PERGAMON

Expert Systems with Applications 23 (2002) 1–8

Expert Systems with Applications

www.elsevier.com/locate/eswa

Knowledge-based data mining of news information on the Internet using cognitive maps and neural networks

Taeho Hong*, Ingoo Han

Korea Advanced Institute of Science and Technology, Graduate School of Management, 207-43 Cheongryangri-Dong, Dongdaemun-Gu, Seoul 130-012, South Korea

Abstract

In this paper, we investigate ways to apply news information on the Internet to the prediction of interest rates. We developed the Knowledge-Based News Miner (KBNMiner), which is designed to represent the knowledge of interest rate experts with cognitive maps (CMs), to search and retrieve news information on the Internet according to prior knowledge, and to apply the information, which is retrieved from news information, to a neural network model for the prediction of interest rates.

This paper focuses on improving the performance of data mining by using prior knowledge. Real-world interest rate prediction data is used to illustrate the performance of the KBNMiner. Our integrated approach, which utilizes CMs and neural networks, has been shown to be effective in experiments. While the 10-fold cross validation is used to test our research model, the experimental results of the paired *t*-test have been found to be statistically significant.

Keywords: Data mining; Internet; Cognitive maps; Neural networks

1. Introduction

Nowadays, the capability to both generate and collect data has been expanded enormously and provides us with huge amounts of data. Millions of databases are being used in business data management, scientific and engineering data management, as well as other applications. Data mining has become a research area with increasing importance with the amount of data greatly increasing (Changchien & Lu, 2001; Chiang, Chow, & Wang, 2000; Fayyads, Piatesky-Shapiro, & Smyth, 1996; Park, Piramuthu, & Shaw, 2001). Furthermore, data mining has come to play an important role since research has come to improve many methods used in data mining applications including statistical pattern recognition, association rules, recognizing sequential or temporal patterns, clustering or segmentation, data visualization, and classification.

Although most data is stored in a database from which it can readily be applied to a data mining application, some kinds of data such as news information is not. As the popularity of the World Wide Web increases, many newspapers expand their services by providing news information on the web in order to be more competitive and increase benefits. The web disseminates real time news to investors. News information includes articles on the political situation, social conditions, international events, government policies, trader's psychology, and all those topics, which we see and understand through the Internet. Such information is formulated in the form of texts, referred to as documents, and thus text mining is required if the information is to be applied in data mining applications.

Many researchers attempt to predict interest rates by using the time series model (Bidarkota, 1998), neural networks model (Hong & Han, 1996), the integrated model of neural networks and case-based reasoning (Kim & Noh, 1997). Meanwhile another approach was attempted in the prediction of the stock price index where Kohara, Ishikawa, Fukuhara, and Nakamura (1997) took into account non-numerical factors such as political and international events from newspaper information. They insist that, with event information acquired from newspapers, this method improves prediction ability of neural network. Although they personally read newspapers and rated each political and international event according to their judgment, it is, however, not easy for people to search and retrieve the vast amount of news simply through his/her knowledge and capacity. So we propose a means of applying news information from the Internet for the prediction of interest rates. The system discussed here, named the Knowledge-Based News Miner (KBNMiner), is designed to adopt a prior knowledge base, representing expert

* Corresponding author. Tel.: +82-2-958-3131; fax: +82-2-958-3604.
E-mail addresses: hongth@kgsm.kaist.ac.kr (T. Hong).

0957-4174/02/$ - see front matter
PII: S0957-4174(02)00022-2

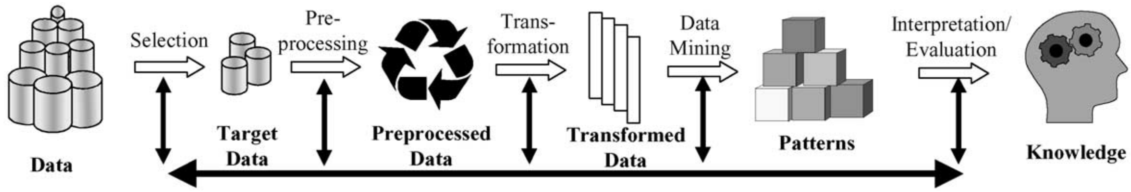

Fig. 1. Overview of the steps constructing the KDD process 0.

knowledge, as a foundation on which to probe and collect news and then to apply this news information to a neural network model for interest rate predictions.

A cognitive map (CM) is used to build the prior knowledge base. CM is a representation perceived to exist by a human being in a visible or conceptual target world. CM manages the causality and relation of non-numeric factors mentioned earlier. The KBNMiner retrieves the event information from news information on the web utilizing CM and prior knowledge. Event information is divided into two types in the KBNMiner. One is positive event information, which affects the increase of interest rates, and the other is negative event information, which affects the decrease of interest rates. A neural network model is developed and experimented on using event information.

This study focuses on the effect news information can have on the prediction of interest rates. As discussed earlier, the event information, which is acquired by the KBNMiner, is applied into a neural network model for the validation of our suggested method. More, specifically, the following research question is addressed:

- What is the effect of the event information on the neural network performance when compared to other prediction models with no event information such as the neural network and random walk models?

In Section 2, we provide a brief overview of data mining and discuss the CM method employed in KBNMiner and the way to build prior knowledge with CMs. Section 3 introduces the architecture of KBNMiner and presents a detailed description of KBNMiner. In Section 4, interest rate prediction data is used to illustrate the performance of KBNMiner. And we present the results of our approach and analyze the results statistically. Finally, the conclusion is presented.

2. Data mining and knowledge engineering

2.1. Data mining

Data mining has become a research area with increasing importance (Changchien & Lu, 2001; Chiang et al., 2000; Fayyads et al., 1996; Park et al., 2001). Berry and Linoff (1997) defines data mining as the exploration and analysis, by automatic or semiautomatic means, of large quantities of data in order to discover meaningful patterns and rules. Frawley, Piatesky-shapiro, and Matheus (1991) refer to the entire process involving data mining as knowledge discovery in database (KDD). They view the term data mining as referring to a single step in the process that involves finding patterns in the data. However, Allen (1996) notes 'data mining is the entire process of knowledge discovery'. Fayyads et al. (1996) outline a practical view of the data mining process emphasizing its interactive and iterative nature in Fig. 1. The KDD process is summarized as: (1) Learning the application domain, (2) Creating a target data set, (3) Data cleaning and preprocessing, (4) Data reduction and projection, (5) Choosing the function of data mining, (6) Choosing the data mining algorithms, (7) Data mining, (8) Interpretation, (9) Using discovery knowledge. The core of the knowledge discovery process is the set of data mining tasks used to extract and verify patterns in data. However, this core typically composes only a small part (estimated at 15–25%) of the effort of the overall process (Brachman, Khabaza, Kloesgen, Piatesky-Shapiro, & Simoudis, 1996).

Specific techniques used in data mining applications include market basket analysis, memory based reasoning, cluster detection, link analysis, decision trees and rule induction, neural networks, and genetic algorithms, etc. However, most of existing algorithms are primarily data-driven and do not fully exploit domain knowledge and intuition that decision makers in business environment have (Padmanabhan & Tuzhilin, 1999). Data mining with prior knowledge is expected to exhibit superior performance than data mining without. This suggests a need for methods to initially adopt a prior knowledge base in data mining applications and this study thus develops a framework and the KBNMiner system to address this issue.

2.2. Knowledge engineering and cognitive maps

Knowledge is an interesting concept that has attracted the attention of philosophers for thousands of years. In more recent times, researchers have investigated knowledge in a more applied way with the chief aim of bringing knowledge to life in machines. Artificial intelligence has contributed to the perceived challenge by developing new tools to produce knowledge from data. However, knowledge is a complex concept and is itself, invisible. These two factors lead to difficulties in the attempt to manage knowledge. One of the more serious problems is that knowledge is built differently among human beings corresponding to their common experiences. People have knowledge consisting

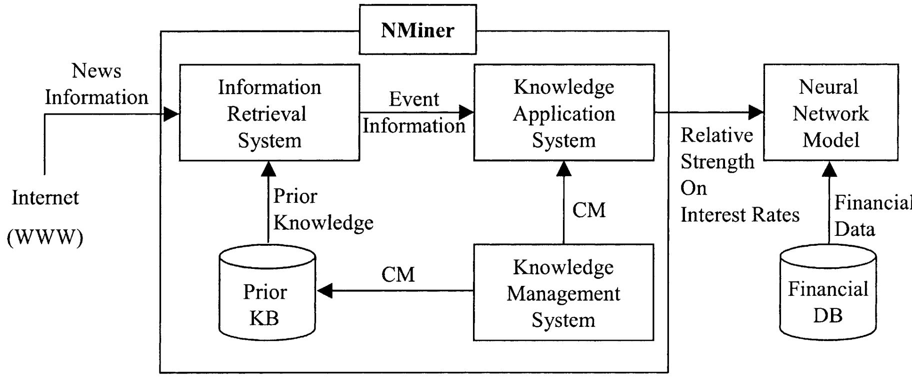

Fig. 2. System architecture of KBNMiner.

of their own views of things and events. It may be difficult to communicate the full meaning of such views to others. The other problem is that knowledge is invisible. Knowledge is represented differently in the process of visualization even though it comes from the same concept. Despite the difficulties faced in artificial intelligence, it is the most effective means to discover knowledge in data or human beings. From the perspective of knowledge representation, the CM is the proper tool by which the perception of human beings can be captured (Park & Kim, 1995; Taber, 1991; Zhang, Chen, Wang, & King, 1992).

CM, introduced by Axelrod (1976) for representing social scientific knowledge, has originally been used for representing knowledge in many studies, representing the cause—effect relationships which are perceived to exist among the elements of a given environment. Though the term 'CM' is used in many different ways, all CMs can be categorized by their target worlds. One category is physical and visible. Another is conceptual and invisible (Zhang et al., 1992). Thus, a CM is a representation perceived by a human being to exist in a visible or conceptual target world.

Integration of knowledge and machine learning has been extensively investigated, because such integration holds great promise in solving complicated real-world problems. One method is to insert prior knowledge into a machine learning mechanism and to refine it with learning through examples (Frasconi, Gori, Maggini, & Soda, 1991; Giles & Omlin, 1993; Towell, Shavlik, & Noordeswiser, 1990). Kohara et al. (1997) suggested the method by which the prediction ability can improve with prior knowledge and event information. They take into account non-numerical factors such as political and international events with newspaper information. In stock price prediction, they categorize event information influencing stock price into two types: negative event information, which tends to reduce stock price, and positive event information, which tends to raise them.

Knowledge-based expert systems usually employ domain experts. And the knowledge engineer is responsible for converting experts' knowledge into the knowledge base. Therefore, the knowledge engineer extracts CMs and has two or more maps of the domain. He generally tries to combine CMs into one. But domain experts sometimes cannot agree with one other. Taber (1991) notes that experts have varied credentials and experience. There is little justification for assuming that experts are equally qualified. Although combined CMs are would always be stronger than an individual CM because the information is derived from a multiplicity of sources and make point errors less likely, it is however not easy for them to have equal weight. Even when experts address the same topic, a map will differ in content and edge weight. For example, three experts estimate (+0.8, +0.8, 0.2), making the average 0.6. This is the error for resulting from weights. On the other hand, the direction of arches among nodes can be derived more easily through their agreement since it has polarity. In this case, the direction is (1, 1, 1) and the result is 1.

3. The proposed knowledge-based data mining system

3.1. System overview

The KBNMiner is designed to utilize CMs as a representation of expertise on interest rate movements, to search and retrieve news information on the Internet according to such prior knowledge and expertise and to further apply this news information to a neural network model in order to achieve more accurate interest rate predictions. KBNMiner consists of several subsystems including a prior knowledge base, information retrieval (IR) system, knowledge application system, and knowledge management system, as illustrated in Fig. 2. KBNMiner is developed by Microsoft® visual basic 6.0 and Microsoft® Access 97 tools. And neural network module is integrated to KBNMiner with the package, NeuroShell® 2 release 4.0 provided by Ward Systems Group, Inc.

The procedure of the KBNMiner is shown in Fig. 3. Prior knowledge is built by using CMs of specific domains as its primary source of solving problems in that domain. The Knowledge Management System (KMS) receives the knowledge built by using CMs and deposits it in the prior

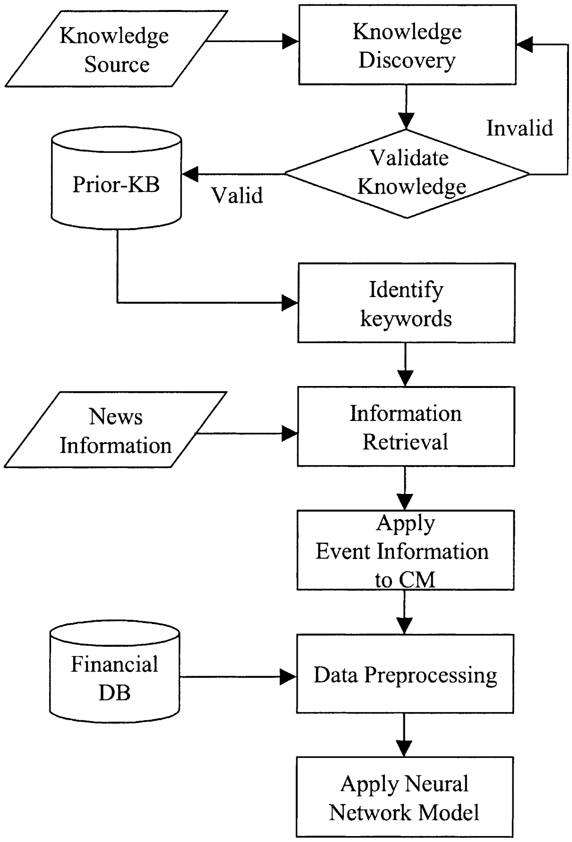

Fig. 3. The procedure of KBNMiner.

knowledge base. The IR System is used to retrieve news information on the Internet by drawing on prior knowledge. The results of the retrieved information are applied to CMs. Knowledge Application Systems apply the retrieved event information to CMs and perform the causal propagation with a causal connection matrix. The final result of the causal propagation is input into a neural network model as positive or negative information along with other financial variables. The details of subsystems are described in Sections 3.2–3.4

3.2. The knowledge management system

Representing an application domain involves much effort, in which experts of the application domain provide knowledge for knowledge engineers, who then have to represent it in an appropriate form suitable for the application. Although domain experts' needs are kept in mind, they only seldom take part in the construction of the knowledge base. The knowledge base cannot ceaselessly be updated, so the updating process becomes a discrete process where the time interval between updates can be lengthy. To overcome this problem, we employed the KMS. The KMS takes on the role of discovering from theory of interest rates and experts' learning and experiences. In addition it also takes on the role of converting the acquired prior knowledge into information to be further applied in IR systems. Thus knowledge is converted into symbolic types in a prior knowledge base after it is discovered by human experts and theory of interest rates The KMS employs CMs for the purpose of (1) knowledge engineering and; (2) storing a prior knowledge base with knowledge acquired.

Human experts consider their experiences and learning about the specified domain and convert their knowledge to a prior knowledge base. The human expert knowledge is deposited into the prior knowledge base and used for retrieving news information in the IR System. We built the prior knowledge base using CM for the prediction of interest rates. CM is constructed in two phases. In the first phase, we define candidate concept nodes affecting the movement of interest rates without any direction among concept nodes after reviewing the theory of interest rates such as loanable fund hypothesis, liquidity preference hypothesis, income effect hypothesis, Fisher's hypothesis and rational expectations hypothesis. The second phase is to determine the final nodes and the direction among them for the CM. The five domain experts who are fund managers in trust and investment-company determined the concept nodes of CM through discussion. They selected the final nodes for the candidate nodes through brainstorming. After the concept nodes are finally determined, they are discussed among the experts and the direction is modified for each node until the conclusions are passed unanimously. No weight was used here. As mentioned in Section 3, we use only polarity among nodes in our CM to avoid the biased weights which result from the diversity in the experience of experts. To acquire a credible CM, we made a decision for the consensus of experts in interaction polarity and did not, the priority or degree of which can hardly be agreed upon among experts. The CM is shown in Fig. 4.

3.3. The information retrieval system

Although there is much research on IR and improved IR techniques such as neural networks and genetic algorithms, we deployed and adopted here the classical IR model which is called *n*-gram matching and modified it to apply Korean characters because our study is focused on the application of the results of IR with prior knowledge, not on IR models. The examples of *n*-gram are explained in the following.

The new printer does not work.

is represented in the form of the following set of 3-grams:

{the, new, pri, rin, int, nte, ter, doe, oes, not, wor, ork}

To retrieve information from Korean news, we defined a keyword set representing concept nodes in CMs (Fig. 4) and built thesauri according to the

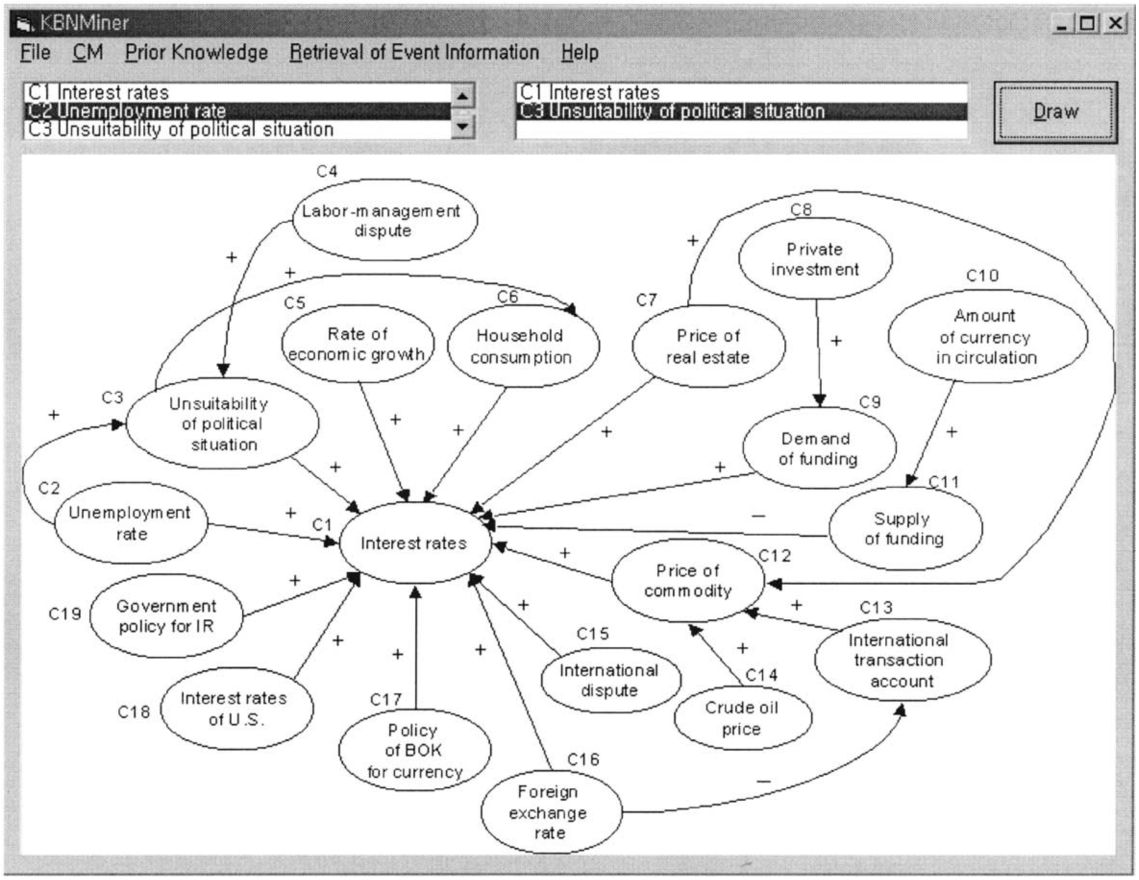

Fig. 4. CM for Korean interest rates.

keyword set. The thesauri are compared to the words in texts in order to find information according to the meaning of concept nodes and we regard those texts, which completely match the thesauri as having the same meaning as those concept nodes. In the earlier case, we represent the set of thesauri as {the, new, printer, does not, work} according to the keyword set representing concept, 'The new printer does not work'. This approach is applied in the Korean language. For example, C_2 node in Fig. 5 means the unemployment rate. The positive information of C_2 increased the C_1 (interest rates) and C_3 (unsuitability of political situation). And the negative information of C_2decreases the C_1 and C_3. Thus Fig. 5 shows an illustrated example that the keyword set for positive events are defined as {unemployment and increase, unemployed person and increase, the state of unemployment and decrease} and the keyword set for negative events are {unemployed person and decrease} according to C_2 node defined here as prior knowledge. This example is merely a translation from Korean into English to assist in the understanding of our IR method.

Our study uses four major newspapers in Korea as the information source to validate our proposed system (KBNMiner). These are the newspapers stored on database in the form of documents.

3.4. The application system

We consider the 19-by-19 causal connection matrix E that represents the CM in Fig. 5 (Fig. 6). With this causal connection matrix, we can apply to the causal propagation (Kosko, 1986; Zhang et al., 1992). Event information, which is gathered by the IR system, is applied to the causal propagation by a causal connection matrix. Let us consider the input vector D. For example, the IR results of the news at 6 January, 1998 is represented by the vector, $D=$ (1 0 0 1 1 1 1 1 1 0 0 1 0 1 0 $-$1 0 0 0), then the output vector is (1 0 1 0 0 0 0 0 1 0 0 1 $-$1 0 0 0 0 0 0) as the results of causal propagation. For more information on casual inference, see Kosko (1986) and Zhang, Chen, and Bezdek (1989).

Finally, we gathered the positive and negative information each 7 days. And the results of causal propagation of positive and negative event information are converted into the relative strength of effects on interest rates. If the relative strength, $\mathrm{EK}_t = \mathrm{Pk}_t/(\mathrm{Pk}_t + \mathrm{Nk}_t)$ is over 0.5, then it can be stated that the positive effect on interest rate is stronger than the negative effect. If the relative strength is under 0.5, then the negative effect on interest rates is weaker than the positive effect. The relative strength is input to neural networks as meaningful signals in application systems.

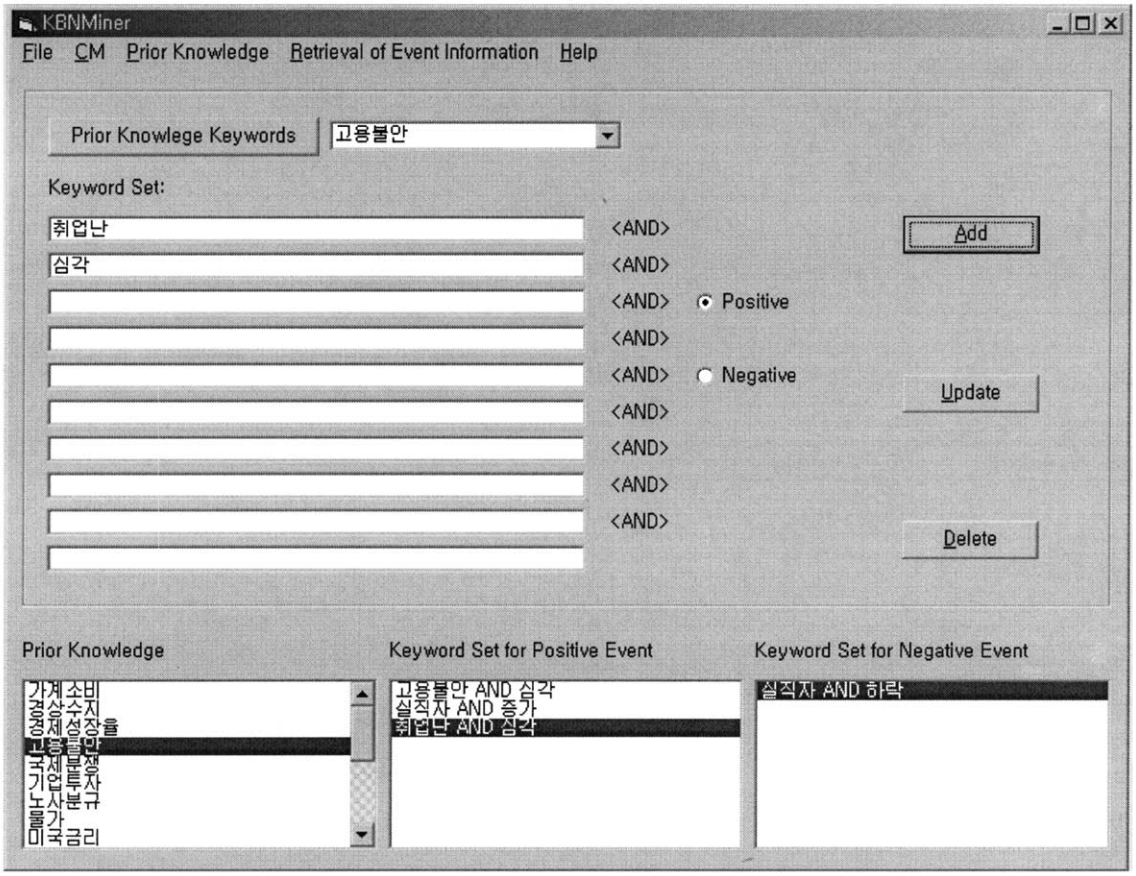

Fig. 5. KBNMiner—setting screen for prior knowledge keywords.

4. KBNMiner for interest rate predictions

4.1. Data and event knowledge

KBNMiner is applied to Korean newspapers in the way mentioned earlier. We selected four major newspapers in 1998. There are about 180,000 articles related to national politics, business and international affairs in the news database. KBNMiner founded 3731 events in all articles matching prior knowledge. We compared the daily Corporate Bond Yields (CBY) with event knowledge on each date and finally got 252 samples in 1998. Our research model was developed for predicting 1 month into the future.

$$
\mathrm{E} = \begin{array}{cc}
 & \begin{matrix} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 & C_8 & C_9 & C_{10} & C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} & C_{17} & C_{18} & C_{19} \end{matrix} \\
\begin{matrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \\ C_8 \\ C_9 \\ C_{10} \\ C_{11} \\ C_{12} \\ C_{13} \\ C_{14} \\ C_{15} \\ C_{16} \\ C_{17} \\ C_{18} \\ C_{19} \end{matrix} &
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

Fig. 6. The causal connection matrix from CM.

4.2. Design of the neural network model

We utilized a neural network model to illustrate our approach with cases and to test the validity statistically for the result of KBNMiner. Three-layer feedfoward neural

Table 1
Variable description

Variable	Description
CBY_{t+n}	N days ahead of the corporate bond yield
CBY_t	Average of corporate bond yield in the previous 7 days
$KOSPI_t$	Average of Korea stock price index in the previous 7 days
FX_t	Average of the foreign exchange rate for Korean Won/US dollar in the previous 7 days
EK_t	Relative strength in the previous 7 days $EK_t = Pk_t/(Pk_t + Nk_t)$, where PK: Number of positive events, NK: Number of negative events

Table 2
Out-of-sample MAE

Model	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	Average
RW	1.676	1.743	1.876	1.617	1.874	1.404	1.812	1.891	1.597	1.754	1.725
NN1	0.557	0.627	0.624	0.503	0.665	0.652	0.626	0.528	0.569	0.506	0.586
NN2	0.652	0.520	0.586	0.479	0.579	0.483	0.585	0.465	0.448	0.469	0.527

Table 3
Out-of-sample MAPE (%)

Model	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	Average
RW	12.06	14.50	15.31	12.78	14.82	11.00	14.75	14.29	12.80	14.05	13.64
NN1	4.03	4.99	4.85	3.71	5.44	4.89	5.01	4.12	4.63	4.18	4.59
NN2	4.36	4.34	4.60	3.51	4.63	3.72	4.65	3.50	3.59	3.85	4.08

networks are used to forecast the Korean interest rates. Logistic activation functions are employed in the hidden layer and linear activation is utilized in the output layer. The number of output nodes is one that is targeted in the neural network forecaster. The number of hidden nodes are selected through experimentation with $n/2$, n, and $2n$ of nodes (n is sum of input and output node) by fixing the input and output nodes.

The input variable of the neural network is described in Table 1. The prediction model is designed to predict 30 days ahead of time. We set the input variables as CBY, KOSPI and FX while considering autoregressive characteristics and correlation to the target variable. These variables are averaged in the previous seven days.

We designed three models to compare the performance of event knowledge in the neural network forecaster: (1) RW is the random walk model. (2) NN1 is the neural network model without event information. (3) NN2 is the neural network model with event information.

4.3. Empirical results

Table 2 shows the results using MAE as the performance measure. The results show the effect of event information comparing 10-fold validation. The NN2 model with event information has a figure of 0.527 in average error, which is the minimum error in comparison to other models such as RW and NN1. We found that the neural network forecaster is greatly superior to random walk and that the effect of event information does exist. It illustrates that information from news mentioned earlier is useful in the prediction of interest rates. It is easy to understand why the same results appear in Table 3, measuring the errors by MAPE.

We attempted to test the results of our experiments statistically. Absolute percentage error (APE) is commonly used (Carbone & Armstrong, 1982) and is highly robust (Armstrong & Collopy, 1992). As the forecasts are not statistically independent and not always normally distributed, we compare the APEs of forecast using the pairwise t-test. Paired t-test results for NN1and NN2 show a significant difference at the 1% level in Table 4. We conclude that event information affects our forecaster model statistically at a significant level. This supports our suggested method, by which the event information is integrated into neural networks with CMs. And our integrated method provides a decision maker with meaningful knowledge to aid in the effective decision making related on the movement of interest rates.

5. Conclusions

The KBNMiner was developed as a means of applying current information acquired on the World Wide Web in the prediction of interest rates. The KBNMiner provides traders or those who are concerned about the movement of interest rates with more relevant knowledge from data and aids in effective decision-making. It is designed not only to apply expert specialist knowledge, but also knowledge of events and conditions influencing interest rate dynamics. The process involves the formation of a prior knowledge base, derived from the CMs of professional experience and learning, upon which the system draws in the search and retrieval news information to further be applied in a neural network model capable of predicting interest rates. The empirical results of our experiments show improvements in performance, when the information news is applied to the neural network. The pared t-test was performed and we attained

Table 4
Paired sample t-test for prediction results of RW, NN1, NN2

Comparison	Paired-t statistics	Error measure	
		MAE	MAPE
RW vs. NN1	Difference	14.310	14.044
	p-value	0.000[a]	0.000[a]
RW vs. NN2	Difference	15.000	14.627
	p-value	0.000[a]	0.000[a]
NN1 vs. NN2	Difference	3.272	3.722
	p-value	0.001[a]	0.000[a]

[a] Significant difference at the 1% level.

significant improvement of the neural network performance statistically.

The research question 'what is the effect of applying event information on the performance of data mining applications' is answered here, in one form, as we have explained how data mining systems with prior knowledge are statistically more effective. We have also described how to apply prior knowledge to data mining systems. Furthermore, our study shows that the CM is a useful tool for representing knowledge and reflecting the causality of knowledge.

Our methods need to refine CMs and to improve the algorithm of the IR system for acquiring more correct results. The more progressive approach should be considered in future although our methods are designed and developed from conservative points, which accompany minimal errors and risks.

References

Allen, L. E. (1996). Mining gold from database. *Mortgage Banking*, *56* (8), 99–100.

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: empirical comparisons. *International Journal of Forecasting*, *8*, 69–80.

Axelrod, R. (1976). *Structure of decision: the cognitive maps of political elites*, New Jersey: Princeton University Press.

Berry, M. J. A., & Linoff, G. (1997). *Data mining techniques for marketing, sales and customer support*, New York: Wiley.

Bidarkota, P. V. (1998). The comparative forecast performance of univariate and multivariate models: an application to real interest rate forecasting. *International Journal of Forecasting*, *14* (4), 457–468.

Brachman, R. J., Khabaza, T., Kloesgen, W., Piatetsky-Shapiro, G., & Simoudis, E. (1996). Mining business databases. *Communications of ACM*, *39* (1), 42–48.

Carbone, R., & Armstrong, J. S. (1982). Evaluation of extrapolative forecasting methods: results of academicians and practitioners. *Journal of Forecasting*, *1*, 215–217.

Changchien, S. W., & Lu, T. (2001). Mining association rules procedure to support on-line recommendation by customer and products fragmentation. *Expert Systems with Applications*, *20*, 325–335.

Chiang, D., Chow, L. R., & Wang, Y. (2000). Mining time series data by a fuzzy linguistic summary system. *Fuzzy Sets and Systems*, *112*, 419–432.

Fayyads, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD processes for extracting useful knowledge from volumes of data. *Communications on the ACM*, *39* (11), 27–34.

Frasconi, P., Gori, M., Maggini, M., & Soda, G. (1991). An unified approach for integrating explicit knowledge and learning by example in recurrent neural networks. *Proceedings of International Joint Conference on Neural Networks* (pp. I-811–I-816). Seattle.

Frawley, W. J., Piatesky-Shapiro, G., & Matheus, C. J. (1991). Knowledge discovery in database: an overview. In Fayyad, G. Piatesky-Shapiro & P. Smyth, *Knowledge discovery in database* (pp. 1–27). Cambridge MA: AAAI/MIT Press.

Giles, C. & Omlin, C. (1993). Rule refinement with recurrent neural networks. *Proceeding of International Conference on Neural Networks* (pp. 801–806). San Francisco.

Hong, T.H. & Han, I.G. (1996). The prediction of interest rates using artificial neural network. *Proceedings of The First Asia Pacific Conference of Decision Science Institute* (pp. 975–984). Hong Kong.

Kim, S. H., & Noh, H. J. (1997). Predictability of interest rates using data mining tools: a comparative analysis of Korea and the US. *Expert Systems with Applications*, *13* (2), 85–96.

Kohara, K., Ishikawa, T., Fukuhara, Y., & Nakamura, Y. (1997). Stock price prediction using prior knowledge and neural networks. *Intelligent System in Account, Finance and Management*, *6*, 11–22.

Kosko, B. (1986). Fuzzy cognitive maps. *International Journal of Man–Machine Studies*, *24*, 65–75.

Padmanabhan, B., & Tuzhilin, A. (1999). Unexpectedness as a measure of interestingness in knowledge discovery. *Decision Support Systems*, *27* (3), 303–318.

Park, K. S., & Kim, S. H. (1995). Fuzzy cognitive maps considering time relationships. *International Journal of Human–Computer Studies*, *42*, 157–168.

Park, S. C., Piramuthu, S., & Shaw, M. J. (2001). Dynamic rule refinement in knowledge-based data mining systems. *Decision Support Systems*, *31*, 205–222.

Taber, R. (1991). Knowledge processing with fuzzy cognitive maps. *Expert Systems with Applications*, *2*, 83–87.

Towell, G., Shavlik, J., & Noordewiser, M. (1990). Refinement of approximate domain theories by knowledge-based neural networks. *Proceeding of National Conference on Artificial Intelligence* (pp. 861–866). Boston.

Zhang, W. R., Chen, S. S., & Bezdek, J. C. (1989). Pool2: a generic system for cognitive map development and decision analysis. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*, 31–39.

Zhang, W. R., Chen, S. S., Wang, W., & King, R. S. (1992). A cognitive-map-based approach to the coordination of distributed cooperative agents. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*, 103–114.