#### PSCharacterizing (Oil Sands) Ore Body with Process Performance Indicators using Pattern Recognition\*

#### Yang Wang<sup>1</sup>, Dean Wallace<sup>2</sup>, and Leonard Slipp<sup>1</sup>

Search and Discovery Article #40699 (2011) Posted February 25, 2011

#### **Abstract**

One of the biggest challenges for oil sands extraction operations is to minimize the number of undesirable events caused by the variation of ore feed. Information on the properties of the incoming ore feed should be the basis to establish mitigation and control strategies so that operators can proactively deal with the dynamics of the feed. Common properties such as bitumen concentration of the ore reveal very little about how the ore will process in the extraction circuit. The process is governed by the interactions of many factors and understanding how these combine to impact behaviour in the extraction circuit is often more valuable from a control perspective. This is made possible by assigning a processability performance indicator for each block in the ore body on a granular basis.

It is generally accepted that many other characteristics of the ore such as connate water chemistry and mineralogy combine with the physical factors to influence the overall performance of the ore. The relationship is a complicated high dimensional nonlinear interaction among many variables. This paper proposes a data driven method for associating ore properties with plant performance indicators using statistical pattern recognition techniques.

Comparing with numerical approaches such as regression and neural networks, the proposed method has a number of advantages:

- 1. Both of the learning and prediction steps are transparent to the investigators. Engineers can easily understand and interpret the outcomes;
- 2. It does not require that the learning data set completely describes the domain;

<sup>\*</sup>Adapted from poster presentation at AAPG International Conference and Exhibition, Calgary, Alberta, Canada, September 12-15, 2010

<sup>&</sup>lt;sup>1</sup>Pattern Discovery Technologies Inc., Waterloo, ON, Canada (wang,yang@patterndiscovery.com)

<sup>&</sup>lt;sup>2</sup>Dean Wallace Consulting Inc., Beaumont, AB, Canada

- 3. It does not interpolate the data to fit a pre-defined model;
- 4. No bias is assumed before the data analysis phase, making this approach good at detecting unexpected associations;
- 5. The proposed method is also capable of discovering high dimensional nonlinear relationship;
- 6. Works well with noisy data;
- 7. Acts upon different data types, continuous or categorical.

Yang Wang, Dean Wallace, and Leonard Slipp

PATTERN DISCOVERY
TECHNOLOGIES IN C
Discover Patterns That Matter™

# 1. The Challenge

How do you minimize the number of undesirable events caused by the variation of ore feed?

### Overview

One of the biggest challenges for oil sands extraction is to minimize the number of undesirable events caused by the variation in ore feed. That variance is due to the diverse combinations of various chemical and mineralogical properties that vary throughout the feed. Common properties, such as the bitumen concentration of the ore, on their own reveal very little about how the ore will process. Rather it is the combination of various elements that determine the efficiency of the system; understanding how these various combinations impact process-behaviour is pivotal to reducing undesirable events, such as plant upsets and diminished product quality.

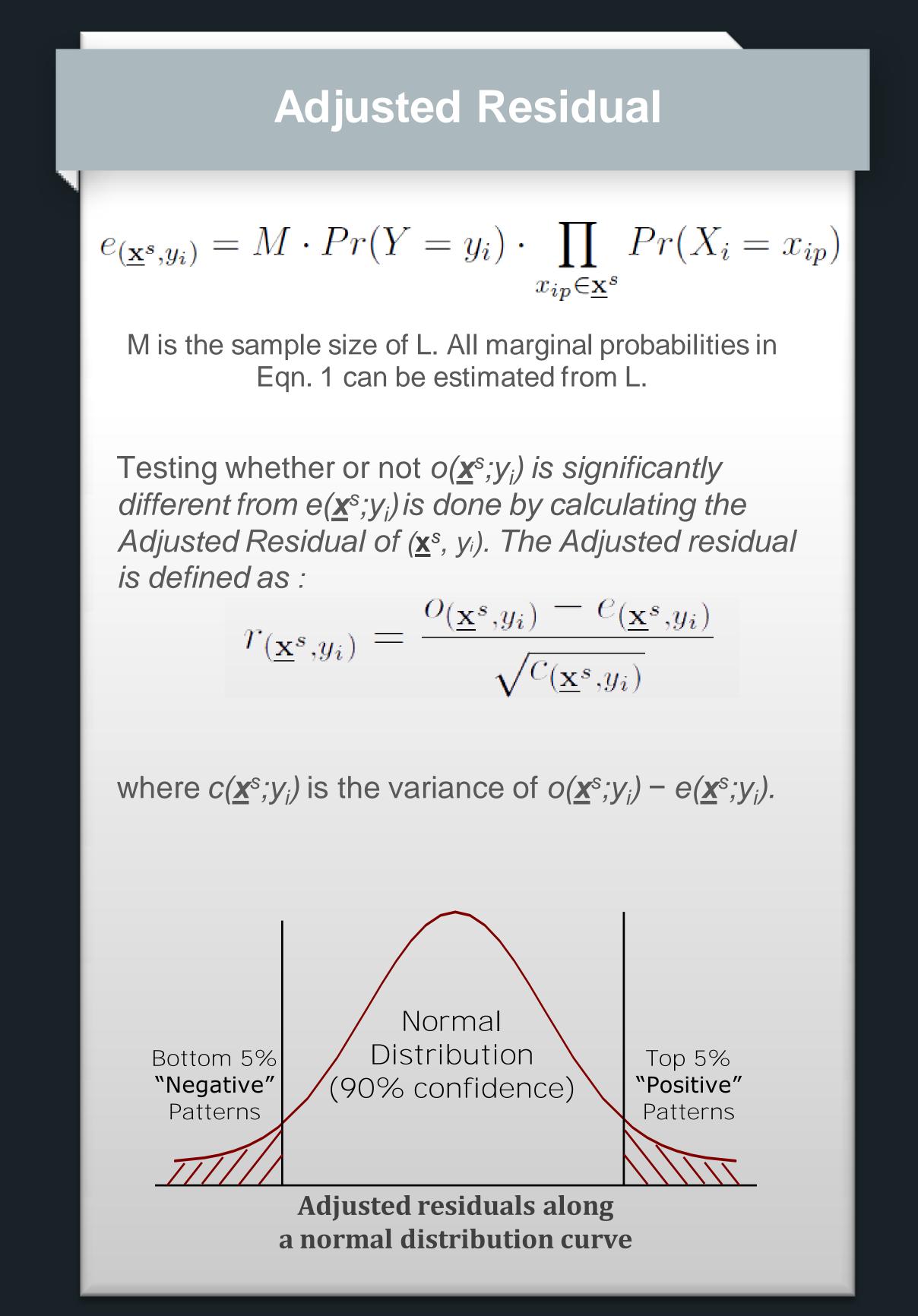
However, the relationship between ore feed constituents is a complicated, high dimensional nonlinear interaction among many variables [1]. What is proposed here is a data driven method for associating ore properties with plant performance indicators using statistical pattern recognition techniques.

### Advantages to Using Statistical Pattern Recognition Techniques

 Both of the learning and prediction steps are transparent to the investigators;

- It does not require that the learning data set completely describes the domain;
- It does not interpolate the data to fit a predefined model;
- No bias is assumed before the data analysis phase, making this approach adept at detecting unexpected associations;
- The proposed method is also capable of discovering high dimensional nonlinear relationship;
- Performs well with noisy data;
- Acts upon different data types, continuous or categorical.

# 2. Method



### Discover Association Rules among Properties and Performance

The rule discovery process consists of two steps. First, the statistically significant event associations between a combination of different ore properties X and the performance indicator Y are detected. Then these event associations are converted into production rules suitable for prediction and classification.

#### Rule Discovery as Event Association

The rule discovery process is to statistically test if the association between a combination of ore properties **x**<sup>s</sup> and a performance category  $y_i$  is significant using the learning data set L. This test is based on residual analysis.

As stated earlier, the  $i_{th}$  record in L can be represented in the form of  $(x_{1i},$  $\cdots$ ,  $x_{Ni}$ ,  $y_i$ ). We call  $(x_{1i}, \cdots, x_{Ni})$ , or  $\underline{x}$ , an observation (corresponding to an ore type) of order N (in total N ore properties are of interest). If, from L, we can confirm a significant association between <u>x</u> and y<sub>i</sub>, we obtain an Nth order rule. In reality, due to the limited number of records in L and the incomplete observations of different ore types and performance combinations, hardly any Nth order rules can be discovered statistically. Lower order rules have to be detected to make use of partial information. We denote **x**<sup>s</sup> as a subset of **x** that contains s values (ore properties) where  $s \le N$ . If  $(\underline{\mathbf{x}}^s, \mathbf{y}_i)$  is a confirmed significant association, we obtain an s order rule of  $\underline{\mathbf{x}}^s \rightarrow \underline{\mathbf{y}}_i$ .

# Method cont.

Rule Strength by Weight of Evidence

Using weight of evidence, we can now give every rule "If  $\underline{\mathbf{x}}^{s}$  then  $Y = y_{i}$  a measure of strength, which has some unique features when the rule is applied to determine the possible performance category of an ore feed.



The discovered rules form a rule model which can be applied to determine the likely performance  $y_i$  of a given ore type  $\mathbf{x}$ which has not been seen before. Assigning a performance indicator to a given ore type is a classification process using

Assigning a Performance Indicator to an Ore Type

$$I(Y = y_i : \underline{\mathbf{x}}^s) = \log \frac{Pr(Y = y_i \mid \underline{\mathbf{x}}^s)}{Pr(Y = y_i)}$$

Clay Activity (e.g. methylene blue index)

•Particle Size (e.g. fines content and average particle

Ore Properties

Connate Water Chemistry

Dean Stark

### Performance Indicators

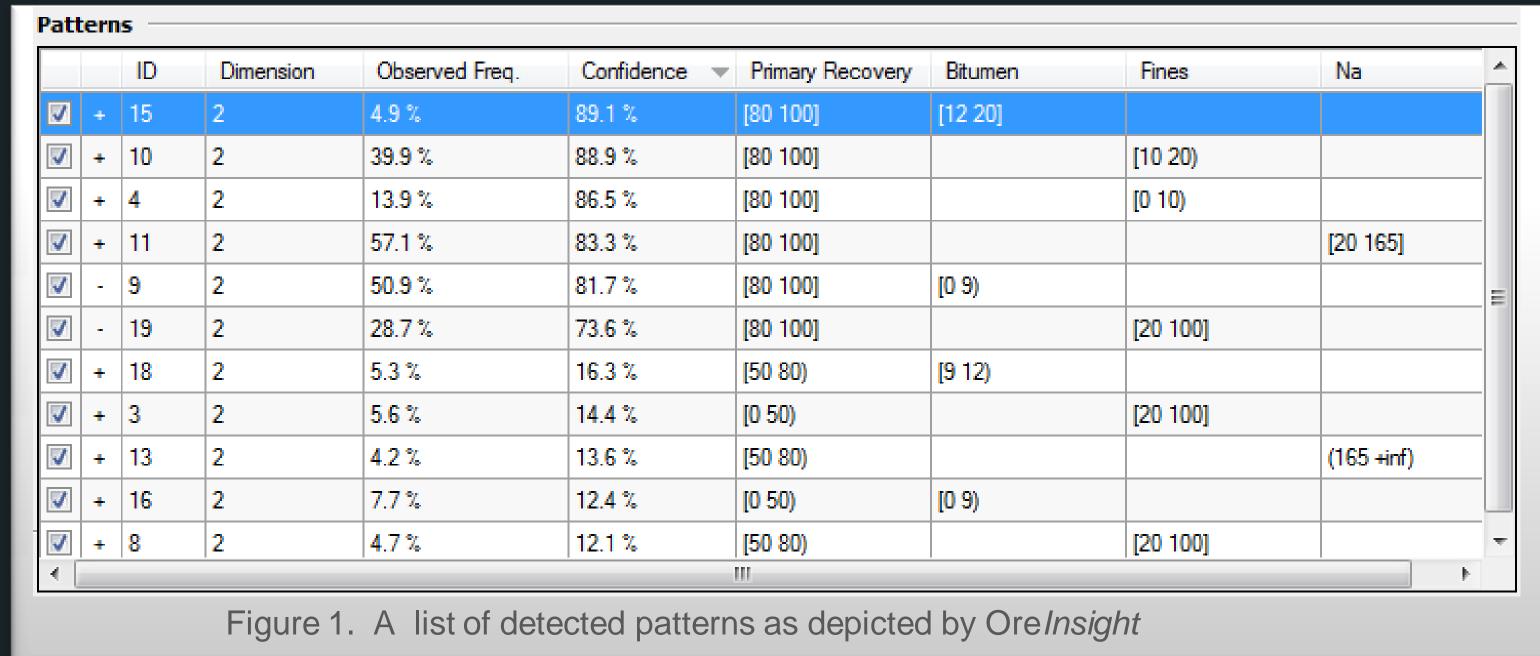
the rules discovered earlier in the process.

•Recovery Percentage: primary recovery, total recovery

•Froth Quality: bitumen percentage in primary froth, water percentage, bitumen/water ratio, etc.



Statistical Pattern Recognition Techniques provide a data-driven approach to obtain knowledge of ore processing performance from its chemical and mineralogical properties. The resulting performance-predicting rules clearly indicate their underlying statistical foundation, producing results that facilitate interpretations.



weight of evidence calculated from the underlying statistics. These patterns have a corresponding confidence level (displayed as a percentage) as well as an observed frequency. Individual rules relate Primary Recovery with statistically relevant relationships between chemical and mineralogical concentrations.

Patterns are reported and ranked based on the

# 3. Results

Domain experts have the opportunity to examine the rules before putting them into action. Performance predictions assigned to geological blocks are also transparent in that the rules contributing to the conclusion can be easily examined. If there is not enough evidence, no prediction will be made.

The transparent nature of the approach makes it possible to easily interpret the results of the rule discovery process and the predictions generated so that risk mitigation strategies can be developed for poor performing ores in a feed forward fashion.

# Application

### Implementation for Oil Sands Processing

1. Obtain a learning data set containing known ore types and their performance values

2. Convert any real numbered variables into categories

3. Perform the rule discovery learning process to extract a set of significant association-rules from the training data

4. Assign the most likely performance indicator to every block in the geological block model using the discovered rules

Once performance indicators are incorporated into the geological block model, doors to other applications are open. Mine planning based on realistic processability targets and where to locate shovels to optimize ore feed to the plant are just two possibilities.

To date, the proposed method has been successfully applied in different forms to several surface mining oil sands projects.

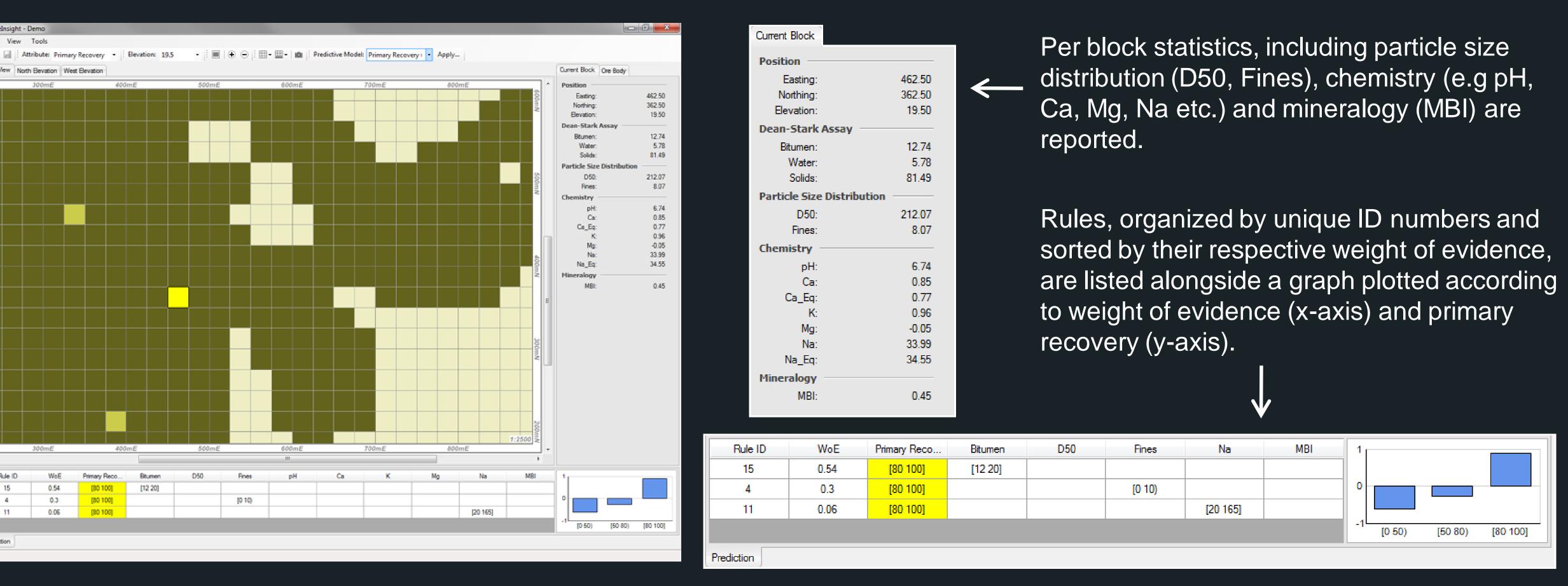


Figure 2. A topographical view of an ore body model including a processability prediction as depicted by Ore *Insight* 

Figure 2 gives the recovery performance map at a particular elevation of a demonstration mine. Different shades of brown represent a different potential recovery rate for each block.

References 1] D. Wallace, H. Pinto, D. Fisher, and Y. Wang, "Application of knowledge discovery techniques to improve efficiency of oil sands characterization and extraction operations", Tech. Rep., Alberta Research Council and Pattern Discovery Software Systems Ltd., 2002. [2] S. J. Haberman, "The analysis of residuals in cross-classified tables", Biometrics, vol. 29, pp. 205–220, 1973. [3] S. J. Haberman, Analysis Of Qualitative Data, vol. 1, Academic Press, 1978. [4] A. K. C. Wong and Y. Wang, "High order pattern discovery from discrete-valued data", IEEE Trans. On Knowledge and Data Engineering, vol. 9, no. 6, pp. 877–893, 1997. 5] D. B. Osteyee and I. J. Good, *Information, Weight Of Evidence, The Singularity Between Probability Measures and Signal Detection, Springer-Velag, Berlin, Germany,* [6] Y. Wang and A. K. C. Wong, "From association to classification: Inference using weight of evidence", IEEE Trans. On Knowledge and Data Engineering, vol. 15, no. 3, pp. [7] L. Liu, A.K.C. Wong, and Y. Wang, "A global optimal algorithm for class-dependent discretization of continuous data", Intelligent Data Analysis, vol. 8, no. 2, pp. 151–180,