Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Identifying the potential for Failure of Businesses in the Technology, Pharmaceutical and Banking Sectors using Kernel-based Machine Learning Methods

Yashodhan Athavale
Department of Electrical and Computer Engineering
Ryerson University
Toronto, Canada
yashodhan.athavale@ryerson.ca

Pouyan Hosseinizadeh
Department of Mechanical and Industrial Engineering
Ryerson University
Toronto, Canada
pouyan.hosseinizadeh@ryerson.ca

Sridhar Krishnan
Department of Electrical and Computer Engineering
Ryerson University
Toronto, Canada
krishnan@ee.ryerson.ca

Aziz Guergachi
Ted Rogers School of Management
Ryerson University
Toronto, Canada
a2guerga@ryerson.ca

***Abstract*— The objective of this paper is to analyze the performance of a kernel-based method in identifying the potential for collapse (or survival) of a firm operating in three different sectors of the economy – Technology, Pharmaceutical and Banking. The analysis uses the actual stock market data, collected on a weekly basis in a common time-series interval for the active and dead companies in each of the three sectors. The basic idea is to apply the concept of Fisher kernels and visualization to reduce the data from a time-series format to two-dimensional plots that can be visually inspected and potentially segregate the 'collapse' class from the 'survival' one. From our experiments we observe that our method fits well for the Technology and Banking sectors, but is not able to provide a visually clear classification for the Pharmaceuticals sector. Depending on the range of data we use as input, and its distribution, the classification pattern varies from an ideally separable case to a non separable one, in a two dimensional feature space.**

***Keywords*— Fisher kernels, financial time series, Gaussian probability model, collapse versus survival**

I. Introduction

The stock market, be it any one in this world, is probably one of the few places, where we can actually observe a random behavior. There are various factors which affect the stock market behavior such as political events and decisions, influx-deflux of businesses, and also the 'market emotions' like rumors in the stock market floor. Consequently it's very difficult or even impossible to predict accurately what will be the actual value of a stock price in a week, a month or a year.

The challenge of making accurate and useful predictions has been discussed by Taleb [1] in his recent book "The Black Swan" that showed a great deal of contempt to finance academics and economists. According to Taleb [1], the Black Swan refers to a large-impact unpredictable event, which is beyond the realm of normal expectations.

This paper does not intend to predict the value of a stock price, but rather to identify the possibility of collapse of a firm, based on historical stock price data processed over a certain time window. The question that this paper proposes to address is thus: In the long term, will this company survive or collapse?

In recent times, trade analysts have come to take the assistance of several intelligent systems to make the stock trading decisions on the market floor. Many of the intelligent systems have been automated to ease the tasks of the end-user, and perform the necessary complex calculations, in order to give good forecasts. To facilitate the separation of the 'collapse' and 'survival' classes, this paper makes use of a set of methods from the area of machine learning that were successful in several applications: kernel-based methods and, more specifically, the implementation of Fisher kernel for processing the stock market time series and reducing them into two-dimensional plots that segregate the 'collapse' and 'survival' classes. Fisher kernels have been extensively used in the areas of speech recognition [2], protein homology detection [3], web audio classification [4], object recognition [5], image decoding [6] [7] and other applications such as PLSA analysis.

Most of the intelligent systems used for analyzing the dynamics of the stock market, have been mostly applied in predicting stock prices, credit risk and bankruptcy. But in this study, we are developing a method to monitor the dynamics of an economic sector by constructing a classification algorithm to predict the companies that will remain active in this sector from the ones that will "collapse". An incident of Bankruptcy is one of the primary reasons leading to the collapse of firms in an economic sector, but other incidents such as mergers and acquisitions are equally possible. Significant amount of work has also been done in the field of bankruptcy prediction through the application of intelligent systems such as Neural Networks [8] [9], Genetic algorithms [8] [9], Support Vector Machines [10], Self-organizing maps [11] and other machine learning models [12] [13] [14]. These works however make little or no use of stock market data. They instead use various financial ratios available from the company's income statements. The use of financial ratios is beneficial in the sense that, more information is available, for a particular company in order to predict its possible bankruptcy. The challenge, however, is that these financial ratios may not always be readily available, and thus the monitoring of the dynamics of economic sectors in the world's countries may not be fully automated. This is why the study presented herein focuses on the use of stock prices as main indicators for assessing the overall health of companies and making judgments about the corresponding economic sectors in various countries.

A major significance of the stock price is that, it is directly related to the earnings and dividends of the company. Usually it is observed that the higher the stock price of a particular firm, the better is its performance, and hence more profitability with higher returns. If a company is profitable, it means that the stocks for that company are likely to fetch higher returns in the market, thus increasing the net income for the stakeholder as well as the company. The more profit the company makes, the more dividends it would pay out to the stakeholders, and the more easily it would be able to pay off its liabilities. This consequently leads to lesser chances of bankruptcy or a total collapse of the company. So instead of looking at the individual company performance & data, we are analyzing the stock price dynamics over an entire economic sector. As discussed before, stock prices are affected by many socio-political factors and others such as factor of gambling, fake information etc. Since we are using weekly closing stock prices here, all the factors affecting the market in a particular week inherently become a part of the closing stock price. Hence we need not explicitly consider the above factors for analyzing a stock at this stage of research.

II. FINANCIAL APPLICATION USING FISHER KERNELS

In this paper, we have analyzed the behavior of the major stock markets in the North American region, namely, the Toronto Stock Exchange (TSX), NASDAQ and the New York Stock Exchange (NYSE).

The process of collecting stock market data has been done with the help Datastream software. Using this software we collected the weekly stock prices for various active and dead companies, from the Technology, Pharmaceuticals and Banking sectors, in a common time frame of about 2 years. This data is further read processed using Fisher kernels, thus giving us the Fisher Scores.

III. GENERAL DESCRIPTION OF THE CONCEPTS AND MODELS USED IN THE STUDY

In this study, we are dealing with a binary classification of business enterprises, sorting them into 'collapse' and 'survival' categories. Usually, whenever we do a binary classification, we tend to separate the two categories of data using a hyperplane [2]. In this study, the classification involves *clustering* of data points in the Fisher scores space, which is obtained by applying the Fisher kernel method. However, the two classes involved in the classification (active/dead or survival/collapse) have been defined from the start, while collecting the data. Thus, the learning method that is used in this study is not completely unsupervised, as the data is already labeled.

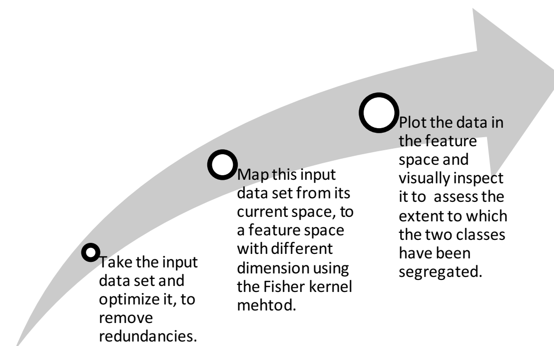

Figure 1 Step-wise procedure for classification

Fisher kernel [3] is among the most versatile kernels in pattern classification. They are known for their wide range of applications. In this study, we propose to use it to analyze financial time series and infer statements about the risk of collapse of a publicly-traded firm.

Figure 1 depicts the general steps that make up the procedure used in this study. The following describes these steps in more details.

- In this study we have two classes of data, namely: active companies, and dead companies.
- The weekly stock price data has been collected in a common time interval for both the active and dead companies.
- We organize this data into two different matrices C_1 and C_2, corresponding to the active and dead categories respectively.

$$\begin{bmatrix} x_{11} & \cdots & x_{1j} \\ \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ij} \end{bmatrix}$$

  where, x_{ij} is the stock price of the firm i in week j. Hence we have i number of rows and j number of columns for each class.
- Next, we calculate the mean and variance for the weekly stock prices of each active and dead company.
- These mean & variances are then inputted into the Gaussian probabilistic model:

$$S(x_{ij}) = \frac{1}{\sigma_j\sqrt{2\pi}}\, exp\left(-\frac{1}{2}\left(\frac{x_{ij}-\mu_j}{\sigma_j}\right)^2\right)$$

- Therefore for each company's time series, we get,

$$G_\theta(X_i) = \prod_{j=1}^{k} S(x_{ij})$$

where k is the number of weeks over which the data has been collected, G_θ is the generative model, θ is the model parameters, and X_i is the financial time series. In this model, we are assuming that the current and past stock prices are independent of each other.

- Using the score-operator ∇_θ, as a vector-gradient of the log-likelihood expression $\log(G_\theta(X_i))$, we calculate the Fisher scores as:

$$g(\theta, X_i) = \nabla_\theta \log(G_\theta(X_i))$$

- The Fisher score equation can be expanded as:

$$g(\theta, X_i) = \left(\frac{\partial \log(G_\theta(X_i))}{\partial \mu}, \frac{\partial \log(G_\theta(X_i))}{\partial \sigma}\right)$$

$$or\ , g(\theta, X_i) = (F_1, F_2)$$

$$where, F_1 = \left(\frac{\partial \log(G_\theta(X_i))}{\partial \mu}\right)$$

$$and, F_2 = \left(\frac{\partial \log(G_\theta(X_i))}{\partial \sigma}\right)$$

- The abscissa and ordinate are actually the summation of all the Fisher Scores F_1 and F_2 and are given as:

$$SF_1 = \sum_{j=1}^{k}\left(\frac{x_{ij} - \mu_j}{\sigma_j^2}\right)$$

$$SF_2 = \sum_{j=1}^{k}\left(-\frac{1}{\sigma_j} + \frac{(x_{ij} - \mu_j)^2}{\sigma_j^3}\right)$$

Consider that the total number of active & dead companies is N in a given sector. So we get a matrix of size Nx2. Consequently for every company in each class of data, we get two unique values F_1 and F_2, which are then plotted against each other, as shown in Figures 5, 6 and 7.

IV. Experiments & Results

We consider two to four years of stock price data for various active and dead companies falling under technology, pharmaceuticals and banking sectors, in the North American stock market. The various stock exchanges we have considered are Toronto Stock Exchange (TSX), NASDAQ and New York Stock Exchange (NYSE).

Figures 2, 3 and 4 illustrate the stock price charts for active and dead businesses in the sectors under study. It is evident from these stock price charts that not only is it difficult to predict the future stock price for the active companies , but it is also hard to visually distinguish the companies that will likely to fail from those that will survive by looking at the stock price time-series. Some of the trends below show a constant value after a particular week. Such a constant indicates that the company is dead, i.e. it no longer exists as a standalone entity in the market. In this context the term 'dead' refers not only to the companies that have shut down due to bankruptcy or other reasons, but also those firms which have merged or merged with others.

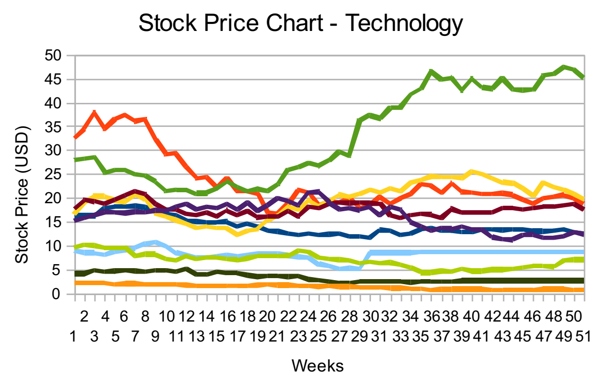

Figure 2

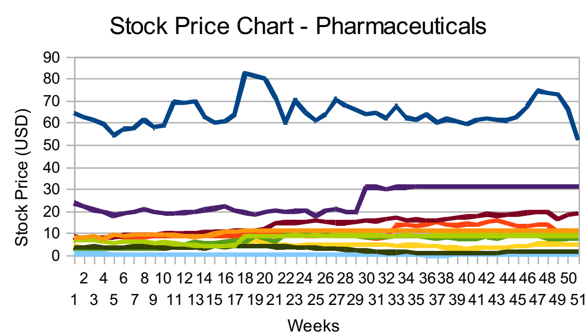

Figure 3

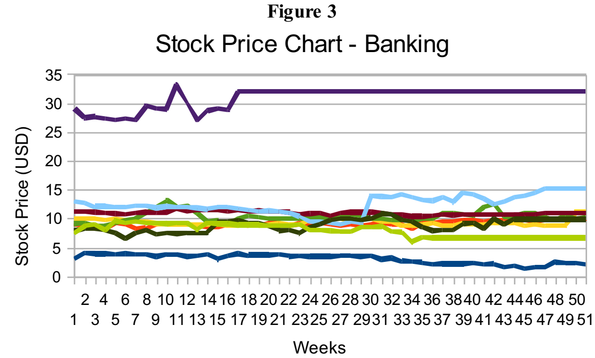

Figure 4

The application of Fisher kernel to the stock price data generates the respective Fisher scores. These Fisher scores (F_1 and F_2) are actually the values of the gradients of the stock price probability model with respect to the mean and variance of this model. The plots are shown in Figures 6, 7 and 8 for technology, pharmaceuticals and banks respectively.

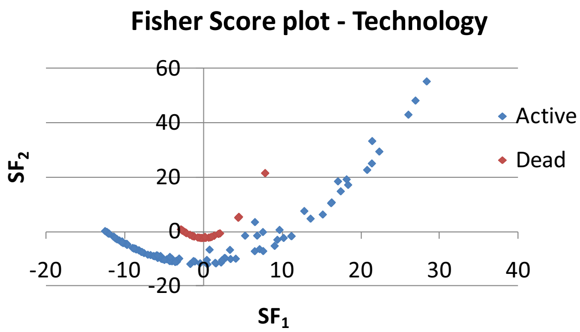

Figure 5

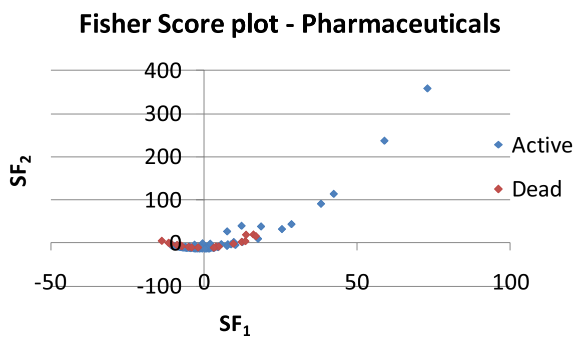

Figure 6

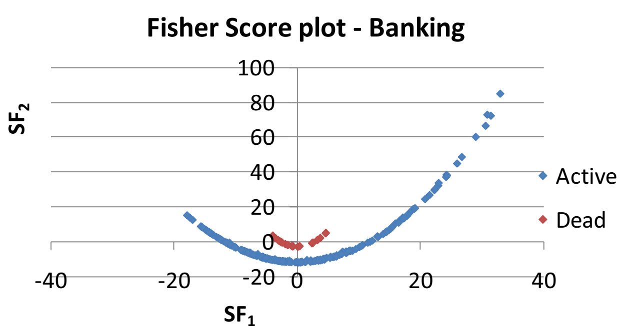

Figure 7

From plots shown (fig. 5, 6 and 7), the following observations can be made:

- ✓ Fisher Score plot for Banks in Figure 7 shows an ideal classification case, wherein there is a clear demarcation between the active and dead companies.
- ✓ The plot for Pharmaceuticals in Figure 6 indicates that the active and dead companies cannot be easily separated using a separating hyperplane.
- ✓ In Figure 5, for the Technology sector, we can see a case where separation is possible, even though the data points for each series do not represent a perfect parabolic curve; we can still separate the two of classes and make a judgment about the risk of collapse.

V. Discussions & Future Works

The Fisher kernel was successful in classifying data in the Technology sector, whereas it could not classify the Pharmaceutical industry data. This is quite an interesting situation, where we can't make any concrete statements as to why the classification fails for certain sectors of the economy and work for others. But certain observations can certainly be made from the method used and the stock price charts above (Figures 2, 3 and 4).

We can make some observations from the stock price charts (Figures 2, 3 and 4), and the corresponding Fisher score plots (Figures 5, 6 and 7). The stock price chart for technology shows some sort of even distribution of data, as compared to those of pharmaceuticals and banking. The stock data in the pharmaceutical sector tends to concentrate mostly between 0 to 15 USD; whereas in the Banking sector, the data concentrated more between 4 to 15 USD. Probably because of the spread out distribution in the technology sector, we obtained, in this sector, a clear classification of active and dead companies. In the pharmaceuticals sector, however, the stock price data for active and dead companies tend to overlap each other.

Paradoxically, the stock price data for active and dead companies also overlap each other in the banking sector; but in this case we manage to obtain a clear separation between the two classes of companies. Such a situation leads us to the following question: How does the data distribution affect the Fisher score plots? In other words, what is remarkable and different about the data in the technology sector, as compared to the pharmaceutical sector, which contributed to a clear classification? It becomes quite evident that the data distribution in the feature space not only depends on the kernel selected, but there also exists some dependency between any two consecutive stock prices, which, if explored might lead to better results, for a wide range of data in a large number of economic sectors.

The data used for analysis in this experiment is a fixed-length time series. Experimenting with variable-length time series, might be able to give us more insight in to the performance of each company. Our further work would be targeted towards the analysis of the performance of Fisher kernels implemented using the Gaussian Mixture Model (GMM) [15]. We can further also implement the GMM as a kernel in Support Vector Machines, for better results, and a faster learning machine.

Experiments have already been performed for market predictions in bankruptcy, future stock prices and credit risks using other machine learning tools such as neural networks, fuzzy logic, SOMs. Our future research intends to look at how the Fisher kernel method combined with SVMs performs, as compared to other machine learning algorithms.

Also, the data used in the above study are the weekly closing prices for each sector. Using daily closing prices might improve the results, especially for the pharmaceuticals sector.

VI. CONCLUSIONS

In this paper, we analyzed the performance of a Fisher kernel-based system, using a Gaussian probabilistic model, to classify the stock price data into active and dead categories. In order to implement the model, real stock market data was used from three different sectors. The method was effective in classifying the companies in the technology and banking sector, but was unable to classify the same in the pharmaceutical sector.

ACKNOWLEDGMENT

Sincere thanks to NSERC Granting Council, CFI/OIT and Canada Research Chairs program for funding the research.

REFERENCES

[1] Taleb, N., The Black Swan: The impact of the highly probable, Random House, 2007.

[2] Vincent Wan and Steve Renals, Speaker Verification Using Sequence Discriminant Support Vector Machines, IEEE Transactions on Audio & Speech Processing , Vol.13,No.2, March 2005

[3] T. Jaakkola, M. Diekhams, and D. Haussler, Using the Fisher kernel method to detect remote protein homologies, Proceedings of the Seventh International Conference on Intelligent Systems for Molecular Biology, 1999.

[4] P. J. Moreno and R. Rifkin*, Using the Fisher Kernel method for Web Audio Classification

[5] A. D.Holub, M. Welling and P. Perona, Combining Generative Models and Fisher Kernels for Object Recognition, Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05).

[6] Jingdong (JD) Chen and Y. Wang, Exploiting Fisher Kernels in Decoding Severly Noisy Document Images, Ninth International Conference on Document Analysis and Recognition (ICDAR 2007).

[7] F.Perronnin and C. Dance, Fisher Kernels on Visual Vocabularies for Image Categorization, Xerox Research Center Europe.

[8] Y. Sai; C. Zhong; P. Nie; A Hybrid RST and GA-BP Model for Chinese Listed Company Bankruptcy Prediction Natural Computation, 2007. ICNC 2007. Third International Conference on; Volume 2, 24-27 Aug. 2007 Page(s):513 – 517

[9] H. Fu-yuan; A Genetic Fuzzy Neural Network for Bankruptcy Prediction in Chinese Corporations , Risk Management & Engineering Management, 2008. ICRMEM '08. International Conference on 4-6 Nov. 2008 Page(s):542 – 546

[10] Z. Gao; M. Cui; Lai-Man Po; Enterprise Bankruptcy Prediction Using Noisy-Tolerant Support Vector Machine, Future Information Technology and Management Engineering, 2008. FITME '08. International Seminar on; 20-20 Nov. 2008 Page(s):153 – 156

[11] Nakaoka, I.; Tani, K.; Hoshino, Y.; Kamei, K.; A Bankruptcy Prediction Method Based on Cash flow Using SOM, Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on, Volume 4, 8-11 Oct. 2006 Page(s):2790 – 2795

[12] Yeung, D.S.; Ng, W.W.Y.; Chan, A.P.F.; Chan, P.P.K.; Firth, M.; Tsang, Bankruptcy Prediction Using Multiple Intelligent Agent System via a Localized Generalization Error Approach , E.C.C.;Service Systems and Service Management, 2007 International Conference on; 9-11 June 2007 Page(s):1 – 6

[13] Chan, A.P.F.; Ng, W.W.Y.; Yeung, D.S.; Tsang, E.C.C.; Firth, M.; Bankruptcy Prediction Using Multiple Classifier System with Mutual Information Feature Grouping ,Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on Volume 1, 8-11 Oct. 2006 Page(s):845 – 850

[14] Tang Baojun; Qiu Wanhua; Sun Xing; Business Failure Prediction Using Exponential Smoothing Forecasting and Pattern Recognition ,Risk Management & Engineering Management, 2008. ICRMEM '08. International Conference on 4-6 Nov. 2008 Page(s):576 – 581

[15] R. O.Duda, P. E.Hart and David G.Stork, Pattern Classification, © John Wiley & Sons 2001

[16] B.W. Broreson and S. Yang, Nonlinear dynamics and the distribution of daily stock index returns, The Journal of Financial research, 1994

[17] Vapnik N, The Nature of Statistical Learning Theory, 2000.

[18] Jaakola T. and Haussler D., Exploting generative models in discriminative classifiers, NIPS, 1998

[19] K. Tsuda, S. Akaho, M. Kawanabe, Klaus-Robert Müller, Asymptotic Properties of the Fisher Kernel, Letter Communicated by John Platt, Neural Computation 16, 115-137 (2004), © Massachusetts Institute of Technology.