The following publication Yang, Y. C., Ke, Y. S., Wu, W., Lin, K. P., & Jin, Y. (2019, July). Recommendation as a Service in Mergers and Acquisitions Transactions. In International Conference on Human-Computer Interaction (pp. 151-159). Springer, Cham is available at https://doi.org/10.1007/978-3-030-22338-0_12

Recommendation as a Service in Mergers and Acquisitions Transactions

Abstract

"Mergers and acquisitions (M&A)" is a very common business activity between the corporations to combine and/or transfer their ownerships, operating units and assets, which put separate companies together to establish larger ones. M&A happens frequently in the high-technology industries because these IT companies are motivated for the speedy innovation and required to extend their resources and capabilities through the M&A transaction. It is of interest to study and analyze which kinds of firms will be selected as the M&A target. The purpose of the study is to provide a method that automatically determine a feasible M&A deal. However, few studies attempts to use the techniques of support vector machine (SVM) to evaluate a M&A deal. The study aims to automatically determine a M&A deal by applying the SVM model. We further integrate the SVM model with three different kernels, including Gaussian, polynomial, and financial. The proposed technique is empirically evaluated, and the result shows the effectiveness of the technique.

Keywords: Mergers and acquisitions, machine learning, support vector machine, kernel

1. Introduction

The term "Mergers and acquisitions" (abbreviated M&A) usually refers to a common business activity in which companies transfer their ownership, operating units and assets to other business organizations, which put separate companies together to establish larger ones. The major objective of M&A is to improve companies' financial and operating performances with potential synergies, such as market share, profits, economies of scale, influence in the industry etc. Shown as Figure 1, Thomson Reuters reports that the value of worldwide M&A deals in the first nine months of 2018 reached $3.3 trillion, increased by 39% from 2017, and almost half of the deals worth more than $5 billion[1]. Thomson Reuters also shows that the largest M&A market is in the United States, following by Europe, Asia Pacific, Japan, and Africa-Middle East[2]. The volume of global M&A is continuing to grow rapidly, and M&A is one crucial trend of business behavior.

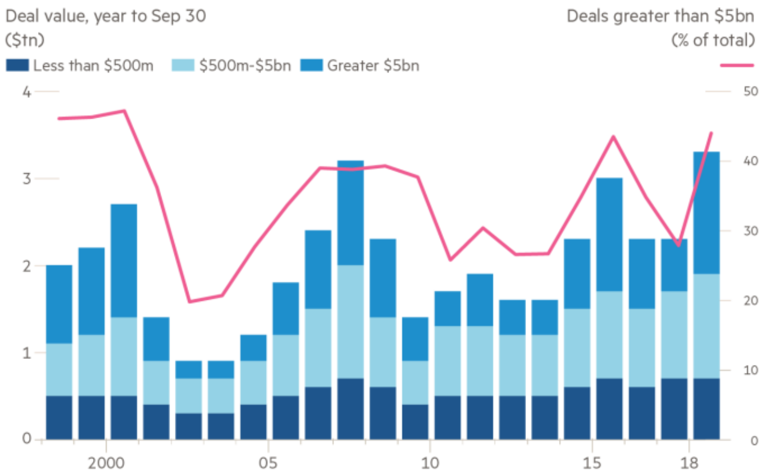

Figure 1: Global M&A activities from 1998 to 2018.
Source: Financial Times & Thomson Reuters.

M&A happens frequently in the high-technology industries. Hagedoorn & Duysters (2002) and King et al. (2008) further indicate that high-technology companies may need to expand their resources and capabilities through the transactions of mergers and acquisitions due to the motivations for the speedy innovation. A significant example is the acquisition of Skype for $8.56 billion from Microsoft in

[1] https://www.ft.com/content/b7e67ba4-c28f-11e8-95b1-d36dfef1b89a
[2] https://www.nytimes.com/2018/07/03/business/dealbook/mergers-record-levels.html

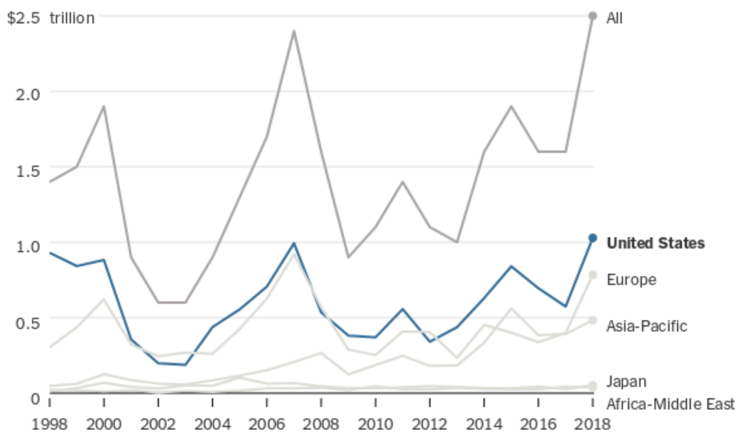

Figure 2: The Major M&A Markets.
Source: New York Times & Thomson Reuters.

2011[3]. The acquisition increases Microsoft's accessibility of real-time video and voice communication. There are still many other well-known examples, just to name a few, *e.g.* Microsoft bought Nokia's mobile phone unit for $7.2 billion in 2013[4], Facebook spent $19 billion to acquire WhatsApp in 2014[5]. The M&A transactions can even be international, *e.g.* Taipei-based company Gogolook, well-known for the online caller ID and number management app "Whoscall", was the subject of a $17.6 million purchase by the Korean Internet giant Naver in 2013 [6]. Hence, It is apparently important for businesses to consider combing another company, and furthermore, how can we analyze and judge a M&A target company more effectively and precisely? For example, some companies may attempt to manipulate the financial statements in order to sell at a higher price; other potential companies may be offered a lower price due to its inaccurate financial records. It is of interest to study and analyze which kinds of firms will be selected as the M&A target.

The purpose of the study is to provide a method that help determine if a firm is suitable to be the M&A target. Prior studies attempts to use various techniques of machine learning to evaluate a M&A firm, for example, logistic regressions (Meador et al., 1996; Barnes, 2000; Ali-Yrkkö et al., 2005; Pasiouras & Gaganis, 2007), discriminant analysis (Barnes, 2000; Pasiouras & Gaganis, 2007) , rule

[3]https://news.microsoft.com/2011/05/10/microsoft-to-acquire-skype
[4]http://venturebeat.com/2013/09/02/welcome-to-microkia-microsoft-buys-nokias-devices-and-services-biz/
[5]https://techcrunch.com/2015/02/19/crazy-like-a-facebook-fox/
[6]https://techcrunch.com/2013/12/26/gogolook-confirms-its-acquisition-by-naver-the-owner-of-line/

induction (Ragothaman et al., 2003), neural networks (An et al., 2006) and decision tree (Tsagkanos et al., 2007; Yang et al., 2014) etc. However, few studies implement the support vector machine (abbreviated SVM) algorithm as M&A prediction models. Comparing with the above prediction models, the SVM model is very efficient for binary classification, including quickly finding hyperplanes to separate data and shorter training time (Cristianini & Shawe-Taylor, 2000). Furthermore, our work incorporate three different kernels, including a Gaussian kernel, a polynomial kernel (Cristianini & Shawe-Taylor, 2000) and a financial kernel (Cecchini et al., 2010). Finally, as we mention in the first section, the largest M&A market is in the United States, and we focus mainly on the market. Previous studies such as Yang et al. (2014) provide a novel technique to evaluate the M&A deals, but they only work on the Asia Pacific market, which may not be representative of the whole M&A markets.

The remainder of the paper is organized as follows. Section 2 lists the related work to our study. Section 3 formulates the proposed M&A forecasting model based on the integration of SVM and kernels. Section 4 presents the evaluation of the proposed technique. Finally, the conclusions and possible directions for future research are provided in Section 5.

2. Literature Review

2.1. M&A Predictions

The popular analysis techniques applied to developing M&A predictions includes logistic regression (Meador et al., 1996; Barnes, 2000; Ragothaman et al., 2003; Ali-Yrkkö et al., 2005; Pasiouras & Gaganis, 2007), discriminant analysis (Barnes, 2000; Ragothaman et al., 2003), rule induction (Ragothaman et al., 2003), and decision tree (Tsagkanos et al., 2007; Yang et al., 2014). First, logistic regression is the most common model. Meador et al. (1996) use logistic binary regression analysis to examine the accounting, financial, and market variables to predict the M&A target companies as well as horizontal and vertical subsamples of merged companies over the period 1981 to 1985. Their model shows the strongest predictive ability for horizontal acquisitions. Pasiouras & Gaganis (2007) also employ the model of logistic regression to examine the financial characteristics of Asian banks during the period of 1998 to 2004. They further indicate that high asset risky portfolios and high liquidity increase the probability of being involved in an acquisition. Ali-Yrkkö et al. (2005) adopt the idea of patent-driven M&A and test their ideas with a large sample of small and private Finnish firms. The multinomial logistic estimations show that the number of patents owned by a Finnish firm will positively influence the probability of being acquired by a foreign firm. Second, discriminant analysis and rule induction are also popular for predicting M&A targets. Barnes (2000) compares the logit model with linear discriminant analysis, while Ragothaman et al. (2003) apply the techniques of artificial intelligence (AI)-based rule induction to identify acquisitions targets. Finally, decision tree is a new application in the section of M&A. Tsagkanos et al. (2007) apply the decision tree models to predict Greek M&A targets, and the results show that their proposed technique is better than the traditional regression tree. Yang et al. (2014) propose a M&A prediction technique that incorporates a comprehensive set of technological indicators, the technological profiles of both the bidder firm and a candidate target firm. Different from prior studies, they derive some technological indicators derived via patent data analyses. The work of Yang et al. (2014) is the latest study that explores the M&A predictions so far. They employ the model of decision tree learning to evaluate the M&A deals and aggregate some key financial indicators related to M&A. However, their model may not be generalized since they only apply the technique in the East Asia market. Hence, we will extend their work and the related indicators in the study.

2.2. Support Vector Machine and Financial Kernels

The support vector machine is a popular classification method created by Vapnik and colleagues (Boser et al., 1992; Cortes & Vapnik, 1995; Cristianini & Shawe-Taylor, 2000). SVMs are supervised learning models that are used for binary classifications and regression analyses. Given a set of training data points, each point is marked as either one or the other of two categories, *i.e.,* $\{\mathbf{x}_i, y_i\}$, $i = 1, \ldots, l$, where $\mathbf{x}_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$. SVMs distinguish the data points into two groups by determining a separating hyperplane $\{\mathbf{x} : \langle \mathbf{w}, \mathbf{x} \rangle + b = 0\}$ in the feature space, where $\mathbf{w}$ is normal to the hyperplane, $\frac{|b|}{||\mathbf{w}||}$ is the perpendicular distance from the hyperplane to the origin, and $||\mathbf{w}||$ is the Euclidean norm of $\mathbf{w}$. The feature space is an abstract space where each pattern sample is represented as a point in n-dimensional space and its dimension is determined by the number of features used to describe the patterns.

Sometimes, we may have to deal with high-dimensional feature space. In order to reduce the dimensionality, some techniques can be employed, and one of the popular techniques is “kernel methods” – providing a powerful tool for learning non-linear relations with a linear machine. Furthermore, it is helpful to produce better outcomes if kernels are built based on application-specific information (Cristianini & Shawe-Taylor, 2000). Kernel methods are used for pattern discovery, dealing with general types of data (*e.g.* strings, vectors or text), and find out general types of relations (*e.g.* rankings, classifications, regressions, clusters). Kernel methods are widely used in many fields, such as finance (Cecchini et al., 2010), image recognition (Chapelle et al., 1999), bioinformatics (Hua & Sun, 2001; Hu & Pan, 2007), etc.

The basic philosophy is that a certain type of similarity measure, *i.e.* a kernel, maps the data set into a high-dimension feature space, in which linear methods are used for learning and estimation problems. We use the symbol K to represent a kernel matrix such that $K(\mathbf{u}, \mathbf{v}) = \langle \phi(\mathbf{u}), \phi(\mathbf{v}) \rangle$, where $\phi : X \to F$ means an implicit mapping ϕ from an input attribute space X onto some feature space F, and $\mathbf{u}, \mathbf{v} \in X$. A kernel matrix K is required to satisfy these conditions: symmetric, positive semidefinite, and the Cauchy-Schwarz inequality (Cristianini & Shawe-Taylor, 2000). Several common types of kernel functions are listed in Cristianini et al. (2002) and the number is ever growing. One typical example is the polynomial kernel, which is defined as $K(\mathbf{u}, \mathbf{v}) = (K_1(\mathbf{u}, \mathbf{v}) + R)^d$, where $K_1(\mathbf{u}, \mathbf{v})$ is the normal inner product kernel, d is a positive integer, and R is fixed. Another typical example is the Gaussian kernel (or called the radial basis function kernel): $K(\mathbf{u}, \mathbf{v}) = \exp\left(\frac{\|\mathbf{u}-\mathbf{v}\|^2}{2\sigma^2}\right)$, where σ is a free parameter and determines the width of the kernel.

Cecchini et al. (2010) propose a useful financial kernel to determine management fraud. The financial kernel is denoted as $K_F\{\mathbf{u}, \mathbf{v}\}$ and computes all ratios of input attributes as well as year-over-year ratio. It begins with n attributes and produces $3n(n-1)$ features, which can be broken into six feature types. The mapping ϕ is represented as:

$$\phi(\mathbf{u}) = \left(\frac{u_{i1}}{u_{j1}}, \frac{u_{j1}}{u_{i1}}, \frac{u_{j2}}{u_{i2}}, \frac{u_{i2}}{u_{j2}}, \frac{u_{i1}u_{j2}}{u_{j1}u_{i2}}, \frac{u_{j1}u_{i2}}{u_{i1}u_{j2}}\right)', \quad i, j = 1, \ldots, n,\ i < j.$$

The financial kernel is also working on other financial analyses even though initially it is proposed for detecting management fraud. We are following the work of Cecchini et al. (2010) and apply the financial kernel into the M&A prediction.

3. The Proposed M&A Forecasting Technique

In the section, we detail the design of our proposed M&A forecasting technique. Following the work of Yang et al. (2014), we replace the forecasting technique with the SVM model integrated with a Gaussian kernel, a polynomial kernel, or a financial kernel presented by Cecchini et al. (2010). Figure 3 shows the details of our proposed technique, where a training phase and a forecasting phase are involved.

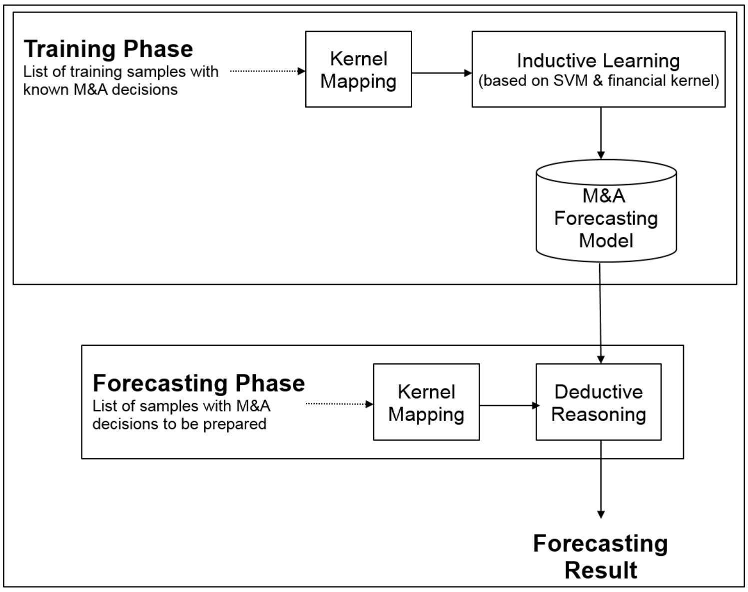

Figure 3: Overall process of the proposed M&A forecasting technique

The training phase involves two major steps: kernel mapping and inductive learning. First, we follow the previous studies and extract the values of the related financial variables for each training sample (see Table 1). Then all the values of the financial variables will be mapped via a kernel function such as a polynomial kernel, a radial basis kernel or a financial kernel. Following kernel mapping is the inductive learning step, and we choose the R package “e1071” (Hornik et al., 2006; Dimitriadou et al., 2008), a supervised learning technique that provides computational efficiency and excellent interpretability. The package “e1071” offers a powerful function “svm()” with flexible parameter tuning methods. In the inductive learning step, we employ SVM integrated with a kernel to induce M&A forecasting models from the set of training instances. In the forecasting phase, we use the forecasting model induced by SVM in the training phase to make an M&A prediction for each candidate target company.

Table 1: Definitions of Financial Indicators of M&A

Indicator	Definition
3-year average dividend (DVT3)	A company's dividend payments to its shareholders over the last three years
Capital-expenditures-to-total-asset ratio (CETA)	(Current assets − current liabilities)/(total assets)
Cash flow (CF)	Amount of money moving into and out of a business
Common shares traded divided by common shares outstanding (CSTRCSHO)	(Shares traded) / (shares outstanding)
Cost of goods sold (COGS)	Carrying value of goods sold during a particular period
Cost of goods sold divided by average inventory (COGSNI)	(Costs of goods sold) / inventory
Current ratio (CURRENTRATIO)	(Current assets) / (current liabilities)
Debt-to-assets ratio (DEBTTOASSETS)	(Sum of long-term and short-term debt) / (total assets)
Dividend (DVT)	A company's dividend
Debt-to-equity ratio (DEBTTOEQUITY)	(Sum of long-term and short-term debt) / (book value of equity)
Earnings before interest and taxes or operating income after depreciation (EBIT)	Revenue minus expenses, excluding tax and interest
Tobin's Q (Q)	(Sum of short-term and long-term debt) / (total assets)
Price-to-earning ratio (PE)	A company's share price to its per-share earnings
Profit margin (PROFITMAT)	Net income divided by revenue, or net profits divided by sales
Ratio of tangible (fixed) assets to total assets (TANGIBLEAT)	(Tangible assets) / (total assets)
Return on assets (ROA)	(Net income) / (total assets)

Return on equity (ROE)	(Net income) / (book value of equity)
Return on investment (ROI)	(Gain from investment − cost of investment) / (cost of investment)
Sales to total assets or asset turnover (ASSET-TURNOVER)	(Net sales) / (total assets)
Tax shield effects (TAXSHIELD)	Taxable income reduces claiming deductions

4. Empirical Evaluation

In this section, we express how we collect the data and design the evaluation. We then show the results from the evaluation of the proposed M&A forecasting technique.

4.1. Data Collection

The M&A cases are collected from the SDC Platinum database, which is available at *https://financial.thomsonreuters.com/en/products/data-analytics/market-data/sdc-platinum-financial-securities.html*. Totally 5,804 cases are collected, and the M&A cases are within 2000 and 2011 in technology-related industries of North America. These industries includes hardware (with first-two-digit SIC code 35), software (with first-two-digit SIC code 36), and computer related business service (with first-three-digit SIC code 234). We further check whether these cases are available for our empirical evaluation purpose. If the M&A case is not likely technology-oriented, we remove it from the our data set. As a result, we end up with retaining a data set consist of 83 M&A cases and 680 non-M&A cases. The ratio between M&A cases and non M&A cases in the data set is $83/680 = 0.12205$.

For each case, the values of the corresponding financial variables (see Table 1) are collected from the Compustat database in the Wharton Research Data Services (WRDS, available at *https://wrds-web.wharton.upenn.edu/wrds/*). In order to transforming values taken from different sources into a consistent format, we further standardize the range of financial variables by this way:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)},$$

where x is an original value, and x' is the standardized value. Furthermore, the decision on M&A deals in a constantly changing business environment is usually dynamic and oscillated over time. We use lagged variables to incorporate feedback over time and capture the year-over-year effect. Hence, for each case, the values of 20 pairs financial variables are extracted from WRDS, and each pair of variables include one variable for the current year and one corresponding lagged variable for the previous year. Each M&A case (or each training instance) is expressed as:
(*DVT3, DVT3_lag, DVT, DVT_lag, CETA, CETA_lag, CF, CF_lag, CSTRCSHO, CSTRCSHO_lag, COGS, COGS_lag, COGSNI, COGSNI_lag, CURRENTRATIO, CURRENTRATIO_lag, DEBTTOASSETS, DEBTTOASSETS_lag, DEBTTOQUITY, DEBTTOEQUITY_lag, EBIT, EBIT_lag, Q, Q_lag, PE, PE_lag, PROFITMAT, PROFITMAT_lag, TANGIBLEAT, TANGIBLEAT_lag, ROA, ROA_lag, ROE, ROE_lag, ROI, ROI_lag, ASSETTURNOVER, ASSETTURNOVER_lag, TAXSHIELD, TANGIBLEAT_lag*).

4.2. Evaluation and Results

In this section, we measure the effectiveness of our proposed M&A forecasting technique on the basis of the complete data set. The criteria used to measure the performance evaluation are "accuracy", "precision", "recall", and "F_1". First, we compare with different kernels in order to determine the optimal model based on the proposed M&A prediction technique. Totally three different kernels are considered, including Gaussian, polynomial, and financial. We also attempt to fine tune the parameters after finding the optimal kernel. Second, we report the results of the test sample using the proposed technique. Finally, we also measure the performance of the proposed technique in comparison to another technological-based benchmark (*i.e.*, logistic regression).

In order to determine the optimal model, we fine tune the parameters, measure the effectiveness, and compare the proposed technique with different kernels. Tables 2 and 3 show the tuning results of the technique integrated with the Gaussian kernel and the polynomial kernel, respectively. Apparently, the Gaussian kernel performs best as the parameter $\gamma = 0.5$, and the polynomial kernel is best with $\gamma = 0.5$ and 0.7. After detecting the proper parameter values, we further measure the performance of these three models, *i.e.* the proposed technique with three different kernels: Gaussian ($\gamma = 0.5$), polynomial ($\gamma = 0.5$) and financial. The results are shown as Table 4. First, the Gaussian kernel beats the financial kernel. It shows higher accuracy ($84.81\% > 83.54\%$), higher precision ($50\% > 33.33\%$), higher recall

($16.67\% > 8.33\%$), and higher F_1 ($25\% > 13.33\%$). Second, although the recall values are low, the Gaussian kernel's precision is 50%, much higher than the polynomial kernel's precision (23.53%). The result further indicates that the Gaussian kernel still has a 50% opportunity to correctly identify the M&A cases, while most predictions made with the polynomial kernel are incorrect (23.53%). For the above two reasons, the Gaussian kernel is much more valuable than the other two and used for further evaluation.

Table 2: Tuning results of the proposed prediction model with the Gaussian kernel

Gaussian Kernel	Accuracy	Precision	Recall	F_1
$\gamma = .3$	83.54%	33.33%	8.33%	13.33%
$\gamma = .5$	84.81%	50.00%	16.67%	25.00%
$\gamma = .7$	84.81%	50.00%	8.33%	14.29%

Table 3: Tuning results of the proposed prediction model with the polynomial kernel

Polynomial Kernel	Accuracy	Precision	Recall	F_1
$\gamma = .3$	70.89%	21.05%	33.33%	25.81%
$\gamma = .5$	73.41%	23.53%	33.33%	27.59%
$\gamma = .7$	73.41%	23.53%	33.33%	27.59%

Table 4: Evaluation of different kernels

Kernel	Accuracy	Precision	Recall	F_1
Gaussian	84.81%	50.00%	16.67%	25.00%
Polynomial	73.41%	23.53%	33.33%	27.59%
Financial	83.54%	33.33%	8.33%	13.33%

Then we report the results of the test sample using the SVM-Gaussian methodology. First, we separate the dataset, ranging from 2000 to 2011, into two sets (train and test). The first-n-year dataset are treated as training samples and the remaining $12 - n$ are treated as testing samples. For example, the ratio

“$7:5$” implies that the training samples contain the dataset of Year 2000-2006, while the rest (2007-2011) are used for test. Table 5 shows that at the ratio $9:3$, the SVM-Gaussian reaches the best results. Second, we further consider how the lagged effect works. We use the same variables, expressed in Section 4.1, to validate the prediction technique. Then the lagged variables are removed, and we re-train the model using the updated variables. As Table 6 shows, the model with the lagged variables excels. The predictions to M&A cases are time-sensitive, and the lagged variables help capture the features of the year-over-year financial changes.

Table 5: SVM-Gaussian results on the test set

Weightings (Train:Test)	Accuracy	Precision	Recall	F_1
6:6	42.86%	10.71%	17.14%	86.51%
7:5	66.67%	8.00%	14.29%	85.19%
8:4	50.00%	11.11%	18.18%	84.35%
9:3	50.00%	16.67%	25.00%	84.81%

Table 6: Evaluation of the lagged effect

Year	Accuracy	Precision	Recall	F_1
Lagged Effect	84.81%	50.00%	16.67%	25.00%
w/o Lagged Effect	71.08%	12.50%	16.67%	14.29%

Finally, we measure the effectiveness of the proposed technique by comparing with the benchmark. The model using the proposed technique is SVM integrated with the Gaussian kernel. Logit regression is a popular and classical methodology, and prior studies employ the logit model to make the predictions, including Meador et al. (1996); Barnes (2000); Ali-Yrkkö et al. (2005); Pasiouras & Gaganis (2007). Table 7 indicates that the SVM-Gaussian model has higher precision (50%) and higher F_1 (25%) even though they have the similar recall value. Furthermore, the SVM-Gaussian model also excel in the term “accuracy” (84.81%). Hence, we can conclude that the proposed SVM model is better than the benchmark.

Table 7: Evaluation results (SVM-Gaussian versus technological-based benchmark

Model	Accuracy	Precision	Recall	F_1
SVM-Gaussian	84.81%	50.00%	16.67%	25.00%
Benchmark (Logit Regression)	79.75%	25.00%	16.67%	20.00%

5. Conclusion

M&A is a very common business activity and happens frequently in the high-technology industries because these IT companies are motivated for the speedy innovation and required to extend their resources and capabilities through the M&A transaction. In this study, we aim to provide a method that automatically determine a feasible M&A deal. Based on the SVM model, we extend the work of Ke (2017) and propose the M&A forecasting technique. We also derive 40 financial variables and develop a training and forecasting method for the technique. The M&A cases in the U.S. market are collected to validate the effectiveness of the proposed model.

The study also indicates some interesting findings. First, the empirical evaluation shows that our proposed technique is more effective than the benchmark, *i.e.* the logit model. Using SVM derives a benefit of a non-linear classification, especially when integrated with the kernels. Second, three different kernels are compared, but the financial kernel is not as superior as we expected. Originally, we expected that the year-over-year effects are able to be captured by the financial kernel, and subsequently the technique combined with the financial kernel exhibits better performance. However, the financial kernel's performance is not as superior as that of the Gaussian kernel. It's probably because the combination of variables when constructing a financial kernel offsets the carryover effect. Third, we also find that the lagged variables take effect. It implies that the variables will have influence on its future values. Finally, the proposed technique is validated by the U.S. market, which is the largest and the typical market in the world. Hence, we believe that it is workable to make the M&A predictions using the proposed technique.

However, the study is an initial exploration of the application of SVM in M&A forecasting. Our works still contains some limitations and may need to be improved in the future. First, We plan to show the performance of our proposed technique by incorporating other benchmark models, such as decision tree (C4.5), neural networks, or other machine learning techniques. Second, our proposed technique only focus on the prediction of M&A targets and does not consider the possible negative outcomes, e.g declining market shares and profits. Third, the study only focus on North America and only collect M&A cases from there. However, we need to generalize the proposed technique although most M&A deals are made in North America. We plan to collect the M&A cases from other markets, including Europe and Asia Pacific to verify the generalization of our proposed technique.

References

Ali-Yrkkö, J., Hyytinen, A., & Pajarinen, M. (2005). Does patenting increase the probability of being acquired? Evidence from cross-border and domestic acquisitions. *Applied Financial Economics*, *15*(14), 1007–1017.

An, S., He, Y., Zhao, Z., & Sun, J. (2006). Measurement of merger and acquisition performance based on artificial neural network. In *2006 5th IEEE International Conference on Cognitive Informatics*, vol. 1, (pp. 502–506). IEEE.

Barnes, P. (2000). The identification of UK takeover targets using published historical cost accounting data Some empirical evidence comparing logit with linear discriminant analysis and raw financial ratios with industry-relative ratios. *International Review of Financial Analysis*, *9*(2), 147–162.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, (pp. 144–152). ACM.

Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Detecting management fraud in public companies. *Management Science*, *56*(7), 1146–1160.

Chapelle, O., Haffner, P., & Vapnik, V. N. (1999). Support vector machines for histogram-based image classification. *IEEE transactions on Neural Networks*, *10*(5), 1055–1064.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273–297.

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge university press.

Cristianini, N., Shawe-Taylor, J., & Lodhi, H. (2002). Latent semantic kernels. *Journal of Intelligent Information Systems*, *18*(2-3), 127–152.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., & Weingessel, A. (2008). Misc functions of the Department of Statistics (e1071), TU Wien. *R package*, *1*, 5–24.

Hagedoorn, J., & Duysters, G. (2002). The effect of mergers and acquisitions on the technological performance of companies in a high-tech environment. *Technology Analysis & Strategic Management*, *14*(1), 67–85.

Hornik, K., Meyer, D., & Karatzoglou, A. (2006). Support vector machines in R. *Journal of statistical software*, *15*(9), 1–28.

Hu, X., & Pan, Y. (2007). *Knowledge Discovery in Bioinformatics: Techniques, Methods, and Applications*, vol. 5. John Wiley & Sons.

Hua, S., & Sun, Z. (2001). A novel method of protein secondary structure prediction with high segment overlap measure: Support vector machine approach. *Journal of molecular biology*, *308*(2), 397–407.

Ke, Y.-S. (2017). *Forecast the M&A Targets in IT Industry Using Support Vector Machine*. Master Thesis, National Sun Yat-sen University.

King, D. R., Slotegraaf, R. J., & Kesner, I. (2008). Performance implications of firm resource interactions in the acquisition of R&D-intensive firms. *Organization Science*, *19*(2), 327–340.

Meador, A. L., Church, P. H., & Rayburn, L. G. (1996). Development of prediction models for horizontal and vertical mergers. *Journal of financial and strategic decisions*, *9*(1), 11–23.

Pasiouras, F., & Gaganis, C. (2007). Financial characteristics of banks involved in acquisitions: Evidence from Asia. *Applied Financial Economics*, *17*(4), 329–341.

Ragothaman, S., Naik, B., & Ramakrishnan, K. (2003). Predicting corporate acquisitions: An application of uncertain reasoning using rule induction. *Information Systems Frontiers*, *5*(4), 401–412.

Song, X.-L., Zhang, Q.-S., Chu, Y.-H., & Song, E.-Z. (2009). A study on financial strategy for determining the target enterprise of merger and acquisition. In *Service Operations, Logistics and Informatics, 2009. SOLI'09. IEEE/INFORMS International Conference On*, (pp. 477–480). IEEE.

Tsagkanos, A., Georgopoulos, A., & Siriopoulos, C. (2007). Predicting Greek mergers and acquisitions: A new approach. *International Journal of Financial Services Management*, *2*(4), 289–303.

Yang, C.-S., Wei, C.-P., & Chiang, Y.-H. (2014). Exploiting technological indicators for effective technology merger and acquisition (M&A) predictions. *Decision Sciences*, *45*(1), 147–174.

6. Appendix

Table 8: List of financial indicators

Indicator	Reference
Dividend and 3-year average dividend (DVT and DVT3)	Barnes (2000)
Capital-expenditures-to-total-asset ratio (CETA)	Barnes (2000); Ragothaman et al. (2003)
Cash flow (CF)	Ragothaman et al. (2003); Ali-Yrkkö et al. (2005); Song et al. (2009)
Common shares traded divided by common shares outstanding (CSTRCSHO)	Meador et al. (1996)
Cost of goods sold (COGS)	Meador et al. (1996)
Cost of goods sold divided by average inventory (COGSNI)	Meador et al. (1996)
Current ratio (CURRENTRATIO)	Meador et al. (1996); Barnes (2000); Ragothaman et al. (2003); Tsagkanos et al. (2007)
Debt-to-assets ratio (DEBTTOASSETS)	Barnes (2000); Pasiouras & Gaganis (2007)
Debt-to-equity ratio (DEBTTOEQUITY)	Meador et al. (1996); Ragothaman et al. (2003); Song et al. (2009)
Earnings before interest and taxes or operating income after depreciation (EBIT)	Meador et al. (1996)
Market-to-book-value ratio or Tobin's Q (Q)	Meador et al. (1996); Barnes (2000); Ragothaman et al. (2003); Song et al. (2009)
Price-to-earning ratio (PE)	Meador et al. (1996); Barnes (2000); Ragothaman et al. (2003); Song et al. (2009)
Profit margin (PROFITMAT)	Tsagkanos et al. (2007)
Ratio of tangible (fixed) assets to total assets (TANGIBLEAT)	Ali-Yrkkö et al. (2005)
Return on assets (ROA)	Meador et al. (1996); Pasiouras & Gaganis (2007)

Return on equity (ROE)	Meador et al. (1996); Barnes (2000); Tsagkanos et al. (2007)
Return on investment (ROI)	Ali-Yrkkö et al. (2005)
Sales to total assets or asset turnover (ASSET-TURNOVER)	Meador et al. (1996); Barnes (2000); Tsagkanos et al. (2007)
Tax shield effects (TAXSHIELD)	Song et al. (2009)