

STATISTICAL MODELLING USING DATA MINING TOOLS IN MERGERS AND ACQUISITION WITH REGARDS TO MANUFACTURE & SERVICE SECTOR

S. KALAIVANI*, K. SIVAKUMAR, J. VIJAYARANGAM

ABSTRACT. Many organizations seek statistical modelling facilitated by data analytics technologies for determining the prediction models associated with M&A (Merger and Acquisition). By combining these data analytics tool alongside with data collection approaches aids organizations towards M&A decision making, followed by achieving profitable insights as well. It promotes for better visibility, overall improvements and effective negotiation strategies for post-M&A integration. This paper explores on the impact of pre and post integration of M&A in a standard organizational setting via devising a suitable statistical model via employing techniques such as Naïve Bayes, K-nearest neighbour (KNN), and Decision Tree & Support Vector Machine (SVM).

AMS Mathematics Subject Classification : 65D30, 65D32.

Key words and phrases : Mergers & Acquisition, data analytics, manufacturing sector, service sector.

1. Introduction

Mergers and Acquisitions have become a great choice for the companies around the world to expand its business growth. M & A transactions are creating large interests in the recent years, which have been proved by many business organizations all over the world. This M & A have brought a solution for new markets when two companies with different activities are merged forming a business partner. The growth of web technologies in the recent years supports a huge volume of materials easily in just a click through browsing or search engines. The financial statement of Romanian companies proved that the online availability of that statements eases the analysis of competition and to predict it moves. M & A can be used to combine the transactions of two different companies to

Received December 13, 2021. Revised February 22, 2022. Accepted March 3, 2022.

*Corresponding author.

© 2022 KSCAM.

expand the growth of that companies. In merger, similar size companies can be merged to form a new market. In acquisition, one company can take over the another to expand its business process. The integration of M & A is difficult due to the reasons such as inappropriate integration, inadequate rationale, and lack of pre and post-merger and acquisitions. To enhance the performance of M & A, integration of M & A should be considered. According to the survey of Mergemarket(2017), about 25% of the people responded that the integration of M & A can be much easier than realized. In this paper, we explore the effect of M & A utilizing the computing technologies such as KNN, SVM, Naïve Bayes, and Decision tree

2. Related work

In today's economic scenario, the establishments of various business Organizations aimed to achieve certain key corporate objectives that comprises of corporate growth followed by an increased growth rate, which indicated via gain ability. Growth is attributed as a major yardstick via which, an organizational success in terms of their business firm can be monitored measured. If the business organizations tend to operate under dynamic macroeconomic spheres namely growth has been threatened over the periods of volatility associated with economic instabilities (Weston & Copeland, 1989). From the resulting effects on recent global economic meltdown associated with financial crisis in corporate organizations. One strategy that opens up with corporate organizations under such critical scenario of economic crisis involves the influence of M & A. Companies are actively combining numerous configurations, ever since the advent of business. Nevertheless, by conglomerating two organizations requires a complex process since it comprises of every other aspect for both organizations (Indhumathi, Selvam, & Babu, 2011). One of the popular forms in business combination is M & A consolidations. In its infancy stages with special regards to India, M & A is often compared with other developed nations. It acts as a major financial indicator and is quite significant in terms of its concept that collectively attributes to the national growth in terms of economy via substantial growth in the organizational profitability and productivity. As it is quite apparent over the fact that M& A is significant, and there is a greater need for testifying their effectiveness via imparting data analytics over the number of transactions that are carried out in recent times (Li, Mai, Shen, & Yan, 2020). Also, there is an emerging utilization involving data analytics for actualization of acquisition, as organizations are seeking for new data resources for feeding their analytical based capabilities. These associations generally prefer to maintain and work with traditional M & A methodologies rather than using analytical development. Given the perceived risks, these companies need to use richer and detailed datasets to drive and enhance their analysis. As if the exchanges were made at a particular level and these developments were not available, the number of purchases lost current judgment due to risks and issues that may have been

mitigated by a more detailed analysis. Increase. There is an increased risk of resolving this unacceptable option. Therefore, at this point, it is clear that it is important to integrate the data analysis into the existing hierarchy. Mergers and acquisitions, including big data rooms, different valuation assumptions between buyers and traders, and large buyers who need to adjust their targets shortly after closing, are cutting-edge thinking and data control innovations can benefit from the contributions (Gole & Morris, 2007). Jimjang et.al 2019, proposed a deep learning framework for smart manufacturing. He summarized the emerging trends in manufacturing technologies. He also presented the advantages of deep learning over conventional machine learning algorithms. Even though, the work highlighted the importance of deep learning framework. The challenges in computational complexity were not well addressed. Amish 2016, discussed a wealth analysis of share holder in minimum period investment. He identified top 10 merger and acquisition of companies in India. The work used regression line to identify the returns of the companies. However, the work did not discuss about the long term stability of finance in the market. Today, few organizations implement advanced analytics in a merger and acquisition, so they do not comprehensively describe the benefits, such as a case library that show complex best practices. Obviously, during the engagement phase, most associations apply analysis to investigate potential business issues that may arise during a merger or acquisition. focus on. Most of them were not because some of the correlations from the analysis went through a merger. Focus on four exercises, especially if it has a merge of high-level analyses. First, it is engaged in deciding whether it can support the union of organizations. Affect revenue and cost collaboration, generate insights, and accelerate ideal opportunities to increase resource adequacy. Therefore, presenting a high level of analysis as an important possibility to act in each of the four regions during integration accelerates the ideal opportunity for influence and circulation. It has been suggested that it can grow extensively (Long, Song, & Cui, 2017). Many scholars from finance discipline evaluated pre M & A in terms of pricing, banking, and transactions to generate shareholder value (Sudharsanam, 2012). The performance of pre M & A were evaluated based on abnormal returns with the help of stock market measures. (Brickley, Jarell, 1988; Jensen & Ruback, 1983; Weston & Chung, 1983). While stock market based measures show better performance, accounting based measure indicate negative performance. (King 2004, Steigner, 2011) This review was particularly sensitive because it introduced some of the most popular machine learning tools and strict limits on the type and depth of data that organized deals can share. In addition, many major tests require creating an optimal suite of antitrust laws, including laws that include industry sensitive data such as ratings and acquired data. The study generated a dataset of 58 observed collection and control regions.

3. Problem definition

The main purpose of the work is to investigate the performance of M & A in pre and post integration process. The work uses the data set of 58 companies to investigate the long-term performance of the acquired companies. Since long term is required to analyse the performance of M & A, the study uses the data set of M & A during the period of 2010 to 2018. The analysis of data set proved that the profit of the companies was increased efficiently after involving integration of M & A. The study also observed the improved operating performance of the companies involved in M & A. In this paper, the author focuses on classifying subjects as non-demented or demented using various AI procedures in data mining. The main contributions of the work are as follows: 1. Create records using machine learning procedures that include radial SVMs, radial SVMs, logistic regression, KNNs, decision trees, random forests. 2. Different parameters comparisons 3. Learn about the best classification procedure, depending on the results

4. Architectural Framework

Figure 1 shows a functional flowchart of this paper. The first step is to collect the details of the topic. Then, preparations are made following this. Data preparation is an important issue for data mining, as current actual data is generally inadequate, noisy, and inconsistent

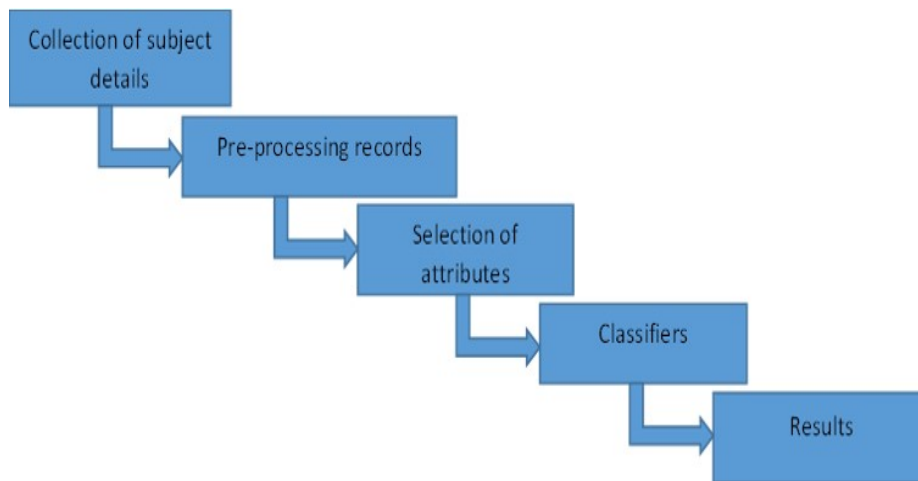


FIGURE 1. Block diagram indicating the working flow

The data collected here is primary data and will not be lost. Therefore, there is no need to pre-process the data. The next step is to select the characteristic. However, there are credits and a certain percentage of scores that have a

significant impact on dynamism. The next step is classification. Sorting is done to see how accurately the data is arranged. The Classification tab shows an overview of your AI account. These usually calculate valuations and are performed frequently. These can control calculation limits and information data weights to improve workbook accuracy. WEKA instruments are used for analytical purposes. Two learning implementation evaluators have been integrated with WEKA. The first is a classifier that divides the dataset into ready-made data and test data. The second is mutual recognition using creases. Ratings are usually drawn accurately.

Below is the collection of datasets.

The dataset contained 58 companies classified into service sector and manufacturing. These datasets have been edited from reliable data accessible from the 2010-2018 financial year annual report. Of these, nine important properties were investigated by M and A.

Pre-processing

Pre-processing is the stage of checking for absence and wrong values. There is no pre-processed data here as there is no risk of data loss.

Selection of Attributes

The dataset contains nine ratios that treat the principal components associated with M and A as markers for the overall implementation of the individual associations incorporated into the review.

Linear SVM

Computed regression is a measurable strategy for double layer prediction. Output or target variables are decomposed. Disconnecting means that there are only two possible classes. This is an exceptional example of direct regression where the objective variable is inherently simple. Opportunity records are used as dependent variables. Strategic regression uses rational functions to predict the probabilities of two events.

SVM radial

This technique includes radial basis function fractions (RBF) and C-related sigma optimization. For example, this strategy, applied by (Davenport, 2013) improved the SVM model for candidate prediction. RBF parts are commonly used for data fit adaptation, but other common parts such as polynomial bits and x bits are also applied and scaled along these lines. Since SVM relies on calculations using data estimation to plot each setup model in a high dimensional space, the actual data is an important part when dividing the data into two classes and the boundaries between the segments and the SVM. Interpretation of high-dimensional space is a duty of work, but the simplicity of partitions depends heavily on the list of accessible features.

Logistic Regression Logistic regression is a realistic double-season forecasting strategy. Output or target variables are decomposed. Disconnecting means that there are only two possible classes. For example, it is often used for problems related to the detection of malignant tumours. Monitor the probability of the appropriate event. This is a rare example of linear regression where the

target variable is inherently simple. Opportunity records are used as dependent variables. The calculated regression uses a Boolean function to predict the probability of a double event.

KNN

K-nearest neighbour (KNN) calculations have been used in many data analysis applications such as design decisions, data mining, databases, and artificial intelligence because of their high clarity and accuracy. It is considered one of the ten major calculations in data mining (Wu et al., 2008). KNN classification commands occur based on similarity. This is a slow learning calculation where the function is approximated locally, and the calculation is assigned to the classification. KNN is primarily used for classification and aggregation. Many scientists track the KNN process to be very efficient in experiments with different datasets. However, data collection for Pima Diabetes India is complicated by lack of quality. The KNN strategy replaces the missing property with comparative estimates from adjacent sections of the Euclidean distance. If there are no additional comparative estimates available from the nearest neighbours, get them from the quick estimates in the next section. This model is very simple and dangerous when compared to various strategies. Inadequate KNN is a lack of random semantics that allows the use of old probabilities. Multiple authors have modified KNN to address its effectiveness. The determination of K in KNN depends on the data. A large gain of k can reduce the perturbation of the classification. A mutual agreement procedure can determine the appropriate k value. By evaluating the k value at ten times the mutual recognition, a classification accuracy of 97.4% was obtained (DHL Solutions & Innovation Trend Research, 2013). The idea for calculating KNN is shown in Figure 2

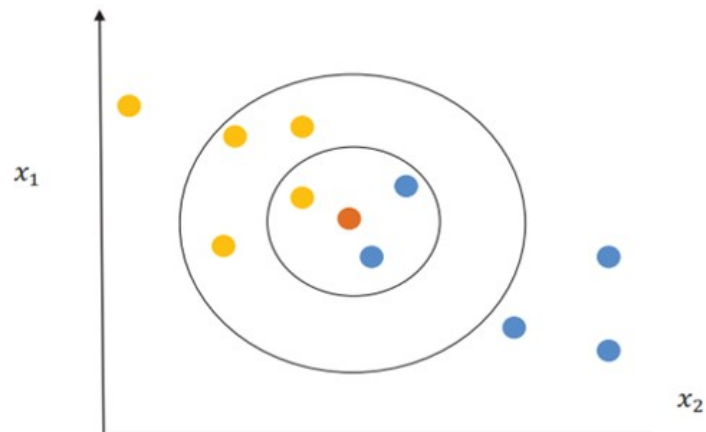


FIGURE 2. K nearest neighbour algorithm

Decision tree Figure 3 shows the decision tree model. A decision tree is a flow

chart-like tree structure, with internal axes processing elements (or quality), branches processing decision rules, and each hub processing the final result. The upper axis of the decision tree is called the root centre. Find a way to package based on quality values. It divides the tree by a recursive method called recursive division. This flowchart-like structure is useful for decision making. This is a graph-like visualization of a flowchart that effectively simulates human-level thinking. This is why decision trees are so simple.

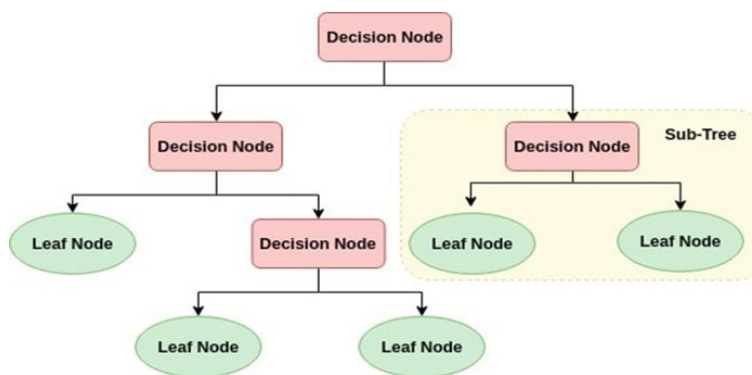


FIGURE 3. Decision Tree Representation

Random forest Random Forest is a structured learning algorithm. They tend to be used for characterization and iteration. It is also the most adaptable and easy-to-use algorithm. The background contains trees. It is said that the more trees there are, the stronger the supporting trees. Irregular forests get decision trees by testing randomly selected data, getting predictions from each tree, and voting to choose the best order. It also demonstrates the importance of components. Every forest area has various uses, including recommended drives, image characterization, and component identification. It can often be used to describe loyal potential applicants, recognize false movements, and predict infection. This is the basis of the Boruta algorithm, which identifies the key components of a dataset. In this regard, the following key attributes are pertinent and essential as a dataset namely-

1. Gross profit ratio
2. Operating profit ratio
3. Expenses ratio
4. Return on Shareholder's funds
5. Capital turnover ratio
6. Debtors turnover ratio
7. Absolute liquid ratio
8. Fixed assets ratio
9. Return on total assets

5. Results & Discursion

The dataset was primarily used for forecasting around M and aggregating and applying to the control area of 58 well-defined associations obtained for review. After parting, various AI models were implemented to clarify expectations. The dataset test results provided a noisy network near the curvature of the discrete AI algorithm and the ROC (Receiver Operating Characteristics) of the model. The recommended batch preparation time is typically 2-4 seconds between algorithm samples. The confusion network is processed together to represent the ML model concerning the width and accuracy of the order of the test data (Gabriel, 2012; Hasan & Curry, 2014a; Hoberg & Philips, 2018; Gentzkow & Shapiro, 2010). Also, note that the error matrix was used to determine the accuracy of the applied model. This helps determine the reality and false rate of the model. At the same time, the ROC curvature aims to handle false-positive rates against the observed true positive rates. Th work uses various data mining algorithms such as SVM, Linear Regression (LR), K-nearest neighbour (KNN), Bayesian classifier, and logistic regression. The performance of the proposed work is analysed in terms of accuracy, and ROC against all the algorithms. From the comparative analysis, we observed that the data collection approaches aids organizations towards M&A decision making, followed by achieving profitable insights as well. This will promote better visibility as well as overall improvements and effective negotiation strategies The resulting improved model accuracy is shown below in the attached case.

$$Accuracy\ score = \frac{correct\ predictions}{All\ predictions} = \frac{TP + TN}{TP + FN}$$

TP and TN suggest positive and true negative values, and FN suggests false negative values. Having this property is the result of approximating the discrete algorithm model used and can be easily obtained via a fuzzy network. Figure 4 depicts the correlated classes together to investigate the performance of the companies in profit after the merge of M & A.

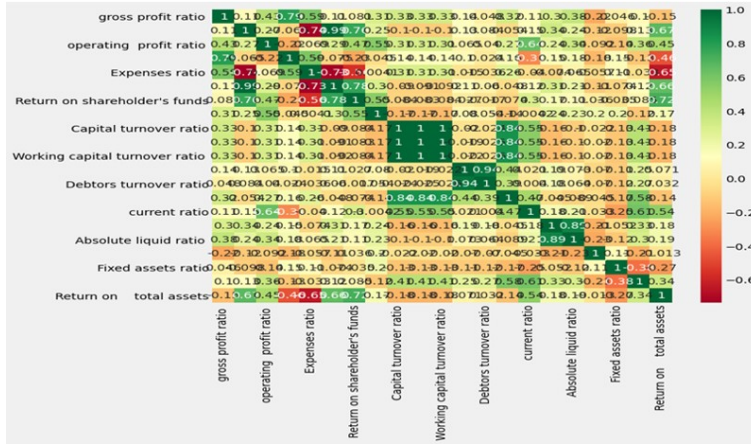


FIGURE 4. Correlation Matrix

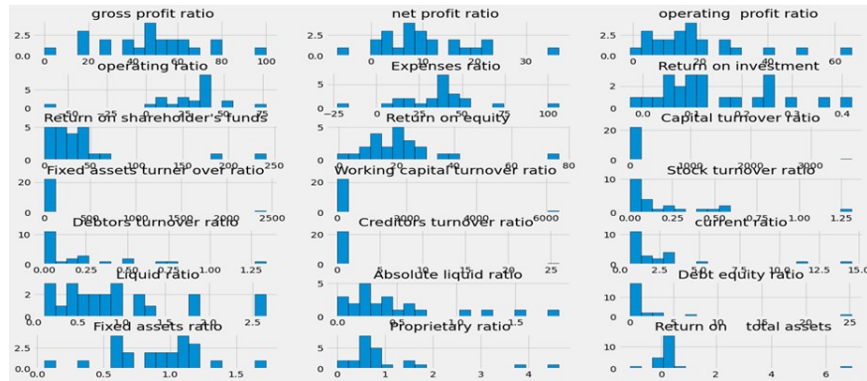


FIGURE 5. The overall statistical outcomes associated with the following attributes achieved from datasets.

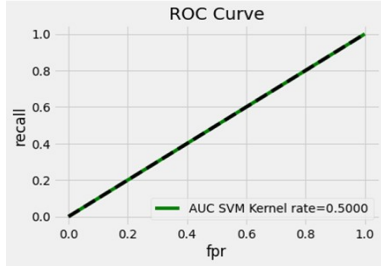


FIGURE 6. ROC Curve (SVM)

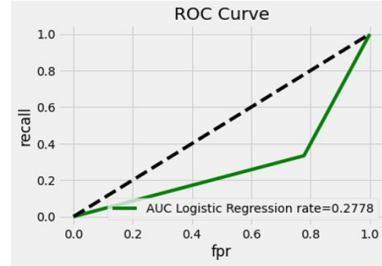


FIGURE 7. Logistic Regression

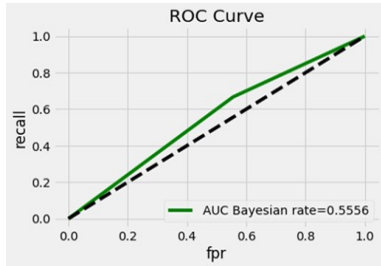


FIGURE 8. ROC Curve (Bayesian)

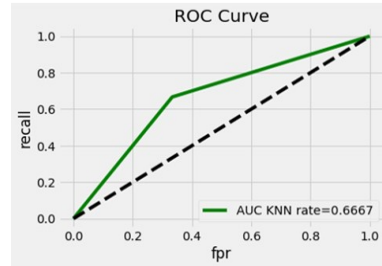


FIGURE 9. ROC Curve (KNN)

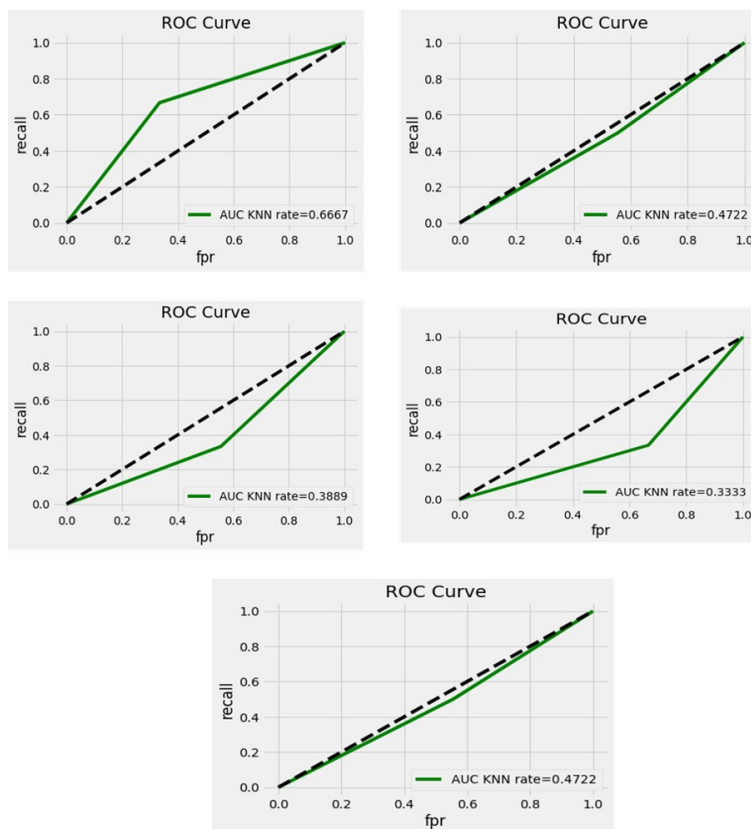


FIGURE 10. ROC for the KNN

6. Conclusion

The main aim of the paper is to investigate the long-term performance of companies involving integration of M & A. The outcome of the study after the analysing the data shows that the acquired companies are seen improved performance after M & A when compared to pre M& A period in terms of profit and operating performance. This paper focused on professional practices that are of great significance. Several other strategies are also shown, including regression, correlation, analysis of factor, clustering, some big data, and many cases that are primarily described. Ideas on how to present recent data and answers to assist in decision making. Big data research can advance care, save lives, and reduce medical costs. It also supports businesses such as collection areas by examining customer records to understand customer needs better. The area of aggregation has a major impact on society, and the use of big data scans in this area will help authorities better understand people’s requirements and thus

further promote departmental improvement. Huge amounts, speeds, speeds, and diverse data variety data management problems with traditional data, creating brings new study technologies that may help overcome these challenges and the need to investigate and investigate the proactive capabilities of big data scanning, as expected under the circumstances. As part of future research, we plan to focus on the challenge of investigating big data related to security and protection issues, given the limitations of retrieving big data from various sources. This research also encourages paying attention to cloud vendors and security breaches that affect many organizations

REFERENCES

1. T.E. Copeland & J.F. Weston, *Asset pricing*, Finance, London, 1989.
2. Dipika Pramanik, Samar Chandra Mondal & Anupam Haldar, *A framework for managing uncertainty in information system project selection: an intelligent fuzzy approach*, Int. J. Mgmt Sci. & Eng. **15** (2020), 70-78.
3. K. Lyko, M. Nitzschke, A.C. Ngonga Ngomo, *Big Data Acquisition*, In: Cavanillas J., Curry E., Wahlster W. (eds) *New Horizons for a Data-Driven Economy*. Springer, Cham., 2016.
4. G. Gabriel, *Storm: the Hadoop of Realtime Stream Processing*, Available online from <https://pyvideo.org/pycon-us-2012/storm-the-hadoop-of-realtime-stream-processing.html> 2012.
5. M. Gentzkow, and Shapiro, *what Drives Media Slant? Evidence from U.S. Daily Newspapers*, *Econometrica* **1** (2010), 35-71.
6. Gholamreza Haseli, Reza Sheikh & Shib Sankar Sana, *Base-criterion on multi-criteria decision-making method and its applications*, In. J. Mgt.Sci. & Eng.Mgt. **2** (2020),79-88.
7. W.J. Gole & J. Morris, *Mergers and acquisitions: business strategies for accountants*, Mergers and acquisitions: business strategies for accountants, Wiley, 2007.
8. S. Hasan & E. Curry, *Approximate semantic matching of events for the internet of things*, *ACM Transactions on Internet Technology* **1** (2014), 1-23.
9. Hashem Vahdani & Aida Shams ., *Multimode capital constrained project payment scheduling model considering bonus-penalty structure*, In. J. Mgt.Sci. & Eng.Mgt. **1** (2020), 17-25.
10. G. Hoberg and G. Phillips, *Product Integration and Merger Success*, Tuck School of Business Working Paper No. 2933283 Available online from <https://ssrn.com/abstract=2933283>, 2018.
11. G. Indhumathi, M. Selvam, and M. Babu, *The effect of Merger on Corporate Performance on Acquirer and Target Companies in India*, *The review of Financial and Accounting Studies*, 2011.
12. G.A. Jarrell, J.A. Brickley, & J.M. Netter, *The market for corporate control: The empirical evidence since 1980*, *Journal of Economic Perspectives* **2** (1988), 49-68.
13. M.C. Jensen, & R.S. Ruback, *The market for corporate control: The scientific evidence*, *Journal of Financial Economics* **11** (1983), 5-50.
14. K. Li, F. Mai, R. Shen, & X. Yan, *Measuring corporate culture using machine learning*, *The Review of Financial Studies* **7** (2020), 3265-3315.
15. W. Long, L. Song, & L. Cui, *Relationship between capital operation and market value management of listed companies based on random forest algorithm*, *Procedia Computer Science* **108** (2017), 1271-1280.
16. Mohamed Abouhawwash, *Hybrid Evolutionary Multi-Objective Optimization Algorithm for Helping Multi-Criterion Decision Makers*, In. J. Mgt.Sci. & Eng. Mgt. **2** (2021), 94-106.

17. Mergermarket, *M&A Integration In association with: A mergermarket report on issues surrounding post-deal integration for European companies*, mergermarket, London, 2007.
18. S. Sudarsanam, *Value creation and value appropriation in M&A deals*, In: Handbook of Mergers and Acquisitions, (eds.) Faulkner, D., Teerikangas, S., & Joseph, R. Oxford: Oxford University Press, 2012.
19. J.F. Weston, & K.S. Chung, *Do mergers make money: A research summary*, Mergers & Acquisitions **3** (1983), 40–48.
20. Jinjiang Wang, Yulin Ma, Laibin Zhang, Robert X. Gao, C. DazhongWu, *Deep learning for smart manufacturing: Methods and applications*, Journal of Manufacturing Systems **48** (2019), 144-156.
21. Amish Bharat kumar Soni., *Mergers and acquisitions in India and its impact on shareholders wealth*, IOSR Journal of Business and Management (IOSR-JBM) **6** (2016), 79-86.

S. Kalaivani is a Ph.D. student at Sathyabama Institute of Science and Technology's Department of Mathematics. She has worked as a researcher in the field of statistical modelling. Mergers & Acquisitions are the focus of her current studies.

Research scholar, Department of Mathematics, Sathyabama Institute of Science and Technology, Chennai, India.

e-mail: kalaivani1503@gmail.com

K. Sivakumar Prof. of Mathematics working at Saveetha School of Engineering, SIMATS, Saveetha University as HOD of Applied Mathematics for Excellence. He obtained his PhD degree in 2011 in the domain of Predicative Modeling. He has published more than 30 research articles and presented in both National and International conferences. He also obtained a patent on Development of IoT based accident detection and smart traffic and signal monitoring in 2021. He has also been Chief Editor for the book titled Recent Trends in Engineering and Management.

Department of Mathematics, Saveetha Institute of Medical and Technical Sciences, Thandalam, Chennai, India.

e-mail: siva111k@gmail.com

J. Vijayarangam M.Sc., Ph.D. Faculty of Mathematics/Data Science since 1999. Specialization is Financial Mathematics-Portfolio Optimization using Stochastic differential equations. Interested areas of research are Finance, Stochastic Modeling, Modeling and Simulation and Machine Learning.

Department of Applied Mathematics, Sri Venkateshwara College of Engineering, Sriperumbudur, Chennai, India.

e-mail: jvijayarangam@yahoo.com