prudsys XELOPES - the packages

Core and infrastructure:

CORE
- CWM core packages,
- metadata handling including basis transformations,
- base model, settings, and algorithm classes,
- event handler,
- PMML handling including XSLT processor to support different PMML versions,
- handling of graphs and taxonomies,
- mathematical libraries including sparse matrix and tensor packages;

DATA AND TRANSFORMATIONS
- dynamic input stream,
- implementations for memory, database, text, log, Excel files, etc.
- multidimensional package to capture DB functionality
- transformations via dynamic filter stream,
- many transformation algorithms (mappings, functions, normalization, etc.);

STATISTICS
- base statistical characteristics,
- correlations,
- multidimensional grouping;

OLAP
- works in memory or on database,
- based on JOLAP standard,
- multiple hierarchy types and complex measures,
- easy handling of multiple cubes,
- high flexibility to define and execute queries and navigate cursors.

Data mining algorithms:

BAYESIAN NETWORKS
- Bayesian and decision networks,
- junction trees for fast inference,
- EM training algorithms,
- special Naïve Bayes subpackage,
- distributed learning;

CLUSTERING
- hierarchical agglomerative algorithms,
- partitioning k-linkage algorithm,
- partitioning k-means algorithm,
- many distance and similarity measures,
- support for multiple compression types;

BASKET AND SEQUENCE ANALYSIS
- association rule algorithms,
- sequence analysis algorithms,
- sequential basket analysis algorithms,
- multiple collaborative filtering algorithms for recommendation engines,
- all algorithms as decompositions algorithms for large data;

CLASSIFICATION AND REGRESSION
- decision and regression trees, rules,
- linear, polynomial and logistic regression,
- neural networks, SVMs,
- Sparse Grid and wavelet algorithms,
- non-linear decision and regression trees combining these models;

FACTORIZATION
- singular value decomposition,
- tensor-based factorization (PARAFAC, HOSVD),
- incremental factorization methods.

Realtime application modules:

TEXT MINING
- preprocessing and transformations, especially phonetic,
- distance calculation algorithms,
- online and offline search algorithms, duplicate search,
- content-based recommendation algorithms;

DISPOSITION
- regression algorithms for disposition,
- adaptive regression algorithms for online disposition,
- special algorithms for new items,
- optimization algorithms for disposition;

PRICE OPTIMIZATION
- offline regression algorithms, barrier methods,
- adaptive regression algorithms for online mode,
- reinforcement learning algorithms,
- special methods for competitor prices and perishable goods,
- simulations for user-defined prices;

REINFORCEMENT LEARNING
- MC and TD learning agents, off- and on-policy;
- dynamic programming agents, value and policy iterations;
- approximation algorithms based on various models,
- hierarchical preconditioners,
- different policies,
- special RL adapters for recommendation engines.

prudsys | The Realtime Analytics Company

prudsys AG | Zwickauer Str. 16 | 09112 Chemnitz | Germany | Phone.: +49 371 27093-0 | Fax: +49 371 27093-90 | E-mail: sales@prudsys.com | URL: www.prudsys.com

prudsys XELOPES

XELOPES

the realtime analytics library

CORE

DATA AND TRANSFORMATIONS

STATISTICS

OLAP

BAYESIAN NETWORKS

CLUSTERING

BASKET AND SEQUENCE ANALYSIS

CLASSIFICATION AND REGRESSION

FACTORIZATION

TEXT MINING

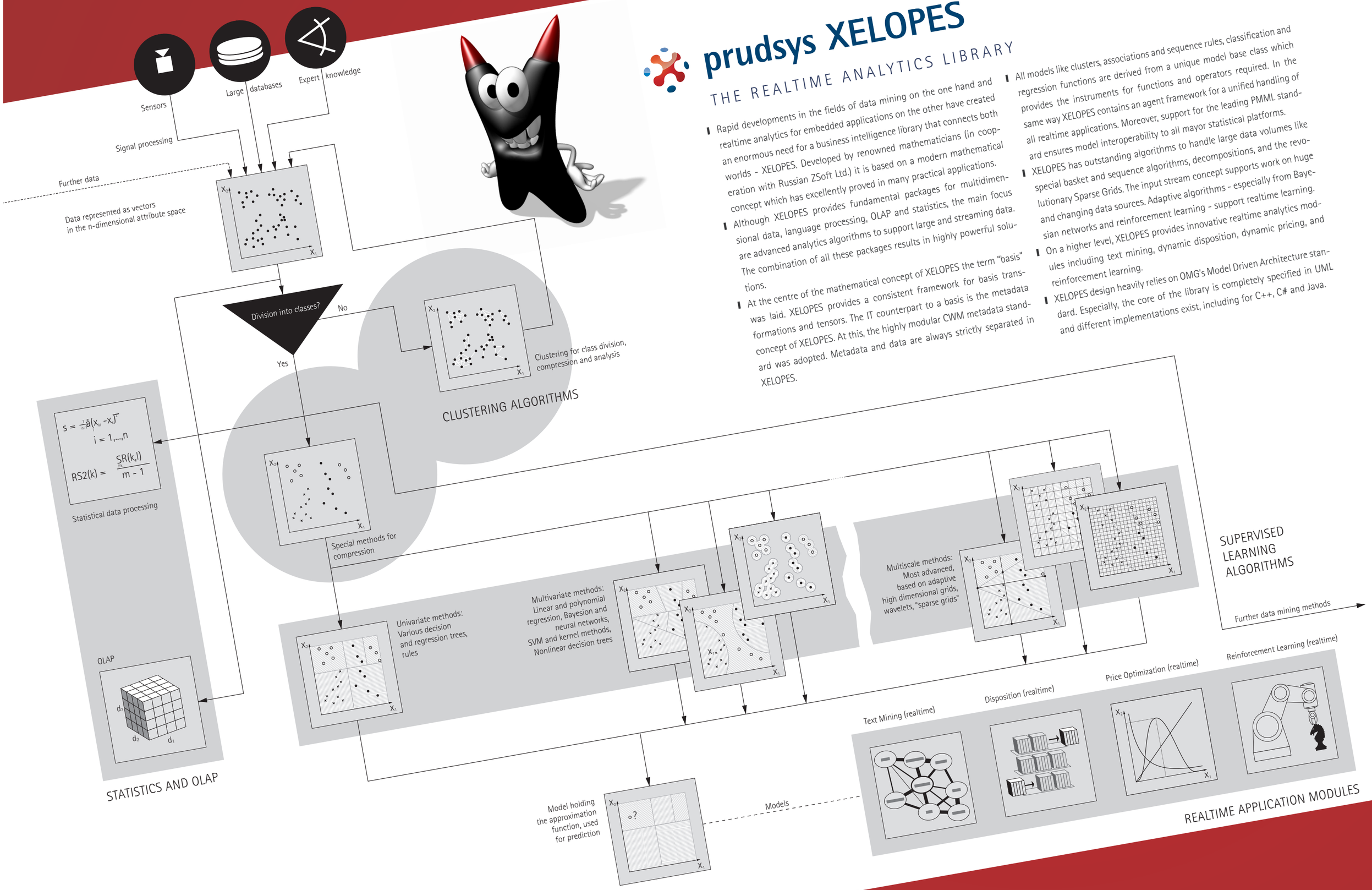
DISPOSITION

PRICE OPTIMIZATION

REINFORCEMENT LEARNING

www.xelopes.com

prudsys XELOPES

THE REALTIME ANALYTICS LIBRARY

- Rapid developments in the fields of data mining on the one hand and realtime analytics for embedded applications on the other have created an enormous need for a business intelligence library that connects both worlds – XELOPES. Developed by renowned mathematicians (in cooperation with Russian ZSoft Ltd.) it is based on a modern mathematical concept which has excellently proved in many practical applications.
- Although XELOPES provides fundamental packages for multidimensional data, language processing, OLAP and statistics, the main focus are advanced analytics algorithms to support large and streaming data. The combination of all these packages results in highly powerful solutions.
- At the centre of the mathematical concept of XELOPES the term "basis" was laid. XELOPES provides a consistent framework for basis transformations and tensors. The IT counterpart to a basis is the metadata concept of XELOPES. At this, the highly modular CWM metadata standard was adopted. Metadata and data are always strictly separated in XELOPES.
- All models like clusters, associations and sequence rules, classification and regression functions are derived from a unique model base class which provides the instruments for functions and operators required. In the same way XELOPES contains an agent framework for a unified handling of all realtime applications. Moreover, support for the leading PMML standard ensures model interoperability to all mayor statistical platforms.
- XELOPES has outstanding algorithms to handle large data volumes like special basket and sequence algorithms, decompositions, and the revolutionary Sparse Grids. The input stream concept supports work on huge and changing data sources. Adaptive algorithms – especially from Bayesian networks and reinforcement learning – support realtime learning.
- On a higher level, XELOPES provides innovative realtime analytics modules including text mining, dynamic disposition, dynamic pricing, and reinforcement learning.
- XELOPES design heavily relies on OMG's Model Driven Architecture standard. Especially, the core of the library is completely specified in UML and different implementations exist, including for C++, C# and Java.