Forecasting systematic risk by Least Angel Regression, AdaBoost and Kernel Ridge Regression

Mahdi Salehi[1], Mahdi Moradi[2] & Samaneh Molaei[2]

[1]Assistant Professor of Accounting, Ferdowsi University of Mashhad, Mashhad, Iran

[2] MS.C. Student of Accounting, Mashhad Branch, Islamic Azad University, Mashhad, Iran

Correspondence: Mahdi Salehi, Department of Accounting, Ferdowsi University of Mashhad, Khorasan Razavi, Iran. E-mail: mehdi.salehi@fum.ac

Abstract

In according to importance of risk in financial decision and investment is one of issue that helps to investors is existing tools and appropriate models in order to predict systematic risk. Aim of this research was forecasting systematic risk of companies admitted at Tehran stock Exchange by Least Angel Regression (LARS), AdaBoost and Kernel Ridge Regression (KRR) and comparing ability of the algorithms in order to find the best methods of the test. In this study the financial data of (1159 observations) during 2005-2014. We used MATLAB software vision (R2013b). Results indicated that Kernel Ridge Regression (KRR) with 9.65% error (90.35% confidence) in comparison with Least Angel Regression (LARS) with 12.15% error (87.85% confidence) and AdaBoost with 28.91% error (71.09 confidence) has more ability for forecasting systematic risk. Moreover, ability of forecasting systematic risk in Least Angel Regression (LARS) is more than AdaBoost.

Keywords: systematic risk, Least Angel Regression, AdaBoost, Kernel Ridge Regression

1. Introduction

Investment is fundamental and necessary in processing grow and economic development of country. Among effective factors in selecting investment is paying attention of investors to risk and investment's return.

Investors try to investment their financial resources in where the highest return and the least risk. Thus, companies should instead of focusing on earnings more pay attention to risk as limited factors as maximize return.

In contrast of return, risk is mental concept and non-quantitative. Therefore, most of financial and economic experts more focus on measuring and identifying risk.

Based on modern portfolio theory, risk is divided into two sections: First section which is related to systematic risk; second section, non-systematic risk which is related to specific condition.

In this theory, risk measures by risk of assets with Beta (criteria of systematic risk). Therefore, beta is one of the most applicable and accepted tools of financial and economic experts in order to evaluate and risk management.

Additionally, beta in field of various financial and accounting sciences like fair value of equity, related research about measuring market reaction to specific decision of a company and related research of price has specific responsibility (Hong and Arker, 2007). Several believe that an accounting as a notification system and some believe that it provides information in order to decision makers can do prior decision. In according to importance risk in financial decision and investment is one of issue can helps investors and it provides tools and prior models for predicting systematic risk. Aim of this research is comparing various method of systematic risk in order to select optimal method. Aim of this research is forecasting systematic risk in companies admitted in Tehran stock exchange by using Least Angel Regression (LARS), AdaBoost and Kernel Ridge Regression (KRR) and comparing ability of the algorithms in order to find the best methods of the test.

Efficiency changes of systematic risk of a particular investment to changes in the market and the investment returns of the index β are measured. The index returns reflect the sensitivity of a stock relative to the market portfolio return. Dividing the covariance between the return on asset i in the market portfolio return and the variance of the portfolio return, the beta of asset i is obtained (βi) and calculate as following (Mahdavi and

Godarzi, 2001).

$$\beta = \frac{Cov\ (R_m, R_i)}{\delta^2 R_m}$$

A central issue in portfolio theory in finance is the maximization of returns for people who invest in assets, that is, in firms' stocks (Lintner, 1965; Mossin, 1966; Sharpe, 1964).

The key idea of portfolio theory is that investors can construct a portfolio of stocks with imperfectly correlated returns and thus eliminate nonsystematic (i.e., individualistic) risk associated with those stocks. The remaining variability, the firm's systematic risk, reflects the extent to which its stock's return responds to movement of the average return on all stocks in the market. Beaver, Kettler and Scholes (1970) related systematic risk to variables that describe the financial position of a firm. Specifically, they suggested that greater systematic risk is related to the following:

•Higher growth because, in a competitive economy, the excessive earnings opportunities may erode when new firms enter the industry;

· Greater leverage because the earnings stream of common shareholders becomes more volatile as debt increases;

· Lower liquidity because liquid or current assets result in less volatile returns than do fixed assets;

· Smaller asset size because smaller firms have greater default risk;

· Lower dividend payout because the need to offer steady dividends causes firms with greater volatility to pay out a lower percentage of earnings;

· Higher levels of earnings variability because this results in a lower payout to stockholders; and

· Higher earnings covariability with the market because this results in higher earnings volatility, again lowering the return on the stocks.

Considering two periods (1947–1956 and 1957–1965), BKS (1970) (1) regressed the aforementioned firm characteristics on systematic risk in the first period and (2) examined whether a model of systematic risk from the first period predicted systematic risk in second period better than did systematic risk in first period.

Two diverse streams of empirical research have emerged from BKS's (1970) study. The first stream of research, which is not related to the current study's objectives, focuses on the prediction of the firm's systematic risk in a future period (Elgers, 1980; Eskew, 1979; Ismail and Kim 1989). The second stream of research, which is more pertinent to the current study, augments the predictor variables in BKS's study with additional firm characteristics that may explain systematic risk. Although there are several studies in this stream, our review indicates a lack of cumulative knowledge building in this area. Rather, each study adds some new variables to a subset of the variables in BKS's study. Variables considered in prior research include dividend policy (Bildersee, 1975), operating leverage (Mandelker and Rhee, 1984), earnings funds flow and cash flow (Ismail and Kim 1989), international diversification (Goldberg and Heflin, 1995), and strategic profiles (Veliyath and Ferris, 1997).

Similarly, two studies (Bharadwaj and Menon, 1993; Kroll, Wright and Heiens, 1999) explore the relationship between aspects of a firm's marketing strategy and its risk.

Using service strategic business units from the Profit Impact of Marketing Strategy database, Bharadwaj and Menon (1993) find that some aspects of marketing (e.g., promotional expenditure, sales force expenditure, relative price) are associated with lower variability in return on investment, whereas other aspects of marketing (e.g., advertising, customization) are associated with higher variability in return on investment. Although variability in return on investment, a surrogate for total risk, confounds systematic risk with nonsystematic risk, these results suggest a relationship between firms' marketing activities and their systematic risk. Kroll, Wright, and Heiens (1999) consider a surrogate for systematic risk, the covariance of firms' cash flows relative to a market portfolio of equities. Bougheas et al., (2014) showed that resent progress in financial economy about measuring systematic risk using applicable programs like data mining for analyzing and pricing systematic risk has substantial advantages in concept and practice. Laeven et al., (2014) showed that in according to bigger banks have more systematic risk due to these organizations are involving in activities of banking or they have more organizational complexity in comparison with smaller banks with less capital, traditional banking rules and they are not sufficient for bigger banks. Based on systematic risk is necessary in order to use some rules for avoiding from bankruptcy and financial failures. This action can lead to overinvestment in bigger bank and

consequently reduce contribution in the market. Gregor et al., (2014) concluded that merging banks and capital rising in market reduce systematic bank of risk. Acharya and Stephen (2013) did a research about analyzing systematic risk from E-banking concluded that sources of debt on the banks of one of the main causes of increasing systemic risk in European banks.

1.1 Kernel Ridge Regression (KRR)

Blades linear regression problem is as follows.

$$\min_{w} \ \lambda \|w\|^2 + \sum_{i=1}^{L} \left(y_i - g(x_i) \right)^2 \quad \Leftrightarrow \quad \min_{w} \ \lambda \|w\|^2 + \sum_{i=1}^{L} \xi_i^2$$
$$s.t. \ g(x) = w^T x \qquad\qquad\qquad s.t. \ \xi_i = y_i - \langle w, x_i \rangle$$

Equitation (1)

In the nonlinear case, it comes in the form below:

$$\min_{f \in H} \ \sum_{i=1}^{N} \left(y_i - \langle f, \varphi(x_i) \rangle_H \right)^2 + \lambda \|f\|_H^2$$

Equitation (2)

Where H denotes the Hilbert space and f coefficients in Hilbert space and function φ mapping from input space to feature space is a Hilbert. If the coefficients can be written in the form of dual $f = \sum_{i=1}^{N} c_j \varphi(x_j)$ and we used Kernel trick $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ the shape of the output function is calculated as follows: (Hastie et al., 2009).

$$f(x) = \langle w, x \rangle = \langle f, \varphi(X) \rangle_H = y^T \left(K + \lambda I_N \right)^{-1} k(x)$$

Equitation (3)

1.2 AdaBoost

The main idea of boosting is to combine rough and moderately inaccurate weak hypotheses to form a very accurate strong one. AdaBoost (YoavFreund and Robert, 1997) named by its capability of adjusting adaptively to the errors of weak classifiers, has been widely used in a variety of object detection applications (Viola et al., 2003; Miyake et al., 2006; Wei and Weiming, 2005) because of its excellent performance in both speed and accuracy since Viola and Jones (VJ) proposed their successful detection framework (Paul et al., 2004; Paul et al., 2001).

There are 3 main contributions in VJ's work. First, the Integral Image representation, obtained in one pass over the original image, helps evaluate features in constant time at any scale or location on the original image; and then among the hundreds of thousands of features, a learning algorithm based on AdaBoost can effectively select a small set of the most discriminative ones with lowest errors, and boost them into a strong classifier; finally, the Cascaded Classifiers Structure combines those strong classifiers and accelerates the detection procedure by quickly rejecting many negative sub-windows with little processing, such that computations are mainly focused on promising regions.

However, it is noticed that a strong classifier learnt by AdaBoost is suboptimal for applications in terms of error rate (Buhman and Yu, 2000). Some improved variants of the algorithm have been proposed. To reduce false positives, VJ proposed a simple extension of AdaBoost called AsymBoost (Viola and Jones, 2002). It balances the asymmetric costs between false negatives and false positives somewhat by reweighting positive and negative samples at each training round. FloatBoost (Li and Zhang, 2004), proposed by Stan Z. Li et al., incorporates the backtrack mechanism from Floating Search, and repeatedly performs a backtrack to remove unfavorable weak classifiers after a new weak classifier is added by AdaBoost, such that a lower error rate and reduced feature set are guaranteed at the cost of about five times longer training time. And another much more complex variant Ent-Boost (Le and Satoh, 2007), partially motivated by Kullback- Leibler Boosting, employs symmetric KL divergence for optimal weak classifier selection. Generally, all these works modified the mechanism of weak classifier selection through error rate evaluation.

1.3 Least Angle Regression

Least Angle Regression (LARS) relates to the classic model-selection method known as Forward Selection, or "forward stepwise regression," described as given a collection of possible predictors, we select the one having largest absolute correlation with the response y, say xj1, and perform simple linear regression of y on xj1. This leaves a residual vector orthogonal to xj1, now considered to be the response. We project the other predictors orthogonally to xj1 and repeat the selection process. After k steps this results in a set of predictors xj1, xj2, . . . ,xjk that are then used in the usual way to construct a k-parameter linear model. Forward Selection is an aggressive fitting technique that can be very greedy, perhaps eliminating at the second step useful predictors that happen to be correlated with xj1. Forward Stagewise, as described below, is a much more cautious version of Forward Selection, which may take thousands of tiny steps as it moves toward a final model. It turns out, and this was the original motivation for the LARS algorithm, that a simple formula allows Forward Stagewise to be implemented using fairly large steps, though not as large as a classic Forward Selection, greatly reducing the computational burden. The geometry of the algorithm, described in Section 2, suggests the name "Least Angle Regression."

It then happens that this same geometry applies to another, seemingly quite different, selection method called the Lasso (Tibshirani, (1996). The LARS–Lasso–Stagewise connection is conceptually as well as computationally useful. The Lasso is described next, in terms of the main example used in this paper.

Least angle regression (Efron et al., 2004) can be viewed as a version of stagewise that uses mathematical formulas to accelerate the computations. Rather than taking many tiny steps with the first variable, the appropriate number of steps is determined algebraically, until the second variable begins to enter the model. Then, rather than taking alternating steps between those two variables until a third variable enters the model, we jump right to the appropriate spot. Figure 3 shows this process in the case of 2 predictor variables, for linear regression.

The first variable chosen is the one which has the smallest angle between the variable and the response variable; in Figure 3 the angle COX1 is smaller than COX2. We proceed in that direction as long as the angle between that predictor and the vector of residuals Y X1 is smaller than the angle between other predictors and the residuals. Eventually the angle for another variable will equal this angle, at which point we begin moving toward the direction of the least-squares fit based on both variables. In higher dimensions we will reach the point at which a third variable has an equal angle, and joins the model, etc.

Expressed another way, the (absolute value of the) correlation between the residuals and the first predictor The LAR algorithm in the case of 2 predictors. O is the prediction based solely on an intercept.

C = ^ Y = ^ _1X1 + ^ _2X2 is the ordinary least-squares fit, the projection of Y onto the subspace spanned by X1 and X2. A is the forward stepwise fit after one step; the second step proceeds to C. Stagewise takes a number of tiny steps from O to B, then takes steps alternating between the X1 and X2 directions, eventually reaching E; if allowed to continue it would reach C. LAR jumps from O to B in one step, where B is the point for which BC bisects the angle ABD. At the second step it jumps to C. The Lasso follows a path from O to B, then from B to C. Here LAR agrees with Lasso and stagewise (as the step size for stagewise). In higher dimensions, additional conditions are needed for exact agreement to hold.

2. Methodology

We used one perspective of two levels by using the algorithms in order to forecast systematic risk. The two levels are as following:

1. Application of stepwise regression to select the input variables to the model
2. LARS algorithms, AdaBoost and Kernel Ridge Regression (KRR) to predict systemic risk

2.1 Hypotheses

H1: Least Angel Regression (LARS) has more ability than AdaBoost in order to forecast systematic risk.

H2: Kernel Ridge Regression (KRR) has more ability than AdaBoost in order to forecast systematic risk.

In first level, after collecting data by using stepwise regression among independent variables; we use significant variables in order to enter in the model. In following step, by using LARS, AdaBoost and Kernel Ridge Regression (KRR) provides a model in order to forecast systematic risk. We used MATLAB and this software is one of the most powerful mathematical software and has various applications in other fields.

2.2 Selected Data

The first level is selecting data. Financial data includes (1159 years-companies) with 30 independent variables along with dependent variable with systematic risk between 2005 and 2014. In following table is shown independent variables

Table 1. Independent variables

Variables	Row	Variables	Row
The percentage change in the sum of non-current assets	16	Gross profit margin	1
The percentage change in total revenue	17	Operating profit margin	2
The percentage change in earnings (losses), after-tax	18	Net profit margin	3
Price to Earnings Ratio	19	Return on assets	4
Interest coverage ratio	20	Return on equity	5
Collection period	21	Earnings per share	6
Maintenance of inventory	22	Current Ratio	7
Current asset turnover	23	Acid ratio	8
Cash interest cover	24	Leverage	9
Cash earnings cover	25	Debt to equity	10
Acquisition of capital	26	Inventory turnover	11
Cash investment finance	27	Fixed asset turnover	12
Size	28	Turnover of total assets	13
Sales to equity	29	Accounts receivable turnover	14
Long-term debt to equity	30	Percent Change Total current assets	15

Cleaning and preparing data:

Second level is cleaning and preparing the data. In this level some of independent variables will be due to dysfunction of information. We used stepwise in order to clean and prepare data. Methods hypothesis underlying this regression is inevitable. Investigate the hypothesis underlying regression testing is done as follows:

A) Homogeneity of variance:

The residual plot against predicted values to evaluate the homogeneity of variance in the model is studied.
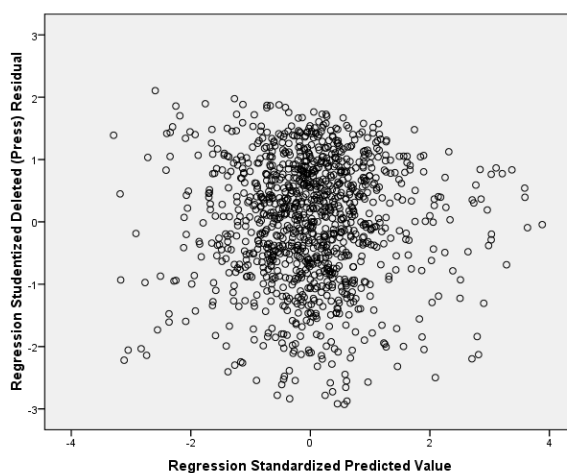

Figure 1. Residual against predicted value

B) Normality: Normality of data is confirmed in according to normality of residual regression.
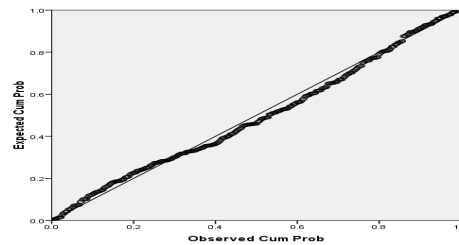
Figure 2. Normality chance of residual

C) Independency of residuals: in according to Durbin-Watson in final model is 1 and 3 and third hypothesis is approved

D) Lack of linearity: Results of VIF statistics is less than 10 and therefore regression is confirmed.

Table 2. Durbin-Watson

Durbin_Watson	Model
1.827	Multiple regression

Table 3. VIF test

Independent variable	VIF	Independent variable	VIF
Percentage change in the sum of non-current assets	1	Gross margin	1
Percentage change in total revenue	1	Operating profit margin OPM	1
Percentage change in profit (loss) after tax	1	Net profit margin NPM	1
Price to earnings ratio P / E	1	Return on assets ROA	3
Earnings before interest and taxes Interest expense - Interest coverage ratio	1	Return on equity ROE	1
Collection period	1	Net earnings per share EPS	1
Maintaining inventory	1	Current ratio CR	1
Circulating current assets	1	QR quick ratio	3
The cash interest coverage (operating cash flow to interest expense(	1	Leverage LEV	1
The cash interest coverage (operating cash flow to net income(	1	Debt to Equity	1
Acquisition of capital (operating cash flow minus dividends paid to fund acquisition of assets(	1	IT inventory turnover	1
Cash investment finance	1	Fixed asset turnover FAT	1
Size (log assets)	1	total asset turnover AT	2
Sales to equity	1	RT receivables turnover	1
Long-term debt to equity	1	Percent change Total current assets	1

After test of regression, we tested research's model. We investigated input variables in order to test exterior variable of the research (systematic research). We used hypothesis of research from analyzing regression in 5% error.

Table 4. Test of Square

Model	Describe	sum of squares	F	Sig
1	Regression	169.175	7.892	0.000
	Residual	5707.675		
	Total	5876.849		

Eventually, among independent variables selected by stepwise of four variables which are shown in following list

Table 5. Independent variable

Row	Independent variable
6	Earnings Per Share EPS
7	Current Ratio CR
17	Percentage change in total revenue
23	Circulating current assets

Results of hypotheses test

In according to related algorithm of LARS, AdaBoost and Kernel Ridge Regression (KRR) and results of the tests are as following:

Table 6. LARS, AdaBoost and Kernel Ridge Regression (KRR) and results

Row	Error of forecasting LARS Algorithm	Error of forecasting Adaboost Algorithm	Error of forecasting Kernel Ridge Regression
1	15.01	38.5	14.03
2	14.05	30.57	12.93
3	12.43	30.36	9.58
4	12.36	31.59	12.07
5	14.65	23.11	9.88
6	7.4	19.79	4.67
7	12.06	28.77	8.02
8	9.42	24.3	4.56
9	14.56	30.64	12.57
10	9.6	31.51	7.98
Mean error	12.15	28.91	9.65

Results of the table above mentioned is Kernel Ridge Regression with 9.65 error percentage (90.35 percent confidence) has higher ability in order to forecast systematic risk in comparison LARS with 12.15 percent error (87.85 percent confidence) and AdaBoost with 28.91 percent error (71.09 percent confidence).

3. Conclusion

In current research, we investigated by using 30 financial ratios in order to forecast systematic risk of companies admitted listed Tehran stock exchange of LARS algorithm, AdaBoost and Kernel Ridge Regression. Financial information of 1159 observations during 2005-2014. Results of research indicated that Kernel Ridge Regression (KRR) with 9.65% error (90.35% confidence) in comparison with Least Angel Regression (LARS) with 12, 15% error (87.85% confidence) and AdaBoost with 28.91% error (71.09 confidence) has more ability for forecasting systematic risk. Moreover, ability of forecasting systematic risk in Least Angel Regression (LARS) is more than AdaBoost.

References

Acharya, V., & Steffen, S.(2013). Analyzing Systemic Risk of the European Banking Sector, in Handbook on Systemic Risk, Fougue – Langsam (eds.), Cambridge University Press, pp, 247-282. Retrieved from http://ssrn.com/abstract=2150616

Beaver, W., Kettler, P., & Scholes, M. (1970). The association between market determined and accounting determined risk measures. *The Accounting Review, 45*, 654-682

Bharadwaj, S. G., & Menon, A. (1993). Determinants of Success in Service-Based Industries: A PIMS-Based Empirical Investigation. *Journal of Services Marketing, 7*(4), 19–40.

Bildersee, J. S. (1975). The association Between Market-Determined Measure of Risk. *The Accounting Review,* 81-89

Bougheas, S., & Kirman, A. P. (2014). Complex Financial Networks and Systemic Risk: A Review, CESifo Working Paper Series 4756. Retrieved from http://ssrn.com/abstract=2436826

Buhlmann, P., & Yu, B. (2000). Invited discussion on Additive logistic regressions: a statistical view of boosting (Friedman, Hastie and Tibshirani). *Annual of Statistics, 28*(2), 377-386.

Efron, B., Hastle, T., Johnstone, L., & Tibshirani, R. (2004). *Least Angle Regression Annals of Statistics, 32*(2), 407–499.

Elgers, P. T. (1980). Accounting-Based Risk Predictions: A Re-Examination. *The Accounting Review, 55*(3), 389–408.

Eskew, R. K. (1979). The Forecasting Ability of Accounting Risk Measures: Some Additional Evidence. *The Accounting Review, 54*(1), 107-118.

Goldberg, S. R., & Frank, H. (1995). The Association between the Level of International Diversification and Risk. *Journal of International Financial Management and Accounting, 6*(1), 1-25.

Gregor, N. F. W., Sascha, N., & Denefa, B. (2014). Systemic Risk and Bank Consolidation: International Evidence. *Journal of Banking and Finance, 40,* 165-181.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction (2nd ed).

Hong, G., & Sarkar, S. (2007). Equity Systematic Risk (Beta) and Its Determinants. *Contemporary Accounting Research, 24*(2), 423-466.

Ismail, B., & Kim, M. (1989). On the Association of Cash Flow Variables with Market Risk: Further Evidence. *Accounting Review, 64*(1), 125-136.

Kroll, M., Wright, P., & Heiens, R. A. (1999). The Contribution of Product Quality to Competitive Advantage: Impacts upon Systematic Variance and Unexplained Variance in Returns. *Strategic Management Journal, 20*(4), 375–384.

Laeven, L., Ratnovski, L., & Tong, H. (2014). Bank Size and Systemic Risk. http://dx.doi.org/10.2139/ssrn.2437729

Le, D. D., & Satoh, S. (2007). Ent-Boost: Boosting Using Entropy Measures for Robust Object Detection. *PRL, 28*(9), 1083–1090.

Li, S. Z., & Zhang, Z. (2004). Floatboost learning and statistical face detection, IEEE Trans. *Pattern Analysis and Machine Intelligence, 26*(1), 1112–1123.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics, 47*(1), 13–37.

Mahdavi, Gh., & Godarzi, K. (2001). Presents an artificial neural network to predict systemic risk using macroeconomic variables (Case Study: Saipa Co.). *Journal of Economic, 4*(11), 219-237.

Mandelker, G. N., & Rhee, S. G. (1984). The Impact of the Degrees of Operating and Financial Leverage on Systematic Risk of Common Stock. *Journal of Financial and Quantitative Analysis, 19*(1), 45–57.

Miyake, N., Takiguchi, T., & Ariki, Y. (2006). Noise detection with multi-class AdaBoost. IEICE Technical Report.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica, 34*(4), 768-783.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance, 19*(3), 425-442.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Royal. Statist. SocB, 58*(1), 267-288.

Veliyath, R., & Ferris, S. (1997). Agency Influences on Risk Reduction and Operating Performance: An Empirical Investigation among Strategic Groups. *Journal of Business Research, 39*(3), 219-230.

Viola, P., & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. *CVPR,* (1), 511–518.

Viola, P., & Jones, M. (2002). Fast and robust classification using asymmetric AdaBoost and a detector cascade.

In NIPS, 14, 1311-1318.

Viola, P., Jones, M. (2004). Robust Real-Time Face Detection. *International Journal of Computer, 57*(2), 137–154.

Viola, P., Jones, M., & Snow, D. (2003). Detecting pedestrian using patterns of motion and appearance. In: ICCV, 734-741.

Wei, H., & Weiming, H. (2005). Network-based intrusion detection using Adaboost algorithm. Proceedings of IEEE/WIC/ACM, *International Conference on Web Intelligence,* 712-717.

Yoav, F., & Robert, D. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences, 55*(1), 119-139.

Copyrights