UCT JOURNAL OF MANAGEMENT AND ACCOUNTING STUDIES 2017(03)

Available online at http://journals.researchub.org

Predicting factors affecting the future stock price crash risk based on support vector machine

Mohammad Reza Razdar[1]* , Ali Mohammad Zahmatkesh[2] , Sanaz Khaleh Oghlizadeh[3]

[1] *Assistant Professor, Department of Accounting, Sanabad Golbahar Institute of Higher Education, Golbahar, Iran.*

[2] *Department of Accounting, Sanabad Golbahar Institute of Higher Education, Golbahar, Iran.*

[3] *Department of Accounting, Sanabad Golbahar Institute of Higher Education, Golbahar, Iran*

ARTICLE INFO

Article history:
Received 03 Jun 2017
Received in revised form 04 Jul 2017
Accepted 15 Jul 2017

Keywords:
Stock prices,
Future stock price crash risk,
Support vector machine,
Predicting factors

ABSTRACT

Objective: The prediction of stock price crash risk is an important and widely studied topic in both accounting and finance, since crash risk has a significant impact on shareholders, creditors, managers, investors, and regulators. The aim of this research is to analyse Predicting factors affecting the future stock price crash risk based on support vector machine. **Methodology:** In this research we study the data of 99 companies listed on the Tehran Stock Exchange (TSE) from 2011 to 2016.And since the Mean Absolute Error in the Testing Sample is less than Training, then the model estimation is possible using the support vector machine method. **Results:** The results shows the summary of support vector machine Model. The results indicate support vector machine includes 3 input layers, and 1 output layer. **Conclusion:** The method used in this study support vector machine is a kind of RBF. The Mean Absolute Error for the Training sample is 0.078 and for the Testing example is 0.065. And since the Mean Absolute Error in the Testing Sample is less than Training, then the model estimation is possible using the support vector machine method.

1. Introduction

Stock market returns are among the most important indicators of financial performance (Ansari and Riasi, 2016; Riasi and Pourmiri, 2015), therefore a stock price crash can have a severe negative impact on a firm's financial stability and competitiveness (Riasi, 2015; Riasi and Amiri Aghdaie, 2013).

Stock price crash risk has been an important and widely studied topic, especially after the 2008 financial crisis. Investors, regulators and academics have paid a lot of attention to firm-level stock price crash risk which is defined as the experience of the extreme negative firm-specific stock returns.

Typically, stock price crash implies that the firm is under a condition of high opacity, weak corporate governance and high agency cost that informed managers withhold bad news from uninformed investors. Once the negative news hoarding reaches a tipping point, all of them is released suddenly at once and results in a stock price crash. Individual investors suffer significant unexpected wealth losses (Merton, 1987) and their confidence in the equity market is jeopardized by stock price crashes. Therefore in this research, we predicting factors affecting the future stock price crash risk by using support vector machine for the companies accepted in Tehran Stock Exchange during 2011-2016.

1.1 Theoretical Framework of Research

The main objective of investors from investing in stocks of the firms is gaining profit that would be achieved by identifying the correct price of the stock market, but sometimes identifying the correct price of firms regardless of market trends does not help shareholder profits. Markets generally take various trends from reaching the highest price bubble until reaching the floor price. The issue of stock price crash is so important that even years after the crash of the market, meetings and seminars are held to investigate the common and specific elements, time periods of stock price crash, the behavior of market participants, and the effect of decisions of regulators and government officials and international organizations. Many articles and studies are published about the crash of the market and are available to users. Superior analysis was based on data from 23 companies of North America, Europe and East Asia

* *Corresponding author:* Mr2razdar@gmail.com
DOI: https://doi.org/10.24200/jmas.vol5iss03pp7-14

where the results show evidence in this regard that the crash in stock prices in different markets varies according to the structure. Moreover, he studied the interdependence of stock markets before, during, and after weeks of crash of stock market in 1987 with a focus on foreign shares traded in the domestic market. He found that foreign stock returns in domestic markets has more correlation with stock returns mentioned in the release place. The price change of such shares after the market crash shows even more correlation before the crash of the stock market. Moreover, in that year, Mervin King explored the capability of the models with full information on the logical explanation of market crash. He discusses the impact of fundamental news for the stock market in 1987 that created a 23 percent crash and made other markets follow it out and make domino effect. According to him, imperfect information theories can better respond to the above questions. One of the applications of this theory is that investors in one market may try to get information from the behavior of other market, so any mistake in each of the markets can easily spread to other markets and cause misunderstanding in mass market participants (Sadat Shojai, 2015). Finally, the main research question in this case is whether predicting future stock price crash based on support vector machine is possible or not. The aim of this study is to find the factors affecting the future stock price crash risk based on support vector machine (Ataei zadeh and Darabi, 2016).

1.2 Research History

Ataei zadeh and Darabi (2016), in this study Predicting factors affecting the future stock price crash risk based on the Neural Network Based Radiological Basis Function of the listed companies in the Tehran Stock Exchange in years 2009 through 2015. Findings of the research predicting factors affecting future stock price crash risk is based on Neural Network based Radial Basis Function is possible.

Liao (2016) in this study The Stock Price Crash Risk Prediction by Neural Network Been paid. In this study compare the performance of the neural network model with the logistic model and random forecast. In this study show that the neural network crash risk prediction model provides a significant improvement in prediction accuracy over logistic regression and random forecast. The results indicate that the neural network methodology is a good alternative to predict stock price crash risk.

Sadat Shojai (2015), investigated the relationship between the effects of real earnings management and competition in the product market on the risk of future stock price crash in listed companies in Tehran Stock Exchange during the period 2008-2014, and concluded there is a significant relationship between abnormal operational cash flow and risk of future stock price crash. There is a significant relationship between abnormal cost of production and risk of future stock price crash. There is a significant relationship between abnormal discretionary cost and risk of future stock price crash. There is a significant relationship between competition in market and risk of future stock price crash. There is a significant relationship between product market competition and operational cash flow and risk of future stock price crash. There is a significant relationship between product market competition and abnormal production cost. There is a significant relationship between product market competition and abnormal discretionary cost.

Meshki and Ghiasmand (2013), examined the relationship between earnings management based on income and future stock price crash risk in companies listed Tehran Stock Exchange during 2007-2012 using panel data and logistic regression. The results showed a positive and significant relationship between fall in stock prices and earnings management based on revenue in future periods in evaluated samples.

2. Materials and methods

2.1 Hypotheses

1) Predicting the factors affecting the future stock price crash risk is possible based on support vector machine.

2.2 Methodology

2.2.1 Research method

This is a practical research in terms of goal-based categorization. The goal of a practical research is to develop knowledge in a specific area. Also, this method is descriptive type in terms of method-based research categorization. Descriptive research describes and anticipates the variables without any manipulation. In addition, this is ex-post facto research i.e. follows discovering and investigating the relations between specific factors and conditions that are occurred in the past.

2.2.2 Analysis method

To analyze information, two methods of inferential statistics and descriptive statistics were applied. In descriptive statistics we analyse mean, middle, standard deviation, skewness, elongation, minimum and maximum and in inferential statistics, support vector machine method are used.

2.2.3 Research model

In this study, the following research model is used to estimate and predict factors affecting future stock price crash risk.

$$CRASH_{t+1} = \alpha_0 + \alpha_1 \ \mathrm{ABCASH}_{it} + \alpha_2 \ ABEXP_{it} + \alpha_3 \ CS_{it} + \alpha_4 \ MB_{it} + \alpha_5 \ SIZE_{it} + \varepsilon_t \quad (1)$$

$CRASH_{t+1}$: The risk of future stock price crash

ABCASH_{it} : Unusual operating cash flows

$ABEXP_{it}$: Abnormal discretionary costs

CS_{it} : Capital Structure

$MB_{..}$: Growth opportunity
$SIZE_{it}$: Firm size

2.2.4. Definition of studied variables and their operational definition

Future stock price crash risk ($CRASH_{t+1}$): To measure this variable, we used the Hutton et al. (2009) criterion. According to Hutton et al., crash course in a given fiscal year is a period during which firm-specific monthly returns of 3.2 standard deviations is lower than the mean firm-specific special monthly returns. This definition is based on the statistical concept that by assuming a normal distribution of firm-specific monthly returns, fluctuations in the mean distance plus 3.2 standard deviation and the mean minus 3.2 standard deviation are considered normal, and fluctuations outside of this are considered as abnormal cases. Given that crash of stock price is an abnormal fluctuation, number 3.2 is considered as the boundary between normal and abnormal fluctuations. In this study, the risk of future stock price crash is a dummy variable that if the company has experienced at least a crash course until the end of the fiscal year, its value is one and otherwise zero. Firm-specific monthly returns are calculated using the following equation.

$$W_{j,\theta} = Ln(1+\xi_{j,\theta}) \quad (2)$$

In the above equation:

$W_{j,\theta}$: is firm-specific monthly returns of the firms j in month θ during the fiscal year.

$\xi_{j,\theta}$: Return of the remaining shares of the firms j in month θ and includes the remaining or residue of the following model:

$$r_{j,\theta} = \beta_0 + \beta_1 r_{m,\theta-2} + \beta_{2j} r_{m,\theta-1} + \beta_{3j} r_{m,\theta,0} + \beta_{4j} r_{m,\theta+1} + \beta_{5j} r_{m,\theta+2} + \xi_{j,\theta} \quad (3)$$

In this equation:

$r_{j,\theta}$: is returns of the firms j in month θ during the fiscal year.

$r_{m,\theta}$: Market return in month θ. To calculate the monthly returns of the market, index of the beginning of the month is deduced from the index of the end of the month and the result is divided into the index of the beginning of the month (Sadat Shojai, 2015).

Abnormal operating cash flows (ABCASH_{it} **):** According to the study by Cohen and Zarowin (2010), the following model is used to estimate abnormal operating cash flows, so that the residual of the model is taken into account as a measure of abnormal operating cash flows.

$CFO_{..}$: Operating cash flows
TA_{it-1} : Total assets of the firm
$Sales_{it}$: Firm sales
$\Delta Sales_{it}$: Change in sales (Cohen and Zarowin, 2010).

Abnormal discretionary costs ($\mathrm{ABEXP}_{\mathrm{it}}$ **):** Costs that management can control by changing their corporate plans are called discretionary spending. According to the study by Cohen and Zarowin (2010), abnormal discretionary costs have been estimated through the following model.

$DISEXP_{it}$: The discretionary cost of the company i at the end of the year t that is equal to the administrative and sales costs (Cohen and Zarowin, 2010).

Capital Structure (CS): Capital structure is obtained from the ratio of total debt to total adjusted assets ratio and shows how much debt has been used to meet the company's assets (Badavar Nahandi and Taghizadeh Khanqah, 2013).

Growth opportunity (MB_{it} **):** Growth opportunity is obtained from market value to book value of shareholders' equity (Badavar Nahandi and Taghizadeh Khanqah, 2013).

Firm size (size**):** Firm size is the natural logarithm of total assets (Badavar Nahandi and Taghizadeh Khanqah, 2013).

2.2.5 The population

The population of the study is all listed companies in Tehran Stock Exchange during the period of 2011- 2016, of the 486 companies listed in Tehran Stock Exchange, which meet all of the following criteria:

1 – The companies are listed in TSE at least from March 2011 to March 2016.

2 - During the desired period, their shares are traded actively in TSE.

3 - Their fiscal year must be ended by March 29th, and it has not been changed during the study period.

4 The companies are not among financial intermediation companies (investment, holding, leasing, and banking and insurance) because of their different financial structure.

5 - The required information is available.

In this study, 99 companies as sample are selected.

3. Discussion and results

3.1 Descriptive statistics

In the section of descriptive statistics, data analysis is performed using central indicators such as mean, middle, and disperse indicators of standard deviation, skewness and elongation. In this regard, mean is the most main central indicator and shows data average such that if data are rowed on an axis, mean value exactly locates on balance point or gravity center. Standard deviation is among dispersion parameters and shows dispersion rate of data. Skewness is also among the parameters to determine deviation from symmetry and is the indicator of symmetry in data. If the socity includes symmetric

distribution, skewness coefficient is equal to zero, if the society has deflection to left, skewness coefficient is negative and if it has deflection to right, skewness coefficient is positive. Elongation is also measurement indicator of society dispersion than normal distribution. In the following table, we observe descriptive statistics of variables after removing outliers.

Table 1. Statistic

	CRASH	ABCASH	ABEXP	CS	MB	SIZE
N	474	474	474	474	474	474
Mean	.66971	1.38320	1.38320	.56555	21.71072	13.94317
Median	.42717	.96185	.96185	.56420	3.45369	13.88652
Std. Deviation	.786516	3.063662	3.063658	.157152	104.053576	1.028313
Variance	.619	9.386	9.386	.025	10827.147	1.057
Skewness	3.171	12.732	12.732	-.046	8.154	.303
Kurtosis	13.082	193.469	193.470	-.179	73.301	.314
Minimum	.008	.166	.166	.221	.858	11.671
Maximum	5.527	53.458	53.458	1.213	1261.699	17.147

3.2 Normality of dependent variable

In order to implement statistical methods and calculate appropriate test statistics and logic inference about research hypothesis, the most important action before any other is to select appropriate statistical method for research. For this purpose, Kolmogorov-Smirnov Test was used in this research for investigating the hypothesis of normality for research data which the results are as following:

Table 2. One-Sample Kolmogorov-Smirnov Test

		CRASH
N		474
Normal Parameters	Mean	.66971
	Std. Deviation	.786516
Most Extreme Differences	Absolute	.210
	Positive	.210
	Negative	-.200
Test Statistic		.210
Sig		.000

With regard to the table of Kolmogorov-Smirnov Test if significance level for dependent variable is lower than test level (0.05), the data distribution is not normal. Since significance level of dependent variable is lower than 0.5. Therefore it doesn't have normal distribution. In this research, we use Box-Cox Plot for normalization.

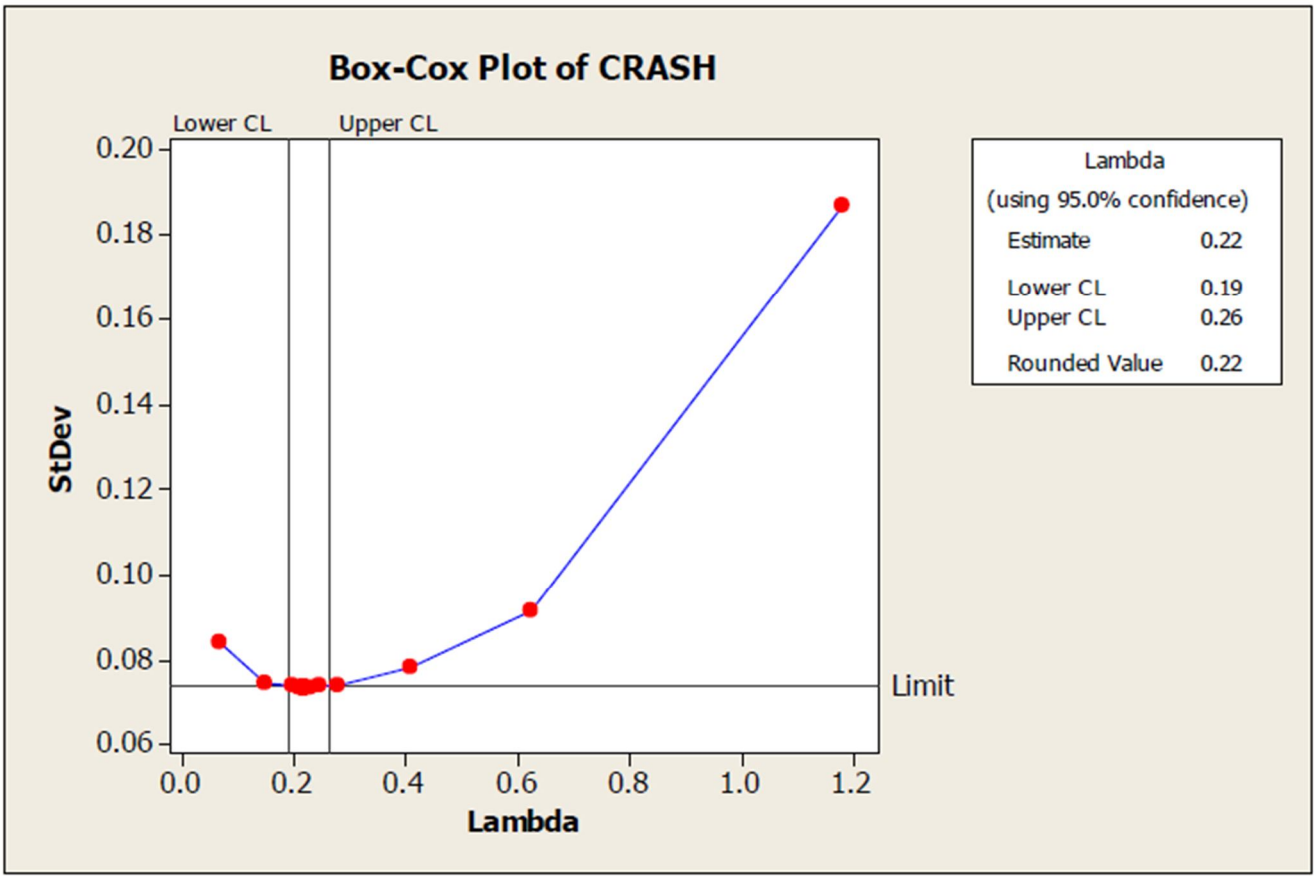

Figure 1. Box-Cox Plot of CRASH

The results from Kolmogorov-Smirnov Test are as following after normalization process:

Table 3. One-Sample Kolmogorov-Smirnov Test

		CRASH1
N		474
Normal Parameters	Mean	.57380
	Std. Deviation	.139396
Most Extreme Differences	Absolute	.077
	Positive	.077
	Negative	-.037
Test Statistic		.077
Sig		.000

According to the Kolmogorov-Smirnov test, after the normalization of the box-cox, the dependent variable is not normal. So, according to the Central Limit Theorem, we consider the dependent variable to be normal. According to the Central Limit Theorem, if the sample number exceeds 30, the variable is normal.

3.3 The findings of the research hypothesis

3.3.1 Allocation of the number of sample members

Allocation of the number of sample members is presented in table 4. The results represent that in this research, 68.78% equivalent to 326 data sample are selected as training sample and 148 data equivalent to 31.22% are selected as testing.

Table 4. Number of Sample

		N	Percent
Sample	Training	326	68.78
	Testing	148	31.22%
Total		474	

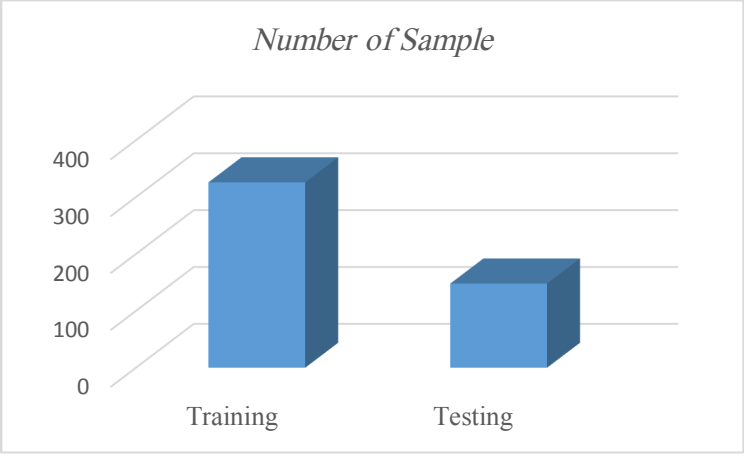

Figure 2. Number of Sample

3.3.2 Summary of model information

The table shows information support vector machine Model related to support vector machine and is applicable for ensure about the correctness of allocated cases. Summary of support vector machine Model is presented in table 5.

Table 5. Summary of Model Information

Input Layer	Covariates	1	ABCASH
		2	ABEXP
		3	CS
		4	MB
		5	SIZE
	Number of Units		5
Output Layer	Dependent Variables	1	CRASH
	Number of Units		1
	Algorithm		SVM
	Kernel type		RBF
	Error Function		Mean Absolute Error

Table 5 shows the summary of support vector machine Model. The results indicate support vector machine includes 5 input layers, and 1 output layer. And the used error function is Mean Absolute Error.

3.3.3 Model Summary

In this table, error Mean Absolute Error are shown since output layer has variables dependent on scale. This is the error function that network tries to minimize it during education operation.

Table 6. Model Summary

'Partition'	Training	Testing
Minimum Error	-0.133	-1.002
Maximum Error	1.144	0.357
Mean Error	0.032	0.005
Mean Absolute Error	0.078	0.065
Standard Deviation	0.133	0.112
Linear Correlation	0.989	0.99
Occurrences	326	148

According to Table 6, the Mean Absolute Error for the Training sample is 0.078 and for the Testing example is 0.065. And since the Mean Absolute Error in the Testing Sample is less than Training, then the model estimation is possible using the support vector machine method.

3.3.4 Independent Variable Importance

This is a sensitive analysis that calculates the importance of each forecaster in determination of support vector machine. The analysis can be performed on the basis of education samples and integrated test or only on education sample if there is no test sample. Finally table7 shows the importance of each forecaster. The results of importance level of independent variable are briefly presented in figure 2 and table 7.

Table 7. Independent Variable Importance

Nodes	Importance
ABEXP	0.0368
ABCASH	0.0368
CS	0.0371
MB	0.3712
SIZE	0.5182

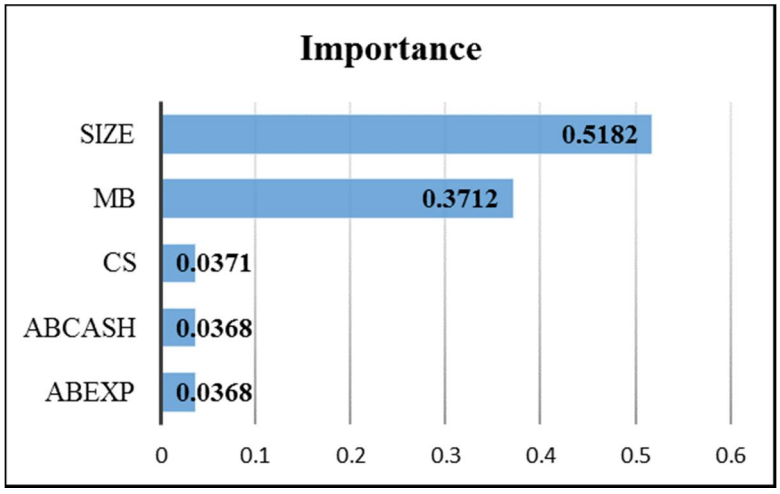

Figure 3. Independent Variable Importance

4. Conclusion

The aim of this research is to Predicting factors affecting the future stock price crash risk based on support vector machine. The data of 99 financial companies listed on Tehran Stock Exchange during the period of 6 years (2011-2016) have been studied. The sample is divided into two categories of testing and training. Allocation of the number of sample members is presented in table 4 and represents that in this research, 68.78% equivalent to 326 data sample are selected as training sample and 148 data equivalent to 31.22% are selected as testing. Also the results of Table 5 shows the summary of support vector machine Model. The results indicate support vector machine includes 3 input layers, and 1 output layer. The method used in this study support vector machine is a kind of RBF. According to Table (6), the Mean Absolute Error for the Training sample is 0.078 and for the Testing example is 0.065. And since the Mean Absolute Error in the Testing Sample is less than Training, then the model estimation is possible using the support vector machine method. Results of this research by research Liao (2016) and Ataei Zadeh and Darabi (2016) and Sadat Shojai (2015) is compatible.

REFERENCES

Ansari, A., & Riasi, A. 2016. An Investigation of Factors Affecting Brand Advertising Success and Effectiveness. International Business Research, 9 (4): 1-11.

Ataei zadeh, R., Darabi, R. 2016. Predicting factors affecting the future stock price crash risk based on the Neural Network Based Radiological Basis Function, 4th International Conference on Applied Research in Management and Accounting, Tarbiat Modares University, 1: 1-12.

Badavar Nahandi, Y., & Taghizadeh Khanqah, V. 2013. The Relationship between Audit Quality and Investment Efficiency, Journal of the Accounting and Auditing Review, 20 (2): 19-42.

Cohen, D., & Zarowin, P. 2010. Accrual-based and real earnings management activities around seasoned equity offerings. Journal of Accounting and Economics, 50 (1): 2–19.

Hutton, A., Marcus, A., & Tehranian, H. 2009. Opaque financial reports and the distribution of stock returns. Journal of Financial Economics, 94 (1): 67-86.

Liao, Q. 2016. The Stock Price Crash Risk Prediction by Neural Network, Accounting and Finance Research, 5 (2): 61-70.

Merton, R. 1987. A simple model of capital market equilibrium with incomplete information. Journal of Finance, 42, 483-510.

Meshki, M., & Ghiasmand, M. 2013. The Relationship Between Earnings Management Based on Income Estimates and Risk of Future Stock Falling in Companies Accepted in Tehran Stock Exchange, Second National Conference on Accounting, Financial Management and Investment, Gorgan, Academic and Professional Association of Golestan Managers and Accountants, https://www.civilica.com/Paper-CAFM02-CAFM02_449.html, 2 (2): 449-464.

Riasi, A. 2015. Competitive Advantages of Shadow Banking Industry: An Analysis Using Porter Diamond Model. Business Management and Strategy, 6(2), 15-27.

Riasi, A., & Amiri Aghdaie, S. F. (2013). Effects of a Hypothetical Iranian Accession to the World Trade Organization on Iran's Flower Industry. Consilience: The Journal of Sustainable Development, 10(1), 99-110. http://dx.doi.org/10.7916/D8HQ3ZK8.

Riasi, A., & Pourmiri, S. 2015. Effects of online marketing on Iranian ecotourism industry: Economic, sociological and cultural aspects. Management Science Letters, 5(10), 915-926.

Sadat Shojai, Z. 2015. "The effects of competition in the product market on the verge of falling real earnings and future stock price of listed companies on Tehran Stock Exchange", a thesis for the degree of Master of Accounting, Islamic Azad University, Branch south Tehran.

How to Cite this Article:

Razdar M., Zahmatkesh A., Khaleh Oghlizadeh S., Predicting factors affecting the future stock price crash risk based on support vector machine, Uct Journal of Management and Accounting Studies 5(3) (2017) 7–14.