Rapid Miner

- RapidMiner is a tool for experimenting with machine learning and data mining algorithms
- An experiment is a set of operators that perform different tasks in the data
  - data input/output, data transformation, preprocessing, attribute selection, learning, evaluation, ...
- The experiments can be described visually as a process
Rapid Miner - First contact

Initializing Operators
RapidMiner 5.0
Licensed under the AGPL version 3
RapidMiner comes with absolutely no warranty
Copyright (C) 2001 - 2010 by Rapid-i and contributors
More information at http://rapid-i.com
Rapid Miner - Setting a repository

New Repository

Enter parameters to create a new local or remote repository.

- [ ] New local repository
- [ ] New remote repository
Rapid Miner - Setting a repository

![New Repository window]

Enter parameters to create a new local or remote repository.

**Alias:** NewLocalRepository

**Root directory:**

[Previous] [Next] [Finish] [Cancel]
Rapid Miner - Setting a repository

Select Directory

Please select a directory.

/home/bejar/RM_Repository

Open  Cancel
Rapid Miner - Perspectives

At the Rapid Miner Community Meeting and Conference (INCOMME 2010), the new RapidMiner extension for the integration of R will be presented for the first time.

However, data analysts do not have to wait until the Rapid Miner conference in order to get a first look at what can be done with the new extension. The following video shows how easy it models and scripts can be integrated into the RapidMiner analysis processes.

“We make free software affordable.”

According to the founder John Gilmore, this slogan was written on the support T-shirt of Cygnus, the first company offering professional services around open source products. Following this motto, the RapidMiner Enterprise Edition now allows you to deploy the most widely used open source data mining solution with confidence, security, and for lower total cost of ownership (TCO) than any other alternative.

Using the free Community Edition of RapidMiner not only increases your risk of failure, it turns out to be more expensive. The RapidMiner Enterprise Edition delivers critical benefits like stabilization software, direct access to product expertise, and committed response times to help you save time and money. In the end, staying RapidMiner maintains the software is cheaper than employing a specialist to do the work for.

Promotion Offer!

RapidMiner Enterprise Edition
Small, Standard, or Developer
Get Info Here!

Courses & Webinars
- Fast Track Introduction into Data Analysis with RapidMiner (Webinar)
- Text Mining with RapidMiner
- Data Management & M尸lification Load and Transform your Data (Webinar)
- Sentiment Analysis: Opinion Mining, and Automatic Market Research
- Time Series Analysis with
Rapid Miner - New project

Repository Browser

Select a repository location.

- Samples (none)
- NewLocalRepository (bejar)
Rapid Miner - Process perspective
Rapid Miner - Repository

- Operators
- Repositories

- Samples (none)
  - data (none)
    - Golf (none &ndash; v1)
    - Golf-Testset (none &ndash; v1)
    - Iris (none &ndash; v1)
    - Labor-Negotiations (none &ndash; v1)
    - Market-Data (none &ndash; v1)
    - Polynomial (none &ndash; v1)
    - Ripley-Set (none &ndash; v1)
    - Sonar (none &ndash; v1)
    - Transactions (none &ndash; v1)
    - Weighting (none &ndash; v1)

- processes (none)
  - 01_Learner (none)
  - 02_Preprocessing (none)
  - 03_Validation (none)
  - 04_Attributes (none)
  - 05_Visualisation (none)
  - 06_Meta (none)
  - 07_Clustering (none)
  - 08_Other (none)

- NewLocalRepository (Dejal)
Rapid Miner - Operators

**Synopsis**
The root operator which is the outermost operator of every process.

**Description**
Each process must contain exactly one operator of this class, and it must be the root operator of the process. This operator provides a set of...
Rapid Miner - Adding operators

![Image of RapidMiner interface with a Retrieve operator highlighted.]
Rapid Miner - Connecting operators

Process is now associated with file //NewLocalRepository/test.
Sep 20, 2010 12:30:48 PM CONFIG: Saving property rapidminer.gui.resolve_relative_repository_locations=true
Rapid Miner - Complex Process

![Rapid Miner Interface]

- **Application Configuration**
  - Parameters: Application Parameter
  - Edits List (Win32)

- **Process Flow**
  - **Main Process**
    - **Read ARFF**
    - **Set Role**
    - **Split Data**
    - **Apply Model**
    - **Decision Tree**
    - **Perform...**

- **Operators**
  - **Filter**
    - Role Initialization (5)
    - Neural Net Training (2)
    - Function Fitting (27)
    - Logistic Regression (3)
    - Support Vector Modeling (6)
    - Discriminant Analysis (1)
    - Merging and Analysis (13)
    - Data Transformation (131)
    - Word (110)
    - Attribute Weighting (35)
    - Clustering and Segmentation (110)
    - Association and Item Set Mining (110)
    - Correlation and Dependency Computation (4)
    - Model Application (111)
      - Thresholds (15)
      - Confidences (12)
      - Apply Model
      - Update Model
      - Group Models
      - Ungroup Models
      - Generate top K Predictions
      - Create Formula
    - Evaluation (15)
      - Validation (7)
      - Performance Measurement (115)

- **Description and Synopsis**
  - **Apply Model**
    - **Synopsis**
      - Applies a model to an example set
      - This might be a prediction or another data transformation model
    - **Description**
      - This operator applies a model to an
Rapid Miner - Results Perspective
Rapid Miner - Results Perspective
Rapid Miner - Results Perspective