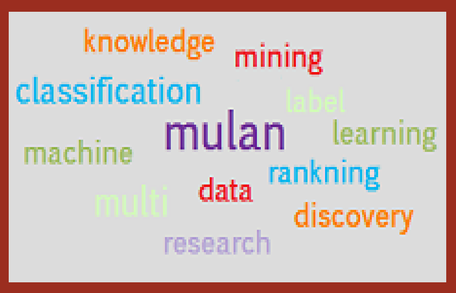

Mulan: A Java Library for Multi-Label Learning

G. Tsoumakas, E. Spyromitros-Xioufis, J. Vilcek, I. Vlahavas

Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece

Multi–Label Learning

Training Set

Training Attributes										Target Concepts				
Frequent Words						Image Features				Politics	Sports	Culture	Science	Health
1	0	0	1	1	0	0.3	0.5	0.2	0.8	True	False	True	False	False

Test Article

World Cup officials asked to ban Vuvuzela noise from games

Wednesday, people around the world are asking what is a **Vuvuzela** and why does it sound so annoying? **World Cup** officials have been asked to ban the Vuvuzela noise makers from the World Cup games in **Africa**, and they have said they will not.

Prediction

Scores

Politics	Sports	Culture	Science	Health
0.3	0.9	0.6	0.2	0.4

Ranking

1st rank	2nd rank	3rd rank	4th rank	5th rank
Sports	Culture	Health	Politics	Science

Bipartition

Relevant	Irrelevant
{Sports, Culure}	{Health, Politics, Science}

Applications

- ✓semantic annotation of images and video
- ✓web page categorization
- ✓direct marketing
- ✓functional genomics
- ✓music categorization into genres and emotions

Mulan at a Glance

The library includes a variety of state-of-the-art algorithms for performing the following major multi-label learning tasks:

- ✓Classification: A bipartition of the labels into relevant and irrelevant for a given instance.
- ✓Ranking: An ordering of the labels, according to their relevance for a given data item.
- ✓Classification and ranking: A combination of the two tasks mentioned-above.

In addition, the library offers the following features:

- ✓Dimensionality reduction: Simple baseline methods are currently supported.
- ✓Evaluation: Classes that calculate a large variety of evaluation measures through hold-out evaluation and cross-validation.

Data Format

ARFF file

```
@relation MultiLabelExample

@attribute feature1 numeric
@attribute feature2 {0,1}
@attribute amazed {0, 1}
@attribute happy {0, 1}
@attribute relaxing {0, 1}
@attribute sad {0, 1}

@data
2.3,5.6,1.4,0,1,1,0,0
```

XML file

```
<?xml version="1.0" encoding="utf-8"?>
<labels
xmlns="http://mulan.sourceforge.net/labels">
   <label name="amazed" />
   <label name="happy" />
   <label name="relaxing" />
   <label name="sad" />
</labels>
```

XML file with label hierarchies

```
<?xml version="1.0" encoding="utf-8"?>
<labels
xmlns="http://mulan.sourceforge.net/labels">
   <label name="sports">
        <label name="football" />
        <label name="basketball" />
    </label>
   <label name="arts">
        <label name="sculpture" />
        <label name="photography" />
    </label>
</labels>
```

- ✓ All labels specified in the XML file must be also defined in the ARFF file with same name
- ✓ Label names must be unique
- ✓ Each ARFF label attribute must be nominal with binary values {0, 1}
- ✓ Data should be consistent with the hierarchy
  - ✓ If any child label appears at an example, then all parent labels must also appear

Multi-Label Learners

- Lazy
  - MLkNN
  - BRkNN
  - IBLR_ML
- Neural
  - BPMLL
  - MMPLearner
- Meta
  - RAkEL
  - RAkELd
  - HOMER
  - HMC
  - ClusteringBased
  - OneThreshold
- Transformation
  - BinaryRelevance
  - CalibratedLabelRanking
  - ClassifierChains
  - EnsembleOfPrunedSets
  - IncludeLabelsClassifier
  - LabelPowerset
  - MultiLabelStacking
  - PPT
  - PrunedSets

Evaluation Measures

	Example – Based	Label - Based
Scores	Coverage IsError Ranking Loss One error	AUC MAP
Bipartition	Subset Accuracy Hamming Loss Accuracy Precision Recall FMeasure Hierarchical Loss	Precision Recall FMeasure

Train – Test Example

```
MultiLabelInstances train, test;
train = new MultiLabelInstances("train.arff", "format.xml");
test = new MultiLabelInstances("test.arff", "format.xml");

Classifier base = new NaiveBayes();
BinaryRelevance br = new BinaryRelevance(base);
br.build(train);

Evaluator eval = new Evaluator();
Evaluation results = eval.evaluate(br, test);
System.out.println(results);
```

Sample Output

```
Hamming Loss: 0.2257
Subset Accuracy: 0.2784
Example-Based Precision: 0.4243
Example-Based Recall: 0.7303
Example-Based F Measure: 0.6574
Example-Based Accuracy: 0.5644
Micro-averaged Precision: 0.6428
Micro-averaged Recall: 0.7224
Micro-averaged F-Measure: 0.6797
Macro-averaged Precision: 0.6426
...
```

Future Goals

- ✓ Command Line / Graphical User Interface
- ✓ Additional Algorithms
  - Classification, Ranking, Classification & Ranking
  - Dimensionality reduction
  - Thresholding strategies
- ✓ Support for extensions of the main learning tasks
  - Active Learning
  - Semi-supervised learning
- ✓ Package with experimental setups from published papers

References

1. Tsoumakas, G., Katakis, I., Vlahavas, I. (2010) "Mining Multi-label Data", Data Mining and Knowledge Discovery Handbook, O. Maimon, L. Rokach (Ed.), Springer, 2nd edition, 2010.