As of June 2011

Two Reputedly Opposed Approaches: On the Integration of Business Rules and Data Mining

White Paper

Bosch Software Innovations

Americas:
Bosch Software Innovations Corp.
161 N. Clark Street
Suite 3500
Chicago, Illinois 60601/USA
Tel. +1 312 368-2500
info@bosch-si.com
www.bosch-si.com

Asia:
Bosch Software Innovations
c/o Robert Bosch (SEA) Pte Ltd
11 Bishan Street 21
Singapore 573943
Tel. +65 6571 2220
info-sg@bosch-si.com
www.bosch-si.sg

Europe:
Bosch Software Innovations GmbH
Ziegelei 7
88090 Immenstaad
GERMANY
Tel. +49 7545 202-300
info-de@bosch-si.com
www.bosch-si.com

Two Reputedly Opposed Approaches: On the Integration of Business Rules and Data Mining

Irene M. Cramer

Business rules and data mining approaches are often treated as competitors, although they have differing fields of application: While business rules provide a valuable insight into the calculation of business parameters and help to trace tacit knowledge, data mining offers powerful methods for processing large data sets and handling complex, potentially dynamic feature spaces. In this paper, we show how the two disciplines team up and are used in a direct marketing and a process control scenario.

Content

Introduction

Are you confronted with data in abundance? Are you undecided how to find out if there is something of interest hidden in your data? Are you looking for relevant patterns? Combine a business rules and a data mining approach. Then you can take advantage of their particular strong points in managing data. In this paper, we discuss two application scenarios and show you:

- How to use a business rules system to feed a data mining toolkit with feature vectors
- How to determine reasonable subgroups if you already know the relevant features
- How to determine the distinctive features if you already know the relevant subgroups
- How to use the results of the data mining steps in business rules

This paper is intended to be used as a cookbook: if the scenarios, i.e. the direct marketing or the process control dishes, are of business interest to you, read it as the recipe. If you would like to cook something else, you can vary the ingredients and read this paper as a description of the method. If your use case addresses any subarea of data mining not discussed in this paper, please get in touch. We will be happy to come up with another recipe to fit your particular use case.

The remainder of this paper is structured as follows: in section "*Scenario 1: Direct Marketing*", we sketch the direct marketing scenario and explain how to combine your business rules with automatically constructed classifiers. In section "*Scenario 2: Process Control*", we sketch the process control scenario and discuss how to explore your data (e.g. in order to test a hypothesis). In section "*Summary*" we recapitulate the main steps you need to perform when integrating a data mining and a business rules approach.

Some Terminology

A feature is (the value of) a variable used to characterize an object. A feature is said to be distinctive if it helps to divide a set of objects into subgroups. A feature vector is an n-dimensional vector of feature values which represents an object.

A subgroup is a set of objects that are equal (or similar) with respect to at least one feature. We call it a class if the set is formed by (intellectual) abstraction. We call it a cluster if the set is formed by similarity (of feature vectors).

Data mining means detecting patterns and extracting them from data. Data mining methods are used to (automatically) handle large numbers of objects and high-dimensional feature vectors. It typically involves four classes of tasks: clustering, classification, regression, and association rule learning. There are also data mining methods that help to find the relevant features or to explore and visualize the data. Popular methods include correlation analysis, feature selection, or feature weighting.

Business rules are statements that formalize some aspects of the business. They are stated in natural language sentences, as a graphical model, or as a mathematical formula. According to the Business Rules Group (Hay, Healy, & Hall, 2000) the rules can define or relate business terms and they can describe constraints or derivations. Business rules are defined, deployed, executed, monitored, and maintained by a Business Rules Management System such as Visual Rules (see http://www.visual-rules.com/ for details).

Scenario 1: Direct Marketing

Direct marketing is a type of advertising that addresses the members of a target group (in a potentially personalized manner) with techniques such as catalog distribution or promotional letters. In contrast to TV, newspaper, or radio advertising, direct marketing implies that the response from consumers is traceable and measurable. In a B2C context, it is typically used by small and medium enterprises that need to maximize the return on investment given a limited marketing budget.

From the marketers' point of view, direct marketing involves the collection of geo-demographic information and the maintenance of a customer-centric data warehouse. The direct responses, in turn, can be used to refine the market segmentation and the consumer profiles. From the consumers' point of view, direct marketing, that is to say possibly unwanted and irrelevant solicitations, violates privacy.

The challenge of a successful direct marketing campaign is to:

- Compile high-quality geo-demographic information and consumer data form various sources
- Carefully segment the market and model target group profiles
- Construct well-fitting solicitations for different types of target group members

Scenario and Data: Since retaining existing customers is more cost-effective than acquiring new ones, the marketing department of a medium-sized online store is assigned the construction of a direct marketing campaign addressing customers that are likely to be lost in the near future. There are several known factors associated with the loss of customers (also referred to as customer attrition), for instance dissatisfaction with the service, price, or support, billing disputes, etc.

The marketers know that customer attrition models can be calculated with data mining techniques (Au, Li, & Ma, 2003). However, the colleagues in the marketing and sales departments are old hands at customer retention. The marketers would like to integrate this experience. So they decide to combine the automatically calculated models with manually constructed rules.

They agree to model customer attrition on the basis of the data stored in the company's data warehouse. This data consists of vectors (see Figure 1) describing the customers in terms of several features, e.g. gender, credit rating, the socio-economic status of the customer and the customer's neighborhood, the number of purchases in the last few months and years, the number of complaints and returns, etc.

Figure 1: Customers Represented as Feature Vectors (Extract)

	B	C	D	E	F	J	AE	AF	
1	GENDER	CUSTOMER_RAT	NEIGHBOR	SES_OF_NEIG	AGE	CREDIT_RAT	PURCHASES	VALUE_YEA	PU
2	F	1	C	1	33	2	11	295,2	
3	M	1	U	1	37	2	3	447,69	
4	M	0	T	2	44	0	10	253,25	
5	F	0	T	4	54	2	5	704,76	
6	F	3	C	2	30	0	10	282,7	
7	M	1	R	1	46	4	10	767,02	
8	F	1	C	2	35	1	10	495,41	
9	F	2	R	1	51	3	9	409,55	
10	F	0	S	0	46	3	5	37,49	
11	M	3	S	0	54	2	11	1.659,96	

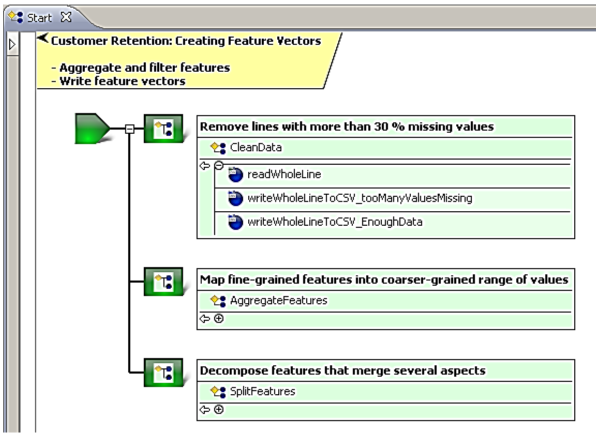
Each customer is also tagged with one of four possible customer ratings: new customer (0), occasional customer (1), regular customer (2), or key customer (3). The rating is determined by, among others, the customer's number of purchases and the amount of turnover.

The marketers are recommended to calculate classification models on the basis of manually labeled, historic feature vectors using the data mining toolkit KNIME (see http://www.knime.org/ for details). These models classify customers as "switchers" or "stayers" and predict if a customer may be lost in the near future. The marketers thus prepare a fraction of the previous year's data: customers who had been rated regular or key customers in recent years are tagged with a "switcher"-flag if they did not order for at least six months last year and did not order this year so far either. Two customer retention experts subsequently revise the data and remove all cases of doubt.

Approach: Since several features need to be aggregated and filtered, business rules are modeled with Visual Rules (see Figure 2), which

- Remove feature vectors with more than 30% missing values
- Map some of the fine-grained features into a coarser-grained range of values
- Decompose some of the features that merge several aspects relevant for this scenario

Figure 2: Construction of Feature Vectors with a Rule Model (Extract)



The business rules also convert the data into the required format. This data is used to train and evaluate the classification models.

Figure 3: Example Process of the Classification Task (with KNIME)

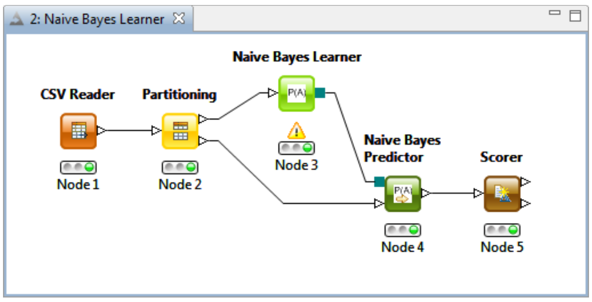


Figure 3 shows the essential steps in the process: first, the data is read (CSV Reader) and partitioned (Partitioning), i.e. split into training and test sets. Second, the classification model is calculated based on the training set (Naïve Bayes Learner). Third, the classification model is used to predict the classes of the test set objects (Naïve Bayes Predictor). Finally, the classification model is evaluated (Scorer), i.e. for each object in the test set the predicted class is compared with the real one (see box "Popular Evaluation Metrics" for details).

Let us assume that among the best-performing classification models are a naïve Bayes and a decision tree classifier. (We use them to simplify matters. Without doubt, there are more sophisticated algorithms; however, these two are known to perform well for certain tasks.)

Popular Evaluation Metrics

The performance of a classification algorithm is measured in terms of e.g. precision, recall, and accuracy given a specific confusion matrix.

$$precision = \frac{tp}{tp + fp} \qquad recall = \frac{tp}{tp + fn}$$

$$accuracy = \frac{tp + tn}{tp + fp + fn + tn}$$

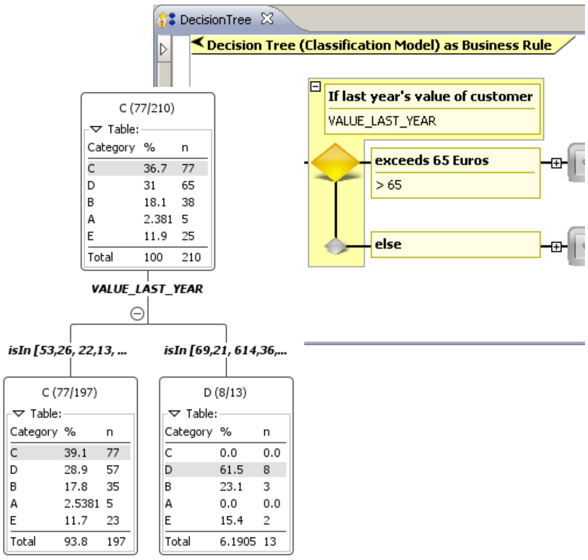
with the confusion matrix (binary case):

		predicted value p'	predicted value n'
actual value	p	true negative	false positive
actual value	n	false negative	true positive

The naïve Bayes is known to be simple and fast – and despite the fact that its simplistic design violates basic statistical assumptions, it often performs surprisingly well for real-world problems. Unfortunately, this classification model is not self-explanatory and its translation into business rules is difficult. Instead, if the marketing and sales experts have constructed business rules that in turn represent a classifier, the two models can be combined. That is to say, the naïve Bayes model is used in the manually constructed model, i.e. the business rules, via service call. Combining classification models can significantly increase the predictive performance of the ensemble. A comprehensive review of possible classifier combination methods, their advantages, and disadvantages is given in (Tulyakov, Jaeger, Govindaraju, & Doermann, 2008).

The decision tree requires little data preparation and can rapidly classify large data sets. Unfortunately, decision trees sometimes create models that do not generalize the data (also referred to as overfitting). This problem can be handled e.g. with pruning mechanisms. Since decision trees are simple to understand and interpret (Figure 4 shows the extract of a decision tree calculated on the basis of the direct marketing feature vectors), they can be combined easily with business rules that are manually constructed by marketing and sales experts.

Figure 4: Simplified Example Decision Tree (Calculated with KNIME) and Corresponding Business Rule (Modeled with Visual Rules)



Given the classification (computed with either a classifier ensemble or a business rule model), the potential switchers can be split into the following groups:

1. Key customers who had no billing disputes with the company in recent years were sent a questionnaire asking for suggestions (e.g. with respect to customer service, range of products, etc.) and a voucher of 20 Euros.
2. Key customers who recently had billing disputes were sent a promotional letter including a promotional gift and a voucher of 15 Euros.
3. Regular customers were sent a promotional letter and a voucher of 10 Euros
4. Occasional customers were sent a promotional letter and a voucher of 5 Euros.

Discussion: As these two examples illustrate for the case of a classification task, the data mining and business rules approaches can be combined in various ways:

- If the classification model is simple to interpret (such as a decision tree, a decision table, or decision rules), it can be modeled, i.e. reproduced, as part of the business rules. In this case, the data mining toolkit helps to find some of the classification rules. It also facilitates the estimation of base-line performance. Nevertheless, if a classifier is included in manually constructed business rules, the performance of the latter can be worse than the performance calculated by the data mining toolkit. In this case, the business rules need to be evaluated carefully (see box "Popular Evaluation Metrics"), even though the modeler was a domain expert.
- If the classification model is not simple to interpret (such as a naïve Bayes, a support vector machine, or a reasonably complex neural network), it can be included in the business rules as a service call. In this case, the classification model constructed via data mining toolkit and the business rules function as a classifier ensemble. Such an ensemble needs to be evaluated just as carefully as the classifier-business-rules combination mentioned above.

What is the use of this combination? Data mining techniques have demonstrated high performance in several real-world applications. As mentioned above, they are able to handle large data sets and complex, high-dimensional feature spaces. Admittedly, automatically computed classification models do not necessarily provide a (comprehensible) insight into the relevant classification parameters. This is why some people are skeptical towards these data mining methods. Business rules, on the other hand, are as high-performance as data mining (though not necessarily in the same fields of application) and in most cases easier to understand. Moreover, if business requirements change fast and frequently, business rules are extremely valuable because they are modeled by experts who best understand the requirements and know how to access and incorporate tacit knowledge. Combining both in a classification task then generates a most welcome synergy effect.

Scenario 2: Process Control

Process control is an engineering discipline that explores and describes the facilities and methods needed to keep the output of a process within a desired range. It typically involves sensors and actuators. For example, the product flow in a plant is a process that has a specific, desired outcome. Its actual value is measured by a flow meter (sensor), whereas its desired value is controlled by the setting of a valve (actuator).

In most plants the sensors and actuators are organized by what is known as a process control system. It typically consists of the field control station, i.e. a component assembling and managing all sensors and actuators, and a control room with monitoring and visualization facilities (see Figure 5). There are different types of control systems for batch and continuous processes. In this paper we focus on the continuous case, which features smooth, uninterrupted variables and a continuous flow of product. Continuous processes are used to manufacture very large quantities of e.g. chemicals per year.

Figure 5: Example of a Control Room (by RobertRED, http://de.wikipedia.org/wiki/Leitstand retrieved February 9, 2011)

Scenario and Data: Because of the (always acceptable but) fluctuating product quality reported by the foreman and the lab assistants, the factory manager of a chemical company decides to implement quality-improvement activities. Concerning the observed fluctuation, two hypotheses have been established:

1. The quality of the product depends on the quality of the raw material, which in turn depends on the supplier or transportation chain.
2. The process control of one or several plant components is in need of improvement.

The continuous production of chemicals in this company's plant is managed by a process control system (e.g. ABB's Freelance, see http://www.abb.de/product/us/9AAC115759.aspx for details), which monitors 400 sensors and operates 200 actuators. The sensors measure temperature, pressure, flow rate, fill level, and pH-value at the various plant components. The settings of the actuators, i.e. the valves, determine the flow rate of the ingredients and hence the production rate in the process steps. Every five seconds the process control system records the measured values and the settings of the valves. On this basis, it can calculate trend lines and display them on the monitoring facilities in the control room. The data is also archived for five years in case an audit needs to be conducted. Every four hours the lab assistants test the quality of the final product and record the results in an Excel sheet. In this scenario, there are three types of test values: residual moisture, particle size, and color scale. The supplier (and transportation chain) of every batch of raw material used in the production is logged in an SAP system.

The factory manager and lab assistants decide to compile data over a period of eight weeks. They agree to use the data mining toolkit RapidMiner (see http://rapid-i.com for details) and to calculate a correlation analysis. (There are without doubt other, more sophisticated algorithms for data exploration. We use this method to show the general idea and to simplify matters.) Although correlation does not necessarily imply a causal relationship, it can give a hint as to which of the two hypotheses mentioned above should be investigated further.

Approach & Results: Because of the temporal aspect of this scenario, the construction of feature vectors is less obvious than in the direct marketing case. In addition, information from various sources (i.e. the process control system, a separate Excel sheet, and an SAP system) needs to be integrated in an appropriate way. For this reason, the approach of a sliding window is employed. The window size is set to four hours, since every four hours the lab assistants test the quality of the final product. The sensor values and the actuator settings are averaged over this period in order to decrease data dimensions and reduce noise. Business rules are modeled that construct the feature vectors as follows:

For every set of quality test results in the Excel sheet,

- Read the values of the 400 sensors and the settings of the 200 actuators of the past four hours from the process control system
- For every sensor (actuator), average the values (settings) over this period, i.e. calculate the mean and variance (see box for details)
- Determine the supplier and transportation chain on the basis of the setting of the storage tank valves, the rate of production, and the code numbers recorded in the SAP system

Why Mean and Variance?

The arithmetic **mean** represents the central tendency (or average) of the 2880 values (i.e. 12 values per minute x 240 minutes = 2880) per sensor or actuator. It condenses the temporal dimension into one feature value instead of about 3000 and thus simplifies the hypothesis testing.

However, the arithmetic mean is known to be sensitive to outliers and does not always represent uneven or skewed distributions adequately. For instance, the values of two four-hours periods featuring the same mean can vary significantly: let us assume that in the first period many of the values spread out far from each other, while in the second one most lie next to the mean. Hence, an additional statistic is needed expressing this difference.

The **variance** is such a statistic. It indicates how a set of values is distributed around the mean. In the case of the two four-hours periods mentioned above the variances differ while the means equal. This suggests that in the first period the production in the process step monitored by the respective sensor had been labile.

In sum, 336 feature vectors are constructed (i.e. 6 per day x 7 days a week x 8 weeks = 336) including the following features (see Figure 6 as illustration):

- 3 quality test values (residual moisture, particle size, and color scale)
- 800 sensor values (mean and variance for 400 sensors)
- 400 actuator settings (mean and variance for 200 actuators)
- 6 supplier and transportation chain data sets (each for 3 raw materials)

Figure 6: Process Represented as Feature Vectors (Extract)

	A	B	C	D	E	F
1	RESIDUAL_MOISTU	PARTICLE_SIZE (%	COLOR_SCALE (KL	KFT10(4) MEAN	KFT10(4) VARIAN	KFT10(V) M
2	0,005	1,4	0,79	2990,2	720	45,1
3	0,002	2,9	0,78	3017,7	930	45,2
4	0,002	1,8	0,82	3036	960	45,2
5	0,002	2,2	0,81	3037,7	930	45,3
6	0,003	1,8	0,77	3000,2	900	45,0
7	0,005	2,2	0,83	2999,3	1230	45,0
8	0,003	2	0,77	2985,6	990	45,2
9	0,003	2,7	0,86	3046,7	1170	45,4
10	0,004	2,1	0,86	3009,2	1410	45,5
11	0,006	2,2	0,86	3005	1290	45,4
12	0,005	2,8	0,84	2959,3	660	44,9
13	0,003	1,4	0,87	2952,2	630	45,2

BR&DM_ProcessControl_2011-

Figure 7 shows an extract of the correlation analysis calculated with RapidMiner. Many consecutive steps in a chemical process are of course correlated to some extent; in this context however, the correlations between the values of the sensor-actuator pairs and the quality scores are of interest. Exploring these correlations indicates that one of the quality scores (namely the color scale) is highly correlated with the values of one of the sensor-actuator pairs (identifier of the sensor: KFP11(1), identifier of the actuator: KFP11(V)). Inspection of the data shows that the values of the KPF11(1) sensor, a flow meter, are gradually rising, ranging from about 44.5 % valve opening at the beginning to about 45.5 % at the end of the eight weeks. At the same time, the color scale values are deteriorating. KFP11(1) specifies the amount of cooling water used for controlled crystallization: the higher the temperature of the product coming from the preceding process step, the more cooling water is needed. And the higher the temperature at the beginning of the crystallization, the more undesired inclusions are there in the product after crystallization. This directly affects the color scale and hence the quality. Although the sensor (identifier: KFT10(4)) of the preceding process step, which is a thermometer, registers routine values, the amount of cooling water rises. The factory manager and the foremen therefore agree to check the sensors of the whole unit. They find that the thermometer in the process step preceding KFP11(1) is indeed damaged. See Figure 8 as illustration.

The pairwise correlation table (see Figure 7) also shows that none of the quality scores is more than moderately correlated with the supplier or transportation chain features. That is to say, there is no evidence supporting the first hypothesis mentioned above, i.e. the quality of the product does not seem to depend on the quality of the raw material.

Figure 7: Results of the Correlation Analysis Calculated with RapidMiner (Extract)

First Attribute	Second Attribute	Correlation
COLOR_SCALE (KLETT)	KFP10(V) MEAN	0.129
COLOR_SCALE (KLETT)	KFP10(V) VARIANCE	-0.028
COLOR_SCALE (KLETT)	KFP11(1) MEAN	0.987
COLOR_SCALE (KLETT)	KFP11(1) VARIANCE	0.033
COLOR_SCALE (KLETT)	KFP11(V) MEAN	0.991
COLOR_SCALE (KLETT)	KFP11(V) VARIANCE	-0.078
COLOR_SCALE (KLETT)	SUPPLIER	0.008
COLOR_SCALE (KLETT)	TRANSPORTATION_CHAIN	0.062

Figure 8: Correlated Values Indicating Damaged Sensor

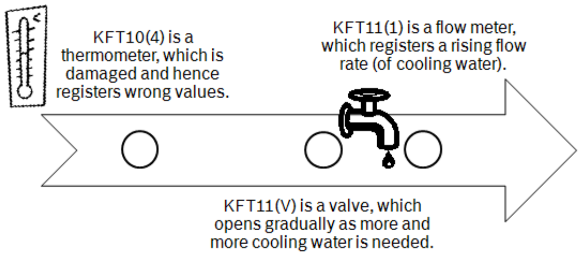


Figure 9 shows an overview of the features, their range, and further statistics. It illustrates that the variance of the KFT17(3) sensor scatters much more than the variance of all the other sensors. Inspection of the data shows an interesting pattern of the KFT17 sensor-actuator pair. (Note that this sensor-actuator pair includes a sensor with the identifier KFT17(3) and an actuator with the identifier KFT17(V)). Whenever the variance of KFT17(3) strongly scatters, the particle size rises, i.e. the quality of the product suffers. Further investigation reveals that every time the process is restarted KFT17(V) oscillates wildly until it reaches the optimal setting (see Figure 10). This causes the observed variance pattern, which corroborates the second hypothesis mentioned above, i.e. the process control of this unit is in need of improvement.

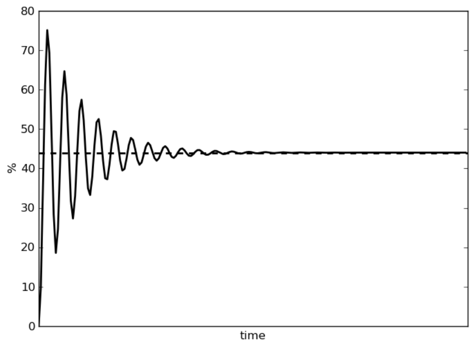
Figure 9: Overview of the Features incl. Statistics (Extract)

ExampleSet (336 examples, 0 special attributes, 117 regular attributes)

Role	Name	Type	Statistics	Range	Mi...
regular	KFP11(V) VARIANCE	integer	avg = 755 +/- 88.033	[600.000 ; 900.000]	0
regular	KFP11(1) VARIANCE	integer	avg = 740.714 +/- 89.203	[600.000 ; 900.000]	0
regular	KFT17(3) VARIANCE	integer	avg = 695.089 +/- 728.504	[150.000 ; 3600.000]	0
regular	KFP6(V) VARIANCE	integer	avg = 1082.679 +/- 258.950	[600.000 ; 1500.000]	0
regular	KFP3x4(2) VARIANCE	integer	avg = 1081.964 +/- 258.250	[600.000 ; 1500.000]	0
regular	KFT13x1(2) VARIANCE	integer	avg = 1075.268 +/- 272.947	[600.000 ; 1500.000]	0
regular	KFT15(4) VARIANCE	integer	avg = 1075.179 +/- 262.607	[600.000 ; 1500.000]	0

After this eight-week experiment, the factory manager, the foremen, and the lab assistants agree to model business rules which constantly examine the sensor and actuator values. The rules, for instance, cross-check the values registered by the sensors (i.e. flow meters) in consecutive units. Moreover, they automatically search for uncommon patterns, such as the oscillating valve opening in Figure 10.

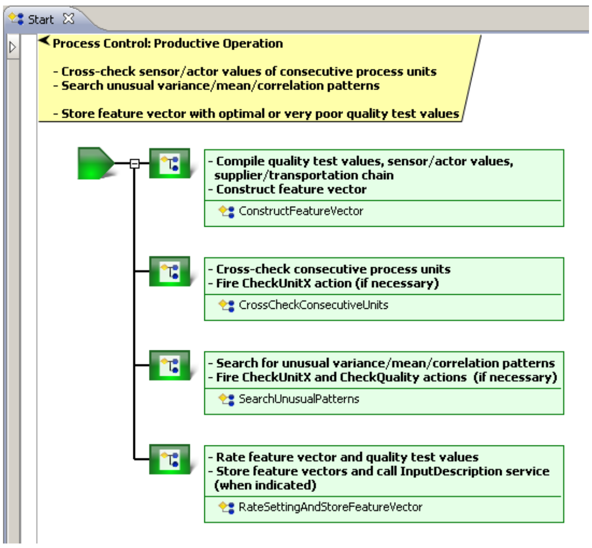
Figure 10: Opening of KFT17(V) Valve after Process Restart



In order to further improve the process and track quality issues, the rules also continuously sample feature vectors with (a) optimal/near optimal and (b) poor quality test values. These can be inspected e.g. every four to six months by a team of quality experts. See Figure 11 as an illustration.

Although the process control system could in principle conduct these consistency checks (at least in part) and sample the feature vectors, using business rules makes it possible to promptly run customized tests: if, for instance, the foremen decide to check a specific unit or process step, they are able to adapt the business rules independently of any IT consultants and run the experiment in no time.

Figure 11: Regularly Executed QA Business Rules Model (Extract)

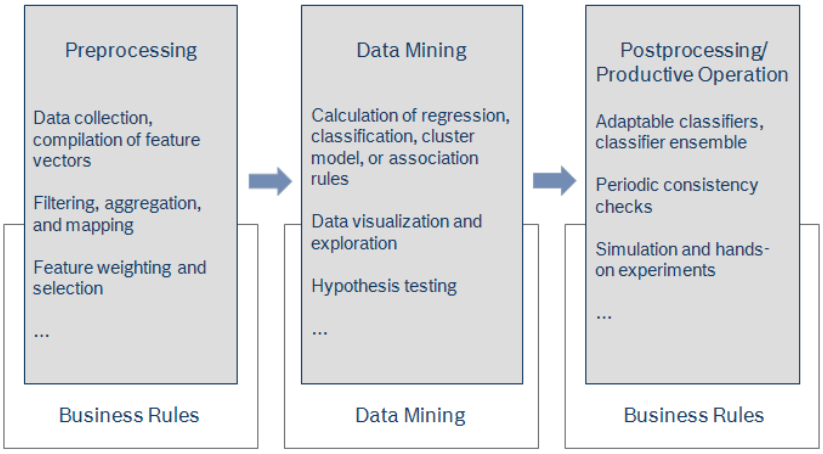


Discussion: You have the impression that exploring data is poking about in the fog? You are right. As the process control example stresses, it is much less straightforward than computing a classifier. But it can be managed if hypotheses are constructed beforehand. Given these hypotheses, it is possible to decide what data mining methods you need to employ, what you can model in terms of business rules, and how to combine both approaches to the best advantage.

Summary

As the two scenarios illustrate, business rules and data mining approaches address two different types of tasks: while data mining helps to explore data and find yet unknown patterns, business rules help to handle the necessary data (pre-/post-) processing steps and calculate known parameters. Many real-world applications can profit from a combination of both approaches as sketched in this paper. See Figure 12 as an illustration of the steps involved. If you want to design such a combination, you need to

Figure 12: Integration of Business Rules and Data Mining Approaches (Process)



- Determine what you know and what you do not know about your data:
  - You can model the first in terms of business rules and
  - You can compute the second with data mining methods
- Define how to explore the data, i.e. which data mining methods might help you to find the relevant patterns
- Decide how to integrate your business rules and the results of the data mining steps

Experience shows that you should discuss your **specific use case** with both a business rules expert and a data mining expert. You will then get a feel for what to model in terms of business rules and which data mining methods to use for data exploration.

Acknowledgements

We would like to thank Andreas Müller, Jan Trnka, Dr. Jürgen Cramer, and Stefan Schacht for their helpful comments regarding the two scenarios described in this paper. We would also like to thank Eric Düll for modeling the business rules.

References & Software

Au, T., Li, S., & Ma, G. (2003). Applying and Evaluating Models to Predict Customer Attrition Using Data Mining Techniques. *Journal of Comparative International Management, 6*(1).

Hay, D., Healy, K. A., & Hall, J. (2000). *Defining Business Rules - What Are They Really? The Business Rules Group. Final Report Version 1.3.* Retrieved January 11, 2001 from http://www.businessrulesgroup.org/first_paper/br01c0.htm.

Tulyakov, S., Jaeger, S., Govindaraju, V., & Doermann, D. (2008). Review of Classifier Combination Methods. In: S. Marinai & H. Fujisawa (eds.). *Studies in Computational Intelligence: Machine Learning in Document Analysis and Recognition*, p. 361-386.

All business rules were modeled with the business rules engine **Visual Rules**. Please refer to http://www.visual-rules.com/technology.html for details on its rich features and intuitive usage. If you want to learn more about modeling business rules and integrating the business rules approach into your architectural environment, please have a look at the following publications: *Full-Power Launch with Your Rules Project* (Javamagazin 5/2010) and *Golden Rules: New and Time-Tested Positions for Rules in Software Architecture* (JavaSPEKTRUM 5/2010). Both papers are available in the Bosch SI media center (see http://www.bosch-si.com/media-download.html). You also find things that are new and noteworthy in our technology blog (see http://blog.bosch-si.com/).

All machine learning experiments were conducted with the machine learning toolkits **KNIME** and **RapidMiner**. Please refer to http://www.knime.org/ and http://rapid-i.com for details on their features and several very illustrative usage examples.

Author

Dr. Irene M. Cramer (irene.cramer@bosch-si.com) holds a Master's Degree and a PhD in Computational Linguistics. Before she started to work at Bosch Software Innovations in 2009, she was a member of the International Post-Graduate College "Language Technology and Cognitive Systems" (Saarbrücken/Edinburgh) and worked as a researcher in the AI projects SmartWeb (Saarland University) and HyTex (TU Dortmund University). She has published various papers and monographs on language technology, data mining, and business rules.