MORET - A Software For Model Management

Ralf Seger   Antony Unwin

Institut für Mathematik
Rechnerorientierte Statistik und Datenanalyse

UseR 2008

Outline

1. **Model Repository**
   - Managing Large Sets Of Models
   - MORET 2006 - fixed model structures

2. **MORET 2008**
   - Further Requirements
   - Configuring Models
   - Other Improvements

Outline

The Task: When A Small Set Of Models grows and becomes hard to handle.

- Models can be computed quickly today
  - Lots of data pile up during this process
- Managing large sets of models by hand is cumbersome work.
  - no bit of information must be lost
  - working with the data (transformation, plots ...)
  - information should be kept at one place (desk, notes. R ...)

The Task: When A Small Set Of Models grows and becomes hard to handle.

- Models can be computed quickly today
  - Lots of data pile up during this process
- Managing large sets of models by hand is cumbersome work.
  - no bit of information must be lost
  - working with the data (transformation, plots ...)
  - information should be kept at one place (desk, notes, R ...)

The Task: When A Small Set Of Models grows and becomes
hard to handle.

- Models can be computed quickly today
    - Lots of data pile up during this process
- Managing large sets of models by hand is cumbersome
  work.
    - no bit of information must be lost
    - working with the data (transformation, plots ...)
    - information should be kept at one place (desk, notes, R ...)

The Task: When A Small Set Of Models grows and becomes hard to handle.

- Models can be computed quickly today
    - Lots of data pile up during this process
- Managing large sets of models by hand is cumbersome work.
    - no bit of information must be lost
    - working with the data (transformation, plots ...)
    - information should be kept at one place (desk, notes, R ...)

The Task: When A Small Set Of Models grows and becomes hard to handle.

- Models can be computed quickly today
    - Lots of data pile up during this process
- Managing large sets of models by hand is cumbersome work.
    - no bit of information must be lost
    - working with the data (transformation, plots ...)
    - information should be kept at one place (desk, notes, R ...)

The Task: When A Small Set Of Models grows and becomes hard to handle.

- Models can be computed quickly today
    - Lots of data pile up during this process
- Managing large sets of models by hand is cumbersome work.
    - no bit of information must be lost
    - working with the data (transformation, plots ...)
    - information should be kept at one place (desk, notes, R ...)

Outline

MORET at UseR 2006

- collects all input from R

- stores the data, input and models in a database

- supported model types

    - lm
    - glm
    - gam
    - rpart

MORET at UseR 2006

- collects all input from R

- stores the data, input and models in a database

- supported model types

    - lm
    - glm
    - gam
    - rpart

MORET at UseR 2006

- collects all input from R
- stores the data, input and models in a database
- supported model types
    - lm
    - glm
    - gam
    - rpart

MORET at UseR 2006

- collects all input from R
- stores the data, input and models in a database
- supported model types
    - lm
    - glm
    - gam
    - rpart

MORET at UseR 2006

- collects all input from R
- stores the data, input and models in a database
- supported model types
  - lm
  - glm
  - gam
  - rpart

MORET at UseR 2006

- collects all input from R
- stores the data, input and models in a database
- supported model types
    - lm
    - glm
    - gam
    - rpart

MORET at UseR 2006

- **provides database support for selection and deletion of models or data**
- allows the comparison global model statistics
  - on a very abstract level (complexity, quality)
  - on a model specific level Model Explorer
- external software can be supplied with stored information

MORET at UseR 2006

- provides database support for selection and deletion of models or data
- allows the comparison global model statistics
  - on a very abstract level (complexity, quality)
  - on a model specific level `Model Explorer`
- external software can be supplied with stored information

MORET at UseR 2006

- provides database support for selection and deletion of models or data
- allows the comparison global model statistics
    - on a very abstract level (complexity, quality)
    - on a model specific level `Model Explorer`
- external software can be supplied with stored information

MORET at UseR 2006

- provides database support for selection and deletion of models or data
- allows the comparison global model statistics
    - on a very abstract level (complexity, quality)
    - on a model specific level `Model Explorer`
- external software can be supplied with stored information

MORET at UseR 2006

- provides database support for selection and deletion of models or data
- allows the comparison global model statistics
    - on a very abstract level (complexity, quality)
    - on a model specific level `Model Explorer`
- external software can be supplied with stored information

Model Explorer

	Coefficients / (Intercept) / Propability	Coefficients / hispanic / Propability	Coefficients / income / Propability
	*	*	*
g2000			
Name			
Value	+6.973E-008	+2.752E-002	+3.507E-004
	+1.508E-005		+1.978E-002
Propability	+4.741E-007	+8.587E-002	+1.546E-003
StdError	+6.547E-007	+9.844E-002	+2.032E-003
TStatistic	+3.302E-007	+4.007E-002	+1.290E-003
	+3.888E-006	+8.747E-002	+9.521E-003
	+3.185E-005		+3.252E-002
k2004			
Name	+9.120E-016	+2.191E-001	
Value	+5.138E-006	+1.131E-001	+1.146E-002
	+3.703E-006		+4.484E-003
Propability	+8.777E-007	+8.881E-002	+2.448E-003
StdError	+5.035E-005		+3.683E-002
TStatistic	+8.173E-006	+7.131E-002	+1.745E-002
	+1.131E-006	+1.017E-001	+2.996E-003
etouch	+7.903E-006	+7.904E-002	+1.682E-002
Name	+5.814E-006	+9.424E-002	+1.182E-002
Value	+1.121E-005	+1.179E-001	+1.950E-002
	+7.884E-006		+7.791E-003
Propability	+4.044E-012	+2.865E-001	
StdError	+2.528E-006	+2.709E-001	+4.750E-003
TStatistic	+1.608E-005	+7.357E-002	+2.638E-002
	+1.516E-005	+8.117E-002	+2.492E-002
v_change	+5.447E-006	+2.802E-001	+8.093E-003
Name			
Value			
Propability			
StdError			
TStatistic			

Synchronize Table

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Outline

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Further Requirements

- There are many more model types and new types emerge regularly
  - Generic Database and Model Configuration
- Assessing model data
  - Feature Query
- External information is not accessible (papers, plots ..)
  - Attachments
- ...

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Further Requirements

- There are many more model types and new types emerge regularly
  - Generic Database and Model Configuration

- Assessing model data
  - Feature Query

- External information is not accessible (papers, plots ..)
  - Attachments

- ...

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Further Requirements

- There are many more model types and new types emerge regularly
  - Generic Database and Model Configuration
- Assessing model data
  - Feature Query
- External information is not accessible (papers, plots ..)
  - Attachments
- ...

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Further Requirements

- There are many more model types and new types emerge regularly
  - Generic Database and Model Configuration
- Assessing model data
  - Feature Query
- External information is not accessible (papers, plots ..)
  - Attachments
- ...

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Further Requirements

- There are many more model types and new types emerge regularly
  - Generic Database and Model Configuration
- Assessing model data
  - Feature Query
- External information is not accessible (papers, plots ..)
  - Attachments
- ...

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Further Requirements

- There are many more model types and new types emerge regularly
    - Generic Database and Model Configuration
- Assessing model data
    - Feature Query
- External information is not accessible (papers, plots ..)
    - Attachments
- ...

Model Repository
MORET 2008
Summary

Further Requirements
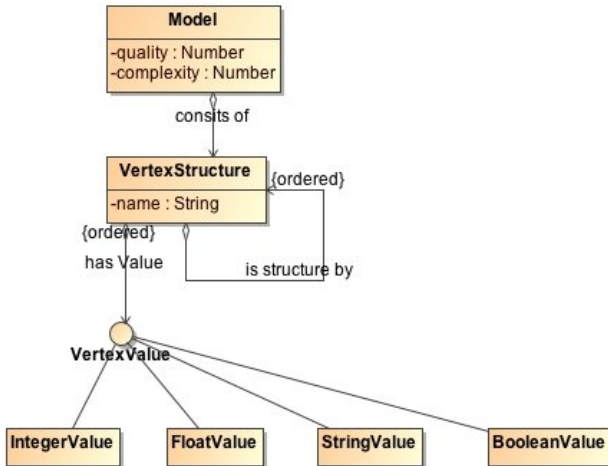Configuring Models
Other Improvements

Further Requirements

- There are many more model types and new types emerge regularly
  - Generic Database and Model Configuration
- Assessing model data
  - Feature Query
- External information is not accessible (papers, plots ..)
  - Attachments
- ...

Outline

Generic Database Structure

Mapping R Model Information

- generic database structure allows storing arbitrary information
- R models come in different data types (List, Vector)
- To store meaningful information the structure from R needs to be filtered/mapped

Mapping R Model Information

- generic database structure allows storing arbitrary information
- R models come in different data types (List, Vector)
- To store meaningful information the structure from R needs to be filtered/mapped

Mapping R Model Information

- generic database structure allows storing arbitrary information
- R models come in different data types (List, Vector)
- To store meaningful information the structure from R needs to be filtered/mapped

Model Repository   Further Requirements
MORET 2008   Configuring Models
Summary   Other Improvements

Full model summary example (loess)

Prerequisites For R Model Mapping

- the returned structure from R must provide
    - a global quality statistic "ResidualDeviance"
    - a global complexity statistic "DF"
- All required information

Prerequisites For R Model Mapping

- the returned structure from R must provide
  - a global quality statistic "ResidualDeviance"
  - a global complexity statistic "DF"
- All required information

Model Repository   Further Requirements
MORET 2008   Configuring Models
Summary   Other Improvements

Prerequisites For R Model Mapping

- the returned structure from R must provide
    - a global quality statistic "ResidualDeviance"
    - a global complexity statistic "DF"
- All required information

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Configurating An R Model Mapping

R–Command	loess
Custom Com...	
Description	smooth regression
Example	samplemodel<-loess(dist~speed,data=cars)
CONFIGURE	

Custom Command

Most R models provide a usable summary command. If still information is missing a custom command can be used to compute the full tree from.

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Using generic models

- After the configuration (wizard) is finished the data will be stored according to the mapping
- The stored information can be accessed
    - via the overview table or
    - the Model Explorer
    - external software

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Using generic models

- After the configuration (wizard) is finished the data will be stored according to the mapping
- The stored information can be accessed
  - via the overview table or
  - the `Model Explorer`
  - external software

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Using generic models

- After the configuration (wizard) is finished the data will be stored according to the mapping
- The stored information can be accessed
  - via the overview table or
  - the `Model Explorer`
  - external software

Using generic models

- After the configuration (wizard) is finished the data will be stored according to the mapping
- The stored information can be accessed
  - via the overview table or
  - the Model Explorer
  - external software

Model Repository   Further Requirements
MORET 2008   Configuring Models
Summary   Other Improvements

Using generic models

- After the configuration (wizard) is finished the data will be stored according to the mapping
- The stored information can be accessed
  - via the overview table or
  - the `Model Explorer`
  - external software

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Outline

Model Repository   Further Requirements
MORET 2008   Configuring Models
Summary   Other Improvements

Feature Query

Feature Query

Retrieve stored models by features instead of relation

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Other Improvements

- **Attachments**
    - Models or data sets relate to other files that can be linked with MORETs database

- Model Management By Group

    - Any model can be added or removed to an administrative group

    - This groups can be used to filter the model table

    - n:m mappings are possible and useful

- Bootstrapping is supported, the index set is persistent

- Database can be exported and merge-process implemented

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Other Improvements

- Attachments
    - Models or data sets relate to other files that can be linked with MORETs database
- Model Management By Group
    - Any model can be added or removed to an administrative group
    - This groups can be used to filter the model table
    - n:m mappings are possible and useful
- Bootstrapping is supported, the index set is persistent
- Database can be exported and merge-process implemented

Model Repository   Further Requirements
MORET 2008   Configuring Models
Summary   Other Improvements

Other Improvements

- Attachments
    - Models or data sets relate to other files that can be linked with MORETs database
- Model Management By Group
    - Any model can be added or removed to an administrative group
    - This groups can be used to filter the model table
    - n:m mappings are possible and useful
- Bootstrapping is supported, the index set is persistent
- Database can be exported and merge-process implemented

Other Improvements

- Attachments
    - Models or data sets relate to other files that can be linked with MORETs database
- Model Management By Group
    - Any model can be added or removed to an administrative group
    - This groups can be used to filter the model table
    - n:m mappings are possible and useful
- Bootstrapping is supported, the index set is persistent
- Database can be exported and merge-process implemented

Other Improvements

- Attachments
    - Models or data sets relate to other files that can be linked with MORETs database
- Model Management By Group
    - Any model can be added or removed to an administrative group
    - This groups can be used to filter the model table
    - n:m mappings are possible and useful
- Bootstrapping is supported, the index set is persistent
- Database can be exported and merge-process implemented

Model Repository
MORET 2008
Summary

Further Requirements
Configuring Models
Other Improvements

Other Improvements

- Attachments
    - Models or data sets relate to other files that can be linked with MORETs database
- Model Management By Group
    - Any model can be added or removed to an administrative group
    - This groups can be used to filter the model table
    - n:m mappings are possible and useful
- Bootstrapping is supported, the index set is persistent
- Database can be exported and merge-process implemented

Summary

- By adapting (mapping) model data, MORET is able to handle all kinds of models.

- The further improvements facilitate the management process.

- Outlook
  - XSLT has been successfully used to map from one XML format to other target formats. Try out if XSLT is capable of mapping models instead of the tree-mapping-wizard.

Summary

- By adapting (mapping) model data, MORET is able to handle all kinds of models.
- The further improvements facilitate the management process.

- Outlook
    - XSLT has been successfully used to map from one XML format to other target formats. Try out if XSLT is capable of mapping models instead of the tree-mapping-wizard.

Project Homepage

Visit and try out the most recent version of MORET at
http://www.rosuda.org