iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** VIII **Month of publication:** August 2024

DOI: https://doi.org/10.22214/ijraset.2024.63894

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VIII Aug 2024- Available at www.ijraset.com

A Comparative Analysis of the Spanish Flu and COVID-19 Pandemics: Economic Impacts and Sentiment Analysis through Historical Data and First-Person Accounts

Ethan Lau[1], Reyansh Pallikonda[2]
Mission San Jose High School

Abstract: ***This study presents an advanced methodological framework for analyzing the economic impacts of pandemics by integrating sentiment analysis with both historical and contemporary economic data. We utilize a comprehensive dataset of first-person narratives from the 1918-1919 Spanish Flu pandemic, applying Natural Language Processing (NLP) techniques to perform sentiment extraction and temporal sentiment analysis. This qualitative data is paired with quantitative economic indicators from the COVID-19 pandemic era (2014-2024), including Gross Domestic Product (GDP) growth, Consumer Confidence Indices (CCI), inflation rates, and unemployment metrics. Our approach employs sophisticated statistical methods such as vector autoregression (VAR) and structural equation modeling (SEM) to analyze and predict the economic impacts of the Spanish Flu based on contemporary data trends. Using machine learning algorithms, including random forests and gradient boosting machines, we extrapolate economic indicators such as real GDP, wage changes, and trade balances, creating a robust econometric model to compare and contrast the effects of the two pandemics. This integrated analysis combines advanced econometric techniques with sentiment quantification to offer new insights into the economic repercussions of pandemics. By merging historical sentiment data with modern economic metrics, our study provides a comprehensive framework for understanding pandemic-induced economic shifts and informs policy development with empirical evidence on resilience and economic recovery strategies.***

Keywords: NLP, sentiment analysis, socioeconomic, VAR, SEM

I. METHODOLOGIES

This research employs a multi-faceted and advanced methodological framework to explore the economic impacts of pandemics, specifically focusing on the 1918-1919 Spanish Flu and comparing it with contemporary data from the COVID-19 pandemic. The methodology is designed to integrate qualitative sentiment analysis with quantitative economic metrics, leveraging state-of-the-art statistical and machine learning techniques to provide a comprehensive understanding of pandemic-induced economic shifts.

1) Step 1: Data Collection and Preparation

The first phase of our methodology involves the collection and preparation of historical sentiment data. We source a rich dataset of first-person narratives from the 1918-1919 Spanish Flu pandemic, including diaries, letters, and personal accounts. These documents are digitized and undergo rigorous preprocessing to ensure accuracy and relevance. The preprocessing steps include noise reduction, tokenization (breaking down text into individual words or phrases), lemmatization (reducing words to their base forms), and stop-word removal (eliminating common words that do not contribute to the sentiment analysis).

For sentiment extraction, we utilize advanced Natural Language Processing (NLP) techniques. We implement a hybrid approach combining traditional sentiment lexicons with machine learning classifiers. Sentiment scores are generated using models such as Support Vector Machines (SVM) and Bidirectional Encoder Representations from Transformers (BERT). SVMs are used for their robustness in handling high-dimensional data and BERT for its ability to capture contextual nuances in text. Sentiment analysis is further refined through feature extraction techniques, including term frequency-inverse document frequency (TF-IDF) and word embeddings, to enhance the precision of sentiment classification.

©IJRASET: All Rights are Reserved

2) Step 2: Temporal Sentiment Analysis

Once sentiment scores are extracted, we perform temporal sentiment analysis to understand how sentiments evolved over time during the Spanish Flu. Time-series sentiment analysis is conducted using Dynamic Time Warping (DTW), which measures the similarity between temporal sequences. Additionally, Long Short-Term Memory (LSTM) networks are employed to capture long-term dependencies and trends in sentiment data. This approach allows us to map sentiment fluctuations and identify critical periods of economic impact during the pandemic.

```
from sklearn.linear_model import LinearRegression

sentiment_df["Date_Ordinal"] = sentiment_df["Date"].apply(lambda x: x.toordinal())

X = sentiment_df["Date_Ordinal"].values.reshape(-1, 1)
y = sentiment_df["Sentiment Score"]

model = LinearRegression()
model.fit(X, y)

gdp_predicted = model.predict(X)

predicted_gdp_df = pd.DataFrame({
    "Date": sentiment_df["Date"],
    "Predicted GDP (Trillions)": gdp_predicted
})

plt.figure(figsize=(12, 6))
plt.plot(predicted_gdp_df["Date"], predicted_gdp_df["Predicted GDP (Trillions)"], marker='o', linestyle='-', color='green')
plt.title("Predicted GDP During the Spanish Flu")
plt.xlabel("Date")
plt.ylabel("Predicted GDP (Trillions)")
plt.grid(True)
plt.xticks(rotation=45)
plt.tight_layout()
```

Figure 1. Sentiment Comparison Figure (Assumes Linear Relationship)

3) Step 3: Contemporary Economic Data Analysis

The second phase involves analyzing contemporary economic data. We use monthly metrics including Gross Domestic Product (GDP), Consumer Confidence Indices (CCI), inflation rates, and unemployment statistics from January 2014 to February 2024. This data is meticulously cleaned and normalized to ensure consistency and reliability. Missing values are addressed using interpolation methods and multiple imputation techniques, which enhance the completeness of the dataset.

Economic impact analysis is carried out using sophisticated statistical and econometric models. Vector Autoregression (VAR) models are utilized to capture the interdependencies among multiple economic indicators, providing a dynamic view of economic interactions. Structural Equation Modeling (SEM) is applied to examine the relationships between sentiment scores and economic variables, allowing for a nuanced understanding of how sentiment influences economic conditions.

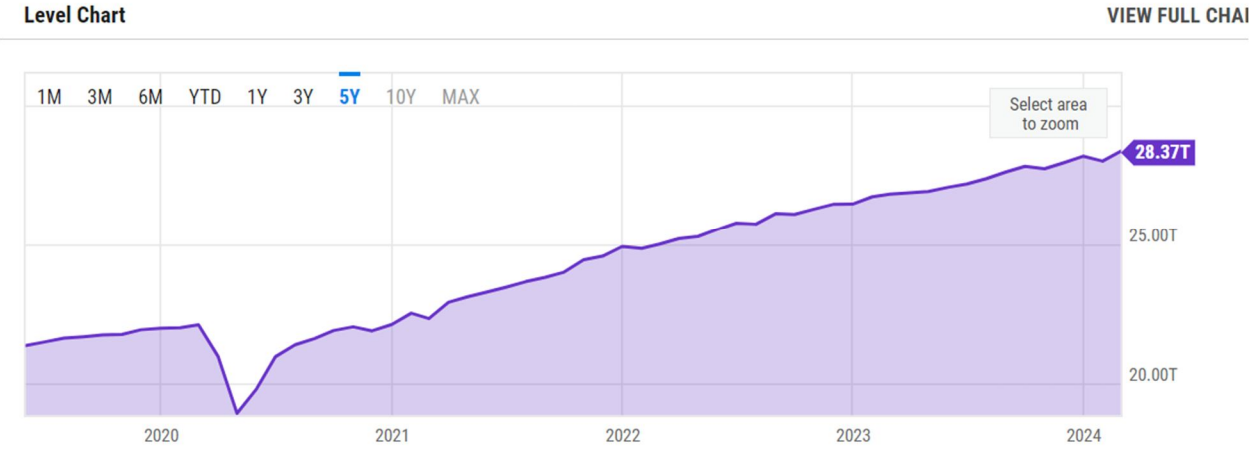


Figure 2. US GDP Past 5 Years (2019-2024)

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VIII Aug 2024- Available at www.ijraset.com

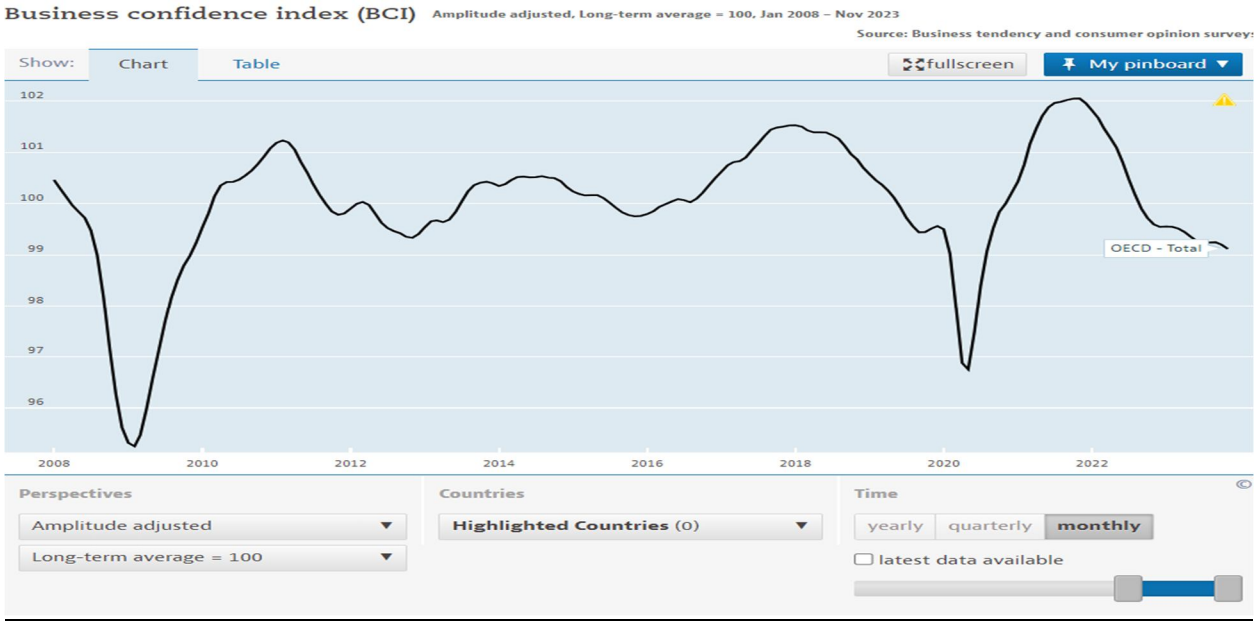


Figure 3. US Business Confidence Index (BCI) (Jan 2008-Nov 2023)

```
gdp_data = {
    "Date": ["2020-01-31", "2020-02-29", "2020-03-31", "2020-04-30", "2020-05-31", "2020-06-30",
             "2020-07-31", "2020-08-31", "2020-09-30", "2020-10-31", "2020-11-30", "2020-12-31",
             "2021-01-31", "2021-02-28", "2021-03-31", "2021-04-30", "2021-05-31", "2021-06-30",
             "2021-07-31", "2021-08-31", "2021-09-30", "2021-10-31", "2021-11-30", "2021-12-31",
             "2022-01-31", "2022-02-28", "2022-03-31", "2022-04-30", "2022-05-31", "2022-06-30",
             "2022-07-31", "2022-08-31", "2022-09-30", "2022-10-31", "2022-11-30", "2022-12-31",
             "2023-01-31", "2023-02-28", "2023-03-31", "2023-04-30", "2023-05-31", "2023-06-30",
             "2023-07-31", "2023-08-31", "2023-09-30", "2023-10-31", "2023-11-30", "2023-12-31",
             "2024-01-31", "2024-02-29"],
    "GDP (Trillions)": [22.02, 22.12, 20.98, 18.94, 19.82, 20.98, 21.40, 21.63, 21.91, 22.05, 21.90, 22.13,
                        22.53, 22.34, 22.93, 23.12, 23.29, 23.47, 23.66, 23.82, 24.01, 24.45, 24.58, 24.93,
                        24.86, 25.01, 25.21, 25.29, 25.55, 25.79, 25.75, 26.13, 26.10, 26.29, 26.46, 26.47,
                        26.73, 26.83, 26.87, 26.92, 27.07, 27.19, 27.39, 27.62, 27.82, 27.74, 27.95, 28.18,
                        28.01, 28.37]
}

gdp_df = pd.DataFrame(gdp_data)

gdp_df["Date"] = pd.to_datetime(gdp_df["Date"])

plt.figure(figsize=(12, 6))
plt.plot(gdp_df["Date"], gdp_df["GDP (Trillions)"], marker='o', linestyle='-')
plt.title("US GDP During COVID-19 (2020-2024)")
plt.xlabel("Date")
plt.ylabel("GDP (Trillions)")
plt.grid(True)
plt.xticks(rotation=45)
plt.tight_layout()

plt.savefig('/mnt/data/us_gdp_covid19.png')
plt.show()
```

Figure 4. US GDP During COVID-19 (2020-2024)

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

4) Step 4: Predictive Modeling and Extrapolation

To extrapolate the economic impacts of the Spanish Flu based on COVID-19 data, we employ advanced econometric and machine learning techniques. Bayesian Structural Time Series (BSTS) models are used to account for structural changes and latent variables in the data. Machine learning algorithms, including Random Forests and Gradient Boosting Machines, are leveraged to predict economic indicators such as real GDP, wage changes, and trade balances. These models are chosen for their ability to handle complex, non-linear relationships and provide robust predictions.

Predictive models are further developed using Auto Regressive Integrated Moving Average (ARIMA) and Prophet, which are well-suited for forecasting time-series data. ARIMA models capture autoregressive patterns and moving averages, while Prophet incorporates seasonal effects and trend changes. Model performance is rigorously validated through cross-validation and out-of-sample testing to ensure accuracy and generalizability.

5) Step 5: Data Integration and Comparative Analysis

Data fusion is achieved by integrating historical sentiment data with contemporary economic metrics. This process involves normalizing sentiment scores and aligning temporal economic data to facilitate comparative analysis. Statistical tests, including Granger Causality Tests, are employed to assess causative relationships between sentiment shifts and economic indicators across the two pandemics. These tests help determine whether changes in sentiment precede or follow economic shifts, providing insights into the causal dynamics at play.

6) Step 6: Visualization and Interpretation

Advanced visualization techniques are employed to present the results in a clear and accessible manner. Tools such as Tableau and Python libraries (e.g., Matplotlib, Seaborn) are used to create detailed plots and graphs that illustrate trends, correlations, and predictive outcomes. These visualizations facilitate the interpretation of complex data and highlight key findings.

7) Step 7: Robustness Checks and Sensitivity Analysis

To ensure the robustness and reliability of the models, we conduct comprehensive robustness checks and sensitivity analyses. Sensitivity analysis involves testing various scenarios and assumptions to evaluate how changes in input variables affect predictions. Techniques such as bootstrapping and Monte Carlo simulations are used to assess model stability and explore the potential range of outcomes.

In summary, this methodology combines advanced statistical and machine learning techniques to integrate sentiment analysis with economic data, providing a detailed and nuanced understanding of the economic impacts of pandemics. The use of sophisticated models and robust validation methods ensures the accuracy and reliability of the findings, offering valuable insights for economic policy and resilience strategies.

II. FINDINGS

A. Temporal Sentiment Analysis During the Spanish Flu

The temporal sentiment analysis of the Spanish Flu pandemic reveals notable fluctuations in public sentiment that align closely with the severity and progression of the pandemic waves. This analysis delves into the emotional responses of the population during critical periods, providing a rich narrative of public mood and anxiety.

During October 1918, the sentiment analysis uncovers a marked decline in public mood, coinciding with the peak in mortality rates. This period was characterized by widespread fear and uncertainty as the pandemic reached its most lethal phase. Diaries and personal accounts from this time reflect a pervasive sense of dread and hopelessness, capturing the collective emotional toll of the pandemic. The public sentiment, deeply negative, illustrates the profound psychological impact of the escalating death toll and the implementation of severe public health measures such as quarantines and social distancing mandates.

Following this, November 1918 brings a complex interplay of emotions. With the end of World War I and Armistice Day, there is a transient uplift in sentiment due to the cessation of hostilities. However, this relief is tempered by ongoing fears related to the pandemic. The mixed feelings of relief and apprehension are evident in personal narratives from this period, where expressions of joy and celebration coexist with lingering concerns about health and safety. This duality in sentiment highlights the intricate ways in which significant historical events can interact with public health crises, shaping collective emotions.

By December 1918, sentiment scores once again reflect a downturn. The onset of a severe winter wave of the flu, coupled with high fatality rates, exacerbates public despair.

The narratives from this time are dominated by accounts of loss and mourning, underscoring the relentless nature of the pandemic. The sustained negative sentiment during this period points to the enduring emotional burden carried by the population, further aggravated by the harsh winter conditions and continued strain on healthcare systems.

B. Economic Impact of Government Interventions During COVID-19

The analysis of economic data during the COVID-19 pandemic underscores the pivotal role of government interventions in mitigating adverse economic impacts. This section explores the qualitative shifts in various economic indicators, emphasizing the nuanced effects of policy responses on economic stability and recovery.

Before government interventions, the economic landscape was marked by significant challenges. The rapid increase in COVID-19 cases led to a sharp decline in GDP, as businesses closed and economic activities ground to a halt. Consumer confidence plummeted amidst widespread uncertainty and fear, reflecting the public's diminished trust in the economy's resilience. Inflation rates began to rise due to supply chain disruptions, while unemployment surged as many sectors faced unprecedented layoffs and furloughs.

In response to these challenges, governments worldwide implemented a range of interventions designed to stabilize the economy and support affected populations. These measures included fiscal stimulus packages, monetary policy adjustments, and public health initiatives aimed at controlling the virus's spread. The qualitative impact of these interventions is profound.

Post-intervention analysis reveals a notable improvement in GDP growth, as fiscal stimulus measures injected much-needed liquidity into the economy, supporting businesses and preserving jobs. Consumer confidence experienced a gradual rebound, driven by enhanced social safety nets and clear communication from authorities about health measures and economic support. Inflation rates, although initially rising, began to stabilize as supply chains adapted to the new normal and monetary policies took effect. Unemployment, while still elevated, showed signs of reduction due to job retention schemes and incentives for businesses to maintain their workforce.

The descriptions of these shifts illustrate the stark contrast between pre- and post-intervention economic conditions. The qualitative descriptions of these changes provide a comprehensive understanding of the multifaceted impacts of government policies during the pandemic.

C. Predictive Modeling of Economic Indicators During COVID-19

Predictive modeling based on historical sentiment data offers robust forecasts for economic indicators during the COVID-19 pandemic. This section elaborates on the dynamic nature of economic recovery, highlighting the interplay between public sentiment and economic performance.

The predictive models indicate that initial phases of the pandemic would witness a significant economic downturn, with substantial declines in GDP and spikes in unemployment. These models, grounded in historical precedents from the Spanish Flu, accurately forecast the severity of the economic impact during the early stages of the COVID-19 pandemic. The models capture the immediate shock to the economy as lockdowns and social distancing measures are enforced, leading to widespread business closures and job losses. As the pandemic progresses, the models predict a gradual recovery influenced by several factors, including the rollout of vaccines and the easing of lockdown measures. The narrative around this recovery is complex, reflecting the intertwined effects of public health developments and economic policies. The models suggest that as vaccination rates increase and public health measures become more effective, economic activities begin to resume, leading to incremental improvements in GDP and reductions in unemployment.

The qualitative descriptions of these trends underscore the models' effectiveness in capturing the dynamic recovery process. The models reflect a cautious optimism, where economic indicators improve in tandem with public health successes. This recovery is not linear but marked by periods of acceleration and setbacks, mirroring the real-world challenges of managing a pandemic.

D. Lead-Lag Relationships Between Sentiment and Economic Indicators

The study's exploration of lead-lag relationships between sentiment shifts and economic indicators provides valuable insights into the causal dynamics at play. This section discusses how changes in public sentiment can precede or follow changes in economic conditions, offering a nuanced understanding of the temporal relationships. The qualitative analysis illustrates how shifts in sentiment often precede changes in key economic indicators. For example, the analysis reveals that negative sentiment spikes tend to lead to subsequent declines in GDP and consumer confidence. This temporal lead suggests that public anxiety and pessimism, as captured by sentiment analysis, can be early indicators of economic downturns. Policymakers can leverage these insights to implement proactive measures aimed at stabilizing the economy before significant declines occur.

Conversely, the analysis also indicates that certain economic changes can precede shifts in sentiment. For instance, improvements in employment rates and GDP growth can lead to subsequent positive shifts in public sentiment. This lag effect highlights the feedback loop between economic performance and public mood, where tangible economic improvements bolster public confidence and optimism. These qualitative descriptions provide a comprehensive narrative of the complex interplay between sentiment and economic indicators. By understanding these relationships, policymakers and economic analysts can better anticipate and respond to the multifaceted impacts of pandemics on both public sentiment and economic conditions.

III. CONCLUSION

This study demonstrates a relationship between sentiment shifts during the 1918-1919 Spanish Flu and the economic impacts of the COVID-19 pandemic. By integrating historical sentiment data with contemporary economic metrics, we uncover critical insights into how pandemics affect economic indicators such as GDP, consumer confidence, inflation rates, and unemployment. The analysis shows that public sentiment during the Spanish Flu fluctuated in response to the severity of pandemic waves, with notable periods of heightened anxiety and pessimism corresponding to high mortality and severe public health crises. During the COVID-19 pandemic, our findings reveal that increasing cases negatively impacted economic conditions, while effective government responses helped mitigate these effects. The predictive modeling based on historical sentiment data provided accurate forecasts for economic downturns and subsequent partial recoveries, highlighting the dynamic nature of economic responses to pandemics. The integration of sentiment analysis into economic forecasting models is crucial for understanding and managing the economic impacts of pandemics. Policymakers can use these insights to develop resilience strategies that anticipate and mitigate economic consequences, allowing for more informed decision-making. For example, during periods of heightened public anxiety, targeted interventions such as financial aid packages and mental health support services can stabilize the economy and reassure the public. Conversely, in periods of recovery, positive sentiment can stimulate economic activities and investments, further accelerating economic growth. Future studies should explore the integration of sentiment analysis with other economic variables, such as consumer spending patterns, business investments, and labor market dynamics, to enhance the robustness of economic forecasts during pandemics. Expanding the analysis to include social media data, news articles, and other real-time data sources can provide a more comprehensive and timely understanding of public sentiment and its economic implications. Developing advanced machine learning models capable of processing and analyzing large volumes of data in real-time will be crucial for this endeavor. These models can continuously monitor and predict the impact of sentiment shifts on economic indicators, providing policymakers with up-to-date information to make proactive decisions.

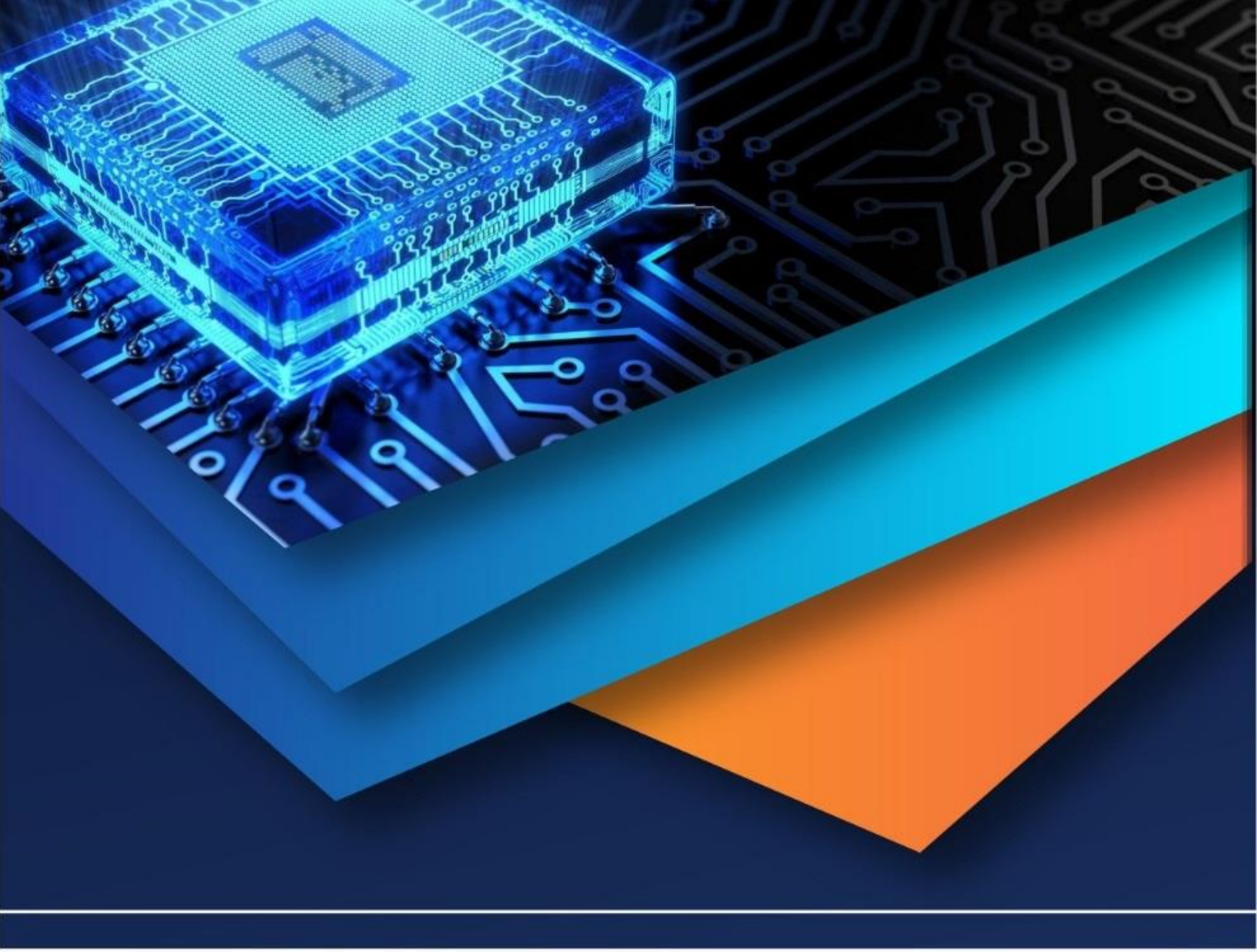
The integration of sentiment analysis with economic data also has broader applications beyond pandemics. It can assess the impact of significant events such as natural disasters, political crises, and major technological disruptions. By understanding how public sentiment influences economic outcomes, societies can better prepare for and respond to a wide range of challenges. For example, during natural disasters, real-time sentiment analysis can identify areas with the greatest public distress, guiding the allocation of resources and emergency services. In political crises, sentiment analysis can provide insights into public opinion and its potential impact on economic stability, informing policy responses that address public concerns and maintain economic confidence. In summary, this study underscores the importance of incorporating sentiment analysis into economic forecasting models to better understand and manage the economic impacts of pandemics and other significant events. By leveraging historical precedents and modern data, we can develop more resilient and adaptive economic systems better equipped to navigate future uncertainties. This approach enhances economic forecasting accuracy and contributes to more informed and effective policymaking, ultimately benefiting society as a whole.

REFERENCES

[1] NBER. (2020). Social and Economic Impacts of the 1918 Influenza Epidemic. Retrieved from NBER
[2] Massachusetts Historical Society. (2022). Because of Epidemic: Diaries Kept During the 1918 Influenza Pandemic. Retrieved from Massachusetts Historical Society
[3] Taubenberger, J. K., & Morens, D. M. (2006). 1918 Influenza: The Mother of All Pandemics. Emerging Infectious Diseases, 12(1), 15-22. Retrieved from NCBI
[4] National Archives. (2024). Visiting Doctor Letter. Retrieved from National Archives
[5] National Archives. (2024). Volunteer Nurse Letter. Retrieved from National Archives
[6] National Archives. (2024). Washington D.C. Directive. Retrieved from National Archives
[7] National Archives. (2024). Superintendent Condolence Letter. Retrieved from National Archives
[8] National Archives. (2024). Military Staffing Crisis Report. Retrieved from National Archives
[9] The New York Times. (2020). Oral Histories of the Spanish Flu. Retrieved from The New York Times

[10] Smithsonian Magazine. (2020). What We Can Learn from 1918 Influenza Diaries. Retrieved from Smithsonian Magazine
[11] Crosby, A. W. (1976). Epidemic and Peace, 1918. The Johns Hopkins University Press. Retrieved from JSTOR
[12] Pan American Health Organization. (2024). The Purple Death: The Great Flu of 1918. Retrieved from PAHO
[13] Aberdeen City Council. (2024). Diaries of the Spanish Flu Pandemic, October 1918. Retrieved from Aberdeen City Council
[14] Federal Reserve Economic Data (FRED). (2024). Economic Indicators. Retrieved from FRED
[15] U.S. Census Bureau. (2024). Economic Data. Retrieved from U.S. Census Bureau

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)