ECML PKDD 2021
VIRTUAL
13-17 September

Landscaping MLC through meta learning on the empirical results

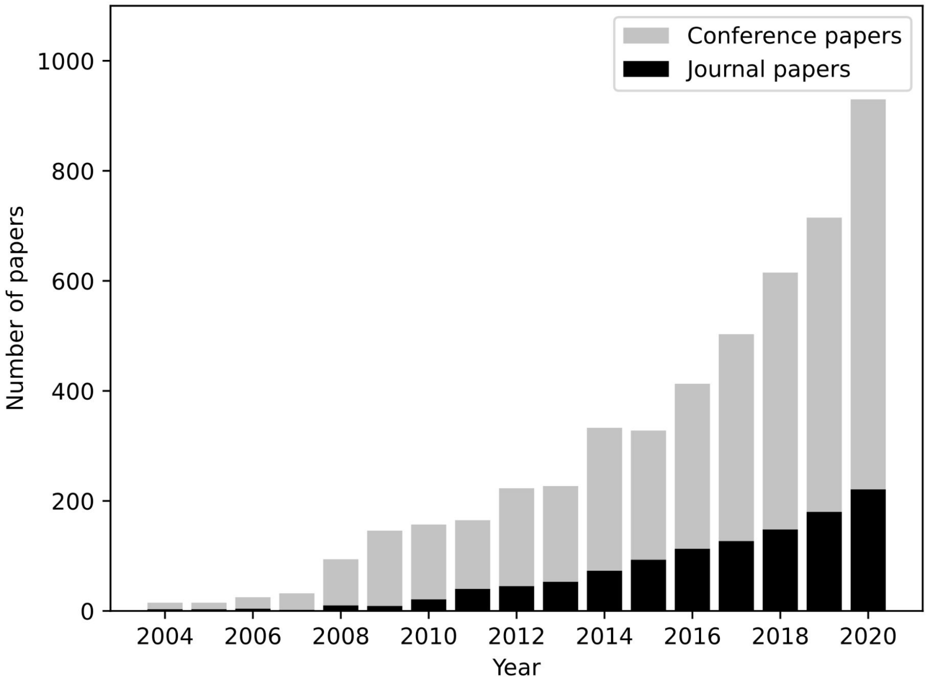
Dragi Kocev, Jasmin Bogatinovski, Ana Kostovska, Panče Panov

ECML PKDD 2021 Tutorial: FAIR multi-label classification
17 September 2021

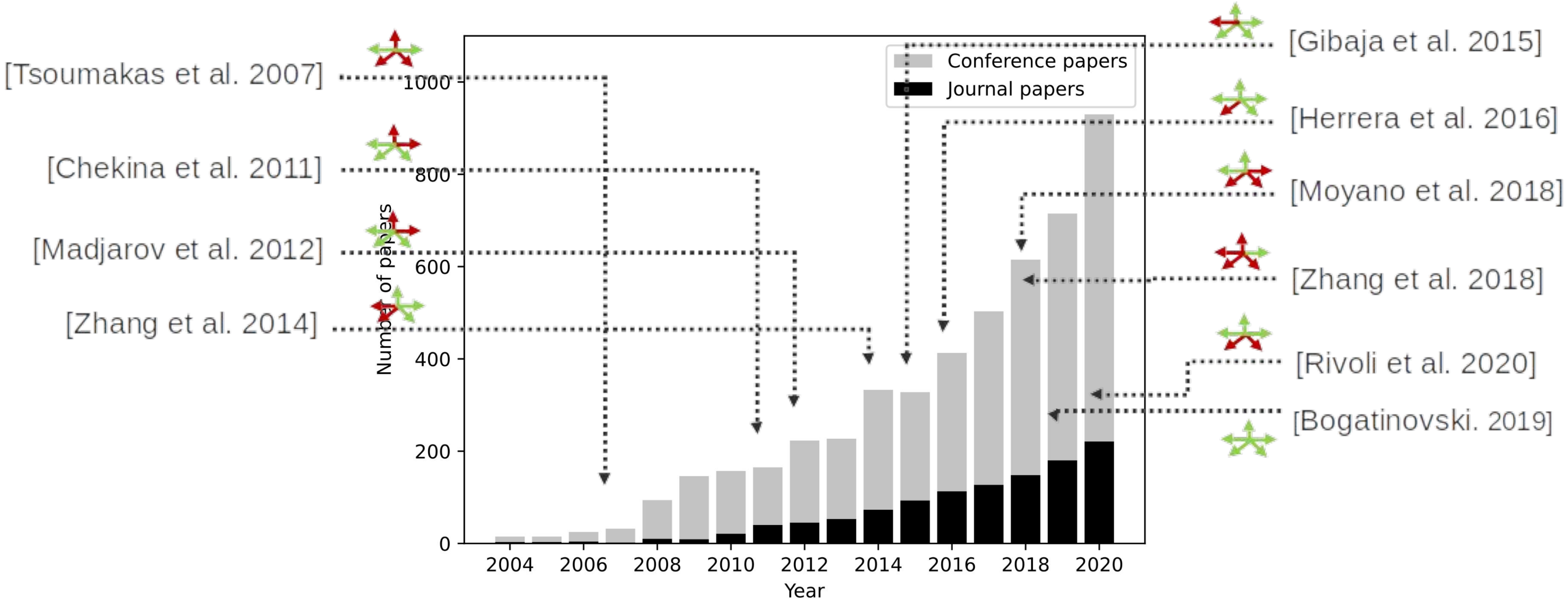
outline

- Motivation and introduction
- Meta-learning for MLC
- Do meta features outline the dataset space?
- Are meta features indicative of (predictive) performance?
- To tune or not to tune?
- Summary

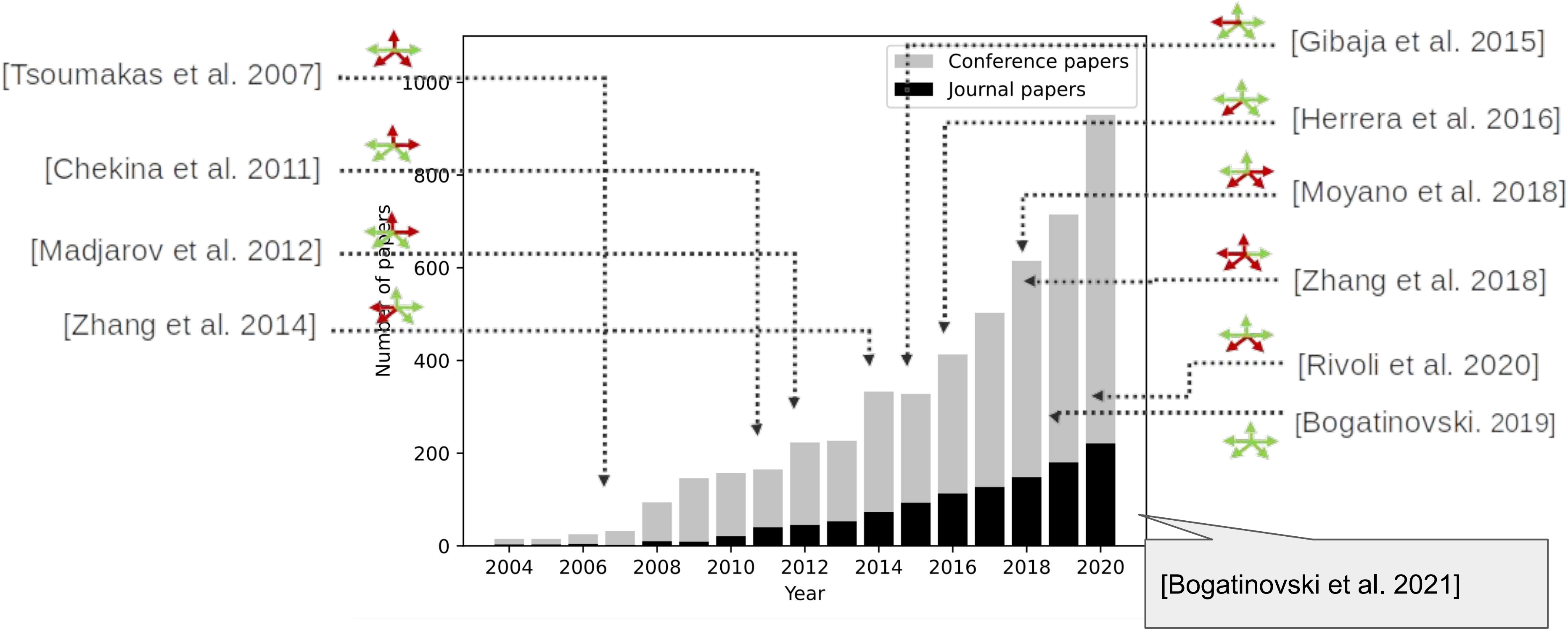
Growing body of work targeting MLC



Growing body of work targeting MLC



Growing body of work targeting MLC



The exponential explosion of MLC papers requires

1. Proper **benchmarking**,
2. **Reusability** of previous results and
3. Better **understanding** of the proposed novel methods and the problems addressed with them.

The exponential explosion of MLC papers requires

1. Proper **benchmarking**,
2. **Reusability** of previous results and
3. Better **understanding** of the proposed novel methods and the problems addressed with them.

Understanding through meta-learning!

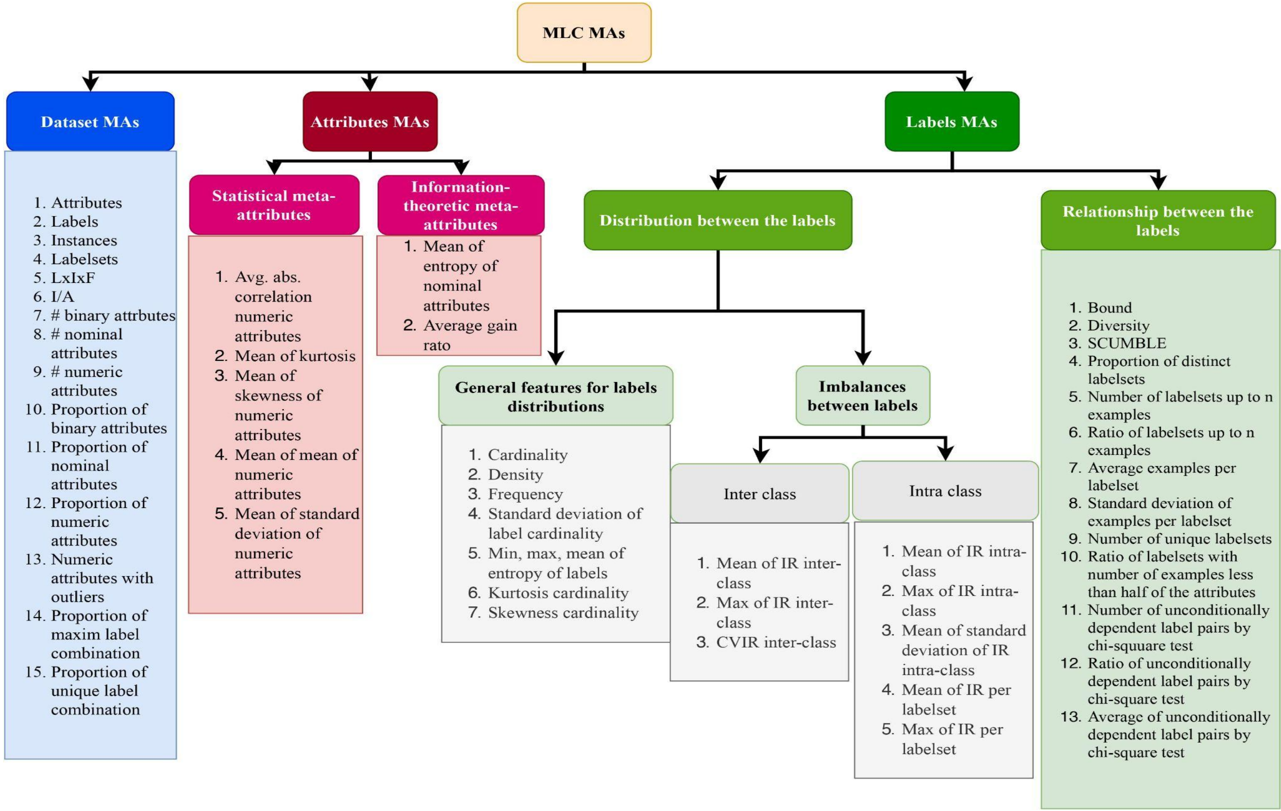
"*A meta-learning system must include a learning subsystem, which adapts with experience. Experience is gained by exploiting meta knowledge extracted: (a) in a previous learning episode on a single data set and/or (b) from different domains or problems*" (Lemke et al., 2015)

- **Meta knowledge** is typically presented with **meta data** describing the data sets and the performance of the methods on past and available data sets (Brazdil et al, 2009).
- The body of meta knowledge is then enriched with the **new experience** gained with the application of the meta system to new data sets (Brazdil et al., 2018).

In a nutshell, meta learning allows for transferring the experience obtained from available problems to a novel problem by learning meta models from the meta knowledge.

Meta-learning for MLC

1. Descriptions of the datasets through meta-features
2. Performance assessment of methods over datasets
3. Learn meta models encapsulating the meta knowledge

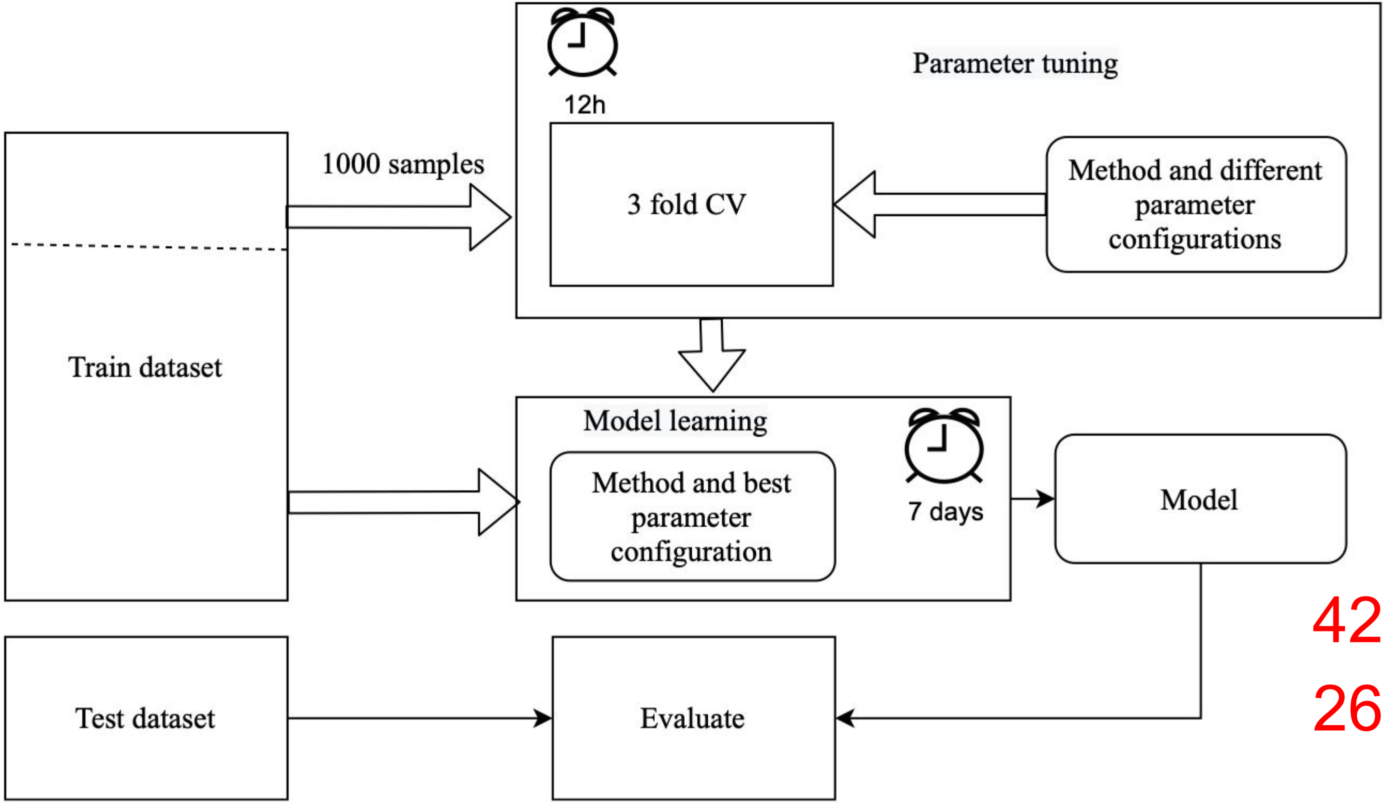


Approaches to meta-learning for MLC

- Moyano et al. 2017, 2018: definition of a set of meta-features for MLC, and analysis of ensembles of MLC methods (12 methods over 20 datasets) using 4 meta features
- Chekina et al. 2011: looking for the most suitable method for a new unseen MCL data set. Experimental study with 12 MLC datasets augmented to 640 variations of datasets. Study of 7 single and 4 ensemble methods. k-NN as a meta learner.

- Beyond the existing body of work:
  - **Size**: Much more comprehensive study in terms of datasets and methods,
  - **Scope**: Parameter selection of the base methods,
  - **Understanding**: Multi-target trees as meta learners.

Meta-analysis of the experimental study



42 datasets and
26 methods

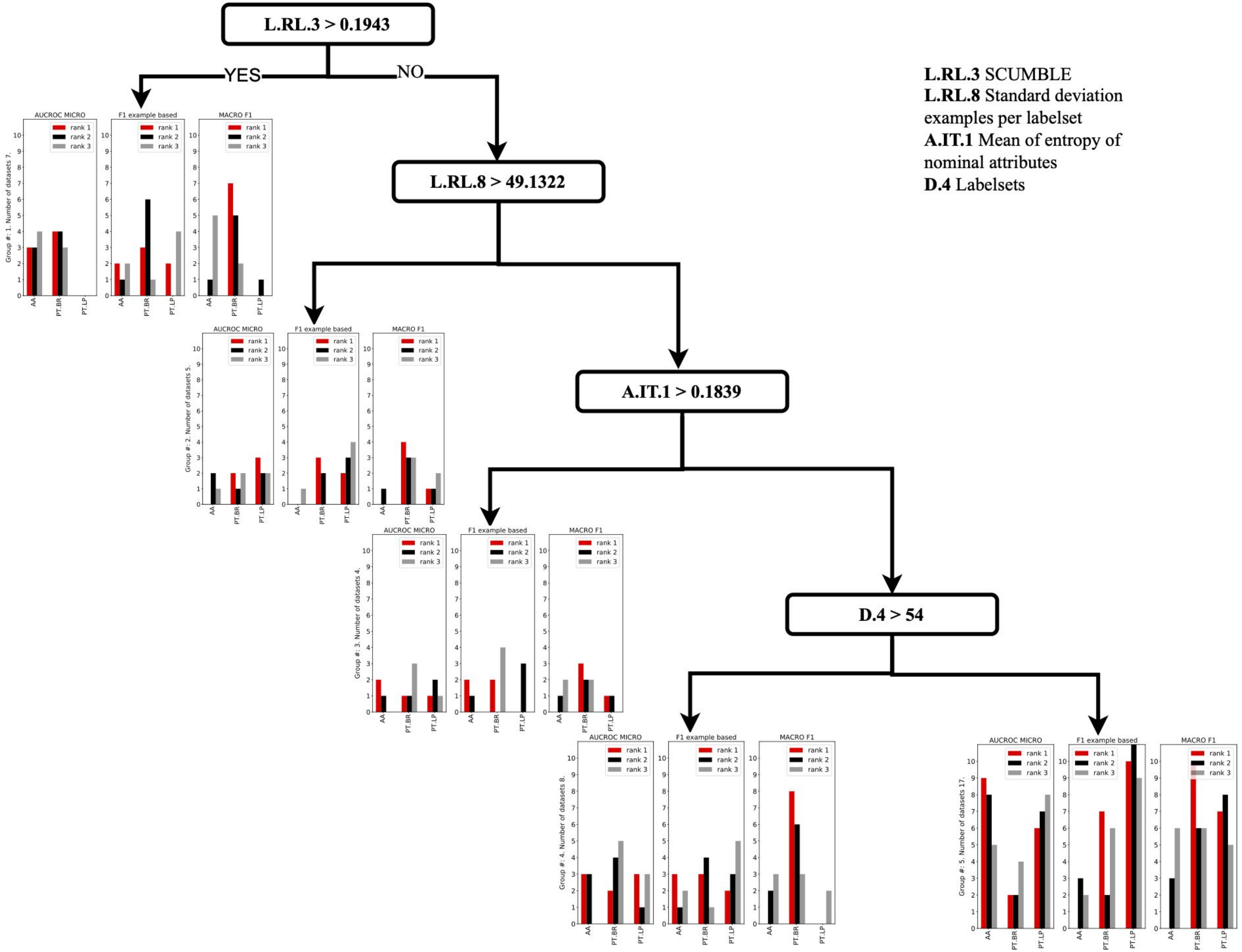
Meta-learning questions of interest

1. What is the potential of the meta features to describe the space of MLC datasets?
2. Whether and how the meta features are related to the predictive performance of the MLC methods?
3. Does tuning of MLC methods improves their predictive performance?

Description of the space of MLC datasets

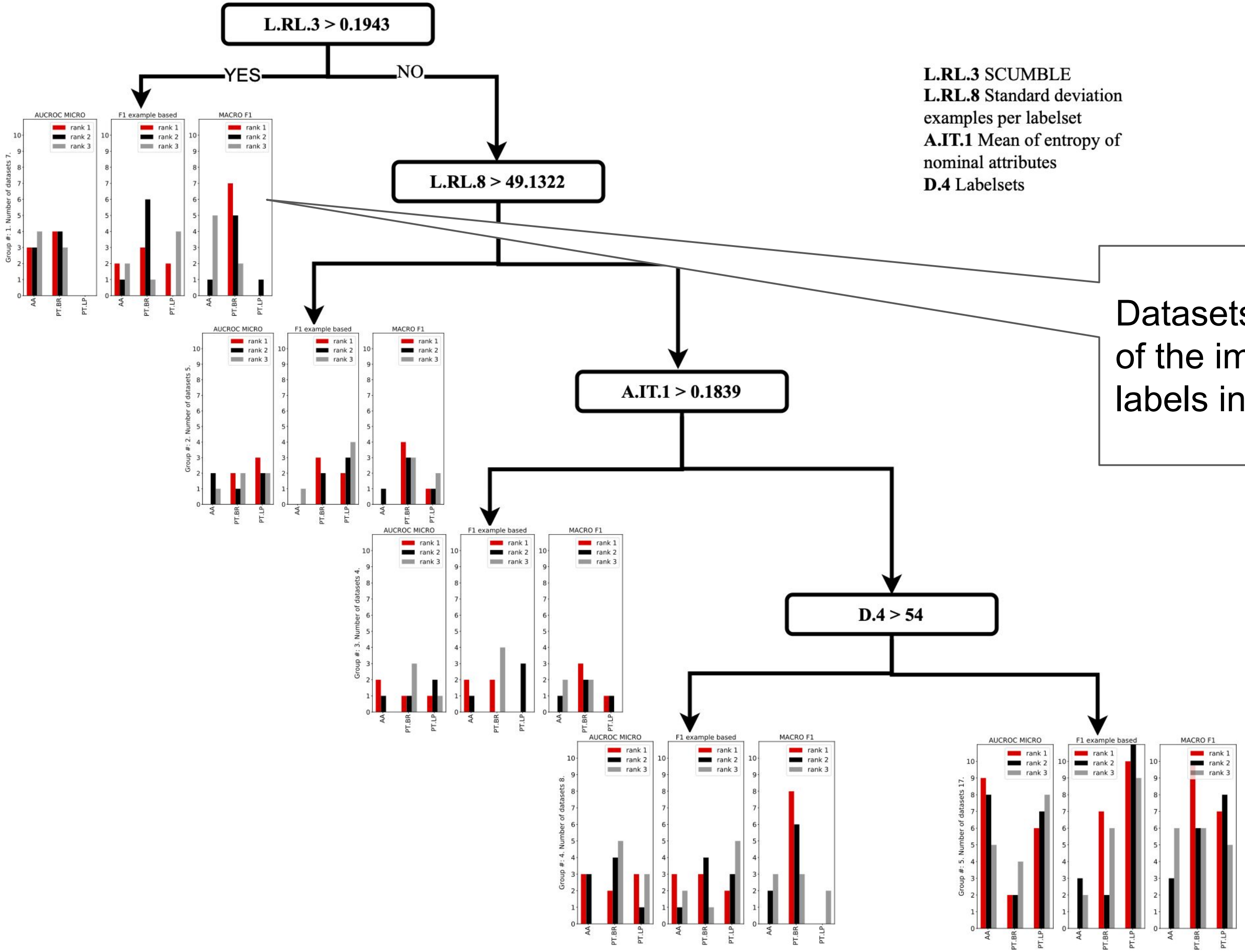
- Descriptive power of the meta features
    - 50 meta-features
- Use them in an unsupervised setting
- Goal: What are the main meta features distinguishing the different datasets!

Divisive clustering tree of the 40 data sets.



L.RL.3 SCUMBLE
L.RL.8 Standard deviation examples per labelset
A.IT.1 Mean of entropy of nominal attributes
D.4 Labelsets

Divisive clustering tree of the 40 data sets.

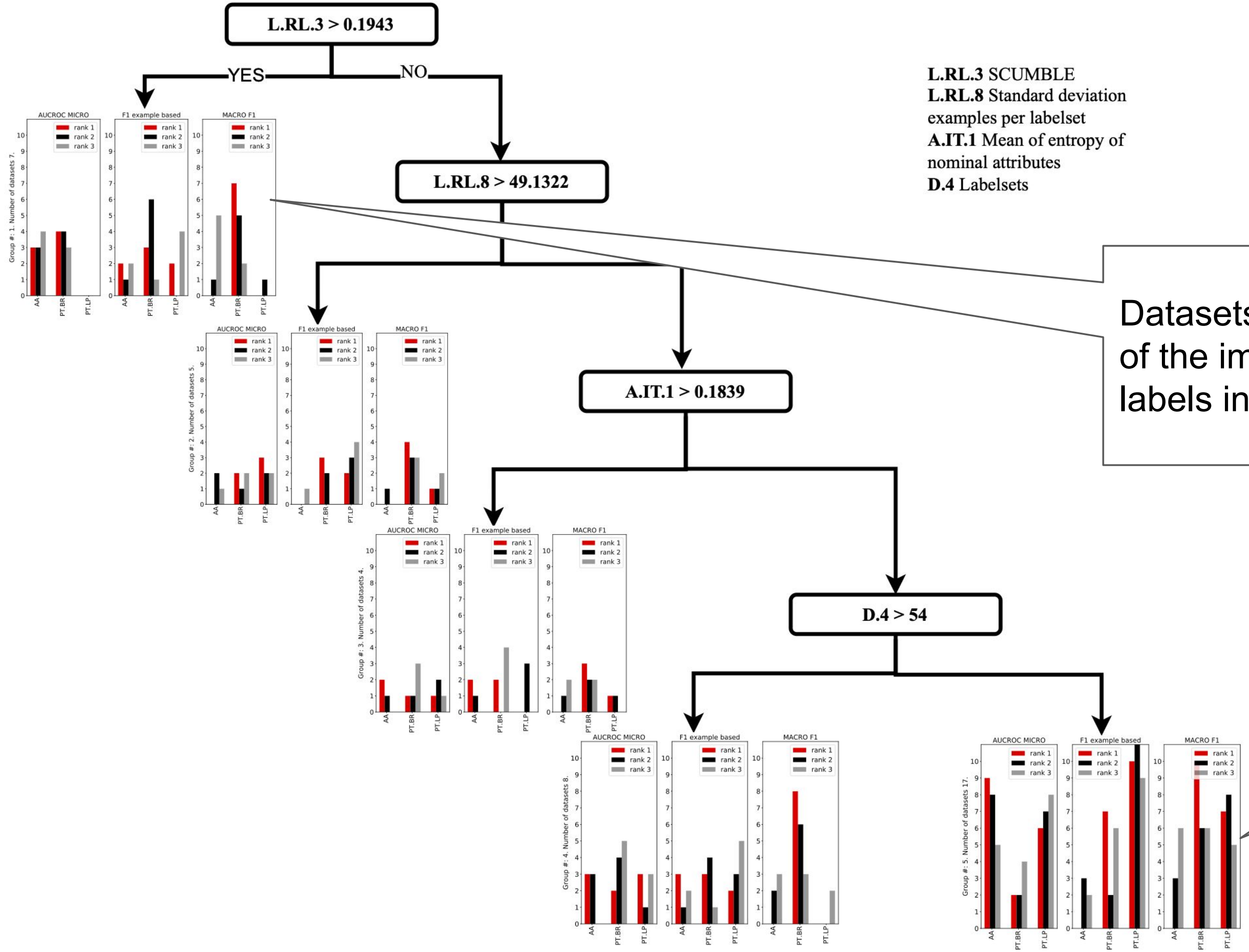


L.RL.3 SCUMBLE
L.RL.8 Standard deviation examples per labelset
A.IT.1 Mean of entropy of nominal attributes
D.4 Labelsets

Datasets with larger variation of the imbalance among the labels in the examples.

Divisive clustering tree of the 40 data sets.

L.RL.3 SCUMBLE
L.RL.8 Standard deviation examples per labelset
A.IT.1 Mean of entropy of nominal attributes
D.4 Labelsets



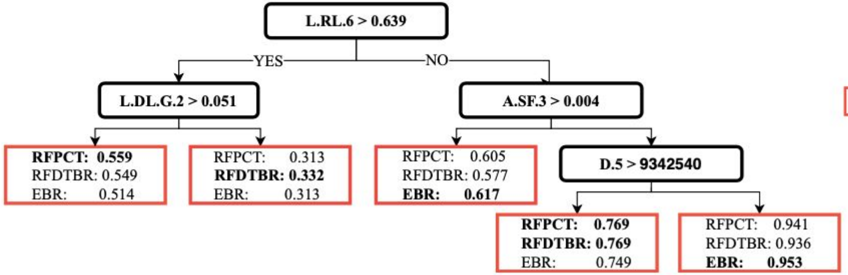
Datasets with larger variation of the imbalance among the labels in the examples.

Datasets with small number of label sets and well balanced distribution of labels.

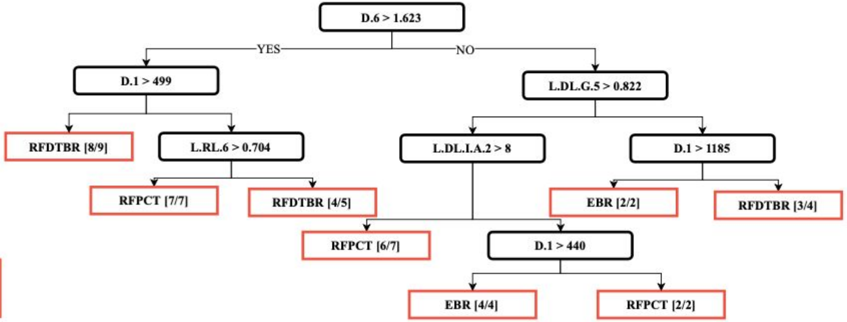
Relating meta-features with performance

- Selected 3 methods for performance analysis
  - RF-PCT, RFDTBR, EBRJ48
- Selected 5 evaluation measures
  - AUROC.micro, F1.example-based, Hamming Loss, F1.macro and F1.micro
- Meta models: Multi-target regression trees
- Learning scenarios
  - Predict performance of the selected methods methods
  - Predict the best performing method

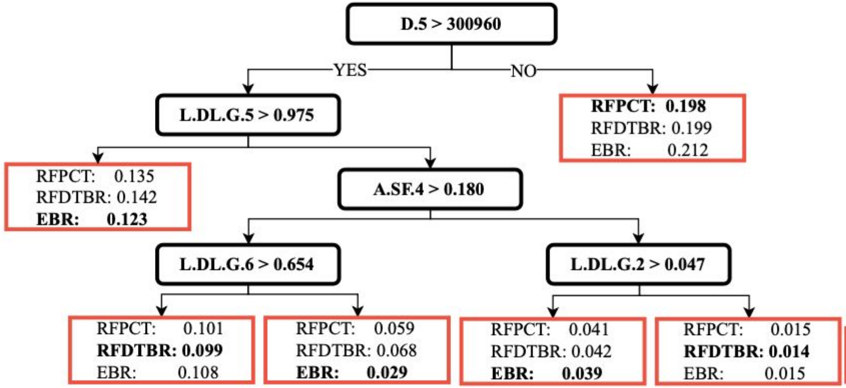
Features provide insights into the intricate interplay of dataset properties and methods



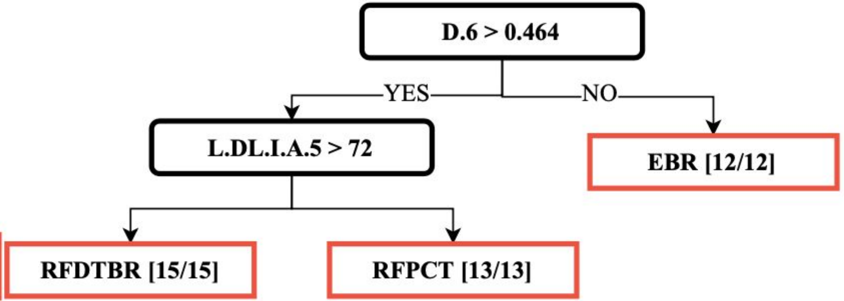
(a) F1.example-based (MTR)



(b) F1.example-based (ST)



(c) Hamming Loss (MTR)

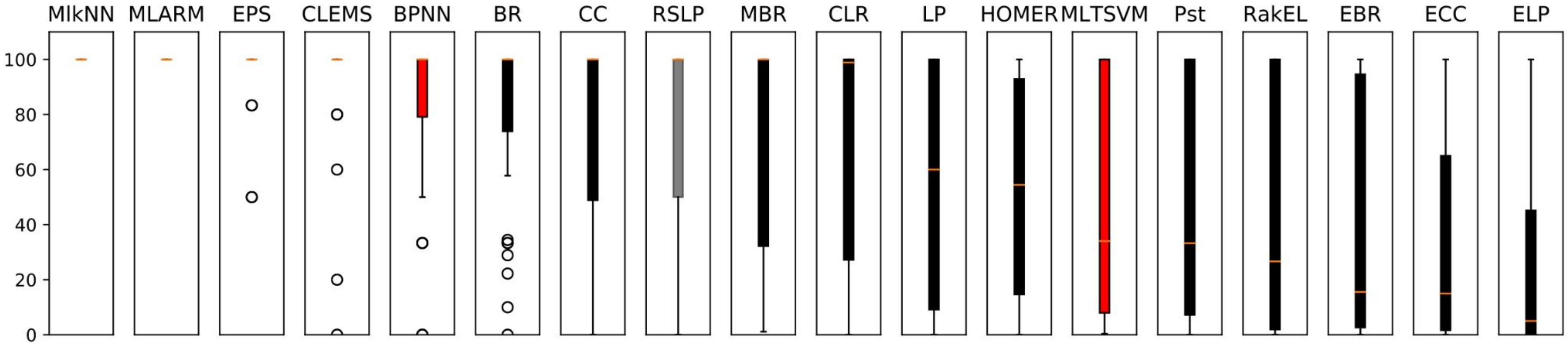


(d) Hamming Loss (ST)

The need for tuning of the parameters

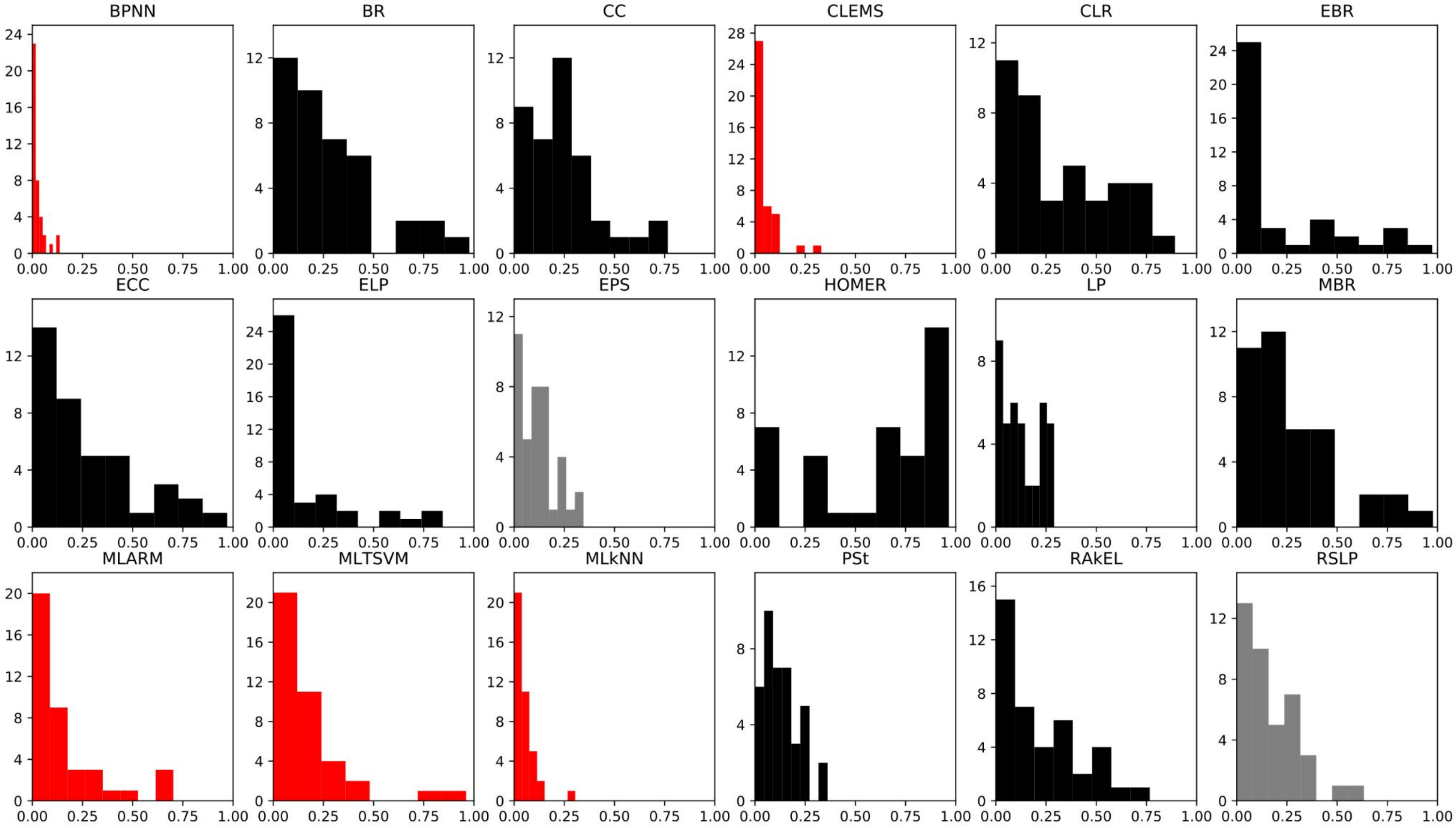
1. Coverage of the experimental space
2. Sensitivity to the parameter tuning
3. To tune or not to tune

Ratio of successful experiments

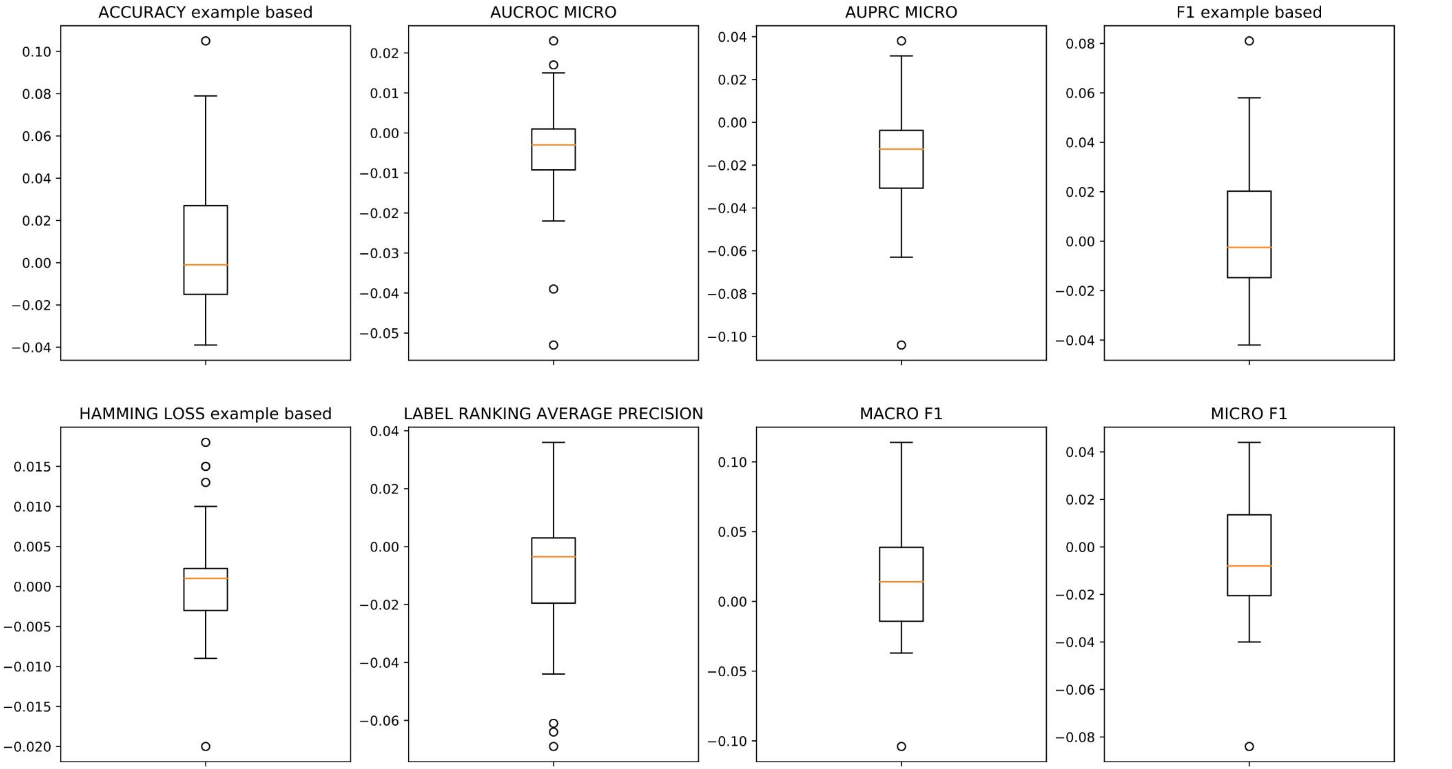


- Coverage of the available experimental space
- Algorithm adaptation methods explore more than problem transformation

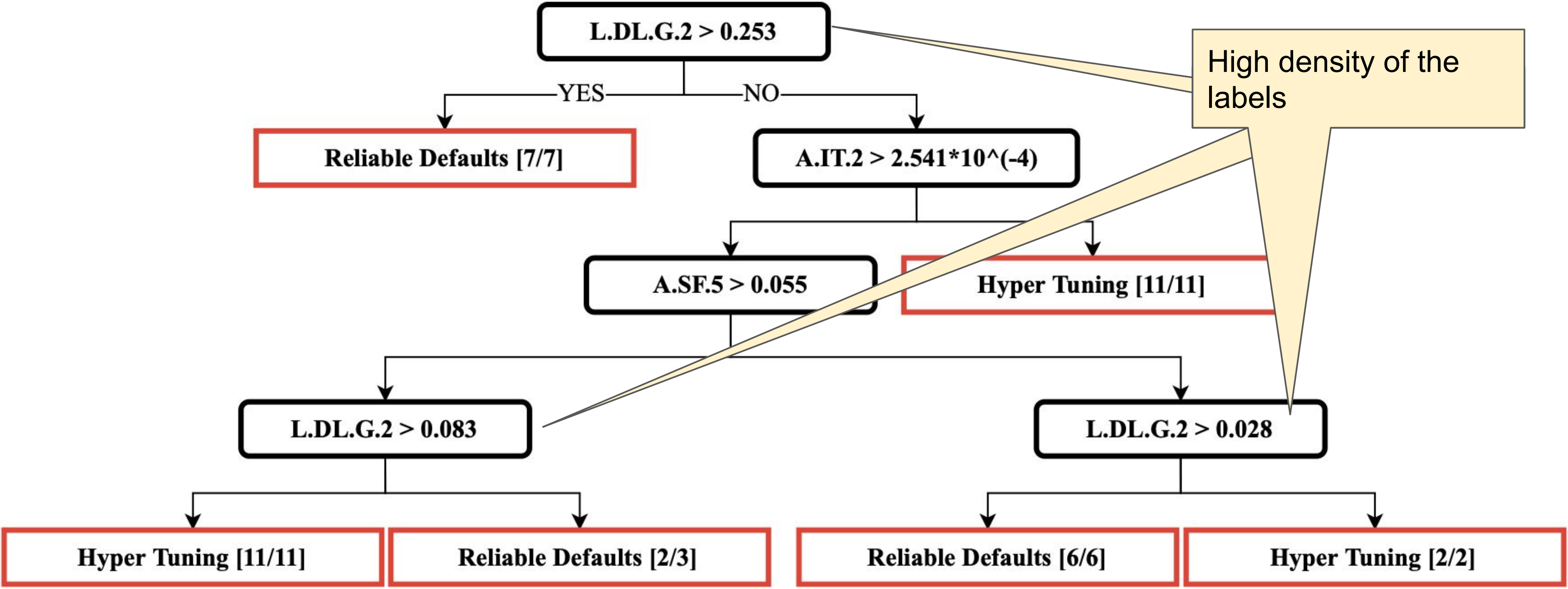
Sensitivity to the parameter tuning (Hamming Loss)



To tune or not to tune:
Absolute difference between reliable defaults and tuned



To tune or not to tune: reliable defaults vs tuned



Summary

1. The meta features paint a very interesting landscape of the MLC datasets and identify the “domains of expertise ” of the MLC methods
2. The meta models obtained in the study are easily understandable and can be used for making predictions for novel datasets
3. The Meta-features related to the labels are the driving force behind the landscape of MLC methods and datasets
4. Methods containing base models sensitive to parameters (e.g., SVM) should always be tuned

Read more...

1. Comprehensive Comparative Study of Multi-Label Classification Methods, Jasmin Bogatinovski, Ljupčo Todorovski, Sašo Džeroski, Dragi Kocev, 2021, https://arxiv.org/abs/2102.07113

2. Explaining the Performance of Multi-label Classification Methods with Data Set Properties, Jasmin Bogatinovski, Ljupčo Todorovski, Sašo Džeroski, Dragi Kocev, 2021, https://arxiv.org/abs/2106.15411

Thank you!

For more details, visit:

- http://mlc.ijs.si
- http://mlc.ijs.si/fair-mlc-ecml-2021/
- http://semantichub.ijs.si/MLCdatasets

Reach out to dragi.kocev@ijs.si