NATIONAL CONFERENCE 2022

PARADIGM SHIFT TOWARDS SUSTAINABLE MANAGEMENT PRACTICES

11th March, 2022

Proceeding & Book of Papers

ISBN: 978-81-956407-0-6

Conveners: **Mr. Sumanta Bhattacharyya**
Dr. Arindam Banerjee

J.D. BIRLA INSTITUTE
Department of Management
Affiliated with Jadavpur University
Satellite Campus
1 Moira Street, Kolkata-700 017 • Ph.: 033-22811438
E-mail: principal@jdbimanagement.in • Website: www.jdbimanagement.in

NATIONAL CONFERENCE 2022

PARADIGM SHIFT TOWARDS SUSTAINABLE MANAGEMENT PRACTICES

11th March, 2022

Proceeding & Book of Papers

ISBN: 978-81-956407-0-6

J. D. BIRLA INSTITUTE
Department of Management
(Affiliated with Jadavpur University)

1 Moira Street, Kolkata - 700 017
Ph.: (033) 2281 1438 ■ Fax: (033) 2281 7479
E-mail: principal@jdbimanagement.in
Website: www.jdbimanagement.in

NATIONAL CONFERENCE 2022

PARADIGM SHIFT TOWARDS SUSTAINABLE MANAGEMENT PRACTICES

11th March, 2022

Proceeding & Book of Papers
ISBN: 978-81-956407-0-6

Edited by

Prof. Deepali Singhee
Principal

Dr. Sraboni Dutta
Dean (Management)

Dr. Manjishtha Sur Roy Chowdhury
Asst. Professor (Dept. of Management)

Dr. Oindrila Chakraborty
Asst. Professor (Dept. of Management)

Address

J. D. Birla Institute
(Department of Management)

1 Moira Street, Kolkata - 700 017
Ph.: (033) 2281 1438 ■ Fax: (033) 2281 7479
E-mail: principal@jdbimanagement.in
Website: www.jdbimanagement.in

First Published	:	March, 2022
Price	:	NIL
Cover Designed by	:	Sanjib Adak Graphic Designer, JDBI

Printed by CDC Printers (P) Ltd., Tangra Industrial Estate-II, 45, Radhanath Chowdhury Road, Kolkata-700015 (WB), India

Published by:

J.D. Birla Institute, Kolkata (West Bengal), India

Effect of Covid-19 Pandemic on BSE Sensex using Logistic Regression

Tinni Chaudhuri[1], Shoumit Saha[2], Banhi Guha[3], Pankaj Kumar[4] & Pulak Konar[5]

[1]Assistant Professor, Amity University Kolkata; tinnistat1991@gmail.com
[2]Student, 3rd year B.Sc. (Hons) Statistics, Amity University Kolkata; mahisumonking19@gmail.com
[3]Assistant Professor, St. Xavier's University, Kolkata; banhi.guha@gmail.com
[4]Associate Professor, Amity University Kolkata; pkumar@kol.amity.edu
[5]Associate Professor, The ICFAI University, Tripura; pulakkonar@gmail.com

Abstract

The Business activity was in standstill when COVID-19 was in full swing in India during 2020, which in turn had hampered the performance of most of the companies. In this paper logistic regression was used to measure the impact of COVID-19 on Indian stock market. Various financial indicators like Net Sales (NS), Earnings per share PS (EPS), Net profit margin (NPM), return of equity values (RoE), PE Ratio (PE), and Price to Book value (PtBV) of BSE SENSEX 30 Companies were considered for five years from 2017 to 2021 to predict their performance (Satisfactory or Unsatisfactory). Initially, the BSE Sensex 30 companies were segregated into two clusters based on their Net sales, then logistic regression model was built (using four years) for both the clusters with six financials pointers and the model so obtained was tested on 2021 for validation. The classification result for the first cluster is 57% but for the second set of companies (major cluster) it was as low as 44% which reveals that most Indian companies were affected by the global pandemic as their stock performance were not as same as estimated for the year 2021. Though we haven't considered any macroeconomic factors influencing the growth however this model will be useful to investors and fund managers to identify and select the outperforming stocks. The primary objective of our paper is to use the fitted model to validate stock performance after the first and second wave of the coronavirus pandemic hit the world economy and Indian financial market in particular.

Keywords: BSE Sensex, stock performance, logistic regression, Indian Stock market, sales growth, prediction.

Introduction

Global crashes do not occur all of a sudden but they are due to local and regional crashes in emerging markets. The investors should pay attention to these markets, as the collapse of local market can affect developed stock markets even if they are not exposed to emerging stock markets. In addition, interdependence is also relevant, as interest rates, bond returns and volatility also affect the probability of the various types of stock market crashes.

Shareholders and future investors will need to use relevant financial information to enable them to make appropriate investment decisions in the stock market. Predicting stock performance is a complex task and is almost always uncertain. A comprehensive accurate model for predicting stock market performance has not been proposed in the stock performance history. Equity performance can be analyzed to some extent based on the financial indicators reported in the company's annual report. Earlier literature suggests that financial ratios are recognized as an important tool for assessing the future stock performance. Analysts, investors, and researchers use financial indicators to predict future stock price trends. Financial indicators are widely used in stock valuation because the ratio analysis has become one of the key parameters for fund managers and investors to determine the intrinsic value of a stock.

Despite all of these factors, there came a time in March-April 2020 when the Covid pandemic came and changed everything. The whole world came to a standstill and Indian BSE Sensex was not different in any case from the rest of the world. The market crashed drastically and changed the whole scenario. The world did not recover yet properly from that change. Healthcare companies, Reliance, banking sectors did well after the pandemic but entertainment, travel & tourism, hospitality sectors were at a constant loss.

BSE Sensex is one of the most common terms to the people investing in the stock market and to those who are related to it one way or the other. The Bombay stock index (BSE) has Sensex as one of the important Index which comprises of 30 companies from different sectors which represents the Indian economy. The Sensex is the most actively traded stocks on the BSE. It is the oldest stock index in India. Other than Sensex, BSE has nineteen Sectoral indices. One of the investment preferences is investing in sectoral indices, which emulate numerous sectoral stocks. Many people invest directly in Sensex instead of individual companies. But this scenario changed after the pandemic situation. Now it is generally more effective to invest in some particular companies instead of the whole Sensex.

The main objective of this paper is applying statistical methods to survey and analyze financial data to develop a simplified model for interpreting the effect of the Covid-19 pandemic and find out which stock is profitable on BSE Sensex. This study aims at developing a model for classifying into two categories based on their rate of return with respect to Sensex. In this study, Logistic Regression method has been applied for classifying the selected companies based on their performance. Logistic regression method is used here for the prediction of the probability of good stock performance by fitting the variables to a logistic curve.

Review of Literature

Logistic regression is a process of modelling the probability of a discrete outcome given an input variable. The most common logistic regression models a binary outcome which can take two values such as true/false, yes/no and good/bad and so on. The application of logistic regression is very large. One way it is used in NLP to determine the sentiment of movie reviews and also in medicine it can be used to determine the probability of a patient developing a particular disease and some people is it in finance, banking and investing as well.

In 1968, Altman, who was the pioneer in this work while Logistic Regression was used by the Ohlson to construct the default-prediction model in 1980. His earlier

research which was on default prediction focuses on classifying firms as in two types which was either defaulters or non-defaulters. Ohlson identifies that thedefault prediction assumption is an equal payoff state. Clearly, misclassifying a firm will be resulted in severe for an investor or a loan officer. As such, this research focuses on the ability of the models to accurately rank the firms based on their default probability. In predicting financial bankruptcy and distress which was widely applied as the evaluation models providing credit-risk information, Logistic Regression was used by Ohlson (1980) which was then followed by several authors such as Zavgren (1985). Subsequently the same trend opted by Zmijewski (1984) for a Probit Analysis.

Logistic Regression technique yields coefficients for each independent variable based on a sample of data (Huang et.al., 2007). The parameters of the logistic regression model are commonly estimated by maximum Likelihood (Pardo et.al., 2005). The advantage of logistic regression is that, through the addition of an appropriate link function to the usual linear regression model, the variables may be either continuous or discrete, or any combination of both types, and they do not necessarily have normal distributions (Lee, 2004).

The predictor values from the analysis can be interpreted as probabilities (0 or 1 outcome) or membership in the target groups (categorical dependent variable). It has been observed that the probability of a 0 or 1 outcome is a non-liner function of the logit (Nepal, 2003). Logistic Regression is useful in which it is required to predict yes or no of a characteristic or outcome based on a set of predictor variables. Where the dependent variable is dichotomous, it is more proficient than linear regression model. Logistic Regression coefficients can be used to estimate odd ratios for each of the independent variables in the model. Logistic Regression helps to form a multivariate regression between a dependent variable and several independent variables (Lee & Kim, 2007). It is designed to estimate the parameters of a multivariate explanatory model in situations where the dependent variable is dichotomous, and the independent variables are continuous or categorical.

Bandyopadhyay et al., 2012 worked in a same type of work on forecasting stock performance in Indian market using Multinomial Logistic regression. They worked on a study consisting of 30 large market capitalization companies and using various financial ratios, they determine the financial indicators that significantly affect the share performance by using multi logistic regression method and concluded that ratio methods have the capability to reveal maximum information content, if variables are chosen very carefully with regard to the purpose at hand.

Shah et. al., 2019 tried to research on stock market analysis as well and applied various approaches. According to them, hybrid approaches that combine statistical and machine learning techniques such as SVM, sentiment analysis might prove to be more useful for stock prediction.

Data Description

Data is collected from moneycontrol.com

In this data the values of SENSEX and the 30 companies present in BSE Sensex from 2016 till 2021 are available. From this data, at first the growth is calculated for all of them in each year. As 2015 data is not taken in this work, the growth rate for the year 2016 is not available to us and the growth in each year from 2017 till 2021 are found. Then, this growth for all the 30 companies in each year is compared with respect to

the growth of Sensex. If the growth for the company is more than that of Sensex then it is satisfactory and encoded as "1" and if the growth for the company is lower than that of Sensex in that particular year then it is not satisfactory and is encoded as "0".

Next, we have the sales value for the 30 companies in the years from 2016 till 2021 and at the same way, the net sales (NS) growth is calculated for these 30 companies from the year 2017 till 2021. Then we have the Earning per share (EPS), Net profit margin (NPM), Current ratio, Return of equity values (RoE), Debt to equity, PE ratio (PER), Price to book value (PtBV) values for all of these 30 companies from 2017 to 2021. As Current ratio and debt to equity have some null values, these might be bad for our model. That is why, the current ratio and debt to equity variables are ignored from this paper.

Later two datasets are created where the variables are Year, Performance (Perf), Company, net sales (NS), Earning per share (EPS), Net profit margin (NPM), Return of equity values (RoE), PE ratio (PER) and Price to book value (PtBV). In the first dataset we have the data of all the 30 companies from 2017 till 2020 which is treated as a training data set and the data for these 30 companies in the year 2021 only is taken in the second dataset which is treated as the test dataset here.

The dependent variable we have here is Performance and the six independent variables are net sales, earnings per share, net profit margin, return of equity values, PE ratio and price to book value.

Methodology

Regression is one of the most popular and effective tools in analyzing various kinds of data. Logistic regression model is fitted here with the filtered data.

Given a dependent variable Y and a set of independent variables X, they can be modelled and instead of fitting a straight line or hyperline, the logistic regression model squeezes the output of a linear equation between 0 and 1.

$$logistic(\eta) = \frac{1}{1 + \exp(-\eta)}$$

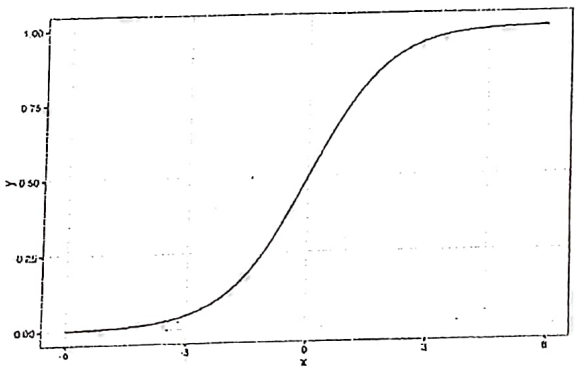

Figure 1: Logistic regression curve

Sigmoid function is applied on the model output and provides the ability to predict with probability.

$$sigmoid\ function: g(z) = \frac{1}{1 + e^{(-z)}}$$

$$Hypot\square esis: \square_\theta(x) = \frac{1}{1 + e^{(-\theta^T x)}}$$

Least Squared Error as loss function is used in Linear regression which gives a convex graph which can completethe optimization by finding its vertex as global minimum.

$$Cost(\square_\theta(x), y) = -ylog(\square_\theta(x)) - (1 - y)\log(\square_\theta(x))$$

So,the summation from all training data samples results in the cost function of the model:

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} Cost(\square_\theta(x^{(i)}), y^{(i)})$$

$$J(\theta) = \frac{1}{m}[\sum_{i=1}^{m} -y^{(i)}\log(\square_\theta(x^{(i)})) + (1 - y^{(i)})\log(1 - \square_\theta(x^{(i)}))]$$

$$m = number\ of\ samples$$

The probability which is the transformed form of the weighted sum by the logistic function.

$$\ln\left(\frac{P(y=1)}{1-P(y=1)}\right) = \log\left(\frac{P(y=1)}{P(y=0)}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$$

The logistic regression model is a linear model for the log odds. The prediction of this changes when one of the features x_j is changed by 1 unit.

$$\frac{P(y=1)}{1-P(y=1)} = odds = \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p)$$

Comparing increase of one of the features values by 1, the ratio of the two predictions:

$$\frac{odds_{x_j+1}}{odds} = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_j(x_j + 1) + \cdots + \beta_p x_p)}{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_j x_j + \cdots + \beta_p x_p)}$$

We apply the following rule:

$$\frac{\exp(a)}{\exp(b)} = \exp(a - b)$$

And we remove many terms:

$$\frac{odds_{x_j+1}}{odds} = \exp\left(\beta_j(x_j + 1) - \beta_j x_j\right) = \exp(\beta_j)$$

We have something as simple as exponential of a feature weight. A change in a feature by one unit changes the odds ratio (multiplicative) by a factor of exponential of β_j. We could also interpret it this way: A change in x_j by one unit increases the log odds ratio by the value of the corresponding weight. Most interpret the odds ratio because thinking about the natural logarithm of something is known to be hard on the brain.

Results

Various results and relations were found from the first dataset to better understand the data and the relations among the dependent and independent variables present here.

A mutual relationship between two or more things is known as correlation. Correlation measures the association between two variables and does not measure causation. The correlation between the dependent and independent variable is very low which is actually good for model fitting. If in one or two cases, the correlation was high, then the model might not perform well with the data as the dependent variable would have been highly correlated.

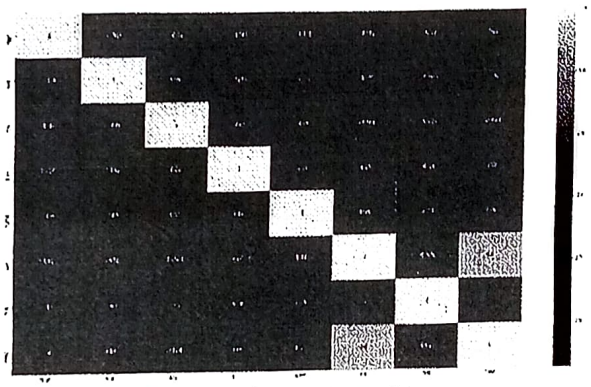

Figure 2: Correlation matrix

The boxplot between the dependent variable performance and the other independent variables gives some outlier. It is not good for our mode.

It can be observed that if we take the whole data for all the companies altogether there seems to be a lot of outliers which is not good for our model. So, 2021 sales data for the 30 companies are taken and clustered.

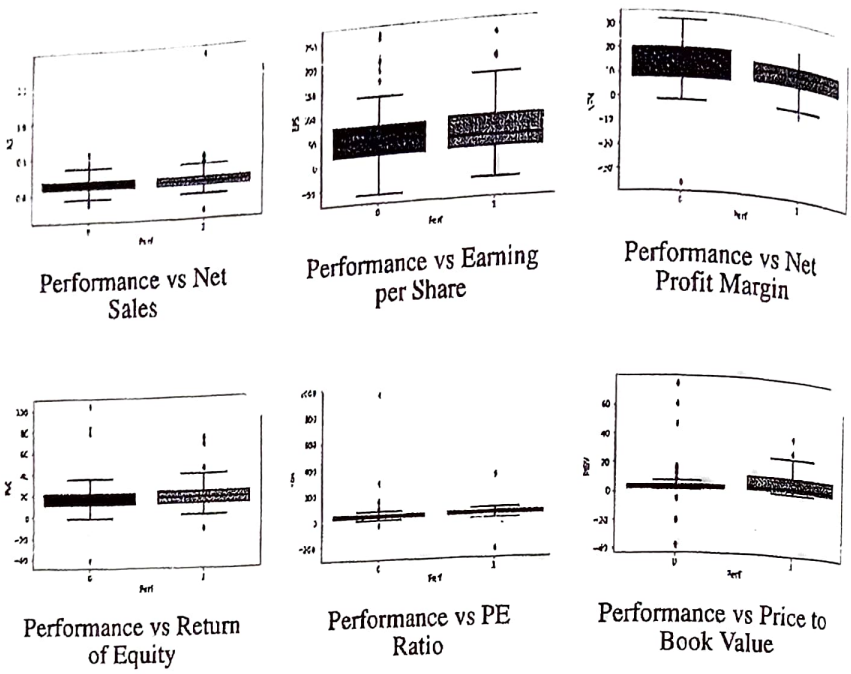

Figure 3: Performance vs independent variables boxplot

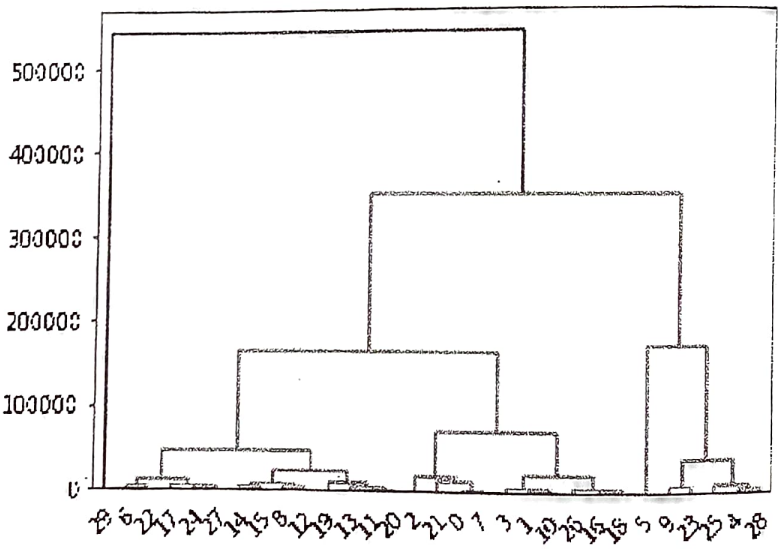

Figure 4: Dendogram

The most appropriate clustering will be if we divide the dataset into 3 parts but as 29th company in this dendogram will be alone in his own group and that is why we decided to cluster our dataset into 2 parts.

Now, we are going to fit the logistic regression model separately inthe data of 2 groups consisting of 7 and 23 companies from the year 2017 to 2020. After running the logistic regression model and fitting the model in the data of the year 2021 for both the groups the output we get is:

	Group 1	Group 2
Accuracy	57.1428%	43.4782%
Positive prediction value (precision)	0.50	0.48
Sensitivity	1.00	0.83
F1 score	0.67	0.61
AIC	39.1961	127.8151
MSE	0.25	0.3369

Group 1 gave an accuracy of 57.14% which is pretty good and was a better model fit than the model fitted in Group 2 which gave an accuracy of 43.47% which is satisfactory. Both group 1 and group 2 gave a decent positive prediction value of 0.50 and 0.48 consecutively. In case of sensitivity, group 1 gave a very good result of 1.00 while group 2 have the sensitivity value of 0.83 which is very good as well. The AIC value of group 1 is 39.19 which is very small value resulting a very good model fit while the AIC value of group 2 is 127.81 which is larger than the value of group 1 but is a relatively smaller value resulting a good model fit as well. Moreover, the MSE value for group 1 and group 2 are 0.25 and 0.34 consecutively which is a lower value resulting the forecast is pretty good in both the cases.

The confusion matrix for the two groups:

GROUP 1		The Truth	
		Satisfactory	Unsatisfactory
Prediction	Satisfactory	3	0
	Unsatisfactory	3	1

GROUP 2		The Truth	
		Satisfactory	Unsatisfactory
Prediction	Satisfactory	10	2
	Unsatisfactory	11	0

Conclusion

This study has employed the Logistic Regression Model to determine the factors which significantly affect the performance of the company in the stock market. Logistic regression method helps the investor to form an opinion about the shares where one can invest. It may be observed that six financial ratios i.e., Performance change in Net Sales, earning per share, net profit margin, return of equity, PE ratio, price to book value can classify them into two categories satisfactory or unsatisfactory.

The prediction rate of 57.14and 43.47 is good. So, it is possible to use it for prediction with better accuracy. From investors' point of view, it is possible to predict outperforming share by examining these six ratios. For data processing and analysis

various methods are available, but in this study, it has been identified that ratio methods have the capability to reveal maximum information content, if variables are chosen very carefully with regard to the purpose at hand. In this study data, for 12 months have been taken into consideration and at the end of 12th month, stock share prices and variances were compared with the previous year and performance was determined. This study uses only financial ratios as only factor affecting share prices. There may various economic and management factor that may also influence the share prices. McConnell, Haslem and Gibson (1986) have identified that qualitative data can provide additional information to forecast stock price performance more accurately.

Future Work

In this work, only Logistic regression is considered to build the model. For further development, in this same dataset, different approaches can be taken like using time series analysis and including neural network and different machine learning algorithms like decision tree classifier, random forest classifier and XG boost in this work to increase the prediction ratio. If we can get more accuracy using different approaches, this work might be able to predict the future stock performances better which might help investors to predict and invest in the companies during the ongoing covid pandemic situation.

References

1. Upadhyay, A.; Bandyopadhyay, G., & A. Dutta, A. (2021). Forecasting Stock Performance in Indian Market using Multinomial Logistic Regression. *Journal of Business Studies Quarterly, 3* (3), 16-39.
2. Altman, E.I., (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*,589-609.
3. Adiana, N. H. A.; Halim, A.; Ahmad, H.; & Rohani, M. R. (2008). Predicting Corporate Failure of Malaysia's Listed Companies: Comparing Multiple Discriminant Analysis, Logistic Regression and the Hazard Model. *International Research Journal of Finance and Economics, 5*, 202-217.
4. Awales, George S. Jr. (1988). Another Look at the President's Letter to Stockholders. Financial Analysts Journal, 71-73, March- April.
5. Beaver, W. H., (1966). Financial Ratios as Predictors of Failure, *Journal of Accounting Research, 4*, 71-111.
6. Bhattacharya, Hrishikes. (2007). Total Management by Ratio's. 2nd ed. New Delhi, India: Sage Publication.
7. Chen, K. H.; & Shimerda, T. A. (1981). An Empirical Analysis of useful Financial Ratios. *Financial Management, Spring*, 51-60.
8. Connor, M.C. (1973). On the usefulness of Financial Ratios to Investors in Common Stock. *The Accounting Review*, 339- 352, April.
9. Ali, S. S.; Mubeen, M.; Lal, I.; & Hussain, A. (2018). Prediction of stock performance by using logistic regression model: evidence from Pakistan Stock Exchange (PSX), *Asian Journal of Empirical Research, 8*(7), 247-258.
10. Shah, D.; Isah, H.; & Zulkernine, F. (2019). Stock Market Analysis: A Review and Taxonomy of Prediction Techniques. *International Journal of Financial Studies, 7*, 26.
11. Abu-Mostafa; Yaser, S.; & Amir F. Atiya. (1996). Introduction to financial forecasting. *Applied Intelligence, 6*, 205-213.
12. Babu, M.; Suresh, N.; Geethanjali; & Satyanarayana, B. (2012). Clustering Approach to Stock Market Prediction. *International Journal of Advanced Networking and Applications, 3*, 1281.

J.D. BIRLA INSTITUTE

Unit of Vidya Mandir Society

(Affiliated with Jadavpur University)

DEPARTMENT OF MANAGEMENT

11, Lower Rawdon Street, Kolkata -700 020 • Website : www.jdbikolkata.in

Certificate of Participants

National Conference on

PARADIGM SHIFT TOWARDS SUSTAINABLE MANAGEMENT PRACTICES

This is to certify that TINNI CHAUDHURI

of Amity University, Rajarhat, New Town, Kolkata has presented

a paper entitled Effect of covid-19 Pandemic on BSE Sensex using logistic regression

at the "National Conference on PARADIGM SHIFT TOWARDS SUSTAINABLE MANAGEMENT PRACTICES organized by Department of Management, J.D. Birla Institute, Kolkata on 11th March, 2022.

Mr. Sumanta Bhattacharyya
(Convener)

Dr. Arindam Banerjee
(Joint Convener)

Dr. Sraboni Dutta
(Dean, Management)

Prof. Deepali Singhee, Ph.D
(Principal, JDBI)