Predicting Corporate Bond Prices in Japan Using a Support Vector Machine

Hiroaki Jotaki
Department of Computational Intelligence and Systems Science, Interdisciplinary Graduate School of Science and Engineering
Tokyo Institute of Technology
Kanagawa, Japan
e-mail: Hiro.jotaki@gmail.com

Yasuo Yamashita
Investment Research Department
Sumitomo Mitsui Trust Bank
Tokyo, Japan
e-mail: Yamashita_Yasuo@smtb.jp

Hiroshi Takahashi
Graduate School of Business Administration
Keio University
Kanagawa, Japan
e-mail: htaka@kbs.keio.ac.jp

Abstract—**Predictability of returns is one of the most important concerns in bond investment. In this study, we analyze the predictability of corporate bond prices after company announcements of financial results using a support vector machine (SVM). This paper will discuss (1) the highest hit ratio found when predicting the movement of corporate bond prices using the four variables of current net earnings, management earnings forecasts, ratings, and a leading composite index, and (2) the highest hit ratio found while using a Gaussian kernel function with a parameter of 0.6 and a slack coefficient of 1.0. In addition to offering captivating insights from the results of this study regarding the mechanism by which financial reports impact prices in the bond market, our results also deepen our understanding of excess returns in asset management.**

Keywords-Corporate Bonds; Prediction; SVM.

I. INTRODUCTION

In the management of corporate bonds, ensuring stable generation of excess profits, identifying sources of excess returns, and predicting credit risk are all critical concerns [1][2][3][4][5].

Research on sources of excess profits is closely related to discussions of market efficiency, and numerous studies focusing primarily on stocks, have been conducted on this topic [6][7]. Among these, the relationship between company financial reports and the stock market is one of the areas where research is most extensive [8][9][10]. Several studies have been conducted on the impact of financial statements on the stock market, but there are few studies that focus on the bond market. Reference [3] focuses on information disclosed in company financial reports in Japan and conducts an event study analysis using the cumulative excess return (CER) to analyze the impact of disclosure information on corporate bond prices. As a result, it found that corporate bond prices tend to exhibit (1) no change in CER if current net earnings are higher than the previous term, while CER tends to become negative if current net earnings are less than the previous term, and (2) CER becomes increasingly negative when current net earnings decrease, management earnings forecasts for the next term are less optimistic, and bond ratings are low. This shows that current net earnings impact bond prices more than management earnings forecasts.

Research regarding price predictability is also a critical stream, and numerous studies have been conducted mainly focusing on stocks. Traditional prediction methods include regression models and auto regression models, but more recently, studies using learning models have also become popular. References [11][12] have predicted stock prices using support vector machines (SVM) and have reported that the accuracy of such predictions is higher compared with traditional models. However, these studies were conducted outside of the Japanese market, and studies regarding the predictability of corporate bond prices in Japan are particularly rare.

To address this gap, this study analyzes the predictability of Japanese corporate bond prices following announcements of financial results using a SVM.

The structure of the rest of this paper is as follows. The analytical methods used are explained in Section II, and the study results are discussed in Section III. A summary of our study is presented in Section IV.

II. ANALYTICAL METHOD

After first characterizing the samples used in the analysis and the corporate bond CER, we describe the four factors used as explanatory variables: current net earnings, management earnings forecasts, the index of business conditions, and ratings.

A. Sample

The sample data used in this study are comprised of reported corporate financial results disclosed between 2002 and 2010. The data used comes from 1,441 companies that satisfied the following five conditions: (1) from company annual reports that had at least two reporting periods between 2002 and 2010, (2) from companies that disclose current earnings and earnings forecasts per share on a consolidated basis (or non-

 ISBN: 978-1-61208-542-5

consolidated basis if unavailable), (3) from companies whose rate of change in the number of shares in comparison to the previous fiscal year was 20% or lower, (4) from the companies which issued one or more bonds with one year or more remaining maturity, and 5) from companies that have been rated (R&I standard).

TABLE I. NO. OF SAMPLE

		Current net earnings		
		Increased	**Decreased**	**Total**
Net earnings forecast	**Increased**	583	474	1057
	Decreased	267	117	384
Total		850	591	1441

B. *Corporate bond CERs*

In this section, we define the annual reporting date as daily 0 (t=0) and then analyze the return on bond j issued by company i. The belief is that information around financial results affects the corporate bond spread, representing corporate credit risk.

Therefore, we focus on bond returns as a function of changes in corporate bond spreads. Corporate bond spreads are calculated based on the difference between the corporate bond yield and the government bond yield whose maturity is the same as the corporate bond and can be converted into returns by multiplying the change in the spread by the price sensitivity (Mdur) against the yield. However, in order to focus on changes in the corporate bond spread resulting from information disclosed by individual companies, it is necessary to calculate the return on bonds (hereafter "excess return") by deducting the effect of changes in the overall market spread. References [13][14] defined excess return as the difference between the total return on corporate bonds and the total return on bond indices with the same rating and maturity as the corporate bonds. Because Japanese bond indices are separated by rating and maturity and therefore have different spreads, in this study, we decided to determine excess returns based on corporate bond spreads with reference to the method described by [13][14].

Equation (1) is used to calculate excess returns on corporate bonds. The excess return on corporate bond j issued by company i is obtained by subtracting the index spread total return (ISR) from the spread total return (SR) of issuer i.

$$er(i, j, t) = SR(i, j, t)-ISR(i, j, t). \quad (1)$$

Equation (2) is used to calculate the SR of corporate bonds used in (1) above. The first variable in (2) represents the capital return coming from the spread, and the second variable represents the income return from the spread. Mdur represents the modified duration.

$$SR(i,j,t) = dSR \times Mdur+Spd \times days/365. \quad (2)$$

Equation (3) is used to calculate the ISR of the index used in (1). In this analysis, NOMURA-BPI data are used as index data. The R&I standard rating (AA, A, BBB, BB), the maturity (short term(less than 1-3 years), middle term (>3, <7 years), or long term (>7 years)) can be obtained as index attribute information by spread. We calculated the returns for each category based on information from these 12 types of spreads.

$$ISR = dSpd \times Mdur+Spd \times days/365. \quad (3)$$

Additionally, if company i is issuing J bonds, the excess return on individual bond j issued by the same company is weighted based on market value of bond j at time t, and the excess return on the bond issued by company i is then calculated. W represents the weighted market value of the bond j.

$$Er(i, t) = \Sigma w(i, j, t) \times er(i, j, t). \quad (4)$$

The average excess return at time t (at time of reporting: t = 0) is as follows.

$$ER(t) = \Sigma\, Er(i, t) / N. \quad (5)$$

CER is defined as the cumulative return of ER(t) obtained in this manner on a daily basis.

C. *Corporate bond CERs on a yearly basis*

During the sample period 2002 to 2010, the economic situation varied depending on the year. Thus, there is a possibility that reactions in the corporate bond market differ depending on economic conditions. For example, based on the economic cycle announced by the Cabinet Office in Japan, it is possible to divide the cycle into two segments: an expansion phase and a recession phase. Based on this schema, 2002 was the bottom of the recession phase and the economy exhibited growth until 2008. However, after the Lehman shock that occurred in 2008, the economy went into a decline. The recession phase continued until 2009, and the economy then entered an expansion phase in 2010.

In this section, we will analyze the CER trend for making predictions of corporate bond CER. Specifically, we divide CERs by fiscal year, and further divide these based on movement in current net earnings (increase, decrease) and management earnings forecasts (increase, decrease). We assume here that investments are made the day following the announcement of financial results, and the cumulative excess return is CER (+1, +30).

We look at these results in the context of management next term earnings forecasts based on an increase in current net earnings. During the economic expansion phase, CERs tend to be positive regardless of the increase or decrease in the next term earnings forecast, but the CER may be negative during a recession phase. Next, we looked at the results in the context of management next term earnings forecasts based on a decrease in current net earnings. When compared on an annual basis, the tendency is that no change in CER is observed, or the CER may be negative. Particularly in 2002 and 2009, during the economic recession, we can see that the CER is strongly negative in the case when the management earnings forecast decreased. This indicates that corporate bond prices may be affected by economic conditions as well as current net earnings and earnings forecasts.

TABLE II. CORPORATE BOND CERS ON A YEARLY BASIS

CER(+1,+30)

	current net earnings	net earnings forecast	2002	2003	2004	2005	2006	2007	2008	2009	2010
(a)	+	+	-0.383%	0.167%	0.033%	0.042%	-0.016%	0.008%	-0.044%	-0.950%	0.211%
			(-3.9)	(2.7)	(1.9)	(4.5)	(-0.9)	(0.4)	(-0.3)	(-4.1)	(3.5)
		N	65	89	102	87	65	63	39	14	59
(b)	〃	−	-0.186%	0.075%	0.029%	0.043%	0.088%	-0.041%	-0.179%	-0.255%	0.287%
			(-2.9)	(2.1)	(1.3)	(2.8)	(1.0)	(-1.8)	(-0.5)	(-0.7)	(1.8)
		N	19	36	34	35	40	36	40	13	14
(c)	−	+	-0.248%	0.092%	0.008%	0.020%	-0.029%	-0.078%	-0.005%	-0.338%	0.177%
			(-5.2)	(1.8)	(0.7)	(1.8)	(-1.2)	(-1.2)	(-0.1)	(-1.5)	(2.8)
		N	96	51	42	42	49	35	49	71	39
(d)	〃	−	-0.379%	0.100%	0.048%	-0.029%	0.045%	0.121%	-0.491%	-1.054%	0.050%
			(-1.9)	(1.3)	(0.6)	(-0.4)	(0.9)	(0.9)	(-1.2)	(-3.6)	-
		N	11	7	5	7	6	13	30	37	1

D. Support vector machines

In this study, we use a SVM, which is one of various types of learning models commonly used for predicting prices. SVMs use a Gaussian kernel function [15]. The Gaussian function has two settings, parameter σ^2 and slack coefficients, which are important factors in measuring the superiority of SVMs. We also present our analysis of these parameters here.

$$y = \beta+\Sigma\alpha \ y \ K(x(i), x). \tag{6}$$

E. Analytical data

In this section, we summarize the results of our analysis of corporate bond price predictability in the Japanese market. As was confirmed in the previous section, when assessing CERs, fluctuations in corporate bond excess returns are generally small. Therefore, it is more important to be able to predict a large negative excess return as seen in 2009, rather than predicting a positive excess return.

Therefore, in this section, we will focus on predicting the negative excess returns on corporate bonds seen in 2009 by designating the training sample as the period from 2002 to 2008, and the prediction sample as 2009.

When predicting abnormal negative excess returns using a SVM, the CER needs to be classified into two types. Specifically, it is assumed that CERs (+1, +30) are divided into two types with -0.01% as a threshold value, where CERs exceeding -0.01% are defined as normal returns, and CERs of -0.01% or less are defined as abnormal negative returns.

$$R(i) = \text{Normal return if } CER(+1,+30) > -0.01\%, \quad \text{Otherwise, abnormal negative return.} \tag{7}$$

F. Explanatory variables

1) Current net earnings

One representative data point disclosed in a financial report is current net earnings. Reference [3] states that of the various types of data disclosed in financial reports, current net earnings may possibly affect bond prices. Thus, in predicting bond prices, we conducted our analysis using current net earnings as an explanatory variable. Equation (8) shows the rate of change in current net earnings per share (A) from the previous term (T-1) to the current term (T). Net earnings per share are treated as current net earnings. If ΔA is positive, current term net earnings increase in comparison to the previous term. If ΔA is negative, current net earnings decline over the previous term.

$$dA(T) = (A(T)-A(T-1)) / A(T-1). \tag{8}$$

2) Management net earnings forecasts

Management forecasts of subsequent terms' earnings are announced at the same time as current net earnings in financial results. In research studies using stock prices, it has also been reported that the influence of the manager next term earnings forecast is greater than the impact of current net earnings. In this study, it is assumed that the management earnings forecast is an explanatory variable and analyzed as such [9]. In our analysis, we focus on net earnings forecast per share and investigate its impact on bond prices. Equation (9) shows the rate of change in net earnings per share (A) during a specific period (T) and the net earnings forecast by management (F) for the subsequent term (T+1). If ΔF is positive, earnings are expected to increase during the T+1 period compared with net earnings during term T. Conversely, if ΔF is negative, earnings during the T+1 period are expected to be lower than net earnings during term T.

$$dF(T+1) = (F(T+1)-A(T)) / A(T). \tag{9}$$

3) Index of business conditions

Reference [3] states that bond price responses may differ depending on the business conditions during the fiscal year that financial results are announced. This study uses two indices, a composite index (CI) and a diffusion index (DI), as indicators capturing the economic trends in this study's sample period 2002 to 2010. The CI measures the magnitude of economic fluctuations and their tempo by compiling the movements of component indicators, while the DI calculates the proportion of these indicators that have exhibited improvement in order to measure diffusion to each component of the economy. There are three types of indices that make up the CI and DI: the leading index that precedes the economic condition, the coincident index that moves in concert with the economy, and the lagging index that moves after (lags) the economic condition. We use the coincident index to understand the current condition of the economy, and because the leading index generally precedes the coincident index by several months, we use this to predict the future movement in the economy. In general, the lagging index lags the coincident index by about a half year, so it is used for ex post factual confirmation.

4) Ratings

Ratings are commonly used as indicators of the financial condition of a company. Reference [3] found that in addition to the impact of financial report data on bond prices, bond prices could also be impacted by ratings and decline significantly when these ratings are low.

While there are five rating agencies, R&I, the Japan Credit Rating Agency (JCR), Standard & Poor's (S&P), Moody's, and Fitch, for this analysis, we adopted R&I, which has the highest coverage rate for our samples.

III. Analysis Results

In this section, we describe how we modeled corporate bond price predictions using a SVM. We analyze the differences between the explanatory variables, then analyze the adjusted parameters, the impact of different models on the kernel functions, and the cross-validation.

A. Analysis of differences between explanatory variables

In this section, we analyze the effects of differences between explanatory variables, which are the current net earnings, management earnings forecasts, ratings, and index of business conditions, as explained in Section II above. Reference [3] indicates that it is possible that corporate bond prices may have a particular impact on current net earnings. Therefore, we add other variables under the assumption that current net earnings were already added.

Using the training data, we look into the suitability of our model based on differences in the explanatory power of the variables for bond prices. First, by combining current net earnings, ratings, and the index of business conditions (six patterns) and looking at the fit for the training model, we can see that a high hit ratio is achieved for all combinations. In particular, the highest hit ratio is for the combination of current net earnings, ratings, and coincident CI at 84.66%. In considering the six trends in the index of business conditions, we can conclude that the CI is more suitable than the DI. Because the CI represents the magnitude and tempo of economic fluctuations and the DI represents the degree of economic diffusion, there is a possibility that corporate bond prices are affected by both the magnitude of economic fluctuations as well as their tempo.

Next, by combining the three variables of management earnings forecasts, ratings, and the index of business conditions (six patterns) and observing the fit for the training model, we can also find that a high hit ratio can be achieved for all combinations of these variables as well. In particular, the hit ratio generated increased to 83.82% when a coincident CI was used. In addition, by comparing the fitness of the six patterns of the index of business conditions, it can be concluded that the CI is more suitable than the DI. The same trend can be observed when using current net earnings and ratings as explanatory variables.

Next, we can find that the hit ratio can reach as high as 78.64% when observing the degree of fitness for the training model exhibited by the three variables of current net earnings, management earnings forecasts, and ratings. This indicates that a certain frequency of correct responses can be obtained only by using the combination of current net earnings, management earnings forecasts, and ratings, even when the index of business conditions is not included in the explanatory variables.

Last, when combining the four variables of current net earnings, management earnings forecasts, ratings, and index of business conditions (six patterns), we found that a high hit ratio is achieved for all combinations. Specifically, we found that the hit ratio was highest when using a coincident CI, at 84.66%.

TABLE III. The prediction performance of variable differences for training data

Number of Variable	Variables[a]		Training Data: Number of Hit /Total Number	Training Data: Hit Ratio
3	NE,	R, Leading CI	985/1193	82.56%
	NE,	R, Coincident CI	1010/1193	84.66%
	NE,	R, Lagging CI	962/1193	80.64%
	NE,	R, Leading DI	955/1193	80.05%
	NE,	R, Coincident DI	936/1193	78.46%
	NE,	R, Lagging DI	936/1193	78.46%
	EF,	R, Leading CI	985/1193	82.56%
	EF,	R, Coincident CI	1000/1193	83.82%
	EF,	R, Lagging CI	957/1193	80.22%
	EF,	R, Leading DI	964/1193	80.80%
	EF,	R, Coincident DI	933/1193	78.21%
	EF,	R, Lagging DI	933/1193	78.21%
	NE,EF,	R	936/1193	78.46%
4	NE,EF,	R, Leading CI	984/1193	82.48%
	NE,EF,	R, Coincident CI	1010/1193	84.66%
	NE,EF,	R, Lagging CI	963/1193	80.72%
	NE,EF,	R, Leading DI	946/1193	79.30%
	NE,EF,	R, Coincident DI	941/1193	78.88%
	NE,EF,	R, Lagging DI	936/1193	78.46%

a. NE: Current Net Earnings, EF: Earning Forecast, R: Rating.

Next, we look at the prediction performance for holdout data. First, when looking at the prediction results when the three variables of current net earnings, ratings, and index of business conditions (six patterns) are combined, the hit ratio when using a leading CI or a coincident CI is very high at 97.86% for abnormal negative returns, and very low for normal returns, at 5.41%. In contrast, when using other indices of business conditions, we find that the hit ratio for normal returns is high and the hit ratio for abnormal negative returns is very low. This suggests that the training model may be overfitting.

Next, looking at the prediction results when using a combination of the three variables of management earnings forecasts, ratings, and the index of business conditions (six patterns), use of the leading CI and the coincident CI resulted in a high hit ratio for abnormal negative returns of 80.61% and 97.96%, respectively, but the hit ratio for normal returns was low. In contrast, when using other indices of business conditions, we find that the hit ratio for normal returns is high and the hit ratio for abnormal negative returns is very low. This suggests the possibility that the training model may be overfitting, similar to what occurs using current net earnings.

Next, when assessing the prediction results when using the three variables of current net earnings, management earnings forecasts, and ratings, the hit ratio for abnormal negative returns is 0, and the hit ratio for normal returns is 94.59%. There is a possibility that the training model is overfitting in this case as well.

Finally, in assessing prediction results when using a combination of the four variables of current net earnings, management earnings forecasts, ratings, and the index of business conditions (six patterns), the hit ratio when using a leading CI is 89.13% for abnormal negative returns, and 29.73% for normal returns. Although the hit ratio for normal returns is not high, there is a possibility that the hit ratio could be improved by adjusting parameters. Results from using other indices of business conditions appear to be strongly biased towards either abnormal negative returns or normal returns, suggesting the possibility of overfitting by the training model.

In this section, we analyze the effects of the differences in the explanatory variables used, and as a result, we find that the highest predictability for corporate bond prices is achieved when using the four explanatory variables of current net earnings, management earnings forecasts, ratings, and the leading CI.

TABLE IV. THE PREDICTION PERFORMANCE OF VARIABLE DIFFERENCES FOR HOLDOUT DATA

Number of Variable	Variables[a]	Holdout Data: Abnormal Negative Return Number of Hit /Total Number	Abnormal Negative Return Hit Ratio	Normal Return Number of hit /Total Number	Normal Return Hit Ratio
3	NE, R, Leading CI	96/98	97.96%	2/37	5.41%
	NE, R, Coincident CI	96/98	97.96%	0/37	0.00%
	NE, R, Lagging CI	5/98	5.10%	31/37	83.78%
	NE, R, Leading DI	0/98	0.00%	35/37	94.59%
	NE, R, Coincident DI	0/98	0.00%	35/37	94.59%
	NE, R, Lagging DI	0/98	0.00%	35/37	94.59%
	EF, R, Leading CI	79/98	80.61%	12/37	32.43%
	EF, R, Coincident CI	96/98	97.96%	0/37	0.00%
	EF, R, Lagging CI	1/98	1.02%	37/37	100.00%
	EF, R, Leading DI	3/98	3.06%	37/37	100.00%
	EF, R, Coincident DI	0/98	0.00%	37/37	100.00%
	EF, R, Lagging DI	0/98	0.00%	37/37	100.00%
	NE,EF, R	0/98	0.00%	35/37	94.59%
4	NE,EF, R, Leading CI	88/98	89.80%	11/37	29.73%
	NE,EF, R, Coincident CI	96/98	97.96%	0/37	0.00%
	NE,EF, R, Lagging CI	16/98	16.33%	30/37	81.08%
	NE,EF, R, Leading DI	0/98	0.00%	2/37	5.41%
	NE,EF, R, Coincident DI	0/98	0.00%	37/37	100.00%
	NE,EF, R, Lagging DI	0/98	0.00%	37/37	100.00%

a. NE: Current Net Earnings, EF: Earning Forecast, R: Rating.

B. *Analysis of parameter differences*

In the previous section, we analyzed the differences between variables, and as a result, found that the best case arises when making corporate bond price predictions using the four variables of current net earnings, management earnings forecasts, ratings, and the leading CI. We then determined the most suitable parameters assuming these four variables.

The SVM has two parameters: one the variance σ^2 of the kernel function (Gaussian) and the other the slack coefficient representing the degree of relaxation of the constraining condition when the discrimination is not possible. By adjusting these two parameters, we are able to look at the suitability for our model.

First, we analyze the hit ratio from the training data and the holdout data after fixing the slack coefficients and the kernel function parameters adjusted from 0.6 to 1. The hit ratio for the training data exceeded 80% in all cases, resulting in a high hit ratio. On the other hand, for the hit ratio for the holdout data, utilized a lower kernel function parameter, as the hit ratio for abnormal negative returns tended to decrease, the hit ratio for normal returns tended to increase. Overall, the hit ratios of abnormal negative returns and normal returns both exceeded 60% when the parameter of the kernel function was set to 0.6.

We next looked at the hit ratio for the training and holdout data after adjusting the kernel function parameter to 0.6 and changing the slack coefficient from 0.5 to 2. The difference in the hit ratio for the training data was not much despite adjusting the slack coefficient, and exceeded 80% in all cases. In contrast, for the hit ratio for the holdout data, which utilized a lower slack coefficient, as the hit ratio for abnormal negative returns tended to decrease, the hit ratio for the normal return tended to increase. In particular, correct responses for both abnormal negative returns and the normal returns exceeded 60% when the slack coefficient was 1.0.

TABLE V. THE PREDCTIONS PERFORMANCE OF PARAMETER DIFFERENCES FOR TRAINING DATA

Variables[a]	Parameter	Slack Coefficient	Training Data: Number of Hit /Total Number	Hit Ratio
NE,EF,R,Leading CI	1	1	984/1193	82.48%
	0.5	1	980/1193	82.15%
	0.75	1	986/1193	82.65%
	0.9	1	986/1193	82.65%
	0.6	1	975/1193	81.73%
	0.6	2	989/1193	82.90%
	0.6	0.5	957/1193	80.22%

a. NE: Current Net Earnings, EF: Earning Forecast, R: Rating.

TABLE VI. THE PREDICTION PERFORMANCE OF PARAMETER DIFFERENCES FOR HOLDOUT DATA

Variables[a]	Parameter	Slack Coefficient	Holdout Data: Abnormal Negative Return Number of Hit /Total Number	Abnormal Negative Return Hit Ratio	Normal Return Number of Hit /Total Number	Normal Return Hit Ratio
NE,EF,R,Leading CI	1	1	88/98	89.80%	11/37	29.73%
	0.5	1	53/98	54.08%	27/37	72.97%
	0.75	1	73/98	74.49%	17/37	45.95%
	0.9	1	77/98	78.57%	15/37	40.54%
	0.6	1	64/98	65.31%	23/37	62.16%
	0.6	2	77/98	78.57%	12/37	32.43%
	0.6	0.5	37/98	37.76%	30/37	81.08%

a. NE: Current Net Earnings, EF: Earning Forecast, R: Rating.

 ISBN: 978-1-61208-542-5

In this section, we analyze differences in parameters and predict the movement of corporate bond prices using the four variables of current net earnings, management earnings forecasts, ratings, and the leading CI. We find that the model fits best when using a kernel function (Gaussian) with a parameter of 0.6 and a slack coefficient of 1.0.

C. Analysis using different Kernel functions

In this analysis, we use the Gaussian function as a general SVM kernel function. In this section, we summarize the results of our analysis of each of three different kernel functions; the linear, polynomial, and sigmoid, other than the Gaussian function, fit with our model.

First, we set the parameters of each kernel function to 1, and then observed the hit ratio for the training data. However, as the hit ratio was low, we changed the parameters and checked the hit ratio again. As a result of the changes, the hit ratio for the training data achieved a 70% range for all kernel functions, lower than the 80% level achieved using the Gaussian function.

The parameters for each Kernel function were adjusted and we looked into the prediction performance for holdout data. As a result, the hit ratio for normal returns increased with all kernel functions, while the hit ratio for abnormal negative returns was as low as the 30% range. Based on these results, we concluded that the Gaussian function is the most suitable for predicting abnormal negative returns.

In this section, by checking the hit ratio for corporate bond prices due to the differences in the SVM kernel functions, we found that the Gaussian function is the most suitable function among the Gaussian, linear, polynomial, and sigmoid functions tested.

TABLE VII. THE PREDICTION PERFORMANCE OF DIFFERENT KERNEL FUNCTIONS FOR TRAINING DATA

Kernel	Parameter	Training Data	
		Number of Hit /Total Number	Hit Ratio
Liner	0.6	933/1193	78.21%
Polynominal	0.6	879/1193	73.68%
Polynominal	0.1	900/1193	75.44%
Sigmoid	0.6	849/1193	71.17%
Sigmoid	0.1	850/1193	71.25%
Sigmoid	2	850/1193	71.25%

TABLE VIII. THE PREDICTION PERFORMANCE OF DIFFERENT KERNEL FUNCTIONS FOR HOLDOUT DATA

Kernel	Parameter	Holdout Data			
		Abnormal Negative Return		Normal Return	
		Number of Hit /Total Number	Hit Ratio	Number of Hit /Total Number	Hit Ratio
Liner	0.6	0/98	0.00%	37/37	100.00%
Polynominal	0.6	38/98	38.78%	24/37	64.86%
Polynominal	0.1	38/98	38.78%	24/37	64.86%
Sigmoid	0.6	38/98	38.78%	24/37	64.86%
Sigmoid	0.1	38/98	38.78%	24/37	64.86%
Sigmoid	2	33/98	33.67%	24/37	64.86%

D. Analysis of cross-validation

The issue of model overlearning has been highlighted in regard to constructive learning models. Therefore, in this section, we summarize our findings when checking for overfitting in our training model using k-fold cross-validation.

Here, we analyze the cross-validation under the conditions tested among the various analyses conducted previously that resulted in the highest rates of correct responses by the training model and prediction model. Specifically, when we predict the movements of corporate bond prices using the four variables of current net earnings, management earnings forecasts, ratings, and the leading CI, we used the Gaussian kernel function with a parameter of 0.6 and a slack coefficient of 1.0. The division method used consisted of separating the data into 10 groups during the 2002 to 2008 training period.

In viewing the results, the hit ratio on cross validation is as high as 80.1%. From this, it seems that such a result indicates that it is highly likely that the model in this study is not overfitting.

IV. CONCLUSION AND FUTURE WORK

In this study, we analyze the predictability of corporate bond prices following company announcements of financial results using a SVM.

From our analysis, we find that we are able to obtain (1) the highest prediction performance when using the four variables of current net earnings, management earnings forecasts, ratings, and the leading CI, and (2) the highest prediction performance when using a Gaussian kernel function with a parameter of 0.6 and a slack coefficient of 1.0 as model conditions.

These results offer captivating insights regarding the predictability of prices in the corporate bond market using a SVM.

In terms of future work, we plan to expand the data to current year and to apply the same structure to other bond markets outside of Japan.

 ISBN: 978-1-61208-542-5

REFERENCES

[1] H. Jotaki,S. Takahashi,H. Takahashi,"The Impact of Headline News on Credit Market in Japan", Nippon Finance Association 17th conference, pp. 113-122, 2009

[2] H. Jotaki,Y. Yamashita,H. Takahashi "The Impact of M&A on the corporate bond Market in Japan ―From the point of view of fixed income investment ―", Market Structure Analysis and New asset management method, Japanese Association of Financial Econometrics and Engineering Journal, pp. 56-74, 2012.

[3] H. Jotaki,Y. Yamashita,H. Takahashi,"The effect of earing announcement on corporate bond market in Japan" , Market Structure Analysis and New asset management method, Japanese Association of Financial Econometrics and Engineering Journal, pp. 56-74, 2012.

[4] G. Levent, H. Dirk,"Corporate bond credit spreads and forecast dispersion", Journal of Banking & Finance, vol. 34, pp. 2328-2345, 2010.

[5] P. Collin-Dufresne, R. Goldstein, and J. Martin, "The determinants of credit spread changes" Journal of Finance, vol. 60, pp. 2255-2281, Dec. 2001.

[6] E. Fama "Efficient Capital Markets: A Review of Theory and Empirical Work", The Journal of Finance, vol. 25,pp. 383-417, 1970.

[7] F. Sharpe, "Capital Asset Prices:A Theory of Market Equilibrium under condition of Risk", The Journal of Finance, vol. 19, pp. 425-442, 1964.

[8] K. Ota, "Reaction comparison of TSE part, TSE II, Osaka, and over-the-counter markets for financial announcement", Kansai University Graduate School Senriyama Commerce, 53, 2001.

[9] R. Conroy,K. Harris,and Y. Park,"Fundamental information and share prices in Japan: Evidence from earnings surprises and management predictions", International Journal of Forecasting, vol. 14, no. 2, pp. 227-244, June 1998.

[10] R. Conroy. K. Eades,and T. Harris,"A test of the relative pricing effect of dividend and earnings : Evidence from simultaneous announcements in Japan", The Journal of Finance, vol. 55, pp. 1199-1227, 2000.

[11] D. Hon,P. Padhy,"Support Vector Machines for Prediction of Futures Prices in Indian Stock Market", International Journal of Computer Applications, vol. 41. no. 3, March 2012.

[12] K. Kyoung-jae Kim, "Financial time series forecasting using support vector machines", Neurocomputing, vol. 55, pp. 307-319, 2003.

[13] A. Warge,W.Ivo,"Bondholder losses in leveraged buyouts", Review of Financial Studies, vol. 6, pp. 959-982, 1993.

[14] M. Billett, T. King, and D. Mauer, "Bondholder wealth effects in mergers and acquisitions: New evidence from the 1980s and 1990s", Journal of Finance, vol. 59, pp. 107-135, Feb. 2004.

[15] M. Buhmann, "Radial basis function", Mathematicshes Institut, Justus-Liebig-Universitat Giessen, 2010.