

# A Next Generation of Data Mining Solution, E-BizMiner

Dongwoo Kim, Jongsoo Kim, Myungchan Kim

*Netinformatics*

*Yeouido-dong*

*Seoul, Korea*

*jskim@netinformatics.co.kr*

Moonyul Huh, Chongsun Park

*Sungkyunkwan University, Statistics*

*53-3 Myungryun-dong, Jongro-gu*

*Seoul, Korea*

*cspark@skku.ac.kr*

## 1. Introduction

The E-bizminer data mining solution is a knowledge discovery and management platform for large data sets in e-CRM. It is an integrated and component-based data mining system to support users from diverse area of CRM, e-business and others.

## 2. Characteristics & Improvements

E-BizMiner has following characteristics in the Hardware aspects:

- E-BizMiner modules for analysis are implemented in pure Java to make it possible to be platform independent.
- The core engine for analysis is placed in the server to process large data set easily and fast. To support multiuser and efficient environment, remote user use client program to communicate with server program in running relevant module and receiving results.
- E-BizMiner is component-based system using EJB and Java beans. Each component is a specialized analysis tool which could be developed into an integrated solution to a specific need by combining several relevant components together.
- Data mining process can be expressed with analysis components and flows connected them in a stream-like fashion. Process-flows which is developed or existing can be used later as well as in other similar problems.
- XML enables information flow and sharing between components and make it possible to save and retrieve results from each component for automated output reporting.

E-BizMiner has lots of improvements compared to other softwares too:

- E-BizMiner provides lots of important analysis tools for many industries from production to e-business. This will include Regression, Neural Network, Decision Trees, Association Rules, and other statistical modeling and descriptive tools.
- Knowledge discovery from a large data set using E-BizMiner through its unique and automated Process-Flow models consists of Situation, Data collection, Data manipulation, Modeling, Assessment, and Feedback.
- Beginners can get results from minimum input without worrying too much about technical details. On the other hand, E-BizMiner provide full of controllable option menus for field experts who want to adjust various optional parameters as they want.
- Easy interpretation of output results is possible from customized reporting which includes core results from the analysis at hand.
- For optimal data preparation, users can integrate necessary data easily through ETT (Extraction, Transformation, Transport) tools.

## REFERENCES

Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (1996). *Advances in Knowledge Discovery and Data Mining. The MIT Press.*

Chapman, P., Clinton, J., Khabaza, T, Reinartz, T, and Wirth, R. (1999). The CRISP-DM Process Model. *Discussion paper.*

## RESUME

Dans cette recherche, on introduira une minimale solution de données, c'est-à-dire E-BizMiner qui est un savoir découverte et une gestion plate-forme de vaste donnée de domaine divers. C'est une donnée intégrée et composante basée sur minimal système qui utilise EJB et Java Beans. Et en même temps, elle est indépendante de plate-forme. Le noyau d'engine pour analyse est encastré facilement et vite dans le server de procédé. L'utilisateur éloigné utilise un client programme pour communiquer avec server programme. Il execute un module nécessaire et obtient un résultat. E-BizMiner contient un grand nombre d'outils d'analyses importantes pour beaucoup d'industrie de production à e-business. Ceci contiendra 'Regression', 'Neural Network', 'Decision Trees', 'Association Rules', et des autres modèles statistiques et des outils descriptifs.