Selected Methods and Applications of Rough Sets in Management and Engineering (Springer, 2012)

Georg Peters, Pawan Lingras,
Dominik Ślęzak, Yiyu Yao

QUICK LINKS

- Home
- Featured Content
- VoIP Call Recording
- Tech Talk
- Videos
- RSS Feeds
- Contact Us

Call Recording Featured Article

September 30, 2011

JDSU Partners with Infobright to Help the World"s Largest Communications Service Providers Ensure the Highest Quality of Service

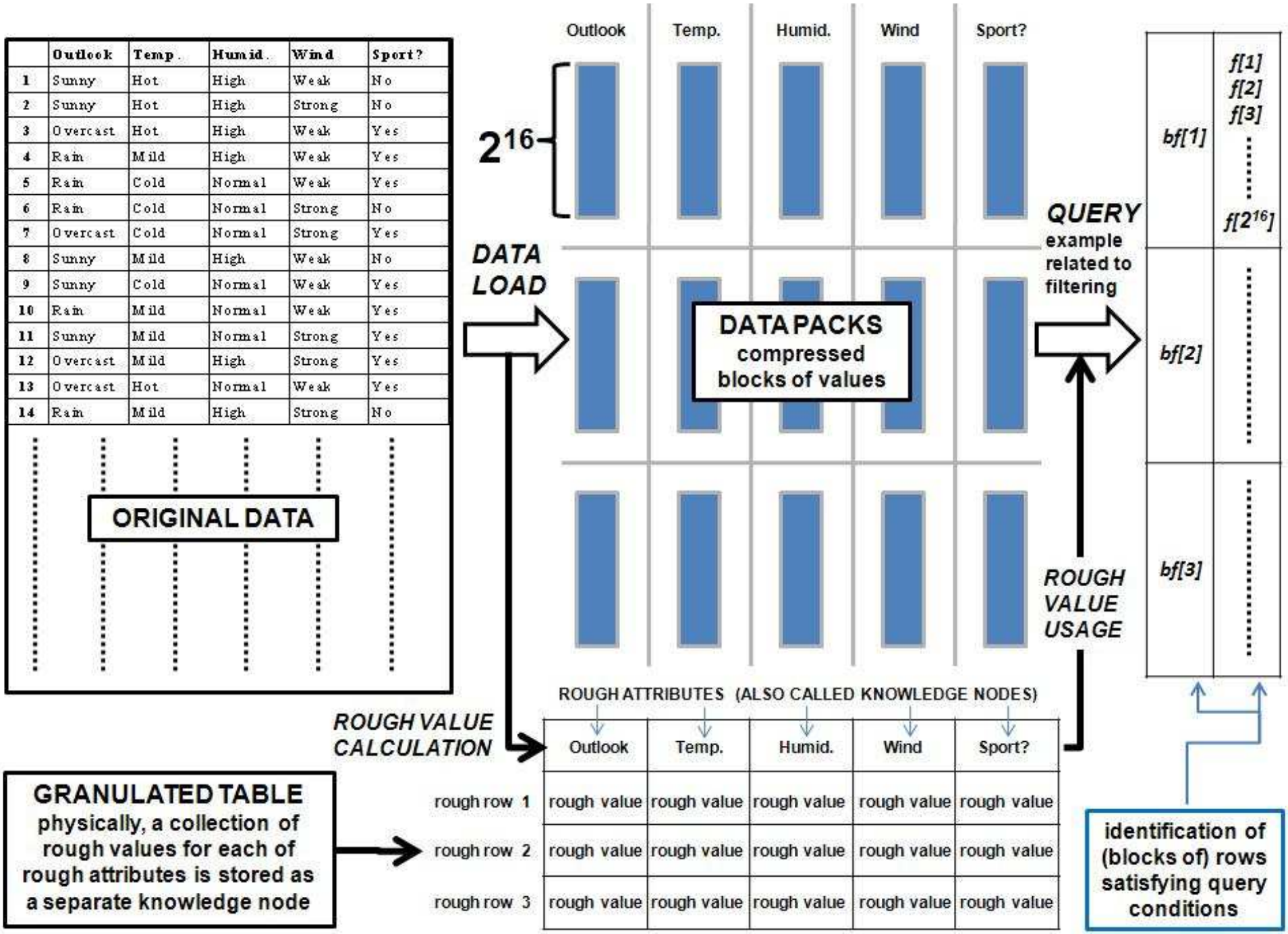
By TMCnet Special Guest

Susan Davis, Vice President, Product Management, Infobright

The growth of smartphones and the increasing volume of data traversing 3G, and now 4G/LTE (News - Alert) networks presents a huge

LATEST INDU

- Atelka Annou SATMAP
- LiveVox Adds Contact Cente
- Sunshine Fec



	Outlook	Temp.	Humid.	Wind	Sport?
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cold	Normal	Weak	Yes
6	Rain	Cold	Normal	Strong	No
7	Overcast	Cold	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cold	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Ensembles of Bireducts: Towards Robust Classification and Simple Representation*

Dominik Ślęzak[1,2], Andrzej Janusz[1]

[1] Institute of Mathematics, University of Warsaw
ul. Banacha 2, 02-097 Warsaw, Poland
[2] Infobright Inc.
ul. Krzywickiego 34 lok. 219, 02-078 Warsaw, Poland
slezak@infobright.com,andrzejanusz@gmail.com

Abstract. We introduce the notion of a bireduct, which is an extension of the notion of a reduct developed within the theory of rough sets. For a decision system $\mathbb{A} = (U, A \cup \{d\})$, a bireduct is a pair (B, X), where $B \subseteq A$ is a subset of attributes that discerns all pairs of objects in $X \subseteq U$ with different values of the decision attribute d, and where B and X cannot be, respectively, reduced and extended without losing this property. We investigate the ability of ensembles of bireducts (B, X) characterized by significant diversity with respect to both B and X to represent knowledge hidden in data and to serve as the means for learning robust classification systems. We show fundamental properties of bireducts and provide algorithms aimed at searching for ensembles of bireducts in data. We also report results obtained for some benchmark data sets.

Keywords: Attribute Subset Selection, Inexact Dependencies, Classifier Ensembles, Discernibility, Decision Rules, Randomized Search.

	Outlook	**Temp.**	**Humid.**	**Wind**	**Sport?**
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cold	Normal	Weak	Yes
6	Rain	Cold	Normal	Strong	No
7	Overcast	Cold	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cold	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

({O,T,W},{1-14})

({O,W},{1-8,10,12-14})

({O,T},{1-3,5,7-9,12-14})

({O,H},{1-5,7-13})

({O,W},{3-7,9-14})

({W},{2-6,9-10,13-14})

({O,H,W},{1-14})

({T,W},{2-3,5-6,8-9,13-14})

({H},{3-5,7,9-13})

({T,W},{1-2,4-5,7,9-10,14})

({T,H,W},{2-3,5,7-13})

({H,T},{1-2,4-5,7,9-13})

({O},{1-5,7-8,10,12-13})

({H,W},{1-2,5-6,8-10,13-14})

({T},{1-2,4,6,10-12})

RapidMiner

Report the Future.

RapidMiner is unquestionable the world-leading open-source system for data mining. It is available as a stand-alone application for data analysis and as a data mining engine for the integration into own products. Thousands of applications of RapidMiner in more than 40 countries give their users a competitive edge.

- **Data Integration**, **Analytical ETL**, **Data Analysis**, and **Reporting** in **one** single suite
- Powerful but intuitive **graphical user interface** for the design of analysis processes
- **Repositories** for process, data and meta data handling
- Only solution with **meta data transformation**: forget trial and error and inspect results already during design time
- Only solution which supports **on-the-fly error recognition** and **quick fixes**
- **Complete and flexible:** Hundreds of data loading, data transformation, data modeling, and data visualization methods

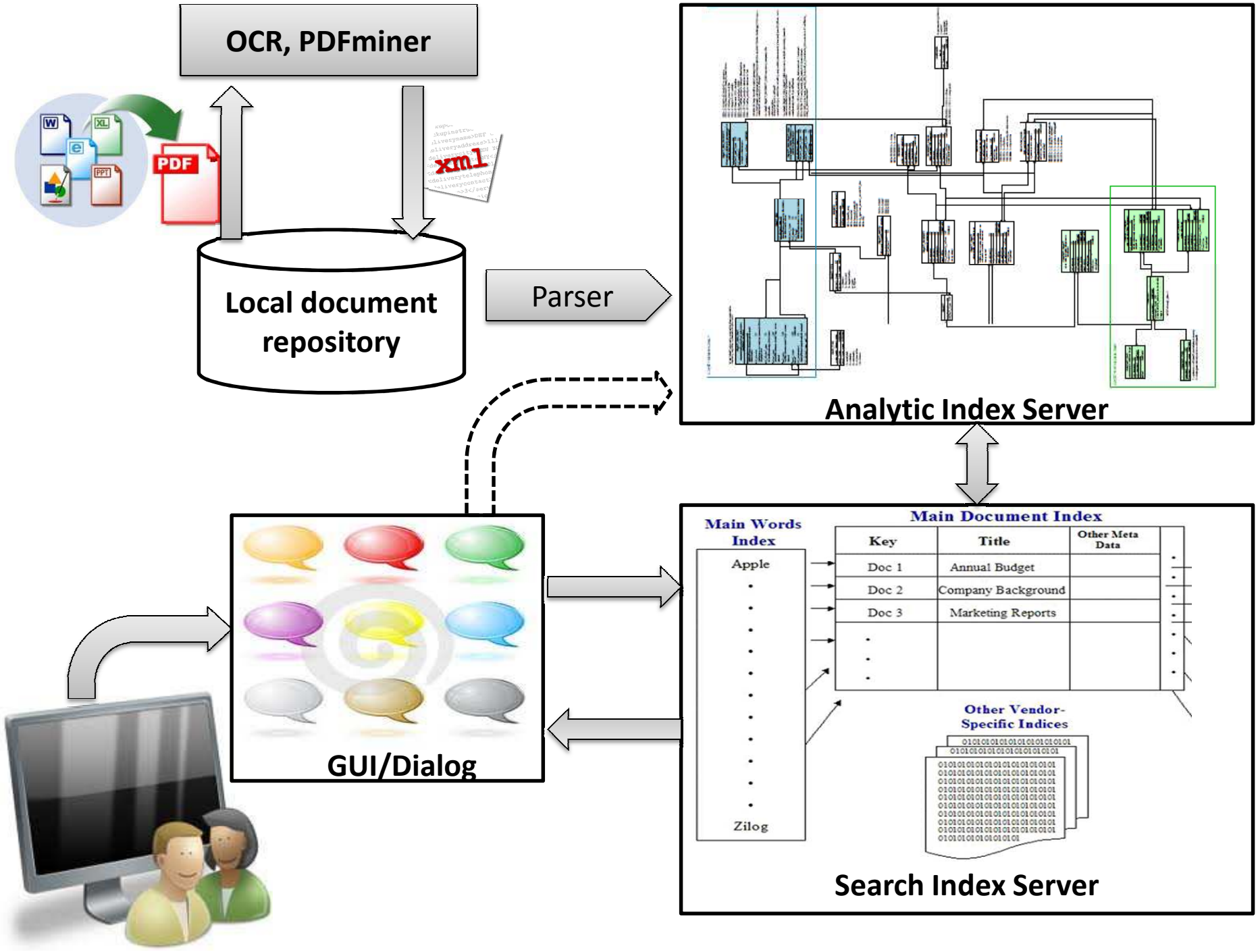
Semantic Analytics of PubMed Content*

Dominik Ślęzak[1,2], Andrzej Janusz[1], Wojciech Świeboda[1],
Hung Son Nguyen[1], Jan G. Bazan[3,1], and Andrzej Skowron[1]

[1] Institute of Mathematics, University of Warsaw
ul. Banacha 2, 02-097 Warsaw, Poland
[2] Infobright Inc.
ul. Krzywickiego 34 lok. 219, 02-078 Warsaw, Poland
[3] Chair of Computer Science, University of Rzeszów
ul. Rejtana 16A, 35-310 Rzeszów, Poland

Abstract. We present an architecture aimed at semantic search and synthesis of information acquired from the document repositories. The proposed framework is expected to provide domain knowledge interfaces enabling the internally implemented algorithms to identify relationships between documents, researchers, institutions, as well as concepts extracted from various types of knowledge bases. The framework should be scalable with respect to data volumes, diversity of analytic processes, and the speed of search. In this paper, we investigate these requirements for the case of medical publications gathered in PubMed.

Keywords: Semantic Search and Analytics, PubMed, MeSH, RDBMS, Document Repositories, Decision Support Systems, Behavioral Patterns.


OCR, PDFminer
PDF
xml
Local document repository
Parser
Analytic Index Server
Main Words Index
Apple
Zilog
Main Document Index
Key
Title
Other Meta Data
Doc 1
Annual Budget
Doc 2
Company Background
Doc 3
Marketing Reports
Other Vendor-Specific Indices
Search Index Server
GUI/Dialog

Thank you!!!!