

ADAMS

Advanced **D**ata mining **A**nd **M**achine learning **S**ystem

Module: adams-meka

MEKA

Peter Reutemann

December 20, 2017

©2014-2016



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato



Except where otherwise noted, this work is licensed under
<http://creativecommons.org/licenses/by-sa/4.0/>

Contents

| | | |
|----------|---------------------|-----------|
| 1 | Introduction | 7 |
| 2 | Flow | 9 |
| 3 | Tools | 13 |
| | Bibliography | 15 |

List of Figures

| | | |
|-----|----------------------------------------------------------|----|
| 2.1 | Flow for cross-validating a MEKA classifier. | 10 |
| 2.2 | The result of a cross-validated MEKA classifier. | 11 |
| 3.1 | MEKA Explorer | 13 |

Chapter 1

Introduction

The *adams-meka* module integrates the MEKA [2] software suite in ADAMS. From the MEKA homepage:

“The MEKA project provides an open source implementation of methods for multi-label classification and evaluation. It is based on the WEKA Machine Learning Toolkit. Several benchmark methods are also included, as well as the pruned sets and classifier chains methods, other methods from the scientific literature, and a wrapper to the MULAN framework.”

Chapter 2

Flow

The following source actors are available:

- *MekaClassifierSetup* – outputs a MEKA classifier setup.

The following transformers are available:

- *MekaClassifying* – generates predictions on Instance objects using a trained model.
- *MekaClassSelector* – can turn any Instances object into a MEKA dataset by choosing which attributes to use as class attributes.
- *MekaCrossValidationEvaluator* – cross-validates a MEKA classifier on the incoming dataset.
- *MekaPrepareData* – prepares Instances objects for the use with MEKA classifiers.
- *MekaResultSummary* – turns the collected statistics from a MEKA evaluation run into a string.
- *MekaResultValues* – picks out selected statistics from a MEKA evaluation and generates a spreadsheet.
- *MekaTrainClassifier* – builds a MEKA classifier on the incoming dataset.
- *MekaTrainTestSetEvaluator* – evaluates a MEKA classifier on the incoming train/test split.

The following sinks are available:

- *MekaGraphVisualizer* – displays graphs obtained from a model.
- *MekaMacroCurve* – displays macro-averaged curve.
- *MekaMicroCurve* – displays micro-averaged curve.
- *MekaPrecisionRecall* – displays precision-recall plots.
- *MekaROC* – displays ROC (receiver operator curve) plots.

Figures 2.1 and 2.2 show a flow and its associated output of cross-validating a MEKA classifier.

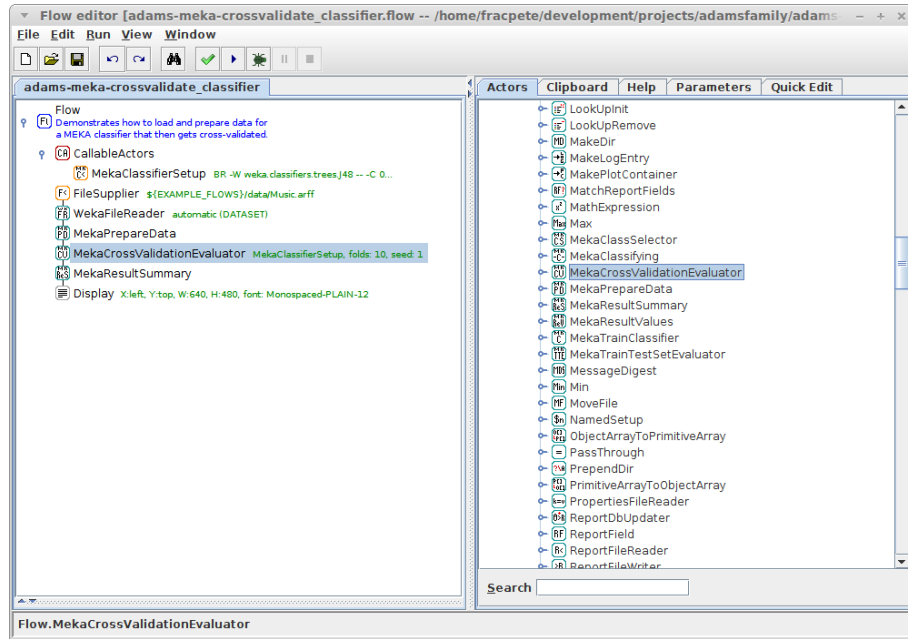
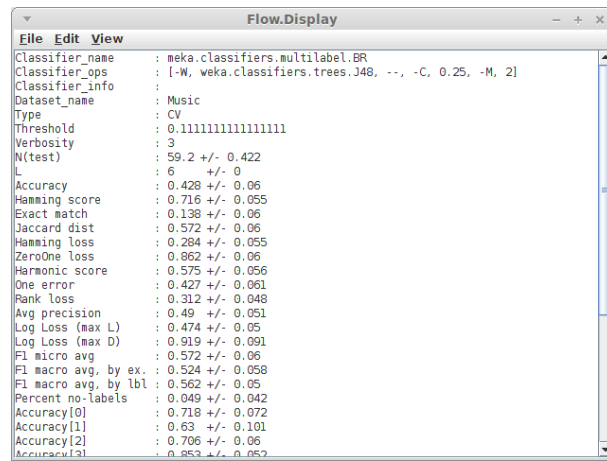


Figure 2.1: Flow for cross-validating a MEKA classifier.



The image shows a screenshot of a software window titled "Flow.Display". It contains a list of performance metrics for a cross-validated MEKA classifier. The metrics are listed in two columns, with the first column containing the metric name and the second column containing the value and standard deviation. The metrics include Classifier_name, Classifier_ops, Classifier_info, Dataset_name, Type, Threshold, Verbosity, N(test), L, Accuracy, Hamming score, Exact match, Jaccard dist, Hamming loss, ZeroOne loss, Harmonic score, One error, Rank loss, Avg precision, Log Loss (max L), Log Loss (max D), F1 micro avg, F1 macro avg, Percent no-labels, and Accuracy for each of the four classes (0, 1, 2, 3).

| Metric | Value |
|----------------------|---------------------------------------------------------|
| Classifier_name | : meka.classifiers.multilabel.BR |
| Classifier_ops | : [-W, weka.classifiers.trees.J48, --, -C, 0.25, -M, 2] |
| Classifier_info | : |
| Dataset_name | : Music |
| Type | : CV |
| Threshold | : 0.1111111111111111 |
| Verbosity | : 3 |
| N(test) | : 59.2 +/- 0.422 |
| L | : 6 +/- 0 |
| Accuracy | : 0.428 +/- 0.06 |
| Hamming score | : 0.716 +/- 0.055 |
| Exact match | : 0.138 +/- 0.06 |
| Jaccard dist | : 0.572 +/- 0.06 |
| Hamming loss | : 0.284 +/- 0.055 |
| ZeroOne loss | : 0.862 +/- 0.06 |
| Harmonic score | : 0.575 +/- 0.056 |
| One error | : 0.427 +/- 0.061 |
| Rank loss | : 0.312 +/- 0.048 |
| Avg precision | : 0.49 +/- 0.051 |
| Log Loss (max L) | : 0.474 +/- 0.05 |
| Log Loss (max D) | : 0.919 +/- 0.091 |
| F1 micro avg | : 0.572 +/- 0.06 |
| F1 macro avg, by ex. | : 0.524 +/- 0.058 |
| F1 macro avg, by lbl | : 0.562 +/- 0.05 |
| Percent no-labels | : 0.049 +/- 0.042 |
| Accuracy[0] | : 0.718 +/- 0.072 |
| Accuracy[1] | : 0.63 +/- 0.101 |
| Accuracy[2] | : 0.706 +/- 0.06 |
| Accuracy[3] | : 0.852 +/- 0.052 |

Figure 2.2: The result of a cross-validated MEKA classifier.

Chapter 3

Tools

The *MEKA Explorer* can be launched from the main menu. It is located in the *Machine learning* section.

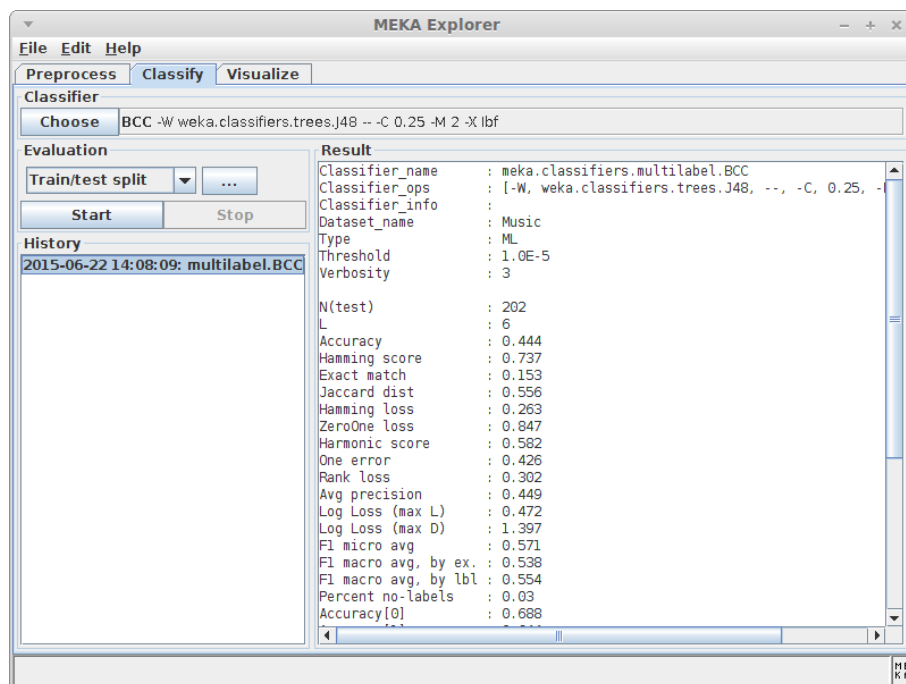


Figure 3.1: MEKA Explorer

Bibliography

- [1] *ADAMS* – Advanced Data mining and Machine learning System
<https://adams.cms.waikato.ac.nz/>
- [2] *MEKA* – A Multi-label Extension to WEKA
<http://meka.sourceforge.net/>
- [3] *MULAN* – A Java Library for Multi-Label Learning
<http://mulan.sourceforge.net/>