INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 05 | MAY - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

Impact of Covid 19 on Employment in India

SHUBHAM KUMAR
SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
Delhi, India
surajpandey9663@gmail.com

SHUBHAM
SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
Delhi, India
shubhamsingh60675@gmail.com

DR. SONIKA DAHIYA
SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
Delhi, India
sonika.dahiya11@gmail.com

***Abstract*— The COVID-19 pandemic severely impacted India's employment, causing a surge in unemployment. We surveyed various sectors and analyzed Twitter sentiments to understand the pandemic's effects on jobs and salaries, aiming to gain insights into individuals' prevailing mindsets during this crisis.**

Keywords—employment, surge in unemployment, twitter sentiments

I. INTRODUCTION

The COVID-19 pandemic significantly impacted global economies, particularly by exacerbating unemployment. In India, unemployment among educated individuals has been a major issue, with those not working at least 2184 hours annually classified as unemployed. The pandemic's rapid spread led to the closure of many industries and small businesses, causing widespread job losses. However, some sectors, like influencer marketing, online teaching, and YouTube content creation, thrived. The pharmaceutical industry, in particular, saw exponential growth, with increased demand for drugs, masks, and sanitizers. According to the IMF, the global GDP is expected to reach only 3%, the lowest since the Great Depression, highlighting the severe economic impact. Despite the challenges, COVID-19 also presented development opportunities, prompting significant investment in pharmaceutical research and development in India.

II. OBJECTIVE

A. *Motivation*

Sentiment analysis of COVID-19-related job tweets provides valuable insights into employment sentiments. This analysis helps identify trends, assess emotional well-being, and inform decision-making. By uncovering themes like job loss and remote work, it aids organizations, policymakers, and researchers in understanding the pandemic's impact and developing targeted support systems.

B. *Problem Statement*

The COVID-19 pandemic profoundly impacted employment, making it crucial to understand social media sentiments about this issue in India. The goal is to develop a sentiment analysis framework to assess the sentiment polarity (positive, negative, neutral) of tweets related to employment during the pandemic. Additionally, the framework should extract key themes and factors, such as job losses, salary impacts, and sector-specific effects. Quantifying sentiment intensity will help differentiate between mild and strong opinions, providing nuanced insights. To achieve a comprehensive understanding, sector-specific data should be collected from diverse sectors like healthcare, tourism, manufacturing, education, retail, and hospitality, ensuring a holistic view of the pandemic's impact on employment..

III. METHODOLOGY

The COVID-19 pandemic has significantly impacted employment globally, with profound effects in India. Sentiment analysis of social media data has emerged as a powerful tool to understand public sentiment and inform policy decisions. The reviewed studies highlight the importance of employing advanced analytical techniques and comprehensive data collection methodologies to capture the nuanced impacts of the pandemic on employment and public sentiment. These insights are crucial for developing targeted support systems and effective policy interventions to mitigate the adverse effects of such unprecedented events in the future.

A. *Google Survey*

A total of 1400 participants from working fields like Engineering, Finance, Marketing, Education, Corporate, General Administration, Professional services from different regions participated in the online questionnaire. Table 1 represent the demographic information of the participants. The mean age is 25-35 years (range, 21-65). The age of the participants was normally distributed ('21-30' year old, 61; '31-40' year old, 40; '41-50' year old,17; '51 years and above',21). 46.8% working in Delhi, 13.7% working in Noida and others from rest part of India.

Table 1: Demographic Data of the survey

AGE GROUP	NUMBER OF PERSONS
21-30	610
31-40	400
41-50	178
51-60	212

B. *Sentiment Analysis*

The Conclusion mining, also known as assumption investigation, is the area of focus in content mining that analyzes how people feel about a particular issue, an occasion, etc. Twitter serves as a persistent repository of data, providing an extensive collection of information that can be utilized for analysis purposes. Twitter API is used to collect data for an investigation. For analyzing the ideas contained in the information posted by diverse clients, used a Dictionary Based technique [4]. Then, this knowledge is arranged in its extreme form. Tweets collected after examinations are categorized into three classes: Positive, Negative, and Neutral.

A. Tools Available for Sentiment Analysis

1) *Tweepy*

Accessing the Twitter API can be accomplished using Tweepy, a Python library that is highly useful for creating Twitter bots and automations. Tweepy's,StreamListener object enables real-time monitoring and capturing of tweets.

2) *Text blob*

A Text blob is a highly regarded Python library for natural language processing (NLP). It is based on NLTK (Natural Language Toolkit) and offers a wide range of functions for various NLP tasks. These include part-of-speech tagging, language translation, text analysis, and sentiment analysis.

3) *Pandas*

Pandas is a powerful data frame library in Python. It offers a range of features, including innovative and flexible data structures, making data manipulation and analysis easy and efficient.

C. Analysis Of Twitter Data

The primary objective of analyzing Twitter sentiment is to classify numerous tweets into different categories based on their sentiment. By building a model, testing it, and refining it, several methods for approaching twitter come into being in this sector.

Following are the steps for sentiment analysis:

1. Different language ancestries: People from various cultures use words from their native tongues in their tweets, which may be promotional or slang.
2. Character Block Size Limitation: the quantity of data content that can be recognized is highly constrained with only 280 characters in scope.
3. Using hashtags: Twitter has hashtags for mentioning emotions, events, etc., however these hashtags must be processed separately from word-based tweets.

- Accept input: Consider any tweet. This tweet might consist of various labels, emotions, tags, and hashtags.
- Evaluation through libraries: In order to establish authentication with the Twitter API and generate access tokens for evaluating tweet functionality and sentiment, it is recommended to utilize libraries such as Tweepy.
- Preprocessing: In [1, 4], this step involves tweets' unstructured data, such as use rids, spaces, digits, and hashtags, is cleaned up. Following this phase, only the primary tweeted text is available for analysis.
- Algorithms: To train and test the datasets produced as a result of the preprocessing stage, a variety of methods can be utilized, including Naive Bayes, SVM etc. Below is an explanation of the aforementioned algorithms.

IV. RESULT

After Gathering of tweets from the internet using python and performed analysis and modelling work as follows: Followed the Data pre-processing and analysing classified collected tweets as positive negative and neutral. Four learning models: Naïve- Bayes, SVM, Random Forest, Adaboost are performed for finding best accuracy for sentiment analysis.

A. Different Learning Model Implementation

The data is collected from the twitter using Jupyter and total of 35000 tweets are fetched in which the only essential tweets are kept and rest are removed only those remain which have following keywords, “covid-19”,”jobs”, “layoff”, “salary”, “job lost during covid-19” and stored in the csv file, Followed the analysis is performed on the file and sentiments are divided such as positive, negative, and neutral.

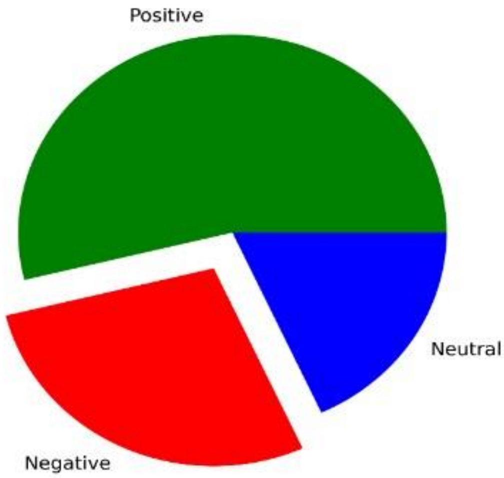

Fig 1. Pie chart depiction of total frequency of positive, negative, and neutral tweets of collected dataset.

Followed these sentiments are categorized as positive for 1 and negative and neutral for 0, after that first data is loaded and split it into training and test sets. After following these Word2Vec, TV-IDF and Count vectorizer embedding class from the library

+Model	Embeddings		
	TV-IDF	Count Vectorizer	Word 2 Vec
Naïve Bayes	71.9%	75.0%	**54.6%**
SVM	**75.7%**	**76.3%**	44.8%
AdaBoost	74.3%	74.6%	45.2%
Random Forest	55.9%	**80.1%**	44.8%

are used to convert the text into numeral vector from for analysing.

Next step, train four different learning models: SVM, Naive Bayes, AdaBoost with SVM base estimator, and Random Forest. use three embeddings as the input features for each classifier.

B. Output

Table 2: Depicts the accuracy of the learning models

Above all the best accuracy in tv-idf embedding was of SVM and in Count Vectorizer was of Random Forest and in Word 2 Vec was of Naïve Bayes. This show that Random Forest exhibits the superior performance when bind with count-vectorizer in impact of covid-19 on employment in India.

C. Conclusion

This thesis aimed to improve sentiment analysis accuracy for COVID-19-related tweets using deep neural networks and various classifiers. It employed natural language processing techniques, including word embedding methods like Count Vectorizer, TF-IDF, and Bag of Words, comparing their effectiveness. Results indicated that Random Forest performed best, particularly for large datasets. The analysis highlighted the pandemic's significant impact on employment and salary satisfaction across sectors, with many losing jobs, notably in tourism. To counteract the negativity on social media, proactive strategies are essential to provide concentrated comfort and support, enhancing user experience and well-being amidst the pandemic's challenges.

ACKNOWLEDGMENT

We extend heartfelt thanks to Dr. Ruchika Malhotra, Dr. Sonika Dahiya, and the Department of Software Engineering at Delhi Technological University for their guidance and support. Our gratitude also goes to everyone who assisted directly or indirectly, and to our well-wishers and parents for their unwavering support.

REFERENCES

[1] Bahrawi Bahrawi (2019): Sentiment Analysis using Radom Forest Algorithm Online Social Media Based PY - 2019/12/18, SP - h.29

[2] N´adia F. F. Silva, Eduardo R. Hruschka & Estevam Rafael Hruschka Jr. (2014): Tweet Sentiment Analysis with Adaptive Boosting Ensemble.

[3] Chirag Kariya & Priti Khodke (2020): Twitter Sentiment Analysis, 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020

[4] Adyan Marendra Ramadhani & Hong Soon Goo, “Twitter Sentiment Analysis using Deep Learning Methods”, Department of Management Information Systems Dong-A University Busan South Korea, 2017.

[5] Huma Parveen & Prof. Shikha Pandey “Sentiment Analysis on Twitter Data-set using Naive Bayes Algorithm”, Dept. of Computer Science and Engineering Rungta College of Engineering and Technology Bhilai, India , 2016.

[6] V. Kharde and S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques", International Journal of Computer Applications, vol. 139, p. 11, 2016.

[7] Nurulhuda Zainuddin & Ali Selamat (2014):Sentiment Analysis using Support Vector Machine SN- 978-3-540-42490-1.

[8] Mohd Akhter Ali, M.Kamraju (2020): A Study on Impact of Covid-19 Pandemic on Unemployment in India, SN - 978-81-945302-8-2

[9] Dimple Rawat, Vivek Dixit, Sarthak Gulati & Arti Gulati (2021): IMPACT OF COVID-19 OUTBREAK ON LIFESTYLE BEHAVIOUR: A REVIEW OF STUDIES PUBLISHED IN INDIA, PMCID: PMC7837201 PMID: 33493852.