## Some Musings on Conjunctive Patterns

Couldn't decide on a subtitle 30/06/08, ML-Meeting

### The End of Comprehensibility

Couldn't decide on a subtitle 30/06/08, ML-Meeting

#### **Ensemble Trees**

- DS '08 submission
- Bastard child of
  - Tree<sup>2</sup>, CG-Clus
- Goal:
  - Stable trees, high accuracy
  - Better comprehensibility than ensembles

#### **Ensemble Trees**

- DS '08 submission
- Bastard child of
  - Tree<sup>2</sup>, CG-Clus
- Goal:
  - Stable trees, high accuracy

#### Better comprehensi

Comparison to bagged/boosted trees
ility than ensembles tions

- typically 100+ nodes!
- no performance jumps!

#### Better comprehensibility than ensembles

- Comparison to bagged/boosted trees
  - I0 iterations
  - typically 100+ nodes!
  - no performance jumps!

#### The competition

- Subsymbolic e.g. SVM
  - more accurate
- But: trees better comprehensibility!
- And: enough iterations (boosting) fix accuracy!



- Learn classifier
- Reweight/Resample misclassifications
- Repeat

 Can approximate classification function to arbitrary degree

### Comprehensible?

- Probably hundreds of nodes
- Trees model different subregions of data

Sounds like a trade-off (2)

### Comprehensible?

- Probably hundreds of nodes
- Trees model different subregions of data

- Sounds like a trade-off (2)
- Stability?
  - small changes in data  $\Rightarrow$  big changes in trees

### Comprehensible?

- Probably hundreds of nodes
- Trees model different subregions of data

- Sounds like a trade-off (2)
- Stability?
  - small changes in data  $\Rightarrow$  big changes in trees
- Let's not even talk about bagging



Ri Re R3









,,,,

Ri Re R'i R3 R"i R'e R"e R'e R"e R's R"s



$$\begin{array}{ccc} \mathcal{R}_{1} & \mathcal{R}_{1}^{'} & \mathcal{R}_{1}^{''} \\ \mathcal{R}_{2} & \wedge \left( \mathcal{R}_{2}^{'} \vee \mathcal{R}_{2}^{''} \right) \\ \mathcal{R}_{3} & \mathcal{R}_{3}^{'} & \mathcal{R}_{3}^{''} \end{array}$$

- Seems kinda hard to understand
- Decision tree much easier, right?

 $(A_1 = v \land A_2 = v) \lor (\neg A_1 = v \land A_3 = v)$ 

 $(A_1 = v \land A_2 = v) \lor (\neg A_1 = v \land A_3 = v)$ 

• Great! Like a rule set (CN2 et al)

 $(A_1 = v \land A_2 = v) \lor (\neg A_1 = v \land A_3 = v)$ 

- Great! Like a rule set (CN2 et al)
- Well not quite
  - DT rules mutually exclusive
  - No negation in CN2-rules

 $(A_1 = v \land A_2 = v) \lor (\neg A_1 = v \land A_3 = v)$ 

- Great! Like a rule set (CN2 et al)
- Well not quite
  - DT rules mutually exclusive
  - No negation in CN2-rules

 $1.A_1 = v \land A_2 = v$  $2.A_3 = v$ 

 $(A_1 = v \land A_2 = v) \lor (\neg A_1 = v \land A_3 = v)$ 

- Great! Like a rule set (CN2 et al)
- Well not quite
  - DT rules mutually exclusive
  - No negation in CN2-rules

 $1.A_1 = v \land A_2 = v$  $2.\neg (A_1 = v \land A_2 = v) \land A_3 = v$ 

 $(A_1 = v \land A_2 = v) \lor (\neg A_1 = v \land A_3 = v)$ 

- Great! Like a rule set (CN2 et al)
- Well not quite
  - DT rules mutually exclusive
  - No negation in CN2-rules

 $1.A_1 = v \land A_2 = v$  $2.A_4 = v \land A_5 = v$ 





 $A_1 = v$ 



 $\overline{A_1 = v \land A_2} = v$ 



 $\overline{A_1} = v \land \overline{A_2} = v$  $A_4 = v$ 



 $A_1 = v \land A_2 = v$  $A_4 = v \land A_5 = v$ 

### Enumeration Behavior (DT)















# **Enumeration Behavior** $(DT) \\ A_1 = v$ $A_2 = v$











 $A_1 = v \land A_2 = v$ 





$$A_1 = v \land A_2 = v$$

$$A = a \wedge A = -$$



$$A_1 = v \land A_2 = v$$













$$A_1 = v \land A_2 = v$$

 $A_1 = v \qquad \land \quad A_2 = v$ 







$$A_1 = v \land A_2 = v$$



















$$A_1 = v \land A_2 = v$$









$$\neg A_1 = v$$





Tree-structure

'bout that Comprehensibility... • Approximate "incomprehensible" classifiers Becoming same in process Trees model only training data Structure needed for interpretation

 $\Rightarrow$  Comprehensibility not selling point

# Towards new ensembles

- Heed the ETs
- Not short hand for DNF



- Same for dendograms, structure-trees
- Organizing local patterns into set/model