# Analysis of Employment-Related Sentiment in Social Media in India

**Ms. Meenakshi[1], Dr. Jaspreet Kaur[2]**
[1]Research Scholar,CT University
[2]Assistant Professor, Department ofComputer Science and Engineering, CT University

***Article Info***
***Page Number: 10277 - 10288***
***Publication Issue:***
***Vol 71 No. 4 (2022)***

***Article History***
***Article Received:*** *15 September 2022*
***Revised:*** *25 October 2022*
***Accepted:*** *14 November 2022*
***Publication:*** *21 December 2022*

***Abstract***

The purpose of this article is to examine the unemployment rate in India and the employment sentiment index. The employment sentiment score, which is gathered from a lexicon-based technique and several social media channels, is utilized to categorize the feelings among three distinct groups depending on the keywords. In addition, the unemployment rate was acquired from the dataset in order to study the variations in the Indian rate of unemployment. From January 2020 to October2022, a total of 734,776 media messages were gathered, and their remarks were transformed into a sentiment score using machine learning algorithms. The employment score is compared to India's unemployment rate. Despite the fact that the unemployment rate has become less variable in recent years, a negative sentiment score towards employment was discovered.

**Keywords**: Unemployment, Employment, Sentiment Score, Unit Root Test, Social Media.

## Introduction:

Numerous definitions of "social networking" exist. Social networking is defined by Market (2010) as "the development, consumption, and exchange of data using social networking interaction and channels." Since the advent of social media, the number of users has grown tremendously. In 2019, there are expected to be approximately 3.47 billion social media users, up from 2.91 billion in 2015. Social media penetration is on the rise. In 2020, 70.3 percent of internet users were users of social media, and it is anticipated that this percentage will increase. Social media are replacing traditional communication methods. In addition to providing a fast, simple, and nearly cost-free method of contact with individuals all over the world, it allows users to share their ideas and feelings on any topic. For instance, consumers may express their dissatisfaction with certain products or services or their enthusiasm about recent positive developments in their lives, such as receiving a position, etc.The emotions stated on social media supply enterprises, governments, financial institutions, and regulators with a lot of information. Businesses, for example, can track the impact of their goods or solutions, people's perceptions of those services or products, and how customer comments on social media sites promote those products or services to a larger market. The collected knowledge may also provide firms with insights for further enhancing their services and products. The government, central banks, and regulators can acquire data on consumer anticipation of future economic circumstances by analyzing social media sentiment. According to Kwan and Cotsomitis (2006), economic results are highly dependent on consumers' predictions of future economic conditions.

Moreover, Puts and Daas (2014) have demonstrated that there is a substantial correlation between shifts in the sentiment of postings on social media and consumer confidence. In addition, the views conveyed via social media enhance the ability of economic forecasters and researchers to know what customers are thinking about the future at any given time. Also shown to have strong explanation power for present changes in household expenditure was customer sentiment (Carroll et al., 1995; Scott andAcemoglu , 1995). In light of the information that can be gleaned from media platforms' sentiment, a lot of studies are conducted on social media sentiment classification. Opinion mining and sentiment analysis are fields of study that generally examine the opinions, assessments, attitudes, and feelings of individuals based on their writing systems (Liu, 2013).

Sentiment analysis can be conducted on three distinct levels: the document-, sentence-, and aspect-levels (Medhat et al., 2015). Methods for analyzing emotions can be broadly divided into three categories: lexicon-based methods, machine learning techniques, and hybrid approaches (Boiy et al., 2008). The lexicon-based technique is based on a collection of well-known and pre-compiled sentiment terms; the approaches to machine learning are derived from the use of various algorithms; and the hybrid approach combines these methods. Several studies have employed such lexicon-based approaches (Liu, 2012.,Arora et al., 2015;Siau, 2014; Hu.,Adeborna and Siau, 2014; Hu .,), machine learning-based supervised (Li, 2011., 2002; Wang, 2010; Shi and Pang et al.,) or unsupervised (Liu and Li , 2010.,Devi and Usha , 2013;Turney, 2002;) approaches, and combined machine learning and lexicon-based approaches (Lu and In addition to sentiment classification based on textual data, Yuan et al. (2016) suggested a method for analyzing social media sentiment using photos. Focusing on emotion annotations, Zhu et al. (2017) have presented an algorithm to track the moods of social media users. Given the vast number of users and the high number of people who frequently share their thoughts on various social media platforms, improper analysis of these data would represent a significant loss. Consequently, this paper will address this gap in the literature by performing research on the attitudes displayed by India social media users.This research was conducted in conjunction with Pittsfield Media. The study examines the sentiment score toward jobs in India. From January 2011 to March 2018, the unemployment rate in India was reported to fluctuate between 3.8% and 4.7%, according to Dataset. Nonetheless, the society is in shock following the release of Bank annual report in 2017, which revealed that India youth unemployment rate has surpassed and more than tripled the national rate of 4.2% due to slower hiring growth. This prompts us to explore the viewpoint of India on the present employment situation. This study aims to (i) investigate the employment sentiment value on social media, (ii) examine the characteristics of the rate of unemployment, and (iii) analyze the employment sentiment value and unemployment rate in India. The structure of this document is as follows: Section 2 provides a concise literature assessment of the unemployment rate in India. The third section describes the study's data sources, sample size, and methodology. The original data and aggregate data analysis for the employment sentiment score are measured and compared to glean further information regarding the employment sentiment score. In addition, exploratory data analysis and the unit root test are used to measure the data in order to comprehend its qualities. Section 4 contains the results and analysis. Section 5 finishes the investigation.

**2. Literature Review:**

Unemployment is a trailing statistic that correlates directly to economic circumstances. The unemployment number is represented as a percentage to indicate the labor force's effectiveness in terms of the unemployed. According to Pasquali (2015), "unemployed" is the economic situation in which an individual deliberately sought employment yet remained unemployed. According to Healthy Survival, there are two distinct types of unemployment: "broad" and "narrow." "Narrow" refers to those who are unemployed and those who are seeking employment, whereas "Board" refers to individuals who have ceased looking for employment and those who are working part-time but want full-time employment.

According to Dato' Sri Abdul et el., the country's average annual unemployment rate was 3% during the first four years of the 10th India Plan (2011-2015) (The Star, 2015).In addition, between January and August of 2011, the Department of Human Resources registered 308,371 bachelors and non-graduates between the ages of 15 and 29. According to figures from the India Statistics Department (2019), the preponderance of unemployed individuals are young people. India youth unemployment rate is much higher than twice the national average. It accounts for 11.4 percent of India overall unemployment rate.Moreover, the youth unemployment rate in India is higher than that of china (8%) and Thailand (4.4%), which are nearby countries (Department of Statistics India, 2019).Totan et al. (2014) explained that being unemployed in India is a long-term chronic condition that causes an employment crisis. Personal unemployment status might have an impact on family cohesion and harmony. Long unemployment, especially among young adults, can result in crime or violence.The unemployed rate is the proportion of able-bodied adults over the age of 17 who have either been laid off or have unsuccessfully sought employment in the span of a month and are currently actively looking for work. The following is the standard formula for calculating the rate of unemployment:

$$Unemploym(cent) = totalnumberottunemployedperson \backslash stotaltotalnumberottlabourttorce\ x\ 100$$

Due to frictional and structural unemployment, the natural rate of unemployment is greater than zero since the unemployment rate is measured while the economy is generating at its maximum potential capacity.According to Trading Economics (2019), the average unemployment rate in India from 1999 to 2019 was 4.28 percent, hitting a maximum of 5.5 percent in March 2000 and a minimum of 3.7% in August 2012. The unemployment rate has recently decreased to 4.4% in March 2019, compared to 3.5% in the same month of the previous year. The number of employed individuals increased from 15401.80 thousand in February 2019 to 16421.70 thousand in March 2019. A high unemployment rate suggests a growing incidence of poverty and a deteriorating standard of living. Figure 1 displays the evolution of India unemployment rate from January 2019 to dec 2022. The series is highly erratic and fluctuating. Several variables contribute to irregular fluctuations in the unemployment rate, including policy execution, the adoption of the Goods and Services Tax in April 2019, inflation, currency depreciation, and demographic dynamics. From 52,835 unemployed graduates in 2019 to 54,285 unemployed graduates in 2020, the number of unemployed graduates increased. This is related to the United States debt-ceiling crisis of 2022, which would indirectly impact the global financial crisis, particularly in nations with trading ties to the India.According to the literature, the Asian Financial Crisis resulted in serious economic problems, including the bankruptcy of multinational corporations and economic contraction, which may have contributed to an increase in the unemployment rate. According to the India Department of Statistics (2019), the labor force participation rate increased to 68.8 percent in July 2019, while the jobless rate was 4.5 percent. In July 2019, the labor force increased by 0.3%, or 15.73 million people, compared to May. When unemployment is high, the currency exchange rate will rise (Kim, 2017). In Pakistan, India, China, Japan, Bangladesh, Argentina, Algeria, Brazil, Colombia, and Sri Lanka, Hina et al. (2012) estimate a positive association between the exchange rate and unemployment by employing a non-linear least square method. There are numerous elements that influence a nation's unemployment rate. For instance, Hanclova et al. (2015) investigated the factors that influenced the level and evolution of long-term unemployment in India countries from 2012 to 2022. The results indicated that labor market flexibility (part-time or contract workers) has the most negative impact on the unemployment rate in figure 1.

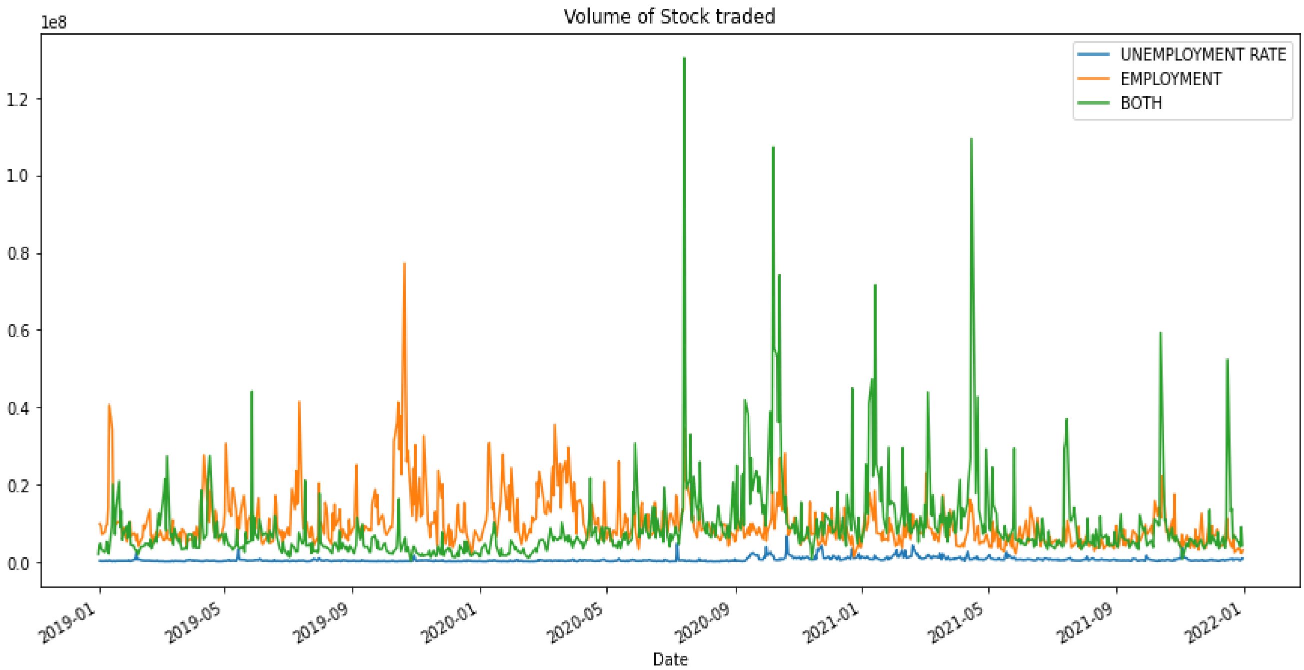


Figure .1 Unemployment, Employment and Both

## 3.Data and Methods:

### 3.1 Data Sources:

Big data has drawn the attention of business, government agencies, and academic institutions on a worldwide scale. An exhaustive study that tracks India citizens' social media posts across multiple platforms produces data having three sentiment categories: positive, negative, and neutral.

The Berkshire News corporation will use social networks, such as aggregators, Facebook, forums, blogs, mainstream media, comments online, YouTube, and, social media, to acquire and Twitter gather communications.

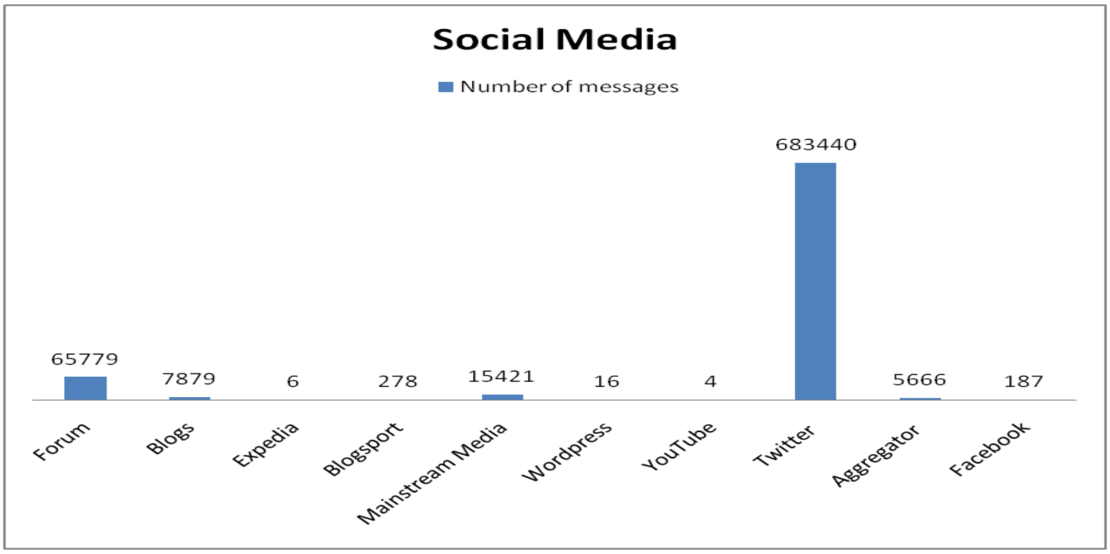

Figure.2 messages that have been compiled from various social channels.

The quantity of messages that were collected from various India media is seen in Figure 2. On Twitter, 693,540 posts, or 89 percent, are about employment-related issues. This number is hardly shocking given that Twitter is a well-known social media site that includes online news, allows users to publish messages, and allows for user interaction. Twitter had over 402 million daily active users as of 2019, was ranked in the top ten most visited websites worldwide, and was dubbed "the SMS of the internet" (Leslie, 2009). After gathering the communications, unstructured data is screened using machine learning techniques. Since natural-language text is frequently inaccurate, this stage is crucial. It includes uncertainties brought on by unpredictable semantics and syntax, including slang, industry-specific terminology, and satirical language. The sentences and phrases in unstructured information can be converted into three sentiment classifications using machine learning technology, which can then be linked with structural information in a database and examined. If a message contains any positive words, it is considered positive; if it includes any negative words, it is considered negative.

**3.2 Sample Size:**

Social media posts are supplementary sources of data that are derived first from the nation's most widely used media platforms. From January 2019 to October 2022, Berkshire Television gathered 7,858,778 media texts from various platforms. Based on a number of terms, a lexicon-based approach is used to categorize the sentiments among neutral, negative, and positive. Messages that are neutral lack any obvious emotion. When categorizing messages into good and negative ranges, with the list of keywords in the posting messages, the similarity measure is utilized to determine if the posts are positive or negative. This method is used because employing machines to learn techniques for sentiment classification can be difficult when there isn't enough data to train a classification, and in those situations, lexicon-based techniques are more practical. According to Taboada et al. (2011), lexicon-based approaches can be reliable between texts and contexts if the phrase dictionary used to categorize the attitudes is appropriately constructed.

### 3.3 Research Methods:

John Tukey initially suggested using data analysis to evaluate and condense the data sets in 1977. In order to discover, comprehend, and analyze the characteristics of the method utilized and the information it contains, exploratory analysis of data uses a number of approaches, including graphical methods like histograms, box plots, and scatter graphs as well as mathematical approaches like descriptive statistics. Big statistics and data analysis both use the exploratory analysis of information methodology. This method's goals are to evaluate the presumptions used to establish confidence intervals and to create a foundation for additional data gathering through experiments and surveys (Behrens, 1997). This study employs exploratory data analysis to identify trends in social media data. Analysis of exploratory data is used because social networking messages have a high level of dimensionality and need to be understood more deeply in order to be converted into valuable information. Data exploration can examine the data from a variety of perspectives, cutting and dicing it along non-trivial or non-orthogonal dimensions and combinations of dimensions. Exploratory analysis can also project the data into a different subspace to use some nonlinear operators, transform the information, and then evaluate the resultant distributions. Furthermore, data from time series is said to be composed of both continuous sequence and mistakes, making it challenging to spot the pattern. To ascertain the stationary properties of the variables used in this study, the unit root test is used. A useful technique for analyzing the characteristics of the series is the Dickey and Fuller (ADF) test. The ADF test can evaluate precise mathematically lag-correlated series with large sample sizes and can construct a parameter correlation for these types of series.

The ADF test's hypotheses are

$H_0 : \beta_j = 0$ (The series has unit root) :

$H_0 : \beta_j \neq 0$(The series does not have unit root)

4.0 Results:

4.1 Exploratory Data Analysis Results for the Employment Sentiment Score:

Analyze India citizens' opinions on the working population in the current labor market using social networking employment messages. The emotion score and descriptive analysis for the post are detailed in Table 1. Descriptive statistical analysis is primarily used to characterize the fundamental aspects of study data. The samples and the measurements are summarized.

Table 1: Employment Sentiment Score Descriptive Statistics:

| Test statistics | |
|---|---|
| 3.3 | Kutosis |
| -39 | Minimum score for sentiment |
| 0.22 | Skewness |
| 32 | Highest level of sentiment |
| 0 | The sentiment score's median |
| -0.06 | The sentiment score's average |
| 8.3 | Standard Deviation |

The term with the highest positive rating based on the term in the posting message is shown by the maximal sentiment that is presented in Table 1. The employment sentiments score ranges from a negative 39 to a positive 32. Additionally, the standard deviation is calculated because it measures data variance. Skewness is an amorphous metric that illustrates the accumulation of facts. In mathematics, a distribution is said to be left- or negatively skewed if the mean is less than the median. When the mean exceeds the median, the data is skewed to the right or positively. A quantity known as kurtosis is used to quantify the height and sharpness of a spike in relation to the remainder of the information. Increasing scores indicate a higher, sharper peak, whereas lower values represent a lower, fuzzier peak. The normal distribution has a kurtosis of three, and any distribution with a kurtosis close to three is referred to as mesokurtic. According to Table 1's results, the kurtosis is reported as 3.3 (less than 4), hence the pattern is referred to as platykurtic.

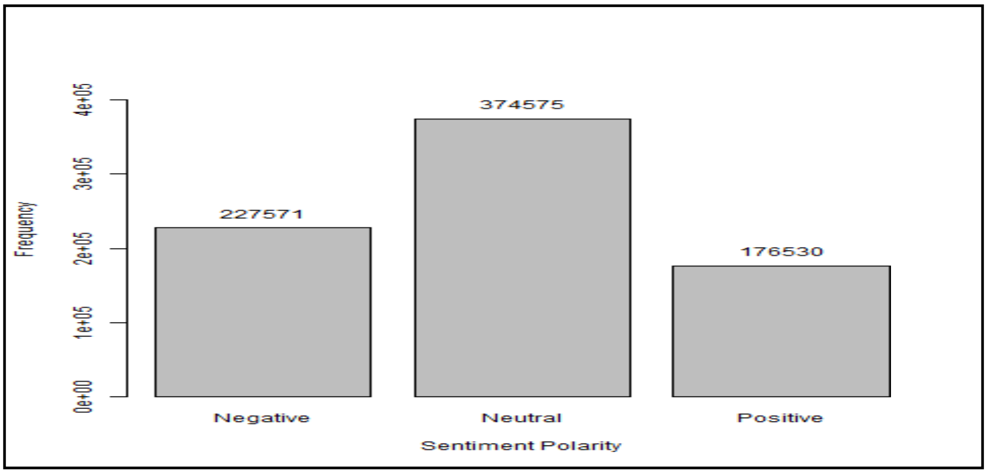


Figure.3 A Sentiment Polarity Bar Chart.

The sentiment classifications for the social media posts are displayed in Figure 3. Negativity is documented in 238,671 messages; neutrality is reported in 385,685 messages; and good sentiment is reported in 185,420 messages. Neutral messages, negative, or Positive can all be posted simultaneously in various media outlets at the level of the particular message. Variables could contain mistakes and incompleteness. In order to collect and express the material in a summarized manner for statistical analysis purposes, received data is employed in this study. A common goal of big data collection is to find the first complete and reliable source of data. According to Regan and Lewls (2008), this information can help measure and legitimize procedures and make a substantial contribution to performance measurement centered on the economy and increasing value.

### 4.2 Exploratory Data Analysis for Aggregate Data Results:

It will be a tremendous waste if all this information is not adequately examined, considering the huge popularity of social media and the huge amount of people who constantly share their comments in various social media. Thus, by carrying out a study on the sentiments expressed by India social media users, this paper will close this research gap. The data was collected monthly, and the sentiment score for that month was determined by averaging the sentiment ratings. The percentage of messages classified as negative was subtracted from the percentage of messages classified as positive to determine the average emotion for every interval in Table.2 and figure.4 .

Table 2: Aggregated Descriptive Statistics for Employment Sentiment

| Score for Employment Sentiment | |
|---|---|
| 0.008 | Probability |
| -12.35 | Mean |
| 10.88 | Jarque-Bera |
| -21.39 | Median |
| 5.99 | Kurtosis |
| 105.68 | Maximum |
| 2.17 | Skewness |
| -78.03 | Minimum |
| 40.34 | Standard Deviation |


Figure.4 Sentiment index for overall data.

Figure 4 displays the daily averages of the sentiment score using social networking sites for the period of two years, from January 2019 to October 2022. Figure 4 shows the duration (per month) on the x-axis and the mean sentiment rating on the y-axis. The series changes as time goes on. The sequence is also erratic or irregular. Series data variations known as "irregular variations" are brief in frequency, random in form, and exhibit no regular pattern of recurrence.

Table 3: The Statistical Measures of the Unemployment Rate.

| Unemployment Rate | |
|---|---|
| 4.225 | Mean |
| 2.73 | Jarque-Bera |
| -21.39 | Median |
| 2.72 | Kurtosis |
| 4.50 | Maximum |
| -0.02 | Skewness |
| 4.00 | Minimum |
| .003 | Standard Deviation |


Figure .5 Data over Time for the Unemployment Rate.

In Table 3 and Figure 5, the characteristics of time series information for the rate of unemployment are presented. According to the quantitative value in Table 3, the mean and median are quite similar. Additionally, the dataset's -0.1 skewness supports the idea that it is approaching standard. This may be connected to a steady adjustment in the unemployment rate between the range of 3.00 and 3.60. Figure 4's scatter plot shows how the unemployment rate changed between January 2019 and December 2022. Less unstable than other series, it reports a maximum available point in March 2019.

**5.0 Conclusions:**

Consumers could react to and share the content on a variety of social data networks and online news websites of today, including blogs, Facebook, and Twitter.It is not surprising that users have posted remarks on the discussion topic linked to employment. Regardless of the fact that the unemployment rate is stable, averaging 4.28 percent from 2019 to 2022, and ranging from 3.0 to 3.6 percent from January 2019 to October 2022, unemployment and employment are hot topics on social networking sites. The sentiment score is negative the findings the results. This indicates that the users' opinions of employment are unfavorable. Additionally, although the unemployment rate and employment sentiment score are not static, they can become static through the process of discretization. Additionally, the employment score is provided with a trend, which means that a pattern immobile procedure can produce stationary properties for the series, whereas the unemployment rate is reported without a trend or seasonal series data component.

**References:**

(1)Acemoglu, D, Scott, A 1994,'Consumer confidence and rational expectations: Are agents' beliefs consistent with the theory?'Economic Journal, vol. 104, pp. 1-19.

(2)Adeborna, E &Siau, K 2014,'An approach to sentiment analysis the case of airline quality rating',Proceedings of the Pacific Asia Conference on Information Systems, China, pp. 363-368.

(3)Arora, D, Li, KF& Neville, SW 2015,'Consumers' sentiment analysis of popular phone brands and operating system preference using Twitter data: A feasibility study',2015 IEEE 29th International Conference on Advanced Information Networking and Applications, Gwangju, Korea, pp. 680-686.

(4)Bland, JM&Altman, DG 1996, 'Statistics notes: measurement error'. BMJ, vol. 312, no. 7047: 1654, doi:10.1136/bmj.312.7047.1654.

(5)Boiy, E, Hens, P, Deschacht, K &Moens, MF 2007,'Automatic sentiment analysis in on-line text', Proceedings of the 11th International Conference on Electronic Publishing, Vienna, Australia, pp. 349-360.

(5)Singha, A. K., Pathak, N., Sharma, N., Tiwari, P. K., & Joel, J. P. C. (2022). Forecasting COVID-19 Confirmed Cases in China Using an Optimization Method. In *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2022, Volume 2* (pp. 683-695). Singapore: Springer Nature Singapore.

(6)Singha, A. K., Pathak, N., Sharma, N., Tiwari, P. K., & Joel, J. P. C. (2022). COVID-19 Disease Classification Model Using Deep Dense Convolutional Neural Networks. In Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2022, Volume 2 (pp. 671-682). Singapore: Springer Nature Singapore.

(7) Siddiqui, S.T., Ahmad, M.O., Khamruddin, M., Gupta, A.K. and Singha, A.K., 2022, January. Blockchain and IoT for Educational Certificates Generation and Verification. In 2022 2nd International Conference on Computing and Information Technology (ICCIT) (pp. 298-303). IEEE.

(8)Daas, PJH, Puts, MJH 2014,‘Social media sentiment and consumer confidence’,Statistics Paper Series, vol. 5, pp. 1-25.
(9) Singha, A.K., PathakS, N., Sharma, N., Urooj, S., Zubair, S., andAreejMalibari. Design of ANN Based Non-Linear Network Using Interconnection ofParallel Processor. Computer Systems Science and Engineering Accepted)
(10)Dickey, D & Fuller, W1979,‘Distribution of the Estimators for Autoregressive Time Series with a Unit Root’,Journal of the American Statistical Association, vol. 74, pp. 427-431.
(11) Singha, A.K., PathakS, N., Sharma, N., Urooj, S., Zubair, S., andLarguech, S., An Efficient Integrated Optimize Method Based on Adaptive Meta Optimizer . Intelligent Automation & Soft Computing,(Accepted)
(12)Engle, R F& Granger, CWJ 1987,‘Cointegration and error correction: representation, estimation and testing’, Econometrica, vol.55, pp. 251-276.
(13) Singha, A.K., PathakS, N., Siddiqui,S.T ., Urooj, S., Mohamed, H.G., and Patil, A.R., 2022. Smart Quarantine Environment Privacy through IoT Gadgets Using Blockchain . Intelligent Automation & Soft Computing, 35(3), pp.3022-3036.
(14)Hanclova,J, Simek, M &Horak, J2012,Factors influencing the long-term unemployment level and development in the European Union.In Mathematical Methods in Economics, Silesian University, School of Business Administration, pp. 249-254.
(15) Hina C, Niaz AB, Falahuddin B, Sherraz AS &Warsha D 2012, ‘The effect of exchange rate on unemployment rate in Asian Countries’, International Conference on Business Management, Spain, ISBN: 978-969-9368-06-6.
(16) Hondroyiannis, G&Papapetrou, E2006,‘Stock returns and inflation in Greece: a Markov switching approach’,Review of Financial Economics, vol. 15, pp. 76-94.
(17) Singha, A.K., PathakS, N., Sharma, N., Gandhar, A., Urooj, S., Zubair, S., Sultana, J. and Nagalaxmi, G., 2022. An Experimental Approach to Diagnose Covid-19 Using Optimized CNN. Intelligent Automation & Soft Computing, 34(2), pp.1066-1080.
(18) Siddiqui, S.T., Singha, A.K., Ahmad, M.O., Khamruddin, M. and Ahmad, R., 2022. IoT Devices for Detecting and Machine Learning for Predicting COVID-19 Outbreak. In Recent Trends in Communication and Intelligent Systems (pp. 107-114). Springer, Singapore.
(19)Kreishan, F.M. (2011) ‘Economic growth and unemployment: An empirical analysis’, Journal of Social Science, Vol. 7, No. 2, pp. 228-231.
(20)Leslie, D 2009, Swine Flu's Tweet Tweet Causes Online Flutter, Business Standard.
(21)Li, G, Liu, F2010,A clustering-based approach on sentiment analysis, 2010 International Conference on Intelligent Systems and Knowledge Engineering,Beijing, China, pp. 331-337.
(22) Singha, A.K. and Zubair, S., 2022. Machine Learning for Hypothesis Space and Inductive Bias: A Review. *AIJR Abstracts*, p.70.
(23)Lu, B&Tsou, B2010, Combining a large sentiment lexicon and machine learning for subjectivity classification.2010 International Conference on Machine Learning and Cybernetics, China, pp. 3311-3316.
(24)Mahadeva, L & Robinson, P2004, Handbooks in Central Banking: Unit Root Testing To Help Model Building. no. 22, London, ISBN 1 85730 138 2.

(25) Zubair, S., Singha, A. K., Pathak, N., Sharma, N., Urooj, S., &Larguech, S. R. (2023). Performance Enhancement of Adaptive Neural Networks Based on Learning Rate. *CMC-COMPUTERS MATERIALS & CONTINUA*, *74*(1), 2005-2019.

(26)Medhat, W, Hassan, A&Korashy, H2014, Sentiment analysis algorithms and applications: a survey. Ain Shams Engineering Journal, vol. 5, no. 4, pp.1093-1113.

(27)Pang, B, Lee, L&Vaithyanathan, S2002, Thumbs up? Sentiment classification using machine learning techniques.Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, USA, vol. 10, ppl. 79-86.

(28)Phoong, SW&Ismail, MT2014, Measuring nonlinear behavior in time series data, AIP Conference Proceedings, vol.1635, no. 837, doi: 10.1063/1.4903680.

(29) Singha, A. K., Kumar, A., &Kushwaha, P. K. (2018). Patient Cohort Approaches to data science using Biomedical Field. *EPH-International Journal of Science And Engineering (ISSN: 2454-2016)*, *1*(1), 457-462.

(30)Shi, HX & Li, XJ2011, A sentiment analysis model for hotel reviews based on supervised learning.2011 International Conference on Machine Learning and Cybernetics, Guilin, China, vol. 3, pp. 950-954.

(31)Taboada, M, Brooke, J, Tofiloski, M, Voll, K,&Stede, M2011,Lexicon-based methods for sentiment analysis.Computational Linguistics, vol. 37, no. 2, pp. 267-307.

(32)Totan, LS, Popescu, BB &Cristache, SE2013, The Impact of Unemployment on Economic Growth in Romania, During the Crisis. Romanian Statistical Review, vol. 6, pp. 32-40.

(33)Tukey, JW1977, Exploratory Data Analysis. Pearson. ISBN 978-0201076165.

(34)Turney, PD2002, Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, Philadelphia, Pennsylvania, pp. 417-424.

(35) Singha, A. K., Kumar, A., &Kushwaha, P. K. (2018). Patient Cohort Approaches to data science using Biomedical Field. *EPH-International Journal of Science And Engineering (ISSN: 2454-2016)*, *1*(1), 457-462.

(36)Tsay, RS2002, Analysis of Financial Time Series. John Wiley & Son, New York.

(37)Usha, M& DeviMI2013, Analysis of sentiments using unsupervised learning techniques.2013 International Conference on Information Communication and Embedded Systems, Tamilnadu, India, pp. 241-245.

(38) Singha, A. K., Kumar, A., &Kushwaha, P. K. (2018). Recognition of human layered structure using Gradient decent model. *EPH-International Journal of Science And Engineering (ISSN: 2454-2016)*, *1*(1), 450-456.

(39)Yuan, J, You, Q& Luo, J2015, Sentiment analysis using social multimedia.Multimedia Data Mining and Analytics, pp. 31-59.

(40)Zhu, C, Zhu, H, Ge, Y, Chen, E, Liu, Q, Xu, T, &Xiong, H2016,Tracking the evolution of social emotions with topic models.Knowledge Information Systems, vol. 47, pp. 517-544.