신경망을 이용한 산업주가수익율의 예측

권 영 삼*, 한 인 구**

Industry Stock Returns Prediction Using Neural Networks

Kwon, Young-Sam, Han, In-Goo

The previous studies regarding the stock returns have advocated that industry effects exist over entire industry. As the industry categories are more rigid, the demand for predicting the industry sectors is rapidly increasing. The advances in Artificial Intelligence and Neural Networks suggest the feasibility of a valuable computational model for stock returns prediction. We propose a sector-factor model for predicting the return on industry stock index using neural networks. As a substitute for the traditional models, neural network model may be more accurate and effective alternative when the dynamics between the underlying industry features are not well known or when the industry specific asset pricing equation cannot be solved analytically. To assess the potential value of neural network model, we simulate the resulting network and show that the proposed model can be used successfully for banks and general construction industry. For comparison, we estimate models using traditional statistical method of multiple regression. To illustrate the practical relevance of neural network model, we apply it to the predictions of two industry stock indexes from 1980 to 1995.

* 교보증권 리서치센터 선임 연구원
** 한국과학기술원(KAIST) 테크노경영대학원 부교수

Ⅰ. Introduction

Providing the arguments against Sharpe's market model, many studies have remarked the existence of industry effects on the stock price behavior [King, 1966: Fertuck, 1976: Fabozzi and Francis, 1983: Roll, 1992: Heston and Rouwenhorst, 1994: Kuo and Satchell, 1998]. After the economic crisis in 1997, distinction among the behaviors of the return on industry stock index is much shaper than it was. To this end, A few of the studies adopted the King's empirical scheme to investigate the industry effects of the listed stocks on Korea Stock Exchange [Park, 1990: Jung, 1995]. In particular, Yoon (1994) reported that industry effects heavily depend on the industrial structure and vary largely with business environments. Na (1996) found that the naive factor model associated with 13 common factors and 20 industry dummy variables increased the explanatory power measured with R^2 by 15%. These literatures show how some industry portfolios can have reasonably negative correlation and produce the industry effects in return generating process. Also, the empirical results from these studies provide the possibilities of identifying the profitable investment strategies using industry effects. Especially, the prediction of the return on industry stock index becomes strategically important for industry rotation in active asset allocation perspective. It is critical for industry specific prediction model to capture the underlying nonlinear dynamics among economic, industry, and market attributes. However, there is no effective analytical framework and methods.

In this article, we propose a neural network-based model for predicting the return on industry stock index. This *nonparametric* model, in which data is allowed to determine both the dynamics and its relation to assets prices, will give the possibility of finding new influencing factors and explaining the structural changes due to the currency crisis. We take as inputs the primary economic variables and industry specific variables that influence the industry asset price, e.g., return on industry stock index or the value of a portfolio containing the stocks within an industry.

Neural networks have several advantages over traditional regression models. First, since they do not rely on restrictive parametric assumption such as normality, they are robust to the specification errors that plague parametric or traditional regression models. Second, they are adaptive and respond to the structural changes in the data-generating processes in ways that regression models can not. Finally, they are flexible to encompass a wide range domain of fixed income, stocks, foreign exchange, and fundamental asset price dynamics. Neural networks have found an important niche in financial market application, especially stock market prediction [Kimoto, 1990: Jang, 1991] and have recently been applied to asset management [Khan, 1996].

However, all these advantages do not come without some costs　insufficient explanations of data-intensive model, requirements of vast data quantity, additional efforts for finding the optimal network architecture, and verification the results under the celebrated theory. Therefore, network-based models must be constructed under the well-understood financial

economics to encapsulate the underlying pricing dynamics. To do this, we employ the multi-factor model scheme proposed by Rosenberg (1974).

In section II, we provide a brief review of industry effects and neural networks. We develop the industry-specific model that decomposes the return on industry stock index into economic, industrial and market attributes to capture the prevailing industry effects. In section III we describe the input variables and draw up the specification of our industrial factor model. We report the results of several prediction experiments in which neural networks yield more profitable results and higher accuracy in section IV. These results are the evidences enough to illustrate the promise of neural network in the application of industry stock index prediction. We find that the neural network-based models outperform the traditional regression models. Also, we apply the trained models to the prediction of monthly industry stock index for both bank and construction. To extend the practical relevance of our resulting model, we suggest several future directions and conclude in section V.

II. Literature Review

2.1. Industry Effects

It has long challenged both academics and professional portfolio managers that knows what factors drive the covariance and how different effects exist in stock returns across industries. Early studies by Grinold, Rudd and Stefek(1989), Drummen and Zimmerman(1992),

and Cavaglia(1995) documented the importance of industrial compositions of international portfolio strategies. In line with these evidences, Roll(1992) suggests that the industrial compositions of a portfolio can explain some of the variance. He finds that industry factors explain approximately 40% of the volatility in stock returns. In contrast, Heston and Rouwenhorst(1994) show that less than 1% of the differences in volatility of national index returns can be explained by their industry compositions.

Nevertheless, with the finer-partitioned industrial classifications and common factors, Na(1996) demonstrates that these industry factors can explain 15% of the volatility in individual stock returns. This finding has important implications for the managers who ignore the industrial mix of their portfolios. In fact, not utilizing an accurate and reliable prediction model, they have no any chance to exploit the investment strategies for industry rotation.

2.2. Neural Networks

Neural networks, arguably the most popular and pertinent tools of *Artificial Intelligence*, are the general category of methods that derive their original inspiration from simple models of biological nervous system. Because of their intrinsic ability to synthesize the models that deal with fuzziness, uncertainty, incomplete data, neural networks have been applied to a wide range of financial domain.

Given the power and flexibility of neural networks to approximate complex nonlinear relations between economy and industry sectors,

network-based models are expected to successfully generate the valuable investment decision supports for stock market prediction. Early experienced with the application of neural networks to financial market, especially since 1990, a number of trading models have been developed for modeling nonlinear statistical relations and predicting the stock market [Chinetti et all, 1993: Choi, 1995]. In particular, neural networks have yielded substantial help in developing and implementing the successful investment system, which is designed to predict the short-term market directions.

Park and Han[1995] adopted a synergistic view on the Korea stock market to predict the short-term trends of KOSPI and identify the best timing for buying and selling the market index. They obtained predictions with more than 60% accuracy in finding the market directions in the next month. Using the forecasts of market index, the only active decision concerns the appropriate allocation of funds between a surrogate market portfolio (KOSPI) and risk free asset(government bonds). Consequently, this investment strategy needs the additional process to determine the appropriate weight in industry groups and individual stocks.

2.3. Industry forecasts and Investment Opportunities

There are many ways to pursue the active management; one of the most popular is to build an equity portfolio based on an industry forecast. After economic crisis in 1997, industry rotation becomes an important invest-

ment theme. Empirical research demonstrates that industry returns follow patterns of persistence as well as patterns of reversal (Sorensen, 1986). This implies that industry return has long term memory to be predicted and may give an opportunity to select the industry group expected to benefit from that forecast. In addition, this forecast can provide the tilted portfolios which match a manager's economic and industry view.

Using this forecast for industry stock index, we are able to differentiate between those groups that represent good value and those that appear expensive. This approach significantly reduces the limitations from the prediction of market timing through providing the implication for undervalued industry. As market goes bearish, this forecast is used to rebalance the current portfolio to create investment opportunities with attractive rate of return.

Recent behaviors of the many industries show that industry do not respond to the same degree over market movement. Economic crisis accelerates the differentiation of industry returns and makes industry category more rigid. This differentiation may translate into falling or rising stock market with different price changes among industry groups. As this phenomenon has been more important in market, the demand for new investment strategy increases among portfolio managers. To this end, we would provide more successful investment opportunities with forecasts generated from the sector-factor model based on industry analysis. Also, our analytical framework can lay the groundwork for developing systematic method and prediction model of

industry stock return.

2.4. Industry sector-factor models

So far, most studies using neural networks do not work out any successful strategy for industry-specific investments because the aim of these is to predict the signs of the monthly or daily difference of market index. Despite the usability of market index, the results of some previous studies report that several industry stock indexes tend to move differently with KOSPI or market index [Farrelll, 1974: Arnott, 1980: Jung and Kim, 1995].

By these findings, we are motivated to develop the stock index prediction model in the context of industry analysis. This model presents the industry that are currently more attractive according to the selected factors. When appraising the relative attractiveness of an industry, we examine several multi sectors including macro economic, industrial, and market features. Eventually, the model can suggest an overweighting and an underweighting signal for two industries(bank and construction). In addition this, we present an analytical framework to evaluate two industries and identify opportunities for industry rotation. To do this, we adopted the typical industry analysis process used by many industry analysts, which relies on top-down analysis to identify industry factors and to evaluate industry effects.

This approach highlights the practical relevance of our model in way that the variable selection steps are differently conducted with previous studies. Furthermore, this approach will offer the additional benefit of interpreting learned knowledge.

In sum, the results from this experiment may be important in that the proposed model is able to provide investors with a better stock selection strategy and a systematic procedure for portfolio construction.

III. Model specification and Neural Networks

3.1. Model for Industry Stock Index

3.1.1. Selected industries: Bank and General Construction

As a market index, KOSPI is widely used for a proxy to predict the future trends of Korean stock market and evaluate whether current index level is overpriced or not. On the other hand, the industry stock index, which represents the value-weighted average of all stocks within an industry, can be useful for industry analysis and industrial mix of portfolio for diversification.

We choose the two popular industry stock indexes to assess the potential value of neural network-based model. We make use of a two-stage approach for industry selection. In the first stage, we performed the extensive data analysis on the industry groups to select the industry, which serves as a leading indicator in stock market. In second stage, we conducted a correlation analysis and checked the trading volume of the selected industries to assess the level of interaction and the degree of coherence across industries. With the aid of industry expertise, we chose two industries containing the valuable implications

and practical relevance. The selected correspond to the bank and general construction.

3.1.2. Variable Selection

A key question for most prediction applications, and especially for financial markets, concerns the influencing factors and their economic values interpreted by the outputs of the derived model. It is an extremely challenging task to extract the important factors from a vast array of studies that report the different results. These kinds of factors spread out in type of quantitative and qualitative data as well as in diverse financial sectors. Although new techniques have been suggested for variable selection, most of them failed to produce the best promising results in predicting the stock price.

The subject of selecting the input variables is very controversial when it comes to the prediction applications. In the previous studies with neural network, there have been another critical issues regarding the input variable selection. The most common way to choose the inputs is to select a rather large group of independent variables and reduce that to a smaller group of statistically significant variables. The usual procedure to filter statistically is stepwise technique.

However, the linear correlation-based and parametric techniques have been criticized due to their masking effects on other variables during selection process. It also encounters difficulties with the incorporation of prominent factors because the important features of the data are likely to be missed by assessing their relative linear correlation. In principal,

they cannot give any clear answer why linear and parametric methods are used for nonparametirc and nonlinear model specification.

Because many previous studies seem to cast doubt on the input variable specification, the sector-factor approach may be important alternative to the traditional methods. In the security analysis perspective, this approach provides the key implications that are crucial to understanding the economic meanings of variables. In particular, our sectorfactor approach can yield valuable insights in at least three contexts. First it provides us with an arbitrage-free method of selecting input variable. Second, our model captures those features of the data which are most salient from industry analysis perspective and which ought to be incorporated into any successful prediction model. Third, it also helps us understand what features are importantly treated by the resulting model.

In the market analysis, portfolio managers need to analyze inter-market and industrial relationships. From this, synergistic market analysis, benefited from both technical and fundamental analysis, was introduced to meet new security analysis challenges of the 1990s by Mendelsohn(1991). We borrow the idea of Mendelsohn to incorporate the technical implications into our model. All candidate factors considered in the model are grouped by four categories; macroeconomic, inter-market, industrial, and technical factors.

Table 1 lists selected potential variables to be examined for selecting the inputs and training networks.

<Table 1> Selected variables, labels and descriptions, categorized into four input sectors

Sector	Label	Description
	FXs	Foreign exchange rate(W/$ and Y/$)
	CPI	Consumer price index
Economic	M2	M2 as a liquidity indicator
	CBY	Corporate bond yields(bank-guaranteed)
	PI	Raw material price index
Industrial		
	CAO	Contract amount ordered
Construction	CAP	Construction area permitted
	CW	Average wage
	CD and CP	Construction Industry dividend and PER
	BTS	Total amount of saving
Banks	BTL	Total amount of loan
	BI	Interest spread
	BD and BP	Bank Industry dividend and PER
Inter-Market	TV and TA	Trading volume and amount
	KO	KOSPI
	SL	Regression coefficient
Technical	RSI	Relative strength index
	DI	Disparity
	MA	Moving average

While the accuracy of the trained network is obviously of great interest, if many independent variables are employed, this is not sufficient to ensure the practical relevance of the neural network model. This additional constraint forced to reduce the candidate variables to the appropriate size. The set is filtered by the results of the interview with industry analysts. These results are taken as inputs for industry sector- factor model.

3.1.3. Multi-Factor Model and Sector-Factor model

In this study, model specification begins with the construction of the possible variable sets. We use multi-factor model[1] Several simplifications of this model have been used historically. For example, if there is only one factor, it becomes a single-factor model; if the single factor is identified with market factor, it becomes the market model presented by Sharp(1963). scheme to specify the return process of industry stock index and to identify the candidate inputs. Multi-factor model, pro-

---

1) Several simplifications of this model have been used historically. For example, if there is only one factor, it becomes a single-factor model; if the single factor is identified with market factor, it becomes the market model presented by Sharp(1963).

posed by Rosenberg(1974), states that the rate of return on any security(asset) is equal to the weighted sum of the rates of return on a set of common factors, plus the specific return on the security, where the weights measure the exposures of the security to the factors. These exposures are identified with the macro-economic characteristics, or descriptors of the firms.

By requiring the multi factor-based variables to our models, we may explicitly capture their impacts on the industry stock returns. Instead of a single factor model, a multi-factor model for industry analysis that considers various influences will be more accurate.

To improve the multi-factor scheme proposed by Rosenberg(1974), we consider the additional industry-specific factors associated with the selected industry. The empirical results by King(1966), where stocks within any industry have returns that more highly correlated with one another than the returns of stocks from different industry, give the insights to extend the factor model for developing the precise industry sector-factor model. Our model is motivated by these independent studies and rely on the their theoretical backgrounds.

Industry effect, a portfolios exposure to an industry group arising from the companies in that group, is a source of common factor exposure. We incorporate industry effects into our prediction model based on sector-factor analytical framework. This extension of multi-factor model will provide more accurate prediction framework to analysis the behavior of industry stock index. Moreover, it will give the opportunities to overcome the traditional multi-factor model that often reveal the limi-

tation to specifying the sector features and predicting the industry sector. And this leads to a systematic procedure for the active management and helps to get to a more detailed insight into the industry sector.

Identification of the relevant factors typically proceeds from an economic and industry analysis of the stocks involved. Aspects of macroeconomics, industrial organization, and fundamental security analysis will play a major role in the process. With these factors specified, we build three primary models based on the economic value of sector-factor group. They correspond to Time Series model, Market model, and Hybrid model.

An explorative scheme (named *composite* model) is also presented in the same way in which a multi factor model obtained from the arbitrage pricing theory. We can now propose the components and the test models of our neural network experiment, which consists of two phases: training and testing. Table 2 shows the input variables of a model by the sector.

The composite model includes the diverse factors ranged from economic features to market ones. The key advantages of the composite model lie in capturing the complexity and nonlinearity regarding the fluctuations in stock return from various sectors.

In this study, we propose an alternative to industry analysis, in which the rate of return is predicted qualitatively, with the sector-factor model through embedding the industry factors. By incorporating the multi-factor model scheme, we are free from the possibility of selecting the variables arbitrarily, which has been the unsolved problem for learning tech-

<Table 2> Prediction models to be trained using multi factors

Sector \ Model	Time Series	Inter Market	Composite
Time Series	Stock Index	Stock Index, Disparity with KOSPI	Stock Index, Disparity with KOSPI
Inter-Market		TV and TA	TV and TA
Technical Data		MA, RSI, SLOPE, DIS	MA, RSI, SLOPE, KOSPI MA
Macro Economic			M2, CBY, CPI, PI
Industrial Data			- Construction: CAP, CAO, CD - Banks FXs, BTS, BD, BP

niques. It may well prove to be a step in the right direction for the financial applications domain, conducted without exact quantitative analysis and greater acceptance of the financial economics. This will provide the great practical value in assessing the component of return arising from the firm specific factors and news. This analytical framework is relatively simple to implement and flexible enough to extend for the prediction of individual stock return.
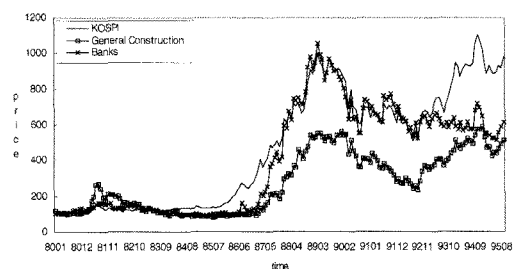
3.2. The data and Experimental Setup

3.2.1. The Data

Our specification is similar to the estimation scheme of general learning networks. Since there is no well-established and systematic method for network construction, we follow the basic process for developing a neural network. This process, suggested by Klimasauskas(1991), consists of four steps; (1) separate the data into training and testing sets, (2) transform the data into network-

appropriate inputs, (3) train, and test the network, (4) repeat steps as required.

The data for our application analysis are monthly closing prices of two industry stock indexes for the 15-year period from January 1980 to June 1995. Industry stock price indexes over this period are shown in Figure 1.

<Figure 1> Overlay of industry stock prices for construction and bank from January 1980 to May 1995

For the simulation data, we divide the industry stock price data into 2 nonoverapping period for training and testing the learning networks. For accuracy comparison, our network-based model generates the expected

movement of industry stock index in the next month. The stock price of one month ahead is naturally chosen to match the demands of portfolio managers in practice.

For both industry stock indexes, the total number of data points is 186 per whole period for construction and 126 for bank industry. To avoid overfitting and data snooping, we trained a separate learning network for the first same period, and tested those networks only on the data from the immediately following validation period.

3.2.2. Data preprocessing

Two widely used preprocessing methods are known as transformation and normalization. The transformation is conducted to each input field to make the information contents obvious to learning networks. As the output values to be predicted is the difference of stock index between next month and current month, the system takes the same type of data to be used in the desired output column. Economic and industry input data are used as the form of ratio. The system also applies the log transformation to the Moving_Average3month calculated from the changes of industry stock price index. The transformation of the specified data is based on the interview with the experts and the recommendations from previous studies.

To improve the learning efficiency, the normalization and scaling transformation is employed to make the input data range from 1 to 1 which is suitable for learning networks.

3.2.3. Set seduction

The training sample contains 140 for con-struction from May 1980 to December 1992 and 96 for bank from March 1985 to December 1993 respectively. In typical neural network experiments, a test set is seduced from the considered database in time order, then remaining examples is to be training set. However, this approach has the crucial drawbacks in generalizing the deployed network. To compensate these, 12 observations are randomly selected from training set. This technique, provided by NeuralWare, Inc.(1995), uses this set to optimize the network and generalize the networks in experiment. This aims to alleviate the memory effects and to reduce the over-fitting possibility of the network.

To assess the practical performance of the network-based model, we set aside some observations for the purpose of validation, which range from January 1993 to June 1995 for construction and from 1994 to June 1995 for bank.

3.3. Learning Architectures and Networks Properties

Neural networks have shown to possess some form of a *universal approximation* property. There may be a choice for the parameters that is better than any other possible choice. However, the "generalization error" of neural network derived from the fixed data points and functional approximation is a critical problem in modeling financial applications. This is to say a specific choice implies a specific assumption about the nature of the nonlinear relation to be approximated.

A general answer does not yet exist and is unlikely to be discovered any time soon. Even if several methods have been proposed to get around the problems of slow convergence and optimal network architecture, these techniques differ in complexity, convertgence speed, and most importantly in their generalization performance. In our analysis, the standard approach is employed to minimize the structural risk of the resulting networks.

The key concerns for neural network-like schemes are the type and the complexity of the networks to be used for learning. Consequently, we adopted the most common prediction architecture for learning networks, which is the three layered-feed forward networks with back propagation learning algorithms.

For the simplicity and generality of the network, we begin with 5 processing elements in one hidden layer. Designing the structure of optimal network is not the primary goal of our article. We take the variation of network configurations through increasing the number of hidden units and layers in networks.

The network runs with changing the learning parameters. To get to the stable state, the proposed model focused on the three learning parameters: learning coefficients, momentum, and training tolerance. As a learning progress, the network makes the learning rate lower and the momentum higher to make the network reach a stable state quickly. To prevent the network from being clamped at the local minimum, the *jog weight module* was used to specify a range for a random value to be added to all the variable weights going into the destination layer from the source processing

elements.

Hyper Tangent transfer function is chosen to smooth the outputs for every processing element in each layer. The networks generate the signals as a type of minus or plus, which is implicitly interpreted as a falling or rising of industry stock index.

IV. An application to Industry Stock Index: General Construction and Bank

We present an application to the prediction and simulation of two industry stock indices to assess the practical relevance of our neural networks. The prediction of industry stock returns will be widely used as inputs in the context of active asset management. It will also serve as a proxy for typical industry analysis and forecasting. We measure RMSE to choose the best network. In addition, the hit ratio of the resulting networks was also evaluated to assess the relative attractiveness and superiority.

4.1. Accuracy Comparison

A meaningful measure of performance for our empirical analysis is the hit ratio of various networks and regression models designed to predict the return on industry stock index. If we have correctly predicted the expected movements of industry stock index, the portfolio more weighted by the industry can yield higher return as much as weighted. With respect to the hit ratio, we count the cases that correctly predict the directions of industry

<Table 3> Hit ratio of three sector-base models for construction and bank industry

Industry \ Models	Time series	Inter Market	Composite
General Construction			
No. of Hidden Units	8	5	5
RMSE	26.9	25.5	22.3
(Ups/Downs)/Total	(15/7)/30	(14/9)/30	(15/9)/30
Hit ratio(%)	73.3	76.7	80.0
Banks			
No. of Hidden Units	11	3	8
RMSE	38.7	37.6	37.4
(Ups/Downs)/Total	(5/7)/18	(1/9)/18	(5/7)/18
Hit ratio(%)	66.7	55.6	66.7

stock index.

Table 3 reports the experiment results of prediction for general construction and bank. In each row, the accuracy measures for predicting stock index are reported for the three model-based networks tested on out-of-sample. The entries in each column correspond to the performance measures against different models.

For comparison, over the same test set the hit ratio for composite model for general construction is 80% as reported in the middle row. The fact that composite model can yield a smaller forecasting error than simple sector model such as time series and inter market may seem intuitive. After all, composite model using neural networks is indeed effective alternative in the context of our experiment.

The hit rate of correct prediction in Table 3 shows that time series model is similar to that of composite model for bank industry. However, since average accuracy obtained from the networks examined is higher than other two models, the benefits of composite
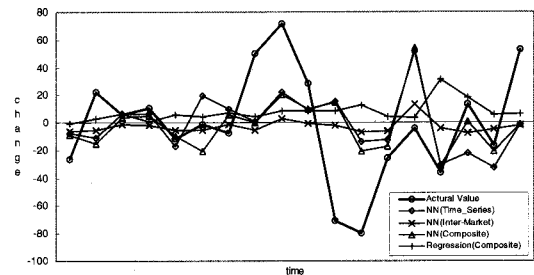
model may be still significant.

Figure. 2 and 3 portray the actual trajectory of the changes of industry stock index in conjunction with forecasts from various models. Apparently, the performance of neural network-based model equipped with sector-factors is better than any other model.

<Figure 2> Actual versus predicted values due to various models for general construction industry

In case of bank industry, there is no significant difference between time series model and composite model in terms of hit ratio.

<Figure 3> Plot of actual versus predicted values due to various models for bank industry

Even the same accuracy rate, we can note that composite model performed better than time series model in learning capability of networks.

4.2. Prediction Error Comparison between methodologies

For more complete comparison, we assess the prediction error of the models deployed according to several methodologies. In many previous studies, multiple linear regression models were selected as the benchmark for the comparison [Jang: 1991, Refenes, 1993: Park, 1995]. Neural network models possess the similar characteristics such as multivariate data analysis and analogue output types which our research model also hold. As such, we compare the performance of neural network model with that of regression model.

As a primary check of methodology performance, we employ the same performance measure used in Section IV.4.1. Table 4 provides similar comparisons for simple regression and learning networks over the same simulation scheme. The performances of learning networks at both industries are good. When learning capability is measured by correlation coefficient, learning networks yield higher coefficient: 97% and 95% at learning phase for construction and bank respectively. Not surprisingly, Table 4 reports that the linear models represented by simple regression exhibit considerable weaker performance than the neural network-based models. In particular, the distinctions in the case of general construction are more conspicuous.

From this evidence, we confirm that network-based approach will be a promising alternative for predicting the return on industry stock index and solving the asset allocation problems.

4.3. Trading Simulations

To complete our performance analysis of industry stock index predictions, we compare the simulated profits using the standard portfolio management strategies: active and passive management. The benchmark portfolio for comparison is the value-weighted market index: KOSPI.

We provide the useful performance evaluation method to assess the practical value through simulating buying and selling the index. Each model re-balances portfolio weights for the industry using the output signals generated from the learned networks and regression equations. For example, if the model

<Table 4> Performance comparison with hit ratio of neural network and linear regression for construction and bank industry

Methods \ Industry	General Construction	Banks
Learning Network		
Learning capability	0.95	0.97
(Ups / Downs)/Total	(15/9) / 30	(5/7) / 18
Hit ratio(%)	80.0	66.7
Regression		
Learning capability	0.48	0.42
(Ups / Downs)/Total	(10/9) / 18	(8/1) / 18
Hit ratio(%)	63.3	50.0

produces the negative(positive) signal, then take a position for short(long).

Over the testing period, the portfolio is re-balanced a monthly basis in way that gives more weights on the industry expressing positive signals. And then, total cumulative returns are calculated from the price changes of the portfolio. This virtual trading simulation without transaction costs may give the chance to assess the practical relevance of the neural networks and to evaluate the performance between methodologies.

Since industry stock index is not actually tradable asset in the stock market, this study has the difficulty for telling the practical value of any improvement in predicting accuracy that neural network might give us. However, we assess the models by assuming that industry stock index is tradable and borrowing the insights form relevant studies conducted by Kimoto, Yoda, and Takeoka(1990) and Kamijo and Tanigawa(1990).
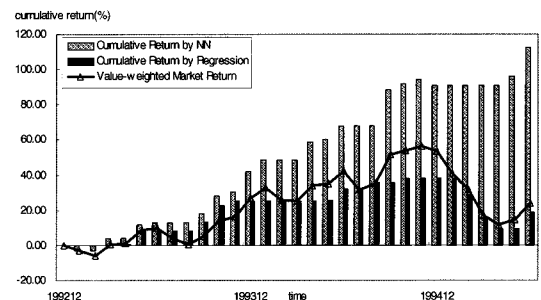
<Table 5> Performance comparisons through buy and sell simulation with cumulative returns for various models and methodologies

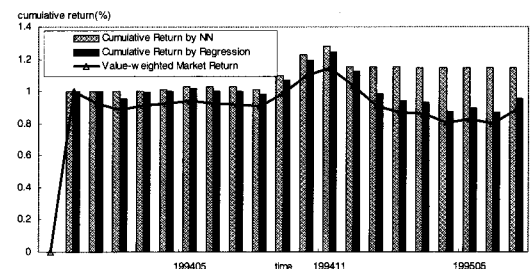| Methods\Industry	General Construction	Banks
Methodologies		
Linear Regression	19%	-4%
Value weighted index	24%	-12%
Neural Networks	112%	4%
Models		
Time Series	54%	14%
Inter Market	58%	11%
Composite	112%	8%

The above Table 5 assures the evidences that active models using learning network outperform the passive portfolio strategy expressed with the value-weighted market index

of KOSPI. In case of construction industry, the buy and hold strategy yields 24% rate of returns, otherwise the neural network model with composite scheme yields 112% cumulative returns over the same period.

In similar performance comparison, the simulation for bank industry report that learning networks present the higher returns than a simple regression does. Figure 4. and 5. illustrate the cumulative returns generated from the simulation for performance comparison by methodologies. As shown in Figure 4 and 5, the case of bank is slightly different with general construction industry. The time series model for bank slightly outperforms the composite model by higher accumulated returns

<Figure 4> Plot of accumulated returns of learning network, regression and market through simulation for construction

<Figure 5> Plot of accumulated returns of learning network, regression and market through simulation for bank

With the results for construction, Figure 4 shows that the neural networks exhibit more profitable returns and may be used as an effective investment tool in practice. However, in case of banks, we must look at this result with caution. It is difficult to infer which methodology performs better. Instead, we observe the evidences that all two composite models using neural networks outperform the specific sector models using regression.

V. Research Contributions and Future Works

We explore the new method to utilize the second-order variables such as consensus estimates by analysts and the outcomes from the independent prediction model. This may provide aggregated information and create the opportunities to obtain the domain specific knowledge and to capture momentum characteristics for each industry. These new effective and useful techniques will facilitate the practical extensions of neural network models that may successfully support the portfolio management and asset allocation.

Hopefully, this will lead to exploratory studies for prediction of sector market or individual stock return. This can give the possibility for analyzing the industry stock index based on investment analysis and the celebrated finance theory. Also, this gives the opportunities to overcome the traditional data-driven model that has been lacking in interpreting the results.

In principal, our multi-factor model for industry stock index can yield valuable insights in at least three contexts. First, it provides us with an reduction of ambiguity in variable selection. Second, our set separation processes give the more reliable results through testing the network and assessing the accuracy and profits. Finally, and most importantly, the prediction model of the return on industry stock index can be used to pursue the tactical asset allocation strategy by industry factor selection.

VI. Conclusions

Although the inconsistent results for two industries are presented, our findings show that neural network model equipped with multi factor framework can be useful alternatives. While our results are promising, we cannot yet claim that our approach will be successful in general due to the limited simulation and only two industry applications.

Nonetheless, we would like to provide some results discovered from this experiment for the purpose of enhancing the generality of findings. First, bank industry is loosely related to the macro economic factors than general construction industry. Second, we require the specification of additional factors that are not readily captured by the proposed model such as firm specific factors, market volatility, and precise trading volume indicators.

In case of bank industry, the published and the quantitative information may not be sufficient to explain the volatility of the short-term stock returns. In particular, the short-term prediction models designed for next month investment decision may consider more technical indicators to explain the variation of

stock return due to the limited data availability of economic and industrial factors. A related issue is the development of the new methods that combine the monthly data with quarterly or weekly. The service industry such as bank is significantly sensitive to qualitative information disclosures such as financial market regulation or government policies. These may make the predictions of industry stock index to be difficult when using only quantitative data.

Finally, the need for better technique to manipulate the missing data is clear. We use a common technique to generate the data that is missing or not available because of the weak data integrity and completeness of the economic and industrial sample.

⟨Reference⟩

[1] Park, S. J, "A study of the effects on stock price by industrial factors: in case of Korea stock market", Master thesis, Yonsei Univ., 1990

[2] Park, J.Y and Han, I.G, "Predicting Korea Composite Stock Price Index using Modular Neural Network", *In proceedings of the Society of Korea Expert System*, 1995, pp 103-131.

[3] Han, I.G, Kwon, Y.S, and Jo, H.K, "A Review of Artificial Intelligence Models in Business Classification", *In proceedings of the Society of Korea Expert System*, No. 1, 1995, pp 91-117.

[4] Yoon. T. S, "A study on Industry effects on Stock Price in Korea stock market", M. S., Seoul Univ, 1994.

[5] Jung, H.Y and H.S, Kim, "Industry Stock Index differentiation in Stock Market", *Monthly Cheil Economic Research*, June, 1995, pp 43-58.

[6] Na, D.K, "Stock price returns by industrial classification using Multi Factor Model", *DWERI unpublished manuscripts*, December, 1997.

[7] Ahamdi, H., "Testability of the Arbitrage Pricing Theory by Neural Network", In *Proceedings of the IEEE International Joint Conference on Neural Networks*, pp. 385-394, 1990.

[8] Arnott, R.D, "Cluster Analysis and Stock Price Comovement", Financial Analyst Journal, 36, no.6(Nov/Dec), 1980.

[9] Cavaglia et all, "Have we underestimated the Gains from International Equity Diversification? An Examination of Industry Portfolios across Countries", City University Business School Working Paper, 1994

[10] D, Chinetti and F, Gardin and C, Rossignoli, "A Neural Network Model for Stock Market Prediction", *In Proceedings of the 2nd Annual International Conference on AI applications on Wall Street*, 1993, pp 64-72.

[11] Choi, J.H, "Trading S&P 500 Stock Index Futures Using a Neural Network", *In Proceedings of the 3rd Annual International Conference on AI applications on Wall Street*, 1995, pp 63-71.

[12] Drummen, M. and H. Zimmerman, "The

structure of European Stock Returns", Financial Analyst Journal, July/August, 1992.

[13] Farrell, L.J, "Analyzing Covariance of Returns to Determine Stock Groupings", Journal of Business, 47, no.2, Apr, 1974.

[14] _____, "Homogeneous Stock Groupings: Implications for Portfolio Management", Financial Analyst Journal, 31, no. 3, May/June, 1973.

[15] Fabozzi, F.J and J.C.Francis, "Industry Effects and the Determinations of Beta", Quarterly Review of Economics and Business, 1983, pp.61-73.

[16] Fertuck.L, "A Test of Industry Indices Based on SIC Code", Journal of Financial and Quantitative analysis. Dec, 1976, pp. 837-848.

[19] Griffin, J. M and Karolyi, G.A, "Another look at the Role of the Industrial Structure of Markets for International Diversification Strategies", Unpublished Manuscripts, Ohio State Univ., 1996

[20] Grinold, R, A. Rudd and D. Stefek, "Global Factors: Fact or Fiction?", Journal of Portfolio Management, fall, 1989, pp. 79-88.

[21] Heston, Steven L and K.G.Rouwnhorst, "Does Industrial structure explain the benefits of international diversification?", Journal of Financial Economics, 1994, pp. 3-27,

[22] Jacob. Boudoukh, M. Richardson and R.Whitelaw, "Industry Returns and the Fisher Effects", Journal of Finance. Vol. XLIX. No 5. Dec, 1994

[23] Khan, R, "Active Portfolio Management", Probus, 1996.

[24] Kimoto, T. et all, "Stock Market Prediction System with Modular Neural Networks", In proceedings of the IEEE International Joint Conference on Neural Networks, 1990, pp. 1-6.

[25] King.B.F, "Market and Industry Factors in Stock Price Behavior", Journal of Business, Jan., 1966., pp.139-190.

[26] Kuo, G.W and Stachell, S.E, "Global Equity Style and Industry Effects: Portfolio Construction via Dummy Variables", Cambridge Univ., 1998

[27] Jang, G. et all, "An Intelligent Stock Portfolio Management System Based on Short-Term Trend Prediction using Dual Modular Networks", Artificial Neural Networks, 1991, pp. 447-452.

[28] Mendelsohn, L., "Preprocessing Data For Neural Networks", Technical Analysis of Stock and Commodities, Oct., 1993, pp. 52- 58.

[29] Richard Roll, "Industrial Structure and the Comparative Behavior of International Stock Market Indices", The Journal of Finance, Vol. XLVIL, No1. Mar, 1992.

[30] Rosenberg, B, "Extra-Market Components of Covariance in Security Markets", Journal of Financial and Quantitative Analysis, March, 1974, pp. 263-274.

[31] _____, "The Prediction Investment Risk: Systematic and Residual Risk", in Proceedings of the seminar on the Analysis of Securities Prices(CRSP, Univ. of Chicago), Nov., 1975.

[32] _____, "Choosing A Multi Factor Model", Investment Management Review, 1988.

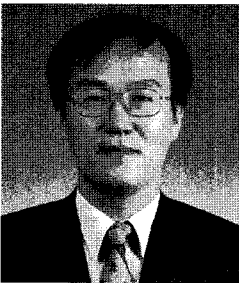[33] Sorensen, E.H, The Salomon Brothers U.S. Attribute Model, The Salomon Brothers Inc., Oct., 1989.

◆ 저자소개 ◆

권영삼 (Kwon, Young-Sam)

현재 교보증권 리셔치센터 기업분석팀 선임연구원으로 재직하고 있다. 중앙대학교 경영학과를 졸업(1994)하고 KAIST에서 경영정보학 및 재무학을 전공하여 공학석사(1996)을 취득하였다. 대우경제연구소에서는 인공지능 및 계량경제학적인 분석기법을 이용하여 자본시장에 대한 계량분석과 투자모형을 연구하였다. 주요 관심분야는 주식선별, 자산배분 및 위험관리 등이다.

한인구 (Han, In-goo)

현재 KAIST 테크노경영대학원 부교수로 재직중이다. 서울대 국제경제학사(1979), KAIST 경영과학석사(1981)을 취득하고 University of Illinois at Urbana-Champaign에서 회계정보시스템을 전공하여 경영학박사(1990)을 받았다. 주요관심분야는 지능형 신용평가시스템, 인공지능을 이용한 주가예측, 지식자산 가치평가, 정보시스템감사 및 보안 등이다. 현재 한국경영정보학회 부회장을 맡고 있다.