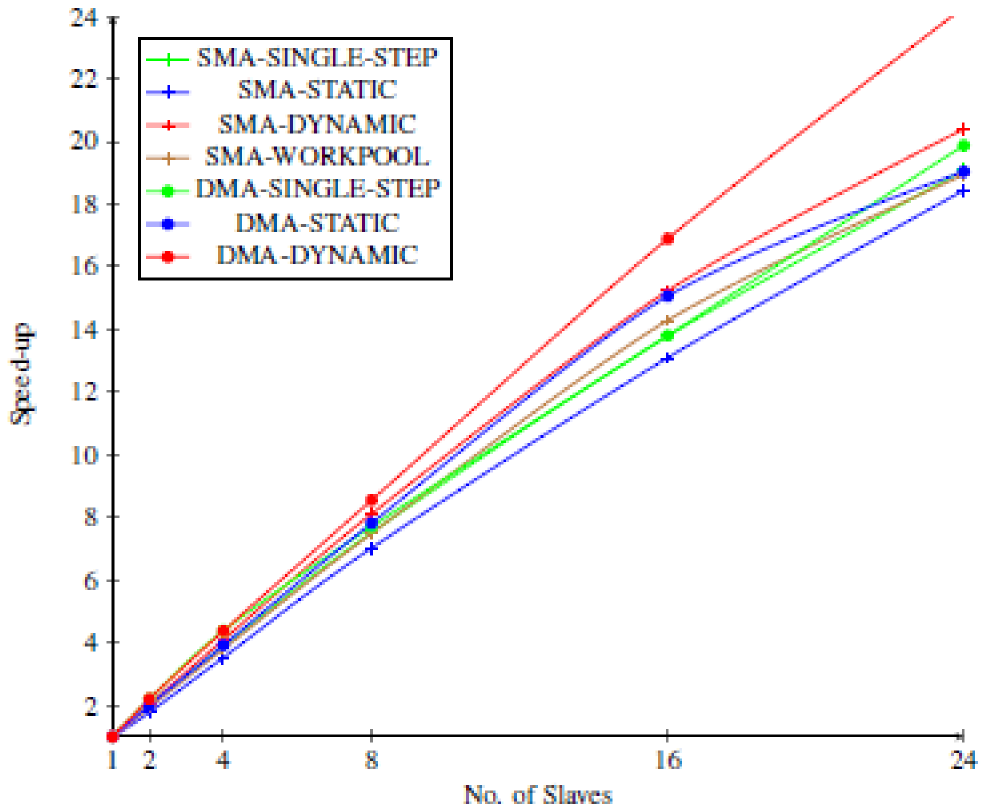


Figure 5.3: Comparison of scheduling methods for MAMMO dataset (600,000 queries and 1,000 elements per chunk)

Parallelisation

- Datalog for GPUs
  - Implemented in CUDA
  - User level library
  - Conversion of Prolog terms to intermediate numeric representation (for GPU efficiency)
  - Join, select, and other database operations

Some Results: Execution times

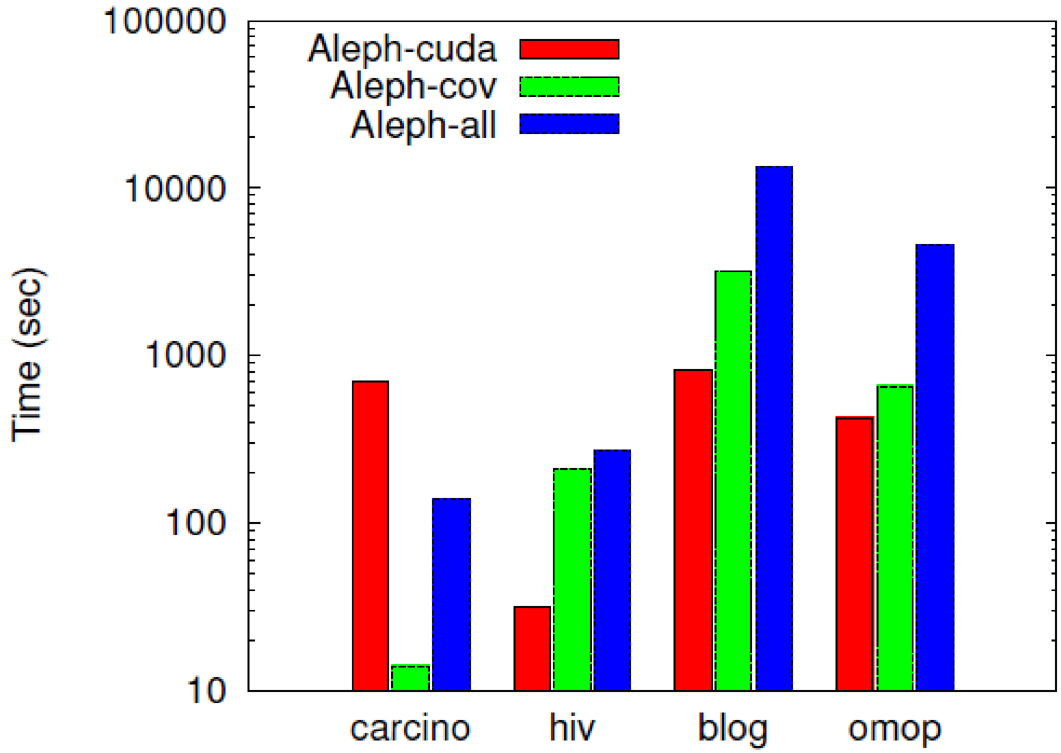


Fig. 5 Applications Total Execution Time for each Aleph version.

People

Universidade do Porto	Cinvestav, Mexico	Universidade of Aizu
Joana Côrte-Real	Carlos Martínez-Angeles	Hitoshi Oi
Ricardo Rocha	Jorge Buenabad-Chávez	
Vítor Santos Costa		
Inês Dutra		
Alípio Jorge		
André Valente		

We acknowledge the *Comissão de Ética da Universidade do Porto* and the Real Hospital Português

Parallelization: grid stuff

- Aleph in grids
- Scheduling algorithms and grid portal
- People:
  - Odair Neves
  - Edgard Neto
  - João Rodrigues
  - Kiran Ali
  - Rafael Nunes
  - Hugo Figueiredo

Wrap Up

- Various machine learning and statistical methods applied to two medical domains
  - Focus on first-order models and graphical models
- Making machine learning efficient
  - Parallelization for multicores and GPUs

We are open to more adventures!

Thank you for your patience!