

Motivation

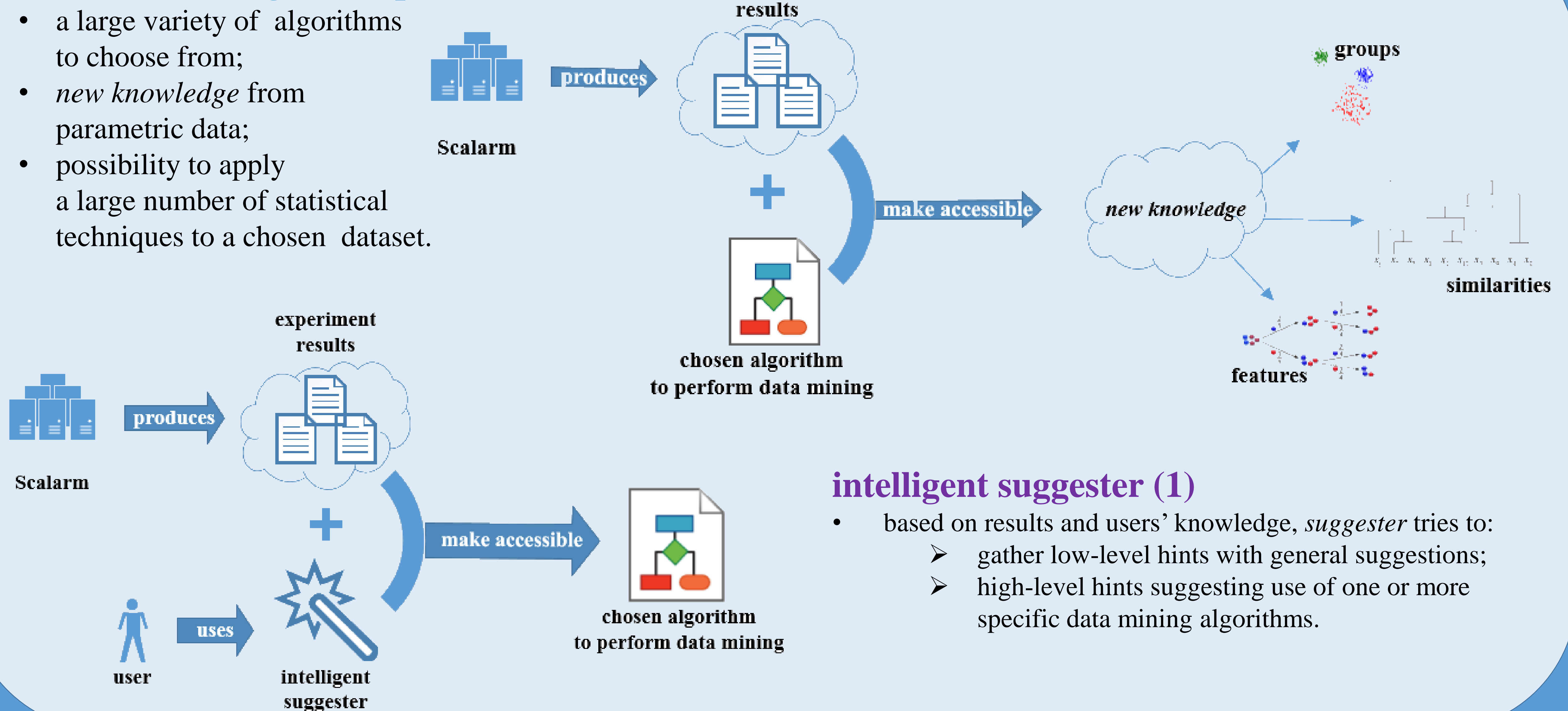
- data farming computations are extremely expensive, Scalarm [1] being not an exception;
- to have an overall view on an experiment we should conduct it using all possible combinations of parameter values and it is no secret we can hardly afford performing such an amount of computations;
- manual analysis of extensive output data is a big challenge and error-prone activity.

Objectives

- research into interpretation of parametric data [2]:
 - results of Scalarm-based computations, simulations, etc.;
- propose a number of approaches to extend Scalarm's functionality to gain new kind of knowledge;
- provide support for discovery of relationships and influence of specific parameters on each other and on the experiment as a whole;
- gain *new knowledge* about parametric data related to its structure, essential features and general similarities.

New knowledge from parametric data (2)

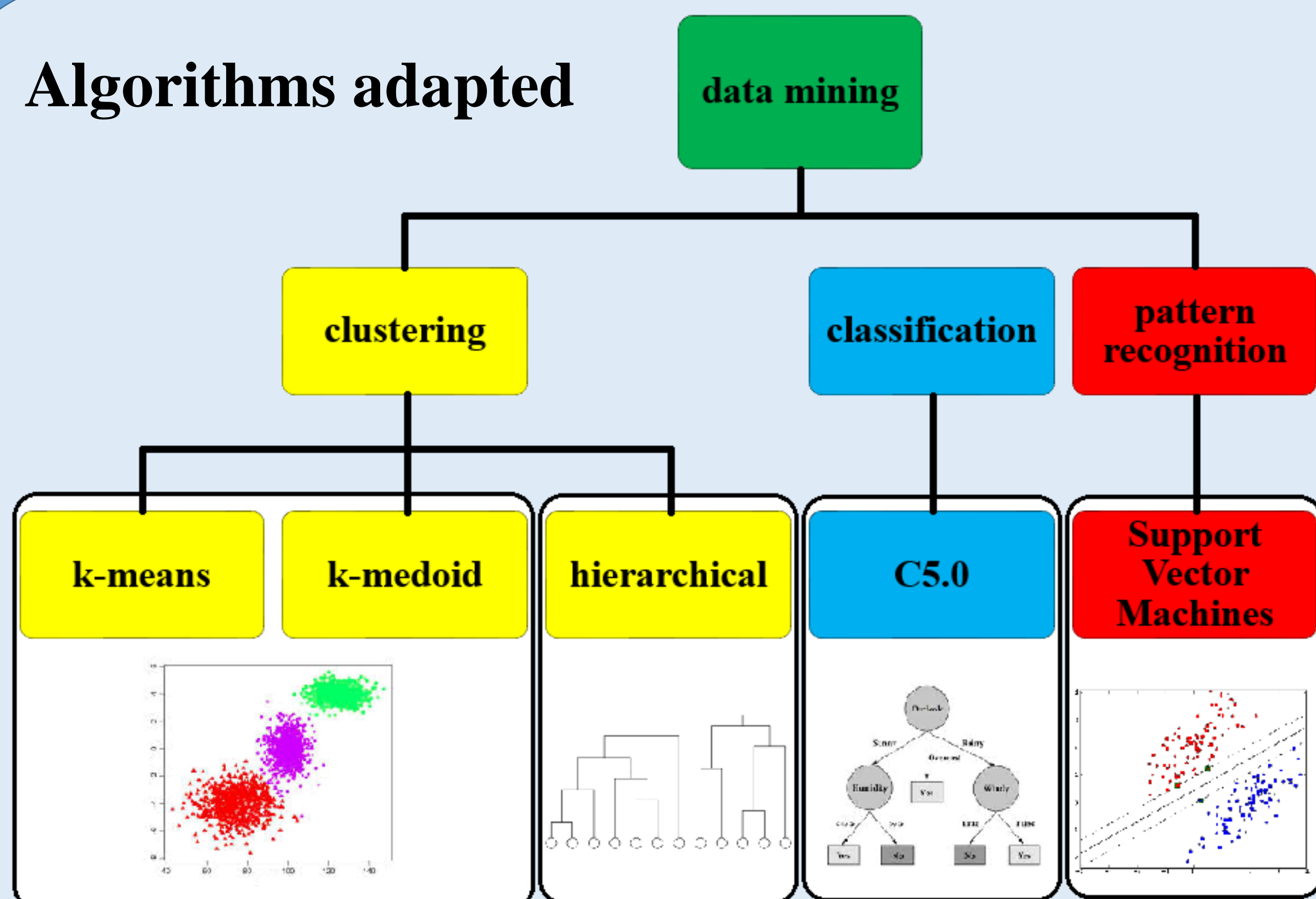
- a large variety of algorithms to choose from;
- *new knowledge* from parametric data;
- possibility to apply a large number of statistical techniques to a chosen dataset.



intelligent suggester (1)

- based on results and users' knowledge, *suggester* tries to:
 - gather low-level hints with general suggestions;
 - high-level hints suggesting use of one or more specific data mining algorithms.

Algorithms adapted



Conclusions and future work

- we can limit the number of experiment runs to the most important ones;
- data exploration and visualization facilitate making crucial decisions related to the choice of input parameters of the experiment and speeding up the process of analysis of an experiment as a whole;
- a number of data mining algorithms adopted;
- the problem of choosing proper algorithms resolved with intelligent suggester;
- extendibility of the suite of data mining algorithms;

Acknowledgments. This research is supported by the European Union within the European Regional Development Fund program as part of the PLGrid Core Project POIG.02.03.00-12-137/13 (<http://plgrid.pl/core>)

References

- [1] Scalarm, Massively Self-Scalable Platform for Data Farming, project's web site: <http://www.scalarm.com/>
[2] Eric Walter and Luc Pronzato. Identification of Parametric Models from Experimental Data. Berlin, Heidelberg, New York: Springer, 1997