ECONOMIC EFFECTS OF RUSSIA-UKRAINE WAR: A SENTIMENT ANALYSIS APPROACH

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

MASTER OF SCIENCE IN MATHEMATICS

by

Garima Singh
(2K23/MSCMAT/18)

Isha Ojha
(2K23/MSCMAT/52)

Under the Supervision of
Mrs. Sumedha Seniaray
Department of Applied Mathematics

Department of Applied Mathematics
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

CANDIDATE'S DECLARATION

We, Garima Singh, 2K23/MSCMAT/18 & Isha ojha, 2K23/MSCMAT/52 students of M. Sc. Applied Mathematics, hereby certify that the project Dissertation titled "Economic Effects of Russia-Ukraine War: A Sentiment Analysis Approach" which is submitted by us to the Department of Applied Mathematics, Delhi Technological University in partial fulfillment of the requirements for the award of the Degree of Masters of Science, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Garima Singh

Date:

Isha Ojha

CERTIFICATE BY THE SUPERVISOR

Certified that Garima Singh, 2K23/MSCMAT/18 & Isha ojha, 2K23/MSCMAT/52 has carried out their research work presented in this thesis entitled "Economic Effects of Russia-Ukraine War: A Sentiment Analysis Approach" for the award of Masters of Science from Department of Applied Mathematics, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the students themselves and the contents of the thesis do not form the basis for the award of any other degree to the candidates or to anybody else from this or any other University/Institution.

Date :

Place: Delhi

Mrs. Sumedha Seniaray

Assistant Professor

Department of Applied Mathematics

Delhi Technological University

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road,

Delhi-42

ABSTRACT

The Russia-Ukraine war has had far-reaching economic consequences, significantly influencing both nations' macroeconomic indicators and public sentiment. This study investigates these economic impacts by integrating multi-source sentiment analysis with macroeconomic data spanning from 2015 to 2025 for Russia and Ukraine. Economic indicators were sourced from the World Bank and the International Monetary Fund, including consumer price index (CPI), inflation rates, GDP growth, unemployment, government debt, trade volumes, foreign direct investment (FDI) inflows, and military expenditures. Missing data for select years were estimated using time series forecasting models: Vector Autoregression (VAR), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Prophet, with SARIMA yielding the most reliable forecasts for Ukraine and a combination of SARIMA and VAR performing best for Russia. In parallel, sentiment data were extracted from over 500,000 social media posts, news articles, and Reddit comments relating to the war. These texts were preprocessed and analyzed using multiple lexicon-based sentiment tools including TextBlob, VADER, AFINN, and SentiWordNet. Annual sentiment scores were calculated and then correlated with economic indicators using Pearson and Spearman correlation coefficients, Granger causality tests, and visual trend analyses. Ukraine's economy shows heightened sensitivity to inflation and trade disruptions, with sentiment reacting more rapidly to economic changes. Russia, on the other hand, exhibits a more delayed sentiment response, aligning with its relatively more controlled economic structure. Granger causality tests confirm a more immediate influence of sentiment on Ukraine's economic variables, whereas in Russia, such impacts surface more gradually. War-related spending is a common inflation driver in both nations, though Ukraine demonstrates a heavier reliance on debt-financed defense efforts. Machine learning models were applied to assess predictive performance, with XGBoost outperforming Random Forest overall, especially in modeling Russia's indicators. In contrast, Random Forest showed a slight edge in predicting Ukraine's economic trends. Lagged correlation analyses reinforce sentiment's predictive value, particularly during periods of conflict. Post-war analysis indicates that Ukraine is experiencing a faster, though more volatile, recovery compared to Russia's steadier but slower rebound. This research highlights the importance of sentiment in economic forecasting during geopolitical crises. Future work could enhance predictive accuracy through the integration of deep learning-based natural language processing models and explore the role of sentiment-informed economic policy in managing shocks during conflicts.

ACKNOWLEDGEMENT

We would like to express our deepest gratitude to our supervisor, Mrs. Sumedha Seniaray, for her invaluable guidance, support, and encouragement throughout the course of this research. Her insights and feedback were instrumental in shaping the direction and quality of this work.

We are sincerely thankful to the faculty and staff of the Department of Applied Mathematics for providing a stimulating academic environment and for their assistance whenever needed. A heartfelt thanks goes to our family and friends for their unwavering support, patience, and belief in us throughout this endeavor. Their love and encouragement kept us grounded and motivated through the challenges of this research.

Finally, I am grateful for the availability of data and tools from institutions such as the World Bank, IMF, and various open-source sentiment analysis libraries, which made this research possible.

This thesis would not have been possible without the collective support of all those mentioned above.

Garima Singh

Isha Ojha

TABLE OF CONTENT

Title	Page No.
Declaration	ii
Certificate	iii
Abstract	iv
Acknowledgement	v
Content	vi-vii
List of Figures	viii-ix
List of Tables	x
List of Abbreviations	x
CHAPTER 1: INTRODUCTION	1-4
1.1 Russia-Ukraine War	1
1.2 Sentiment Analysis	1
1.3 Economic Factors	2
1.4 Correlation between sentiments and economic factors	2
1.5 Motivation and objective	3
CHAPTER 2: LITERATURE REVIEW	5-6
2.1 Sentiment Analysis on the Russia-Ukraine War	5
2.2 Sentiment and Economic Indicator	5
2.3 Sentiment Analysis in Economic Forecasting	6
CHAPTER 3: METHODOLOGY	7-21
3.1 Phase I : Economic data collection and Forecasting for correlation Analysis	7
3.1.1 Data Collection	7
3.1.2 Missing Data Implementation using Time Series Model	8
3.2 Phase II : Sentiment Analysis on Public Relation to War	11

3.2.1 Data Collection 11
3.2.2 Text Preprocessing 11
3.2.3 Sentiment Scoring 12
3.3 Phase III : Correlation Analysis Between Sentiment & Economic Indicators 14
3.3.1 Visualization & Statistical Testing 14
3.3.2 XGBoost and Random Forest for Historical Validation using Sentiments 18
3.3.3 Hypothesis Testing 20
CHAPTER 4: RESULT AND DISCUSSION 22-37
4.1 Filling missing economic data 22
4.2 Aggregated Sentiments 26
4.3 Correlation Analysis: Spearman Heatmaps 26
4.4 Granger Causality Test 28
4.5 Comparing XGBoost and Random Forest for Historical Validation Using Sentiments 29
4.6 Trend Analysis 30
4.7 Performance Matric 34
4.8 Hypothesis Testing results 35
CHAPTER 5: CONCLUSION & FUTURE SCOPE 37
REFERENCE 38- 39

LIST OF FIGURES

Pg. No

Fig. 3.1 : Proposed Methodology Flow chart 7

Fig 3.2 : Code for missing economic data prediction of Ukraine using VAR 8

Fig 3.3 : Code for missing economic data prediction of Ukraine using PROPHET 9

Fig 3.4 : Code for missing economic data prediction of Ukraine using SARIMA 10

Fig 3.5 : Code for calculating yearly aggregate sentiment score using TEXTBLOB 13

Fig 3.6 : Code for calculating yearly aggregate sentiment score using VADER 13

Fig 3.7 : Code for calculating yearly aggregate sentiment score using AFINN 14

Fig 3.8 : Code for calculating yearly aggregate sentiment score using SentiWordNet 14

Fig 3.9 : Code for plotting spearman correlation heatmap 15

Fig 3.10 : Code for Granger Causality 16

Fig 3.11 : Code for plotting Line Plot for trend analysis 17

Fig 3.12 : Code for calculating Pearson Correlation for lagged sentiment on different war phases. 18

Fig 3.13 : Code for Historic validation of economic indicators using XGBoost model and sentiment as feature 19

Fig 3.14 : Code for Historic validation of economic indicators using Random Forest model and sentiment as feature 20

Fig 4.1 : Spearman Correlation Heatmap for Ukraine 27

Fig 4.2 : Spearman Correlation Heatmap for Russia 27

Fig 4.3 : XGBoost Model validation and MAE score for Ukraine 29

Fig 4.4 : Random Forest Model validation and MAE score for Ukraine 29

Fig 4.5 : XGBoost Model validation and MAE score for Russia 29

Fig 4.6 : Random Forest Model validation and MAE score for Ukraine 29

Fig. 4.7 Line plot for trend analysis of economic indicator vs Sentiment for Ukraine 30

Fig 4.8 : Pearson correlation analysis with lagged sentiments on different war phases for Ukraine 31

Fig 4.9 : Pearson correlation analysis with lagged sentiments on different war phases for Russia 32

Fig 4.10 : Line plot for Russia 33

LIST OF TABLES

	Pg. No
Table 4.1 : Final economic data of Russia	24
Table 4.2 : Final economic data of Ukraine	25
Table 4.3 : Yearly aggregated Sentiment Scores	26
Table 4.4 : Granger Causality Test for Ukraine	28
Table 4.5 : Granger Causality Test for Russia	28
Table 4.6: Final performance metric table	34-35

LIST OF ABBREVIATIONS

IMF : International Monetary Fund

VADER : Valence Aware Dictionary and sEntiment Reasoner

SARIMA : Seasonal AutoRegressive Integrated Moving Average

VAR : Vector AutoRegression

XGBoost : eXtreme Gradient Boosting

GDP : Gross Domestic Product

MAE: Mean Absolute Error

RMSE : Root Mean Square Error

MAPE : Mean Absolute Percentage Error

CHAPTER - 1
INTRODUCTION

In today's digitally connected world, economic factors and sentiments shared on platforms like Twitter, Reddit, and news blogs are deeply intertwined, often shaping and reflecting one another. Social media has become a powerful space where public opinion on inflation, unemployment, trade, and government policies is voiced instantly and widely, influencing consumer behavior, investor sentiment, and even policy responses. News blogs add further depth by providing analyses that frame public understanding of complex economic issues. During times of war, such as the Russia-Ukraine conflict, this relationship intensifies. Real-time updates, emotional reactions, and opinionated discussions flood social media and online news, amplifying public anxiety, uncertainty, and speculation. These sentiments can lead to rapid shifts in financial markets, currency fluctuations, disrupted trade patterns, and altered investor confidence. In such scenarios, analyzing online sentiment becomes a crucial tool for understanding the broader economic impact of conflict, offering early signals of economic instability or resilience.

1.1 Russia-Ukraine War

The Russia-Ukraine war, which began with Russia's full-scale invasion in February 2022, has turned into a prolonged and devastating conflict. Initially aimed at a quick takeover, Russia faced fierce Ukrainian resistance, supported heavily by Western countries through military aid and sanctions on Moscow. Over time, the war evolved into a war of attrition, with intense battles concentrated in eastern and southern Ukraine. As of 2025, the front lines have remained largely static, with both sides suffering heavy losses. Ukraine has ramped up its domestic weapons production but still relies on international support. Russia, meanwhile, has strengthened ties with countries like North Korea and Iran for military and logistical aid. Diplomatic efforts to end the war have failed so far, with peace talks stalling over demands for territorial concessions and security guarantees. The war has created a massive humanitarian crisis, displacing millions and crippling Ukraine's economy. It has also reshaped global geopolitics, deepening divides between Russia and the West while drawing new battle lines in international relations. Despite global calls for peace, there is little indication the war will end soon.

1.2 Sentiment Analysis

The collective emotional and attitudinal response of a population—is increasingly being recognized as a significant factor in economic behavior. Sentiment can influence consumer confidence, investment decisions, currency stability, and even government policy. In times of war, when uncertainty is high and information is abundant yet fragmented, sentiment becomes an even more important lens through which economic activity can be understood. The rise of social media and digital communication platforms has made it possible to capture and quantify public sentiment on a large scale, offering researchers new tools to analyze the socio-economic landscape in real time. In this context, sentiment analysis, a technique within natural language processing (NLP) which serves as a powerful method to study the relationship between economic performance and public perception during the Russia-Ukraine war.

Sentiment analysis of platforms like Twitter, Reddit, and news blogs provides valuable insights into public opinion and emotional response surrounding major events such as the Russia-Ukraine war. On Twitter, short, real-time posts often reflect immediate reactions to breaking news, revealing spikes in negative sentiment during major attacks or political announcements. Reddit discussions tend to be more in-depth, capturing a broader range of perspectives, including debates, support for Ukraine, and criticisms of international responses. News blogs, while generally more structured and formal, often carry the tone of the media outlet's editorial stance, ranging from neutral reporting to strongly opinionated commentary. By analyzing sentiment across these platforms, researchers can track how public mood shifts over time, identify key concerns of different communities, and correlate emotional trends with economic or geopolitical developments.

1.3 Economic Factors

Economic factors play a crucial role during wartime, significantly influencing both the immediate and long-term stability of the countries involved. Wars often lead to large-scale destruction of infrastructure, disruption of trade routes, and massive defense spending, all of which strain national budgets and increase public debt. Inflation typically rises due to supply chain disruptions and shortages of essential goods, while unemployment can surge as businesses shut down or relocate. Currency devaluation and reduced investor confidence further destabilize the economy. In the context of the Russia-Ukraine war, both nations have experienced economic shocks, Ukraine due to infrastructure loss and displacement, and Russia due to international sanctions. These factors can have cascading effects, such as reduced GDP growth, increased poverty, and slower post-war recovery. Moreover, global markets can also feel the ripple effects, especially in sectors like energy, food, and commodities, due to the geopolitical importance of the region.

1.4 Correlation between Sentiment Score and Economic Factors

This study aims to explore the intersection of sentiment and economic indicators by leveraging both quantitative economic data and qualitative sentiment data from various online sources. By analyzing macroeconomic indicators such as inflation, gross domestic product (GDP), foreign direct investment (FDI), and military expenditure alongside sentiment data from tweets, news articles, and online forums, this research seeks to construct a comprehensive picture of the economic dynamics in Russia and Ukraine between 2015 and 2025. Notably, the years 2024 and 2025 are forecasted using time series models due to the unavailability of complete data, ensuring a continuous timeline for analysis.

A distinctive feature of this research is its integration of traditional econometric modeling and modern machine learning techniques. While time series models such as Vector Autoregression (VAR), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Prophet are used to forecast economic indicators, multiple lexicon-based sentiment tools: TextBlob, VADER (Valence Aware Dictionary and sEntiment Reasoner), AFINN, and SentiWordNet are used to extract sentiment scores from over 500,000 text entries. This hybrid methodology enables the study to not only forecast economic trends but also investigate the causal and correlational relationships between sentiment and macroeconomic factors through techniques such as Pearson and Spearman correlation matrices, Granger causality tests, and visual trend analysis.

1.5 Motivation and Objective of the Research

The motivation for this research arises from the growing importance of understanding how public sentiment and economic variables influence each other, particularly during periods of extreme crisis such as war. The conflict between Russia and Ukraine has created an environment where economies are pushed to their limits, and public opinion is volatile and reactive. Understanding how these two domains interact can offer important insights for economists, policymakers, and international organizations involved in conflict recovery and economic stabilization. For instance, if sentiment proves to be a leading indicator of economic decline or recovery, then sentiment analysis could serve as an early warning tool for governments and financial institutions.

Furthermore, the comparative aspect of this research offers additional value. Russia and Ukraine have responded to the war under vastly different economic and political conditions. Russia's economy is larger, more centralized, and more insulated due to its energy exports and autarkic policies, while Ukraine has relied heavily on international aid, borrowing, and public mobilization. As such, the economic shocks and public responses in the two countries provide contrasting case studies that can illuminate broader principles of wartime economic resilience and recovery. The differences in how sentiment correlates with economic indicators across the two countries shed light on the underlying structural and institutional factors that govern economic behavior during crises.

The primary objective of this study is to analyze the economic effects of the Russia-Ukraine war through a sentiment analysis approach, integrating large-scale textual data with macroeconomic indicators over a ten-year period (2015–2025). Specifically, the research aims to:

- Compare and contrast the sentiment-economy dynamics between Russia and Ukraine, highlighting differences in responsiveness, volatility, and recovery patterns.
- Evaluate the predictive capacity of sentiment for macroeconomic forecasting using machine learning models such as Random Forest and XGBoost.
- Suggest practical applications of sentiment-informed economic forecasting in policy making, crisis management, and post-conflict economic planning.

By fulfilling these objectives, the study seeks to contribute to the growing body of literature at the intersection of economics, sentiment analysis, and conflict studies. It also aspires to inform future research directions in using natural language processing and machine learning for socio-economic modeling in times of crisis.

CHAPTER - 2
LITERATURE REVIEW

The intersection of sentiment analysis and economic forecasting has gained increasing attention in recent years, particularly in the context of financial markets and macroeconomic indicators. Previous studies have demonstrated that public sentiment, as captured through social media, news articles, and economic reports, can significantly influence economic performance. This section reviews existing research on sentiment analysis in economic modeling, statistical causality tests, and machine learning-based economic prediction.

2.1 Sentiment Analysis on the Russia-Ukraine War

Sentiment analysis has been widely used to assess public perception and its economic implications. Sahi et al. (2025) analyzed Twitter discussions using machine learning. Similarly, Sinha et al. (2024) explored the temporal shifts in sentiment using logistic regression, decision trees, and K-nearest neighbors, highlighting changes in public mood as the war evolved. Extending sentiment analysis to more advanced models, Menaouer et al. (2025) utilized knowledge graph convolutional networks for analyzing war-related tweets, offering a novel approach to sentiment classification.

2.2 Sentiment and Economic Indicators

The correlation between sentiment trends and economic indicators has been explored in various studies. Polyzos (2023) evaluated the influence of over 42 million tweets on 15 key macro-financial variables, revealing that negative war related sentiment triggers immediate declines in European market indices while boosting safe haven assets like the U.S. dollar. Aygun et al.(2025) applied aspect-based sentiment analysis to measure the war's impact on global food security, while Abakah et al.(2023) examined blockchain and FinTech stocks, finding strong correlations between war sentiment and market fluctuations. Our findings further emphasize how public perception correlates with macroeconomic stability, suggesting that sentiment could serve as an early warning system for financial instability. Bollen et al.(2011) demonstrated its influence on stock markets.

Other works have used sentiment and uncertainty metrics as proxies for broader economic indicators. Grebe et al. (2024) examined how uncertainty related to the war influenced the German economy, whereas Sulong et al.(2023) investigated the role of public sentiment toward sanctions on the G7 debt markets. Izzeldin et al compared market responses to the war against the COVID-19 pandemic and the 2008 financial crisis, noting that commodities like wheat and nickel were particularly affected. (Bruhin et al.(2023) further explored these economic impacts within the European context.

2.3 Sentiment Analysis in Economic Forecasting

Beyond war related studies, sentiment analysis has been used for economic forecasting. Algaba et al. provided a comprehensive methodological overview and outlined econometric methods for converting sentiment into economic indicators, while Lukauskas et al. (2022) demonstrated the utility of media sentiment in forecasting inflation and GDP. Chong et al.(2022) and Seki et al.(2022) extended this notion by applying news sentiment indices to national economic forecasting. Classic works such as Baker et al.(2016) and Tetlock & Paul C.(2007) also provide foundational insights into how policy uncertainty and media sentiment influence financial markets. Lastly, Lia et al.(2018) showed the predictive power of Twitter sentiment on cryptocurrency price fluctuations, further emphasizing the financial implications of public opinion in digital contexts.

To the best of our knowledge, this study offers a more comprehensive approach than prior research by integrating sentiment trends and macroeconomic indicators, rather than analyzing them in isolation.While prior research has successfully linked sentiment analysis with economic trends, few studies have examined how different sentiment analysis techniques perform across multiple war phases. Additionally, existing literature lacks a comprehensive comparison of machine learning models for sentiment-driven economic prediction in the context of geopolitical crises. This study bridges these gaps by analyzing the evolving relationship between sentiment and economic indicators across distinct time periods (pre-war, peak war, post-peak war), employing Granger causality to assess short-term and long-term sentiment effects, and evaluating the predictive performance of XGBoost and Random Forest in forecasting economic conditions using sentiment data.

CHAPTER - 3
METHODOLOGY

This study integrates time series forecasting, sentiment analysis, and correlation analysis to examine the economic effects of the Russia-Ukraine war. The proposed method has been implemented into three phases: economic data preparation, sentiment analysis, and correlation analysis depicted by Fig. 3.1

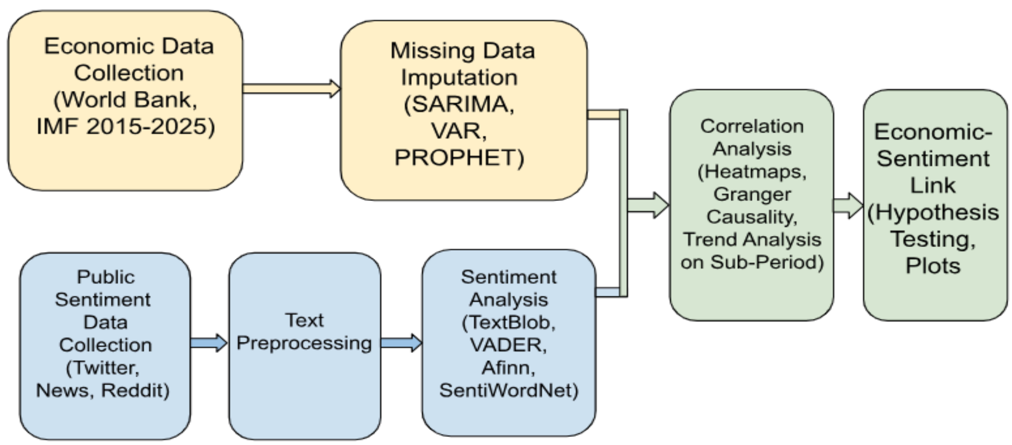


Fig. 3.1 Proposed Methodology Flow chart

3.1 Phase I : Economic Data Collection and Forecasting for Correlation Analysis

Economic data (2015–2025) for Russia and Ukraine were sourced from the World Bank and IMF, with missing values filled using the best-fitting time series models—SARIMA, VAR, and Prophet.

3.1.1 Data collection:-

Yearly economic factor data for Russia and Ukraine (2015-2025) was collected from World Bank and International Monetary Fund (IMF) websites.Each country's dataset was stored in a CSV file with the following columns: Consumer Price Index (CPI), Inflation (%change), GDP Growth (% change), Unemployment Rate (%of labor force), Government Debt (% of GDP), Trade (%of GDP), Foreign Direct Investment (FDI) Net

Inflows (BoP,USD), Military Expenditure (Local Currency Unit - LCU). However, data for Trade, FDI, and Military Expenditure (2024-2025) was missing for both countries, and Russia's CPI (2022-2023) was also unavailable.

3.1.2 Missing Data Implementation Using Time Series Model:-

In the extracted dataset there are few economic factors in both Russia and Ukraine which have missing values. To fill in missing values, three time series models were tested:Vector Autoregression (VAR), Seasonal Autoregressive Integrated Moving Average (SARIMA), and PROPHET , code for which is shown in Fig.3.2, Fig.3.3, and Fig.3.4 respectively.

- Using VAR for Ukraine economic data prediction:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.api import VAR
```

```
# Load data
file_path = "/content/sheetUkr (3).xlsx"
xls = pd.ExcelFile(file_path)
ukraine = pd.read_excel(xls, sheet_name="Ukraine")
```

```
# Selecting relevant columns for VAR model
columns_to_forecast = ["CPI()", "Trade (% of GDP)", "FDI net inflows (BoP, current US$)", "Military expenditure (current LCU)"]
df_var = ukraine[["Year"] + columns_to_forecast].dropna()
# Set year as index for VAR
df_var.set_index("Year", inplace=True)
# Fit VAR model with adjusted lag selection
model = VAR(df_var)

try:
    # Automatically select lag order (reducing maxlags to prevent error)
    selected_lag = model.select_order(maxlags=min(1, len(df_var)-1)).bic  # Ensuring a small lag
    var_model = model.fit(maxlags=selected_lag)
except ValueError:
    print("Reducing maxlags due to small dataset")
    var_model = model.fit(maxlags=1)  # Fallback to maxlag=1 if needed
# Forecast for 2024 and 2025
forecast_steps = 2
forecast_input = df_var.values[-var_model.k_ar:]  # Use only required past values
forecast = var_model.forecast(forecast_input, steps=forecast_steps)
# Create DataFrame for forecasted years
forecast_years = [2024, 2025]
df_forecast = pd.DataFrame(forecast, index=forecast_years, columns=columns_to_forecast)
# Update original DataFrame with forecasts
for col in columns_to_forecast:
    ukraine.loc[ukraine["Year"].isin(forecast_years), col] = df_forecast[col].values
# Save the updated data
ukraine.to_excel("/content/ukraine_forecasted_VAR.xlsx", index=False)
au=pd.read_excel("/content/ukraine_forecasted_VAR.xlsx")
print(au)
```

Fig. 3.2 Code for missing economic data prediction of Ukraine using VAR

- Using PROPHET for Ukraine economic data prediction:

```
import pandas as pd
import numpy as np
from prophet import Prophet
import matplotlib.pyplot as plt
```

```
# Load data
file_path = "/content/sheetUkr (3).xlsx"
xls = pd.ExcelFile(file_path)
df_ukraine = pd.read_excel(xls, sheet_name="Ukraine")
```

```
def forecast_prophet(df, column):
    data = df[["Year", column]].dropna()
    if data.empty or len(data) < 5:
        print(f"Not enough data to forecast {column}")
        return df

    data.rename(columns={"Year": "ds", column: "y"}, inplace=True)
    data["ds"] = pd.to_datetime(data["ds"], format='%Y')

    model = Prophet()
    model.fit(data)

    # Create future dataframe
    future_years = pd.date_range(start=f"{df['Year'].min()}-01-01", periods=len(df["Year"]) + 2, freq='Y')
    future = pd.DataFrame({"ds": future_years})
    forecast = model.predict(future)

    # Debugging: Print forecasted values
    print(f"Forecast for {column}:")
    print(forecast[['ds', 'yhat']].tail())

    # Update only missing values in df
    for year in [2024, 2025]:
        if df.loc[df["Year"] == year, column].isna().any():
            predicted_value = forecast.loc[forecast["ds"].dt.year == year, "yhat"].values
            if predicted_value.size > 0:
                df.loc[df["Year"] == year, column] = predicted_value[0]

    # Plot results
    plt.figure(figsize=(8,5))
    plt.plot(df["Year"], df[column], marker='o', label=f"{column} (Actual & Forecast)")
    plt.axvline(x=2022, color='r', linestyle='--', label='War Start')
    plt.xlabel("Year")
    plt.ylabel(column)
    plt.legend()
    plt.show()

    return df
```

```python
# Apply forecasting function to relevant columns
columns_to_forecast = ["CPI()", "Trade (% of GDP)", "FDI net inflows (BoP, current US$)", "Military expenditure (current LCU)"]
for col in columns_to_forecast:
    df_ukraine = forecast_prophet(df_ukraine, col)
# Save updated dataframe
df_ukraine.to_excel("/content/ukraine_forecasted_PROPHET.xlsx", index=False)
A=pd.read_excel("/content/ukraine_forecasted_PROPHET.xlsx")
print(A)
```

Fig. 3.3 Code for missing economic data prediction of Ukraine using PROPHET

- Using SARIMA for economic data prediction:

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.statespace.sarimax import SARIMAX
```

```python
# Load data
file_path = "/content/sheetUkr (3).xlsx"
xls = pd.ExcelFile(file_path)
ukraine = pd.read_excel(xls, sheet_name="Ukraine")
```

```python
# Selecting relevant columns for SARIMA forecasting
columns_to_forecast = ["CPI()", "Trade (% of GDP)", "FDI net inflows (BoP, current US$)", "Military expenditure (current LCU)"]
df_sarima = ukraine[["Year"] + columns_to_forecast].dropna()
# Set year as index
df_sarima.set_index("Year", inplace=True)
# Forecast storage
forecast_results = {}
# Forecasting each column separately using SARIMA
forecast_years = [2024, 2025]
for col in columns_to_forecast:
    series = df_sarima[col].dropna()  # Ensure no missing values in the series

    # Fit SARIMA model (simple settings: auto-tuning can be added)
    try:
        model = SARIMAX(series, order=(1,1,1), seasonal_order=(1,1,1,4))  # Adjust seasonal period if needed
        sarima_fit = model.fit(disp=False)

        # Forecast for 2024 and 2025
        forecast = sarima_fit.forecast(steps=2)
        forecast_results[col] = forecast.values
    except Exception as e:
        print(f"Error fitting SARIMA for {col}: {e}")
        forecast_results[col] = [np.nan, np.nan]
# Create DataFrame for forecasted years
df_forecast = pd.DataFrame(forecast_results, index=forecast_years)

# Update original DataFrame with forecasts
for col in columns_to_forecast:
    ukraine.loc[ukraine["Year"].isin(forecast_years), col] = df_forecast[col].values
# Save the updated data
output_file = "/content/ukraine_forecasted_SARIMA.xlsx"
ukraine.to_excel(output_file, index=False)

# Load and print the updated file
au = pd.read_excel(output_file)
print(au)
```

Fig. 3.4 Code for missing economic data prediction of Ukraine using SARIMA

Similarly, the missing value in the economic data for Russia was predicted using VAR, PROPHET and SARIMA.

Following the predictions, a comprehensive error analysis was conducted to evaluate and compare the performance of each forecasting model (VAR, Prophet, and SARIMA). This evaluation employed widely accepted statistical metrics: Mean Absolute Error (MAE), which measures the average magnitude of errors in a set of predictions without considering their direction; Root Mean Squared Error (RMSE), which gives higher weight to larger errors and provides a clear indication of the model's accuracy; and Mean Absolute Percentage Error (MAPE), which expresses prediction accuracy as a percentage and is especially useful for interpretability across variables with different scales. By applying these metrics, we were able to quantitatively assess how closely each model's predictions aligned with actual observed values. Based on the comparative results of this error analysis, we selected the model with the lowest error values as the most suitable and reliable for predicting the missing values. This data-driven selection process ensured that the imputed values were not only statistically sound but also appropriate for use in the subsequent stages of analysis in the study.

3.2 Phase II : Sentiment Analysis on Public Reactions to War

Tweets, Reddit comments, and news articles were collected, preprocessed, and analyzed for sentiment using TextBlob, VADER, AFINN, and SentiWordNet.

3.2.1 Data collection:-

A total of 503,696 text records were gathered from three sources: Public sentiment data was collected from Twitter, Reddit and news articles available on Kaggle also some of the news articles were extracted by using Google NewsAPI. The dataset covers textual data from 2015 to 2025, allowing an analysis of sentiment trends before, during, and after the peak war period. Duplicates were removed from each record and were merged together to form a dated text file.

3.2.2 Text Preprocessing:-

The raw text data was cleaned using Natural Language Processing (NLP) techniques.

The script begins by importing necessary libraries and reading a CSV file into a DataFrame. The text data is first converted to lowercase to ensure uniformity. It is then cleaned by removing URLs, hashtags, usernames, and extra whitespace using regular expressions. Punctuation is removed, and non-string values are filtered out. The code further cleans the text by eliminating repeated characters and alphanumeric patterns (A-Z, 0-9). Stopwords (common but uninformative words like "the", "and", "is") are removed using NLTK's stopwords list. The cleaned text is then tokenized and processed through ***stemming*** (reducing words to their root forms) and ***lemmatization*** (converting words to their base or dictionary form) using NLTK's tools. Finally, the code installs and utilizes the `emoji` package to strip emojis from the text. This thorough preprocessing pipeline ensures that the text is cleaned, normalized, and prepared for accurate and meaningful NLP analysis.

3.2.3 Sentiment Scoring:-

Each text entry in the dataset was subjected to sentiment analysis using four distinct models, each offering a unique approach to evaluating sentiment polarity. The first method employed was **TextBlob**, a lexicon-based tool that assigns sentiment polarity scores ranging from -1 (most negative) to +1 (most positive), code for the same is as shown in Fig.3.5 . This method is widely used for its simplicity and effectiveness in capturing general sentiment trends. The second method utilized was **VADER** (Valence Aware Dictionary and sEntiment Reasoner) , a rule-based model specifically optimized for sentiments expressed in social media contexts, code for the same is as shown in Fig.3.6. VADER is capable of detecting nuanced sentiment expressions including slang, emojis, and punctuation-based emphasis (e.g., all-caps or exclamation marks), making it highly suitable for analyzing informal and short texts.

The third sentiment analysis technique applied was **Afinn** as shown in Fig.3.7, which assigns integer-based sentiment scores ranging from -5 (highly negative) to +5 (highly positive). Afinn’s lexicon is particularly effective for brief textual inputs and supports a straightforward interpretation of sentiment intensity. Lastly, **SentiWordNet** Fig.3.8 was incorporated for a more granular, word-level sentiment analysis. Unlike the other models, SentiWordNet provides sentiment scores for individual words based on their part-of-speech and semantic context within the WordNet lexical database. This allows for a detailed classification of each word’s positivity, negativity, and objectivity.

After calculating sentiment scores using these four models, the results were aggregated on a yearly basis, covering the period from 2015 to 2025. This temporal aggregation enabled the identification and analysis of sentiment trends over time, providing insights into how public opinion or emotional tone has evolved throughout the decade, particularly in relation to key socio-economic or geopolitical events.

Sentiment score using TEXTBLOB

```python
!pip install textblob
from textblob import TextBlob
import pandas as pd
# Function to get TextBlob sentiment scores
def get_textblob_sentiment(text):
    blob = TextBlob(text)
    return blob.sentiment.polarity, blob.sentiment.subjectivity

# Apply TextBlob to DataFrame
text_df[['textblob_polarity', 'textblob_subjectivity']] = text_df['text'].apply(lambda x: pd.Series(get_textblob_sentiment(x)))

# Classify sentiment polarity based on polarity score
text_df['textblob_sentiment'] = text_df['textblob_polarity'].apply(lambda x: 'positive' if x > 0 else 'negative' if x < 0 else 'neutral')

text_df
# Convert 'date' column to datetime format (handle errors gracefully)
text_df['date'] = pd.to_datetime(text_df['date'], errors='coerce')
# Extract the year
text_df['year'] = text_df['date'].dt.year
# Aggregate sentiment by year
text_yearly_sentiment = text_df.groupby('year')['textblob_polarity'].mean().reset_index()
print(text_yearly_sentiment.head())
```

Fig. 3.5 Code for calculating yearly aggregate sentiment score using Textblob

Sentiment score using VADER

```python
!pip install vaderSentiment
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
import pandas as pd

# Initialize VADER sentiment analyzer
analyzer = SentimentIntensityAnalyzer()

# Function to get VADER sentiment scores
def get_vader_sentiment(text):
    scores = analyzer.polarity_scores(text)
    return scores['compound'], scores['pos'], scores['neu'], scores['neg']

# Apply VADER to DataFrame
text_df[['vader_compound', 'vader_positive', 'vader_neutral', 'vader_negative']] = text_df['text'].apply(lambda x: pd.Series(get_vader_sentiment(x)))

# Classify sentiment polarity based on compound score
text_df['vader_sentiment'] = text_df['vader_compound'].apply(lambda x: 'positive' if x > 0.05 else 'negative' if x < -0.05 else 'neutral')

text_df
# Aggregate Sentiment by Year
text_yearly_sentiment_vader = text_df.groupby('year')['vader_compound'].mean().reset_index()
print(text_yearly_sentiment_vader.head())
```

Fig. 3.6 Code for calculating yearly aggregate sentiment score using VADER

Sentiment score using AFINN

```
!pip install afinn
from afinn import Afinn

# Initialize Afinn
afinn = Afinn()

# Apply Afinn sentiment analysis
text_df['afinn_score'] = text_df['text'].apply(afinn.score)

# Classify polarity based on score
text_df['afinn_sentiment'] = text_df['afinn_score'].apply(lambda x: 'positive' if x > 0 else 'negative' if x < 0 else 'neutral')

text_df
# Aggregate Sentiment by Year
text_yearly_sentiment_afinn = text_df.groupby('year')['afinn_score'].mean().reset_index()
print(text_yearly_sentiment_afinn.head())
```

Fig. 3.7 Code for calculating yearly aggregate sentiment score using AFINN

Sentiment score using SentiWordNet

```
nltk.download('sentiwordnet')
from nltk.corpus import sentiwordnet as swn
from nltk.tokenize import word_tokenize

# Function to calculate sentiment score using SentiWordNet
def sentiwordnet_score(text):
    tokens = word_tokenize(text.lower())
    pos_score, neg_score = 0, 0
    for token in tokens:
        synsets = list(swn.senti_synsets(token))
        if synsets:
            pos_score += synsets[0].pos_score()
            neg_score += synsets[0].neg_score()
    return pos_score - neg_score

# Apply SentiWordNet sentiment analysis
text_df['sentiwordnet_score'] = text_df['text'].apply(sentiwordnet_score)
text_df['sentiwordnet_sentiment'] = text_df['sentiwordnet_score'].apply(lambda x: 'positive' if x > 0 else 'negative' if x < 0 else 'neutral')

text_df
# Aggregate Sentiment by Year
text_yearly_sentiment_sentiword = text_df.groupby('year')['sentiwordnet_score'].mean().reset_index()
print(text_yearly_sentiment_sentiword.head())
```

Fig. 3.8 Code for calculating yearly aggregate sentiment score using SentiWordNet

3.3 Phase III : Correlation Analysis Between Sentiment & Economic Indicators

This section outlines the methods used to examine the relationship between sentiment and economic indicators through statistical analysis and machine learning-based predictive validation.

3.3.1 Visualization & Statistical Testing:-

To examine the relationship between public sentiment and economic indicators, we used several analytical techniques.

Spearman correlation heatmaps were utilized to visualize the strength and direction of these associations. The Spearman rank correlation coefficient, a non-parametric measure, was chosen due to its ability to capture monotonic relationships between variables, even when the data does not follow a normal distribution. This made it particularly suitable for analyzing complex socio-economic data combined with sentiment scores derived from textual sources. In the heatmaps, positive correlation values indicated a direct relationship with increasing levels of optimistic sentiment, such as positive language or expressions of confidence which tended to align with favorable economic outcomes, including rising GDP, lower unemployment, or reduced inflation. Conversely, negative correlation values suggested increased pessimism in public sentiment such as expressions of fear, dissatisfaction, or uncertainty which coincided with economic downturns, like declining trade activity, rising debt, or economic contraction. These visualizations provided an intuitive and effective way to interpret the underlying relationships between public mood and macroeconomic trends across the studied timeframe. They also offered insights into which sentiment indicators most closely mirrored changes in specific economic variables, thereby supporting further analysis and predictive modeling efforts.

```
import seaborn as sns
# Select relevant columns: sentiment + economic indicators
cols = ['TextBlob', 'VADER', 'AFINN', 'SentiWordNet',
        'Inflation(% change)', 'GDP(% change)', 'Unemployment Rate (% of labour force)',
        'Govt. Debt(% of GDP)', 'Trade (% of GDP)',
        'FDI net inflows (BoP, current US$)', 'Military expenditure (current LCU)']
data = df[cols]
# Compute Spearman correlation matrix
spearman_corr = data.corr(method='spearman')

# Plot it
plt.figure(figsize=(12, 8))
sns.heatmap(spearman_corr, annot=True, cmap='viridis', fmt=".2f")
plt.title(" Spearman Correlation Matrix for Ukraine")
plt.show()
```

Fig. 3.9 Code for plotting spearman correlation heatmap

To evaluate the potential predictive power of sentiment on economic outcomes, the Granger causality test was employed. This statistical method is specifically designed to assess whether one time series can be used to forecast another, making it well-suited for exploring temporal relationships between sentiment trends and macroeconomic indicators. In this study, the test was used to determine whether historical sentiment data, extracted from sources such as public discourse, news articles, and social media contained significant predictive information about future changes in key economic variables, including GDP, inflation, unemployment, and government debt.

By applying the test across two time lags, the analysis aimed to capture both immediate and delayed effects of sentiment on economic performance. A statistically significant result would suggest that sentiment does not merely react to economic conditions but may also anticipate and influence them, thus serving as a leading indicator. These findings have valuable implications for economic forecasting, offering potential improvements in the accuracy and responsiveness of traditional models when sentiment data is integrated.

```
# Granger Causality Test (Checking if Sentiment can predict GDP Growth)
max_lag = 2
test_results = {}
for sentiment in ["Textblob", "Vader", "Afinn", "Sentiwordnet"]:
    test_result = grangercausalitytests(df_merged[["GDP(% change)", sentiment]], max_lag, verbose=False)
    p_values = [round(test_result[i][0]['ssr_chi2test'][1], 4) for i in range(1, max_lag + 1)]
    test_results[sentiment] = p_values

granger_results_df = pd.DataFrame(test_results, index=[f"Lag {i}" for i in range(1, max_lag + 1)])
print("Granger Causality Test Results:")
print(granger_results_df)
```

Fig. 3.10 Code for Granger Causality Test

Time-series line plots were utilized as a visual analytical tool to explore the dynamic relationship between sentiment and various economic variables across the study period. These plots allowed for the simultaneous tracking of yearly sentiment scores generated from text data and corresponding macroeconomic indicators such as GDP, inflation, unemployment, and trade activity from 2015 to 2025. By plotting these variables on a common temporal axis, it became possible to visually assess patterns, trends, and co-movements between public sentiment and economic performance over time.

This visual approach was particularly effective in identifying parallel movements, where shifts in sentiment whether positive or negative appeared to align with similar directional changes in economic indicators. For instance, periods marked by declining sentiment often coincided with economic downturns, while more optimistic sentiment trends aligned with phases of growth or recovery. These time-series plots served not only as a preliminary diagnostic tool but also supported the findings of statistical tests by offering intuitive, interpretable evidence of potential associations and lagged effects between public mood and economic conditions.

```
# Normalize data for better visualization
scaler = MinMaxScaler()
columns_to_normalize = ["GDP(% change)", "Textblob", "Vader", "Afinn", "Sentiwordnet"]
df_normalized = df_merged.copy()
df_normalized[columns_to_normalize] = scaler.fit_transform(df_merged[columns_to_normalize])
# Normalize GDP, CPI, Inflation, and Sentiment Scores
columns_to_normalize = ["GDP(% change)", "CPI()", "Inflation(% change)", "Textblob", "Vader", "Afinn", "Sentiwordnet"]
scaler = MinMaxScaler()
df_merged[columns_to_normalize] = scaler.fit_transform(df_merged[columns_to_normalize])
# Plot GDP, CPI, Inflation along with Sentiment Trends
plt.figure(figsize=(14, 7))
plt.plot(df_merged["Year"], df_merged["GDP(% change)"], label="GDP Growth (Normalized)", color="blue", marker="o", linestyle="-")
plt.plot(df_merged["Year"], df_merged["CPI()"], label="CPI (Normalized)", color="red", marker="s", linestyle="-")
plt.plot(df_merged["Year"], df_merged["Inflation(% change)"], label="Inflation (Normalized)", color="green", marker="^", linestyle="-")
plt.plot(df_merged["Year"], df_merged["Textblob"], label="TextBlob Sentiment (Normalized)", color="purple", marker="d", linestyle="--")
plt.plot(df_merged["Year"], df_merged["Vader"], label="VADER Sentiment (Normalized)", color="orange", marker="x", linestyle="--")
plt.plot(df_merged["Year"], df_merged["Afinn"], label="AFINN Sentiment (Normalized)", color="brown", marker="*", linestyle="--")
plt.plot(df_merged["Year"], df_merged["Sentiwordnet"], label="SentiWordNet Sentiment (Normalized)", color="pink", marker="v", linestyle="--")

# Mark Ukraine-Russia War Start (2022)
plt.axvline(x=2022, color="black", linestyle=":", linewidth=2, label="Ukraine-Russia War Start")

plt.xlabel("Year")
plt.ylabel("Normalized Values")
plt.title("Economic Indicators and Sentiment Trends (Normalized)")
plt.legend()
plt.grid(True)
plt.show()
```

Fig. 3.11 Code for plotting Line Plot for trend analysis

To further investigate the predictive relationship between sentiment and economic performance, A Pearson correlation analysis was conducted to investigate the strength and direction of the linear relationship between sentiment scores and key economic indicators, specifically GDP, inflation, and unemployment across different phases of the Russia-Ukraine conflict. The analysis was performed on three distinct sub-periods: pre-war, during-war, and post-war, to capture how the nature of sentiment-economy relationships may have evolved over time due to geopolitical disruptions.

To enhance the temporal sensitivity of the analysis, sentiment scores were lagged by one year, meaning that the sentiment from a given year was compared with the economic outcomes in the following year. This lagging approach was used to test the hypothesis that shifts in public sentiment could act as early warning signals of upcoming changes in economic conditions. By correlating lagged sentiment with subsequent economic performance, the study aimed to assess the predictive value of sentiment, rather than merely its concurrent association with macroeconomic variables

```python
# Create lagged sentiment columns (only Lag 1)
for col in ['TextBlob', 'VADER', 'AFINN', 'SentiWordNet']:
    df[f'{col}_lag1'] = df[col].shift(1)
# Define lagged sentiment columns and economic indicators
sentiment_lags = [f'{col}_lag1' for col in ['TextBlob', 'VADER', 'AFINN', 'SentiWordNet']]
economic_indicators = [
    'GDP(% change)',
    'Inflation(% change)',
    'Unemployment Rate (% of labour force)'
]
# Split into sub-periods
df_prewar = df[df['Year'].between(2015, 2021)]
df_peakwar = df[df['Year'].between(2022,2023)]
df_postwar = df[df['Year'].between(2024, 2025)]
# Define a function to compute and display correlations
def show_correlations(df_subset, label):
    print(f"\n📅 Correlation during for Ukraine {label} period:")
    for senti in sentiment_lags:
        if senti in df_subset.columns:
            print(f"\n🧠 Correlation for {senti}:")
            valid_cols = [senti] + [col for col in economic_indicators if col in df_subset.columns]
            print(df_subset[valid_cols].dropna().corr().loc[senti, valid_cols[1:]])
        else:
            print(f"⚠️ Skipping {senti}: not in DataFrame.")
# Run for all periods
show_correlations(df_prewar, "Pre-War (2015-2021)")
show_correlations(df_peakwar, "Peak War (2022,2023)")
show_correlations(df_postwar, "Post-Peak (2024-2025)")
```

Fig. 3.12 Code for calculating Pearson correlation analysis for lagged sentiment on different war phases

3.3.2 XGBoost and Random Forest for Historical Validation Using Sentiments:-

Historical validation in machine learning refers to the process of training predictive models on past data to assess their ability to forecast future outcomes. In this study, historical validation was carried out by training machine learning models on data spanning from 2015 to 2023, with the objective of predicting economic indicators for the years 2024 and 2025. This approach ensures that the models are evaluated under realistic conditions, where only historical information is available at the time of prediction, thereby simulating a real-world forecasting scenario.

The primary input features for the models were sentiment scores derived from multiple sentiment analysis techniques, including VADER, TextBlob, AFINN, and SentiWordNet. These models captured varying aspects of public mood and emotional tone expressed in text data related to Ukraine and Russia. The sentiment scores were used to predict key macroeconomic indicators such as GDP, inflation, and unemployment, providing a novel way to assess whether shifts in public sentiment can inform economic forecasting.

To perform the predictions, two powerful machine learning algorithms: XGBoost and Random Forest were implemented. These ensemble learning methods are well-regarded for their robustness, handling of non-linear relationships, and high predictive accuracy. By comparing model performance across different sentiment inputs, the analysis aimed to identify which sentiment model best correlates with and forecasts economic trends.

This historical validation process not only helped assess the predictive reliability of sentiment-based features but also demonstrated the potential of integrating sentiment analysis with machine learning for proactive economic decision-making and policy planning in volatile geopolitical contexts.

```python
import pandas as pd
import xgboost as xgb
from sklearn.metrics import mean_absolute_error
from itertools import combinations
# Define economic indicators as targets
economic_indicators = ["CPI()", "Inflation(% change)", "GDP(% change)",
                       "Unemployment Rate (% of labour force)", "Govt. Debt(% of GDP)",
                       "Trade (% of GDP)", "FDI net inflows (BoP, current US$)", "Military expenditure (current LCU)"]
# Define sentiment analyzers as features
sentiments = ["TextBlob", "VADER", "AFINN", "SentiWordNet"]
# Split data
train = df[df["Year"].between(2015, 2023)]
test = df[df["Year"].between(2024, 2025)]
# Evaluate each sentiment analyzer
results = {}
for sentiment in sentiments:
    errors = {}
    for target in economic_indicators:
        # Train model
        X_train, y_train = train[[sentiment]], train[target]
        X_test, y_test = test[[sentiment]], test[target]

        model = xgb.XGBRegressor(objective='reg:squarederror', random_state=42)
        model.fit(X_train, y_train)

        # Predict
        y_pred = model.predict(X_test)

        # Calculate error
        mae = mean_absolute_error(y_test, y_pred)
        errors[target] = mae

    results[sentiment] = sum(errors.values()) / len(economic_indicators)  # Average MAE
# Find the best sentiment analyzer
best_sentiment = min(results, key=results.get)
print("Best Sentiment Analyzer:", best_sentiment)
print("MAE Scores:", results)
```

Fig. 3.13 Code for Historic validation of economic indicators using XGBoost model and sentiments as feature

```
import pandas as pd
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_absolute_error
# Define economic indicators as targets
economic_indicators = ["CPI()", "Inflation(% change)", "GDP(% change)",
                       "Unemployment Rate (% of labour force)", "Govt. Debt(% of GDP)",
                       "Trade (% of GDP)", "FDI net inflows (BoP, current US$)", "Military expenditure (current LCU)"]

# Define sentiment analyzers as features
sentiments = ["TextBlob", "VADER", "AFINN", "SentiWordNet"]
# Split data
train = df[df["Year"].between(2015, 2023)]
test = df[df["Year"].between(2024, 2025)]
# Evaluate each sentiment analyzer
results = {}
for sentiment in sentiments:
    errors = {}
    for target in economic_indicators:
        # Train model
        X_train, y_train = train[[sentiment]], train[target]
        X_test, y_test = test[[sentiment]], test[target]

        model = RandomForestRegressor(random_state=42)
        model.fit(X_train, y_train)

        # Predict
        y_pred = model.predict(X_test)

        # Calculate error
        mae = mean_absolute_error(y_test, y_pred)
        errors[target] = mae

    results[sentiment] = sum(errors.values()) / len(economic_indicators)  # Average MAE
# Find the best sentiment analyzer
best_sentiment = min(results, key=results.get)
print("Best Sentiment Analyzer:", best_sentiment)
print("MAE Scores:", results)
```

Fig. 3.14 Code for Historic validation of economic indicators using Random Forest model and sentiments as feature

3.3.3 Hypothesis Testing:-

Hypothesis testing is a statistical approach used to assess assumptions about relationships between variables. In this study, we test how sentiment, particularly during crises, influences economic indicators. These hypotheses stem from observed trends during geopolitical events and the growing role of sentiment in shaping economic expectations.

1. **Sentiment-Driven Economic Shocks Are Stronger in Crisis Periods :** The relationship between sentiment and economic indicators intensifies during periods of geopolitical crises or economic downturns, leading to extreme sentiment correlations with key macroeconomic variables.
2. **Sentiment-Based Economic Predictions Vary Across Economic Structures :** The effectiveness of sentiment analysis models in predicting economic indicators depends on a country's economic structure, resilience, and exposure to external shocks.
3. **Military Expenditure Has a Strong Trade-Off with Economic Growth and Trade :** Increased military spending negatively impacts trade and economic growth,with varying effects across different economic and geopolitical contexts.

4. **Economic Recovery Patterns Differ Based on Conflict Impact and Policy Response:** Countries recovering from war or economic shocks exhibit different recovery trajectories depending on the severity of the impact and the effectiveness of policy responses.
5. **The Predictive Power of Sentiment Analysis Models Evolves Over Time :** The effectiveness of sentiment analysis models in predicting economic indicators changes over time, influenced by economic conditions, technological advancements, and shifts in public discourse.

CHAPTER 4

RESULT AND DISCUSSION

Following the proposed three-phase methodology, missing economic data were imputed, sentiment scores were calculated, correlations between sentiment and economic indicators were analyzed, and relevant hypotheses were tested to uncover the impact of public sentiment on economic trends during the Russia-Ukraine war.

4.1 Filling missing economic data

To determine the most appropriate forecasting models tailored to each country's economic characteristics, we conducted a comprehensive evaluation of three well-established time series forecasting techniques: Vector Autoregression (VAR), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Meta's Prophet model. These models were rigorously applied across four critical economic indicators : Consumer Price Index (CPI), Trade (measured as a percentage of GDP), Foreign Direct Investment (FDI) net inflows, and Military Expenditure.

To objectively assess and compare model performance, we employed three widely recognized statistical error metrics: Mean Absolute Error (MAE), which provides a straightforward measure of prediction accuracy; Root Mean Squared Error (RMSE), which penalizes larger errors more severely and is sensitive to outliers; and Mean Absolute Percentage Error (MAPE), which allows for intuitive percentage-based comparisons across different magnitudes of economic data. Each model's output was evaluated against historical data for these metrics to ensure robust and unbiased selection. Through this empirical evaluation process, several important insights were derived regarding the relative strengths and weaknesses of each model in forecasting different types of economic indicators under varying conditions. These insights are summarized as follows:

For Ukraine:

For Ukraine, the results indicate a consistent superiority of the SARIMA model across all economic indicators:

- **CPI**: SARIMA outperforms both VAR and Prophet with significantly lower MAE (189.37), RMSE (191.86), and MAPE (70.70%). VAR performed extremely poorly due to scale mismatch, and Prophet's error was considerably higher.
- **Trade (% of GDP)**: SARIMA again provides the best forecast accuracy with an MAE of 36.10 and RMSE of 37.50, compared to higher errors from VAR and Prophet. Its MAPE of 39.50% is also the lowest among the models.

- **FDI net inflows**: SARIMA offers the most reliable performance with MAE (4.83e+09), RMSE (6.40e+09), and MAPE (540.63%), outperforming the high-error VAR and Prophet models. While MAPE is still large due to scale, it is significantly lower than alternatives.
- **Military Expenditure**: SARIMA produces the most accurate results (MAE: 1.83e+12, RMSE: 2.63e+12, MAPE: 816.84%), with VAR and Prophet yielding much higher errors.

SARIMA is the most effective and consistent model for forecasting all selected economic indicators, making it the best choice for Ukraine.

For Russia :

For Russia, the optimal model varies depending on the specific indicator, suggesting a hybrid model approach:

- **CPI**: SARIMA provides the best forecasts with the lowest MAE (47.01), RMSE (47.37), and MAPE (27.19%). Both VAR and Prophet are less accurate
- **Trade (% of GDP)**: SARIMA also excels in this category, achieving the lowest MAE (4.18), RMSE (5.04), and MAPE (8.75%), indicating high predictive accuracy.
- **FDI net inflows**: VAR performs best with significantly lower MAE (1.07e+10), RMSE (1.31e+10), and MAPE (104.14%) compared to SARIMA and Prophet, which yield much higher errors.
- **Military Expenditure**: VAR again outperforms other models, achieving the lowest MAE (6.34e+11), RMSE (7.20e+11), and MAPE (15.34%), while SARIMA and Prophet result in substantially higher errors.

A hybrid approach to model selection was found to yield the most accurate and reliable results. Specifically, the SARIMA model demonstrated superior performance in forecasting variables such as CPI and Trade, likely due to its strength in capturing seasonal patterns and autocorrelations in univariate time series data. In contrast, the VAR model proved to be more effective for modeling FDI and Military Expenditure, which are better suited to multivariate time series techniques that account for the dynamic interdependencies between variables. This strategic combination of models capitalizes on the unique advantages each offers within their respective domains.

Building on these findings, missing values in the economic dataset were imputed using the model that performed best for each corresponding indicator. For instance, SARIMA was used to estimate missing CPI and Trade data, while VAR was applied to interpolate absent values in FDI and Military Expenditure. After the imputation process, the final version of the economic dataset was manually reviewed, curated, and assembled to ensure consistency and accuracy. The resulting complete dataset, ready for further analysis, is as follows:

Year	CPI	Inflation(%change)	GDP(%change)	Unemployment Rate (% of labour force)	Govt debt(% of GDP)	Trade (% of GDP)	FDI net inflow (BoP, Current US$)	Military expenditure(current LCU)
2015	151.5294641	15.534	-1.973	5.575	15.286	49.35934931	6852970000	4047613000000
2016	162.2008473	7.042	0.194	5.525	14.849	46.51811984	32538900000	4644799000000
2017	168.1752389	3.683	1.827	5.2	14.311	46.87652434	28557440000	3902412000000
2018	173.0158221	2.878	2.806	4.8	13.62	51.58090037	8784850000	3863469000000
2019	180.7502637	4.47	2.198	4.6	13.748	49.22875366	31974770000	4217092000000
2020	186.8626219	3.382	-2.654	5.767	19.156	45.9669082	9478810000	4462125000000
2021	199.3720633	6.694	5.867	4.833	16.397	50.1964411	40450000000	4855157000000
2022	254.0764163	13.75	-1.247	3.95	18.51	43.25870482	-39800944227	7150000000000
2023	271.2560372	5.859	3.649	3.167	19.546	41.82852518	-10045103795	9300000000000
2024	217.0283773	7.861	3.625	2.6	19.87	40.99721779	22571673847	4451256916741
2025	235.8992249	5.915	1.347	3.004	20.412	52.041498	24578675903	4641032185152

Table 4.1 : Final economic data of Russia

Year	CPI	Inflation(%change)	GDP(%change)	Unemployment Rate (% of labour force)	Govt debt(% of GDP)	Trade (% of GDP)	FDI net inflow (BoP, Current US$)	Military expenditure(current LCU)
2015	180.499827	48.684	-9.773	9.143	79.331	107.8066163	-198000000	76516160000
2016	205.6122449	13.913	2.441	9.45	79.513	105.5212049	4128000000	87510000000
2017	235.2992044	14.443	2.36	9.65	71.62	104.0349829	3680000000	96691140474
2018	261.0688343	10.947	3.488	9	60.408	99.19981507	4975000000	129422890647
2019	281.6585956	7.886	3.2	8.5	50.616	90.51123429	5796000000	161789934872
2020	289.3548945	2.74	-3.753	9.15	60.591	79.15645238	304000000	184455207912
2021	316.447596	9.361	3.446	9.835	48.925	82.69796136	7954000000	186718396189
2022	380.318229	20.183	-28.759	24.528	77.721	87.39651399	221000000	1343435116952
2023	429.185403	12.851	5.324	19.072	82.329	78.09811367	4805000000	2382451228053
2024	423.1222132	5.836	3	14.199	95.592	57.65366824	5310355941	3307059971504
2025	446.7566324	8.976	2.5	12.706	106.635	66.10768218	8284492157	4088781151957

Table 4.2 : Final economic data of Ukraine

4.2 Aggregated Sentiments

After preprocessing raw textual data and applying sentiment analyzer (TextBlob, VADER, AFINN, and SentiWordNet), sentiment scores were aggregated yearly resulting in following output:

Year	Textblob_polarity	vader_compound	afinn_score	sentiwordnet_score
2015	0.04631410256	0.01651538462	0.0769230769	0.1538461538
2016	0.06279085498	0.0590260274	0.1917808219	0.09417808219
2017	0.05656555549	0.1056448276	0.0862068965	0.1810344828
2018	0.05177836775	-0.02339791667	-0.6875	0.0078125
2019	0.01615397407	-0.02669579832	-0.6218487395	0.07457983193
2020	0.03022280316	-0.01367435897	-0.5311355311	-0.005494505495
2021	0.01835522071	0.0013249499	-0.2299599198	0.02455561122
2022	0.02485909655	-0.07010751163	-0.4726849689	0.03496869225
2023	0.03129231724	-0.02367098146	-0.342157285	0.03368515389
2024	0.03195506714	-0.02961485641	-0.3786873362	0.02538725255
2025	0.02932714889	-0.01964601393	-0.3529944438	0.02349855225

Table 4.3 Yearly aggregated Sentiment Scores

4.3 Correlation Analysis: Spearman Heatmaps

From the spearman correlation heatmaps for Ukraine and Russia shown in fig below, we can draw insights into how economic factors interact with sentiment scores in both countries. The Spearman correlation heatmaps for Ukraine and Russia highlight key differences in how economic indicators and sentiment analysis models interact. In Ukraine, there is a strong negative correlation between military expenditure and trade (0.95), indicating that increased military spending significantly impacts trade.

Additionally, inflation and sentiment models like SentiWordNet (0.69) show a strong correlation, suggesting that public sentiment in Ukraine is highly reactive to inflationary pressures.

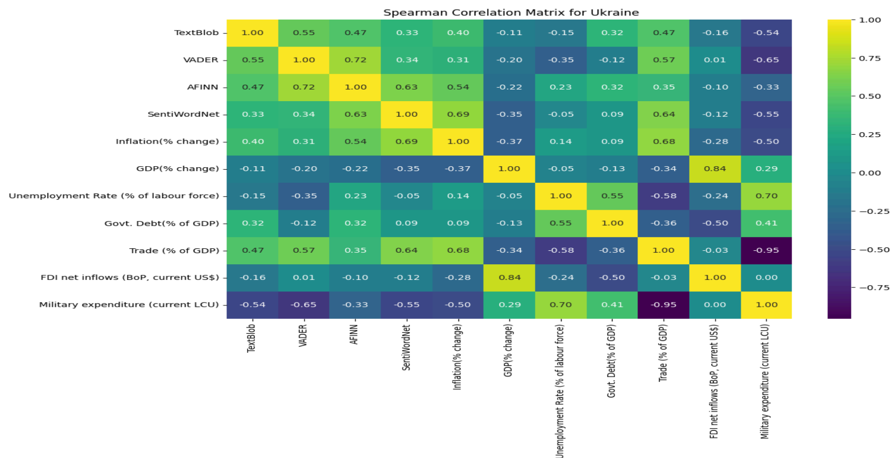


Fig. 4.1 Spearman Correlation Heatmap for Ukraine

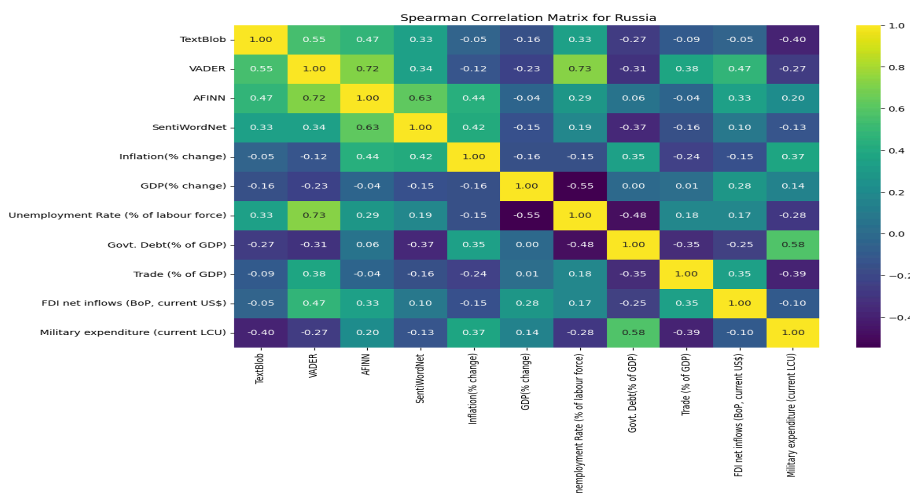


Fig. 4.2 Spearman Correlation Heatmap for Russia

Unemployment and government debt also demonstrate stronger correlations with sentiment metrics, reinforcing the notion that economic instability is closely linked with negative public sentiment. The negative relationship between military expenditure and sentiment models further supports the idea that increased defense spending leads to pessimism in public perception. In contrast, Russia's correlation matrix presents a more stable economic structure, with weaker correlations between sentiment and key economic indicators. While unemployment shows a high correlation with VADER sentiment (0.73), other relationships are comparatively less pronounced. The negative correlation between government debt and sentiment models like TextBlob (-0.27) and VADER (-0.31) suggests that debt accumulation in Russia has a limited but noticeable impact on public perception. Unlike Ukraine, military expenditure does not exhibit as strong a negative correlation with trade (-0.39), indicating a more resilient trade environment despite increased defense spending. Overall, the findings suggest that Ukraine's economic sentiment is more volatile and reactive to economic fluctuations, while Russia's economic sentiment remains relatively stable despite external pressures.

4.4 Granger Causality Test

The Granger Causality test helps determine whether past sentiment scores can predict economic factors. The p-values indicate the statistical significance of this relationship, with lower values (< 0.05) suggesting a stronger causal link. The Granger causality test reveals differences in the influence of sentiment on economic factors between Ukraine and Russia as presented in Table 4.1 and Table 4.2.

For Russia:

	Textblob	Vader	Afinn	SentiWordNet
Lag 1	0.5270	0.9202	0.6476	0.1679
Lag 2	0.0014	0.9541	0.6543	0.5647

Table. 4.1 Granger Causality Test Result for Russia

For Ukraine:

	Textblob	Vader	Afinn	SentiWordNet
Lag 1	0.0390	0.7027	0.9906	0.5549
Lag 2	0.1279	0.7699	0.2058	0.1105

Table. 4.2 Granger Causality Test Result for Russia

In Ukraine, TextBlob sentiment at lag 1 ($p = 0.0390$) shows a significant causal relationship with economic indicators, while at lag 2, SentiWordNet ($p = 0.1105$) and TextBlob ($p = 0.1279$) approach significance, indicating a potential delayed effect. In contrast, Russia exhibits weaker overall causality, with only TextBlob at lag 2 ($p = 0.0014$) showing a significant causal link. This suggests that in Ukraine, sentiment has a more immediate impact on economic variables, whereas in Russia, sentiment-driven economic effects emerge over a longer period. These results align with Ukraine's higher economic volatility and Russia's slower but more controlled economic response to sentiment shifts.

4.5 Comparing XGBoost and Random Forest for Historical Validation Using Sentiments

Results for Historic validation and MAE scores obtained after applying XGBoost and Random Forest on Ukraine economic indicators while keeping sentiment scores as features are depicted in Fig. 4.6 and Fig. 4.7 , while the same for Russia is depicted in Fig. 4.8 and Fig. 4.9.

```
Best Sentiment Analyzer: TextBlob
MAE Scores: {'TextBlob': 229612479926.3604, 'VADER': 370529855330.4612, 'AFINN': 294511063516.1642, 'SentiWordNet': 443061300772.9528}
```

Fig. 4.3 XGBoost Model validation and MAE score for Ukraine

```
Best Sentiment Analyzer: AFINN
MAE Scores: {'TextBlob': 287821116720.9208, 'VADER': 361769605119.1059, 'AFINN': 212804968611.654, 'SentiWordNet': 373510813809.34875}
```

Fig. 4.4 Random Forest Model validation and MAE score for Ukraine

```
Best Sentiment Analyzer: SentiWordNet
MAE Scores: {'TextBlob': 465918915873.51337, 'VADER': 222154379491.45032, 'AFINN': 333403902945.1882, 'SentiWordNet': 75945934915.52464}
```

Fig. 4.5 XGBoost Model validation and MAE score for Russia

```
Best Sentiment Analyzer: VADER
MAE Scores: {'TextBlob': 328491350714.8892, 'VADER': 134300956666.79564, 'AFINN': 499251266688.0025, 'SentiWordNet': 170095053824.7282}
```

Fig. 4.6 Random Forest Model validation and MAE score for Russia

For Russia, XGBoost significantly outperforms Random Forest (MAE: 75.95B vs. 134.30B), indicating better predictive power. Whereas for Ukraine Random Forest (MAE: 212.80B) slightly outperforms XGBoost (MAE: 229.61B), but the difference is marginal.

So, XGBoost is the superior model as it provides more accurate forecasts for Russia and performs comparably well for Ukraine. This analysis highlights that XGBoost's ability to capture complex sentiment-economic interactions makes it more effective for economic forecasting than Random Forest. However, no specific sentiment analyzer could be said best for predicting economic indicator using this method

4.6 Trend Analysis

This section analyzes the temporal relationship between sentiment and economic indicators, highlighting how their interactions shifted before, during, and after the Ukraine Russia war.

Line Plot and Pearson Correlation Analysis with Lagged Sentiment for Ukraine & Russia:-

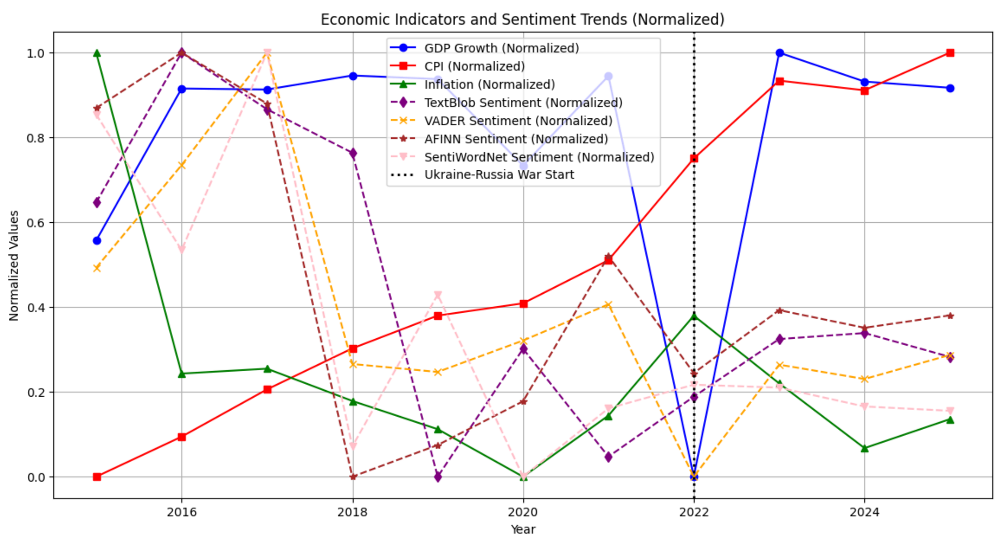


Fig. 4.7 Line plot for trend analysis of economic indicator vs Sentiment for Ukraine

```
Correlation during for Ukraine Pre-War (2015-2021) period:

Correlation for TextBlob_lag1:
GDP(% change)                             0.717460
Inflation(% change)                       0.787502
Unemployment Rate (% of labour force)    -0.147202
Name: TextBlob_lag1, dtype: float64

Correlation for VADER_lag1:
GDP(% change)                             0.414827
Inflation(% change)                       0.604280
Unemployment Rate (% of labour force)     0.085970
Name: VADER_lag1, dtype: float64

Correlation for AFINN_lag1:
GDP(% change)                             0.354932
Inflation(% change)                       0.839577
Unemployment Rate (% of labour force)     0.364423
Name: AFINN_lag1, dtype: float64

Correlation for SentiWordNet_lag1:
GDP(% change)                             0.011871
Inflation(% change)                       0.428780
Unemployment Rate (% of labour force)    -0.002837
Name: SentiWordNet_lag1, dtype: float64
```

```
Correlation during for Ukraine Peak War (2022,2023) period:

Correlation for TextBlob_lag1:
GDP(% change)                            1.0
Inflation(% change)                     -1.0
Unemployment Rate (% of labour force)   -1.0
Name: TextBlob_lag1, dtype: float64

Correlation for VADER_lag1:
GDP(% change)                           -1.0
Inflation(% change)                      1.0
Unemployment Rate (% of labour force)    1.0
Name: VADER_lag1, dtype: float64

Correlation for AFINN_lag1:
GDP(% change)                           -1.0
Inflation(% change)                      1.0
Unemployment Rate (% of labour force)    1.0
Name: AFINN_lag1, dtype: float64

Correlation for SentiWordNet_lag1:
GDP(% change)                            1.0
Inflation(% change)                     -1.0
Unemployment Rate (% of labour force)   -1.0
Name: SentiWordNet_lag1, dtype: float64
```

```
Correlation during for Ukraine Post-Peak (2024-2025) period:

Correlation for TextBlob_lag1:
GDP(% change)                           -1.0
Inflation(% change)                      1.0
Unemployment Rate (% of labour force)   -1.0
Name: TextBlob_lag1, dtype: float64

Correlation for VADER_lag1:
GDP(% change)                            1.0
Inflation(% change)                     -1.0
Unemployment Rate (% of labour force)    1.0
Name: VADER_lag1, dtype: float64

Correlation for AFINN_lag1:
GDP(% change)                            1.0
Inflation(% change)                     -1.0
Unemployment Rate (% of labour force)    1.0
Name: AFINN_lag1, dtype: float64

Correlation for SentiWordNet_lag1:
GDP(% change)                            1.0
Inflation(% change)                     -1.0
Unemployment Rate (% of labour force)    1.0
Name: SentiWordNet_lag1, dtype: float64
```

Fig. 4.8 Pearson correlation analysis with lagged sentiments on different war phases for Ukraine

For Russia :

```
Correlation for Russia during Pre-War (2015-2021) period:

Correlation for TextBlob_lag1:
GDP(% change)                              0.323841
Inflation(% change)                       -0.187892
Unemployment Rate (% of labour force)     -0.490623
Name: TextBlob_lag1, dtype: float64

Correlation for VADER_lag1:
GDP(% change)                              0.194882
Inflation(% change)                       -0.425891
Unemployment Rate (% of labour force)     -0.195336
Name: VADER_lag1, dtype: float64

Correlation for AFINN_lag1:
GDP(% change)                              0.028748
Inflation(% change)                       -0.074177
Unemployment Rate (% of labour force)      0.137740
Name: AFINN_lag1, dtype: float64

Correlation for SentiWordNet_lag1:
GDP(% change)                             -0.344084
Inflation(% change)                       -0.258926
Unemployment Rate (% of labour force)      0.333077
Name: SentiWordNet_lag1, dtype: float64
```

```
Correlation for Russia during Peak War (2022,2023) period:

Correlation for TextBlob_lag1:
GDP(% change)                             1.0
Inflation(% change)                      -1.0
Unemployment Rate (% of labour force)    -1.0
Name: TextBlob_lag1, dtype: float64

Correlation for VADER_lag1:
GDP(% change)                            -1.0
Inflation(% change)                       1.0
Unemployment Rate (% of labour force)     1.0
Name: VADER_lag1, dtype: float64

Correlation for AFINN_lag1:
GDP(% change)                            -1.0
Inflation(% change)                       1.0
Unemployment Rate (% of labour force)     1.0
Name: AFINN_lag1, dtype: float64

Correlation for SentiWordNet_lag1:
GDP(% change)                             1.0
Inflation(% change)                      -1.0
Unemployment Rate (% of labour force)    -1.0
Name: SentiWordNet_lag1, dtype: float64
```

```
Correlation for Russia during Post-Peak (2024-2025) period:

Correlation for TextBlob_lag1:
GDP(% change)                            -1.0
Inflation(% change)                      -1.0
Unemployment Rate (% of labour force)     1.0
Name: TextBlob_lag1, dtype: float64

Correlation for VADER_lag1:
GDP(% change)                             1.0
Inflation(% change)                       1.0
Unemployment Rate (% of labour force)    -1.0
Name: VADER_lag1, dtype: float64

Correlation for AFINN_lag1:
GDP(% change)                             1.0
Inflation(% change)                       1.0
Unemployment Rate (% of labour force)    -1.0
Name: AFINN_lag1, dtype: float64

Correlation for SentiWordNet_lag1:
GDP(% change)                             1.0
Inflation(% change)                       1.0
Unemployment Rate (% of labour force)    -1.0
Name: SentiWordNet_lag1, dtype: float64
```

Fig. 4.9 Pearson correlation analysis with lagged sentiments on different war phases for Russia

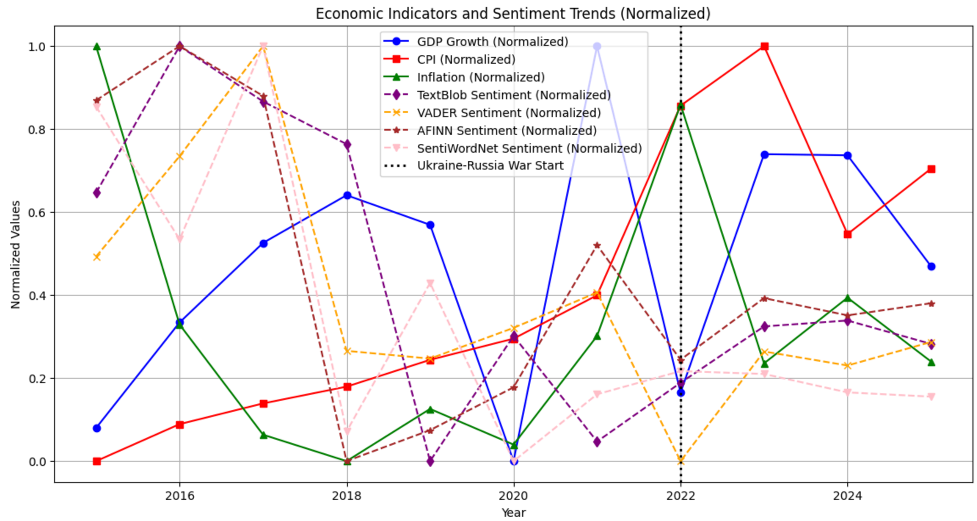


Fig. 4.10 Line plot for trend analysis of economic indicator vs Sentiment for Russia

The line as shown in plots Fig 4.11 and 4.13 illustrate the evolution of GDP growth, CPI, inflation, and sentiment scores over time. Pearson correlation analysis with lagged sentiments on different war phases as shown in Fig 4.10 and Fig.4.12 for both the countries give following observation along with line plots:

- **Pre-War Trends (2015–2021):** Before the war, Ukraine and Russia exhibited economic growth, but with distinct characteristics. Ukraine's economy was volatile, with fluctuating GDP growth, unstable inflation, and cyclical sentiment shifts, showing optimism followed by downturns around 2018–2019. Russia, on the other hand, maintained more stable economic growth with lower inflation, though sentiment exhibited a gradual decline. Correlation analysis reveals that before the war, sentiment had a stronger relationship with economic indicators in Ukraine than in Russia, with TextBlob sentiment highly correlated with GDP (0.71) and inflation (0.78). This suggests that Ukraine's public sentiment was more sensitive to economic changes, while Russia's economic sentiment remained more detached from direct economic fluctuations.
- **Impact of the Ukraine-Russia War (2022):** The war in 2022 triggered a sharp economic downturn in both nations, with GDP falling and inflation surging. Sentiment dropped drastically, reflecting heightened economic distress and uncertainty. The Pearson correlation analysis showed that during this period, sentiment correlations with GDP and inflation became extreme, reaching perfect correlations (1.0) in both countries. This highlights how economic sentiment was directly influenced by war-driven economic shocks. The line plot confirms this with clear downward trends in GDP and sentiment, reinforcing that war led to a

synchronized economic and psychological shock in both nations. Interestingly, different sentiment models reacted differently, with VADER and AFINN showing opposite trends, suggesting divergence in how various sentiment measures capture crisis periods.

- **Post-War Recovery (2023–2024):** Following the war, Ukraine's economy showed a stronger recovery, with GDP rebounding, though inflation persisted as a challenge. Sentiment indicators improved but remained below pre-war levels, reflecting lingering uncertainty. In contrast, Russia's recovery was slower, with continued inflationary pressures and moderate sentiment recovery, suggesting a more gradual stabilization process. Pearson correlation results revealed a reversal in sentiment correlations, with TextBlob and SentiWordNet showing negative correlations with GDP (-1.0) and inflation (-1.0), while VADER and AFINN displayed positive correlations. This reversal indicates that sentiment closely tracked both the economic downturn during the war and the subsequent recovery. While the line plot captures the overall trend of recovery, the correlation analysis provides deeper insight into how sentiment responded differently across various economic conditions in each country.

By combining these insights, the findings emphasize the interconnected nature of economic performance and sentiment, particularly during periods of crisis and recovery. While Ukraine's sentiment appears more reactive to economic changes, Russia's sentiment remains more insulated, aligning with its slower but steadier economic response.

4.6 Performance Matrix

All the metrics used to measure the performance of different models is depicted in the following table with the results

Stage	Method	Country	Metric(s)	Value(s)/Key Findings
Missing Data Imputation	SARIMA	Ukraine	MAE, RMSE, MAPE	Lowest error across all indicators – used entirely for Ukraine.
	SARIMA, VAR	Russia	MAE, RMSE, MAPE	SARIMA best for CPI and Trade; VAR best for FDI and Military Spending.
Correlation Analysis	Spearman Heatmap	Ukraine	Spearman's ρ	Trade–Military Spending: -0.95, Inflation–SentiWordNet: 0.69, Unemployment and

		Russia	Spearman's ρ	Debt correlate negatively with sentiment. Unemployment–VADER : 0.73, Military Spending–Trade: -0.39; overall weaker links.
Granger Causality Test	Granger Causality	Ukraine	p-value (< 0.05)	TextBlob (lag 1): 0.0390 (significant), lag 2 results suggest delayed influence.
		Russia	p-value (< 0.05)	TextBlob (lag 2): 0.0014 - shows delayed sentiment impact.
Machine Learning Models	XGBoost vs. Random Forest	Russia	MAE	XGBoost: 75.95B, Random Forest: 134.30B – XGBoost superior.
		Ukraine	MAE	Random Forest: 212.80B, XGBoost: 229.61B – small difference; both effective.
Trend Analysis	Line Plots, Pearson Correlation	Ukraine	Pearson's r	Pre-war: GDP–TextBlob:0.71, Inflation–TextBlob:0.78. During War: GDP–TextBlob 1.0. Post-war: GDP–TextBlob: -1.0.
		Russia	Pearson's r	Pre-war: stable growth. During War: GDP–TextBlob: 1.0. Post-war: mixed recovery and correlations.

Table 4.6: Final performance metric table

4.7 Hypothesis Testing results:-

Based on the findings, the following hypotheses were observed:

1. **Sentiment-Driven Economic Shocks Are Stronger in Crisis Periods:** Pearson correlation analysis shows that during the peak war years, sentiment correlations

with GDP and inflation reached extreme values (1.0 or-1.0), indicating that economic sentiment became overwhelmingly reactive to crisis conditions. This insight is valuable for policymakers and economists worldwide, as it highlights the need to closely monitor public sentiment during global economic shocks (e.g., financial crises, pandemics, wars) to anticipate economic instability.

2. **Sentiment-Based Economic Predictions Vary Across Economic Structures :** XGBoost vs. Random Forest results show that XGBoost significantly outperforms for Russia, while Random Forest performs slightly better for Ukraine. This suggests that different models are more effective in different economic environments. In economies with strong central control (like Russia), sentiment shifts impact economic variables more gradually, while in more volatile economies (like Ukraine), sentiment-driven fluctuations are immediate and stronger. This has implications for economic forecasting in various countries.
3. **Military Expenditure Has a Strong Trade-Off with Economic Growth and Trade :** Spearman correlation analysis shows a strong negative correlation between military expenditure and trade (-0.95) in Ukraine, while in Russia, this relationship is weaker (-0.39). This suggests that in some economies, heavy defense spending comes at a significant economic cost, while others can sustain military spending with less economic disruption. This hypothesis is crucial for countries with high defense budgets (e.g., the U.S., China, India) as it suggests a potential trade-off between national security and economic prosperity. It also underscores the importance of balancing defense policies with economic growth strategies to maintain long-term stability.
4. **Economic Recovery Patterns Differ Based on Conflict Impact and Policy Response:** line plot results show that Ukraine experienced a sharper post-war economic recovery, with GDP growth rebounding faster, while Russia's recovery was slower and marked by prolonged inflationary pressures. This indicates that the depth of the initial economic shock and government responses shape recovery speed and stability. This hypothesis is applicable to post-conflict or crisis-affected economies worldwide, helping policymakers design strategies for faster economic stabilization and resilience-building in the aftermath of major disruptions.
5. **The Predictive Power of Sentiment Analysis Models Evolves Over Time:** Granger causality results show that in Ukraine, sentiment has a more immediate impact on economic indicators, while in Russia, the effects emerge over a longer period. Additionally, different sentiment models (TextBlob, SentiWordNet) show varying levels of predictive power across different time lags. This suggests that sentiment models used for economic forecasting need periodic recalibration as economic structures, media landscapes, and public sentiment dynamics evolve. It is particularly important for financial markets, policymaking, and central banks relying on AI-driven economic forecasting tools.

CHAPTER - 5

CONCLUSION & FUTURE SCOPE

This study demonstrates the profound and asymmetric influence of public sentiment on the economic trajectories of Ukraine and Russia during the Russia-Ukraine conflict. By integrating sentiment analysis with macroeconomic indicators, it uncovers distinct patterns in how economic activity and public perception interact in times of crisis. The results indicate that Ukraine exhibits stronger and more volatile correlations between sentiment and key economic variables, particularly inflation, trade, and military expenditure. In contrast, Russia's sentiment-economy relationship is more stable and exhibits delayed effects, consistent with its more centralized and controlled economic structure. Granger causality analysis reinforces this temporal divide: sentiment has an immediate impact on Ukraine's economy, whereas in Russia, the effects are lagged.Machine learning model validation further supports these observations. XGBoost consistently outperforms Random Forest for Russia, indicating higher accuracy in capturing its complex but stable economic patterns. Meanwhile, Random Forest provides slightly better predictive performance for Ukraine, reflecting the noisier and more reactive nature of its sentiment-economy dynamics. Time series analysis and line plots reveal divergent post-war recovery trends: Ukraine shows a sharper downturn in 2022 but rebounds rapidly, while Russia's recovery is slower and more controlled. Additionally, Pearson correlation with lagged sentiment scores demonstrates that sentiment becomes highly predictive during crisis periods, confirming its critical role in short- and long-term economic forecasting. These findings underscore the value of sentiment analysis as a forecasting tool in conflict-driven economies and suggest its broader applicability in anticipating and managing economic volatility.

Future research should focus on enhancing the granularity and accuracy of sentiment analysis by incorporating deep learning-based natural language processing (NLP) models such as BERT or LSTM networks. These models could better capture contextual nuances and emotional intensity across various textual data sources. Additionally, exploring real-time sentiment monitoring and its integration into macroeconomic policy tools could offer practical benefits for governments and financial institutions seeking to stabilize economies during conflicts and crises. Further expansion of the dataset to include more diverse linguistic and regional inputs could also improve model generalizability and robustness.

REFERENCE

1. Abakah, E. J. A., Adeabah, D., Tiwari, A. K., & Abdullah, M. (2023). *Effect of Russia–Ukraine war sentiment on blockchain and FinTech stocks* (Vol. 90). International Review of Financial Analysis. https://doi.org/10.1016/j.irfa.2023.102948
2. Algaba, A., Ardia, D., Bluteau, K., Borms, S., & Boudt, K. (n.d.). *Econometrics Meets Sentiment: An Overview of Methodology and Applications* . *Journal of Economic Surveys*,Vol. 34),10.1111/joes.12370
3. Aygun, O., Aygun, I., & Kaya, M. (2025). Using aspect-based sentiment analysis to evaluate the global effects of the food security crisis during the Russia-Ukraine war (Vol. 44). *Global Food Security*. https://doi.org/10.1016/j.gfs.2025.100828
4. Baker, Scott R., Bloom, Nicholas, Davis, & Steven J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593-1636. 10.1093/qje/qjw024
5. Bollen, Johan, Mao, Huina, Zeng, & Xiaojun. (2011). Twitter Mood Predicts the Stock Market. *Journal of Computational Science*, 2. 10.1016/j.jocs.2010.12.007
6. Breiman, & Leo. (2001). Random Forest. *Machine Learning*, *45*. 10.1023/A:1010933404324
7. Bruhin, J. M., Scheufele, R., & Stucki, Y. (2023). The economic impact of Russia's invasion of Ukraine on European economies. *Swiss National Bank.* https://www.snb.ch/public/publication/en/www-snb-ch/publications/research/working-papers/2023/working_paper_2023_04/0_en/working_paper_2023_04.en.pdf
8. Chen, Tianqi, Guestrin, & Carlos. (2016). XGBoost: A Scalable Tree Boosting System. *Association for Computing Machinery.* 10.1145/2939672.2939785
9. Chong, E., Ching Ho, C., Fei Ong, Z., & Ong, H. H. (2022). *Using News Sentiment for Economic Forecasting: A Malaysian Case Study.* Bank for International Settlements. https://www.bis.org/ifc/publ/ifcb57_17.pdf
10. Grebe, M., Kandemir, S., & Tillmann, P. (2024). Uncertainty about the War in Ukraine: Measurement and Effects on the German Economy (Vol. 217). *Journal of Economic Behavior & Organization.* 10.1016/j.jebo.2023.11.015
11. Hall, & Aaron. (2018). Machine Learning Approaches to Macroeconomic Forecasting. *The Federal Reserve Bank of Kansas City Economic Review*. 10.18651/ER/4q18smalterhall
12. Izzeldin, M., Muradoğlu, Y. G., Pappas, V., Petropoulou, A., & Sivaprasad, S. The impact of the Russian-Ukrainian war on global financial markets (Vol. 87). *International Review of Financial Analysis*. https://doi.org/10.1016/j.irfa.2023.102598
13. Lia, T. R., Chamrajnagara, A. S., Fonga, X. R., Rizika, N. R., & Fua, F. (2018). *Sentiment-Based Prediction of Alternative Cryptocurrency Price Fluctuations Using Gradient Boosting Tree Model*. arXiv preprint. 10.48550/arXiv.1805.00558
14. Lukauskas, Mantas, Pilinkienė, Vaida, Bruneckienė, Jurgita, Stundžienė, Alina, Grybauskas, Andrius, Ruzgas, & Tomas. (2022). *Economic Activity Forecasting Based on the Sentiment Analysis of News* (Vol. 10). Mathematics. https://www.mdpi.com/2227-7390/10/19/3461
15. Menaouer, B., Fairouz, S., Meriem, M. B., Sabri, M., & Nada, M. (2025). *A sentiment analysis of the Ukraine-Russia War tweets using knowledge graph convolutional*

networks. International Journal of Information Technology. 10.1007/s41870-024-02357-0

16. Polyzos, E. (2023). Inflation and the war in Ukraine: Evidence using impulse response functions on economic indicators and Twitter sentiment (Vol. 66). *Research in International Business and Finance.* https://doi.org/10.1016/j.ribaf.2023.102044
17. Sahi, G., Khandelwal, A., & Gupta, S. (2025). War of Tweets: Sentiment Analysis on Ukraine Russia Conflict. *Springer Nature Switzerland*. 10.1007/978-3-031-31723-1_5
18. Seki, K., Ikuta, Y., & Matsubayashi, Y. (2022). *News-based Business Sentiment and its Properties as an Economic Index. Information Processing & Management*, 59(59), 102827. 10.48550/arXiv.2110.10340
19. Sinha, A., Rout, B., Mohanty, S., Mishra, S. R., Mohapatra, H., & Dey, S. (2024). *Exploring Sentiments in the Russia-Ukraine Conflict: A Comparative Analysis of KNN, Decision Tree And Logistic Regression Machine Learning Classifiers* (Vol. 235). Procedia Computer Science. https://doi.org/10.1016/j.procs.2024.04.101
20. Sulong, Z., Abdullah, M., Abakah, E. J. A., & Asongu, S. A. (2023). Russia-Ukraine War and G7 Debt Markets: Evidence from Public Sentiment Towards Economic Sanctions During the Conflict. *International Journal of Finance & Economics.* 10.1002/ijfe.2887
21. Tetlock, & Paul C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance.* 10.1111/j.1540-6261.2007.01232.x

Isha and Garima

Final_dissertation thesis.pdf

Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:97425370

Submission Date

May 23, 2025, 10:19 PM GMT+5:30

Download Date

May 23, 2025, 10:20 PM GMT+5:30

File Name

Final_dissertation thesis_removed.pdf

File Size

1.7 MB

39 Pages

8,929 Words

52,846 Characters

10% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text

Match Groups

95 Not Cited or Quoted 10%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

5% Internet sources

4% Publications

8% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **95** Not Cited or Quoted 10%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

5% Internet sources
4% Publications
8% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1 Submitted works

Asia Pacific University College of Technology and Innovation (UCTI) on 2024-10-19 1%

2 Submitted works

Loughborough University on 2024-09-26 <1%

3 Submitted works

Monash University on 2024-08-09 <1%

4 Submitted works

University of Edinburgh on 2024-08-19 <1%

5 Submitted works

CSU, San Jose State University on 2023-05-19 <1%

6 Submitted works

Gilmour Academy on 2025-03-12 <1%

7 Submitted works

Liverpool John Moores University on 2024-04-29 <1%

8 Internet

www.mdpi.com <1%

9 Publication

Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical ... <1%

10 Submitted works

University of Wales Institute, Cardiff on 2025-01-10 <1%

11	Internet	
arxiv.org		<1%
12	Internet	
onlinelibrary.wiley.com		<1%
13	Internet	
studenttheses.uu.nl		<1%
14	Submitted works	
Asia Pacific University College of Technology and Innovation (UCTI) on 2024-07-27		<1%
15	Internet	
spoindia.org		<1%
16	Internet	
www.scilit.net		<1%
17	Submitted works	
Liverpool John Moores University on 2023-10-22		<1%
18	Submitted works	
University of Macau on 2025-04-01		<1%
19	Publication	
"The AI Revolution: Driving Business Innovation and Research", Springer Science ...		<1%
20	Submitted works	
University of Exeter on 2024-08-28		<1%
21	Publication	
Leon, Javier. "Forecasting Imported Fruit Prices in the United States Using Neural...		<1%
22	Publication	
T. Mariprasath, Kumar Reddy Cheepati, Marco Rivera. "Practical Guide to Machin...		<1%
23	Submitted works	
University of Edinburgh on 2025-03-25		<1%
24	Internet	
ro.uow.edu.au		<1%

25	Submitted works	
	College of Professional and Continuing Education (CPCE), Polytechnic University o...	<1%
26	Submitted works	
	Erasmus University of Rotterdam on 2025-04-21	<1%
27	Submitted works	
	Midlands State University on 2025-04-28	<1%
28	Submitted works	
	University of Huddersfield on 2024-09-30	<1%
29	Internet	
	core.ac.uk	<1%
30	Internet	
	www.cosic.esat.kuleuven.be	<1%
31	Internet	
	www.scirp.org	<1%
32	Submitted works	
	University of Edinburgh on 2018-08-22	<1%
33	Internet	
	worldwidescience.org	<1%
34	Internet	
	www.jstage.jst.go.jp	<1%
35	Internet	
	3a4a6dff-f30e-4e65-a2b8-55bcf5c2244c.filesusr.com	<1%
36	Submitted works	
	Khalifa University of Science Technology and Research on 2024-07-11	<1%
37	Submitted works	
	University of Hertfordshire on 2024-12-08	<1%
38	Submitted works	
	University of Hertfordshire on 2024-12-15	<1%

39 Submitted works

University of Kent at Canterbury on 2025-05-08 <1%

40 Submitted works

University of West London on 2024-05-22 <1%

41 Submitted works

WHU - Otto Beisheim School of Management on 2025-05-20 <1%

42 Internet

documents.dps.ny.gov <1%

43 Internet

onlineacademicpress.com <1%

44 Internet

ris.utwente.nl <1%

45 Submitted works

unistgallen-plagiat on 2024-12-11 <1%

46 Publication

Brahami Menaouer, Safa Fairouz, Mohammed Boulekbachi Meriem, Sabri Moham... <1%

47 Publication

Huijian Dong. "Data Analytics in Finance", CRC Press, 2025 <1%

48 Submitted works

Imperial College of Science, Technology and Medicine on 2019-09-10 <1%

49 Publication

Joanna Paliszkiewicz. "Management in the Era of Big Data - Issues and Challenges... <1%

50 Submitted works

Monash University on 2024-05-03 <1%

51 Publication

Verberne, Pieter. "Multiphysics Modelling and Characterisation of the Survivabilit... <1%

52 Submitted works

WHU - Otto Beisheim School of Management on 2025-05-19 <1%

53 Internet

bura.brunel.ac.uk <1%

54 Internet

fau.digital.flvc.org <1%

55 Internet

iimsambalpur.ac.in <1%

56 Internet

repositorio.comillas.edu <1%

57 Internet

www.journals.uchicago.edu <1%

58 Internet

www.researchsquare.com <1%

59 Submitted works

North London Collegiate School Jeju on 2024-01-16 <1%

60 Submitted works

Vrije Universiteit Brussel on 2024-11-17 <1%

61 Publication

Kevin K. F. Wong, Haiyan Song. "Tourism Forecasting and Marketing", Routledge, ... <1%

62 Submitted works

University of East London on 2025-05-09 <1%

ICCTRDA 2025: Paper Notification 320

2 messages

ICCTRDA - 2025@Vietnam <icctrda.congress@gmail.com> Sat, 24 May, 2025 at 08:12
To: Ishaojha <ishaojha08@gmail.com>

International Conference on Communication Technology Research and Data Analytics 2025: ICCTRDA 2025

Dear **Author(s),**

Greetings from **ICCTRDA 2025!**

ICCTRDA-2025 team is pleased to inform you that your paper with submission ID **320** and Paper Title '**Economic effect of Russia-Ukraine War: A Sentiment Analysis Approach'** has been accepted for presentation at "ICCTRDA2025" and for publication in the conference proceedings. The Committee thanks you for your contribution.

The conference proceedings will be published by Springer in Lecture Notes in Networks and Systems series [Indexing: SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago; All books published in the series are submitted for consideration in Web of Science]. This acceptance means that your paper is among the top 15% of the papers received/reviewed. The registrations for the conference are open. **We want to provide you with urgent information and advise you that we have limited slots available, and once they are filled, we will not be able to accommodate any further registrations. To secure your spot at this highly anticipated event, we urge you to complete your registration without delay.**

You are requested to do the registration as soon as possible and submit the following documents to **icctrda.congress@gmail.com** at the earliest.
1. Final Camera-Ready Copy (CRC) as per the springer format. (See https://www.icctrda.com/downloads)
2. Copy of e-receipt of registration fees. (For Registration, see https://www.icctrda.com/registrations)
3. The final revised copy of your paper should also be uploaded via Microsoft CMT.

The reviewers comments are given at the bottom of this letter, please improve your paper as per the reviewers comments.

The paper prior to submission should be checked for plagiarism and AI Plagiarism from licensed plagiarism softwares like Turnitin/iAuthenticate etc. The similarity content should not exceed 15% and AI similarity should not exceed 5%.

Pay registration fees via online portal:
[Kindly note – the conference being organised in Hybrid Mode and you can choose the mode of presentation in either physical (offline) or digital (online) mode; then pay the registration fees]
https://www.icctrda.com/registrations

Once you pay the registration fees, kindly fill the following google form:
https://forms.gle/LUckGUckKnNotQAcA

With Regards
Conference Chair

Reviewer-1
1. Abstract written well, managed. 2 literature review is sufficent. 3 methodology must have flowchart and explanation. 4 presence of algorithm is appreciable. 5 results and explanation required. 6 cite all references in text 7 conclusion has future work and limitations 8 proof read entire paper for gramatical errors and findings. 9 validate entire work with performance meterics. 10 format paper as per conference template 11 update figure resolutions and replace existing ones.

Reviewer-2
1. The paper must be started with introduction section including motivation, main contributions and organization of paper.
2. Literature review section must also be extended.
3. A comparative study may also be shown in graphical form.
4. Language must be improved as there are linguistic errors at some places.
5. The Limitations of the proposed study need to be discussed before conclusion.
6. Add more results and discussion.
7. The paper must be within 10-12 pages (single column); Minimum 8 pages.
8. References must be cited in the text within the paper.
9. Figures must be of higher resolution.

Isha Ojha <ishaojha08@gmail.com> Sat, 24 May, 2025 at 16:06

ADROID
FOUNDATION

Razorpay
Invoicing and payments powered by Razorpay

Payment Receipt Transaction Reference: pay_QZ4FQXw5UhesGA

This is a payment receipt for your transaction on ICCTRDA - 2025

AMOUNT PAID **₹ 15,630.60**

ISSUED TO
ishaojha08@gmail.com
+919318324280

PAID ON
25 May 2025

DESCRIPTION	UNIT PRICE	QTY	AMOUNT
Convenience Fee	₹ 300.00	1	₹ 300.00
Amount	₹ 15,330.60	1	₹ 15,330.60

Total **₹ 15,630.60**

Amount Paid ₹ 15,630.60

Economic Effects of Russia-Ukraine War: A Sentiment Analysis Approach

Isha Ojha[1], Garima Singh[1], and Sumedha Seniaray[3]

Delhi Technological University, Delhi, India
[1]ishaojha08@gmail.com
[2]gswork2308@gmail.com
[3]sumedhaseniaray@dtu.ac.in

Abstract. The Russia-Ukraine war has significantly impacted the economies of both nations. This study integrates public sentiment with the economy. Key indicators were collected from the World Bank and the International Monetary Fund (IMF). Missing values were predicted using Seasonal AutoRegressive Integrated Moving Average (SARIMA) and Vector AutoRegression (VAR). Public sentiment was derived from tweets, news articles, and Reddit comments, analyzed using multiple sentiment analysis tools. Annual sentiment scores were correlated with economic indicators using Spearman correlation and Granger causality tests. Pearson correlation with lagged sentiments and line plots captured sentiment patterns across different war phases. eXtreme Gradient Boosting (XGBoost) and Random Forest were used to assess sentiment's predictive power in economic forecasting. Results indicate Ukraine's economy is highly sensitive to certain factors, while Russia's resilience is state-driven, with sentiment having an immediate impact on Ukraine's economy but a delayed effect on Russia. Findings highlight sentiment analysis's role in economic forecasting, aiding policymakers and analysts.

Keywords: Russia-Ukraine War, Economic Impact, Sentiment Analysis, Time Series Forecasting, Machine Learning.

1 Introduction

Wars cause major economic disruptions, affecting trade, military spending, GDP, inflation, and employment. Since February 2022, the Russia-Ukraine war has worsened inflation, reduced foreign investment, and increased government debt. Sanctions imposed on Russia, along with damage to Ukraine's infrastructure, deepen instability, impacting markets and public sentiment, both of which influences economic decisions and consumer confidence. This study measures economic uncertainty using sentiment analysis of news and social media. Unlike most prior research focusing on either sentiment or economic forecasting alone, we combine both for a comprehensive view of war-driven economic disruptions. To address this gap, this study integrates sentiment analysis with time series economic forecasting. The key contributions of our work are as follows:

- Our findings provide valuable information on how public sentiment can serve as an early indicator of economic downturns or recoveries, aiding in economic decision-making.
- The correlation between sentiment and economic indicators suggests that similar approaches could be used to predict economic trends during future geopolitical crises.
- Our work contributes to the literature on war economics by demonstrating how public perception influences and reflects economic realities.

2 Related Work

The intersection of sentiment analysis and economic forecasting has gained increasing attention, especially in financial markets and macroeconomic analysis. This section reviews key research on sentiment analysis, statistical causality, and machine learning for economic prediction.

Sentiment Analysis on the Russia-Ukraine War: Sahi et al. [16] applied machine learning to Twitter data to gauge public sentiment. Sinha et al. [18] examined temporal sentiment shifts, while Menaouer et al. [14] used knowledge graph convolutional networks for war related sentiment analysis.

Sentiment and Economic Indicators Several studies have linked sentiment to economic variables. Polyzos [15] assessed tweet influence on macro-financial metrics. Aygun et al. [3] used aspect-based sentiment analysis to study global food security. Abakah et al. [1] explored effects on blockchain and FinTech stocks, and Bollen et al. [5]analyzed sentiment's role in stock markets. Grebe et al. [10]studied war-related uncertainty in the German economy, while Sulong et al. [19] examined sentiment on G7 debt markets amid sanctions. Izzeldin et al.[11] compared market reactions to the war and the COVID-19 pandemic. Bruhin et al. [7] analyzed impacts across Europe.

Sentiment Analysis in Economic Forecasting Beyond conflict studies, sentiment aids broader economic forecasting. Algaba et al. [2] proposed methods to convert sentiment into economic indicators. Lukauskas et al. [13], Chong et al. [9] and Seki et al. [17] applied media sentiment to predict inflation and GDP. Foundational works by Baker et al. [4] and Tetlock [20]highlight how the policy and media sentiment shape markets. Ray Li et al. [12] showed Twitter sentiment's predictive power on cryptocurrency prices.

To the best of our knowledge, this study offers a more comprehensive approach than prior research by integrating sentiment trends and macroeconomic indicators, rather than analyzing them in isolation.

3 Background

This section provides the theoretical foundation of the methodologies used in this study and explains how sentiment trends can be linked to economic indicators.

Time series forecasting for economic data completion and sentiment analysis for war-related public opinion Time series forecasting predicts future values using historical data patterns. Vector Autoregression (VAR) models relationships among multiple economic variables, while Seasonal Autoregressive Integrated Moving Average (SARIMA) extends the ARIMA model with seasonality for univariate forecasts. PROPHET handles missing data and trend shifts effectively. Model accuracy was evaluated using Mean Absolute Error (MAE), which measures the average absolute difference between actual and predicted values. A lower MAE indicates better performance.

Sentiment analysis quantifies emotions in text. Models used include TextBlob (a lexicon-based model scoring from -1 to +1), VADER (Valence Aware Dictionary and Sentiment Reasoner, ideal for social media), Afinn (ranging from -5 to +5), and SentiWordNet (which provides word-level sentiment scores).

Correlation Analysis Between Economic Indicators and Sentiments The Spearman heatmap visualizes correlations between sentiment scores and economic indicators, with high values indicating strong links between public perception and economic shifts. The Granger causality test determines whether past sentiment can predict future economic changes. Line plots were used for trend analysis, while lagged Pearson correlations assessed sentiment's predictive power with positive values signaling growth and negative values indicating decline.

To validate predictive accuracy using historical data, XGBoost (Chen and Guestrin) [8] and Random Forest (Breiman)[6] were employed. XGBoost captures nonlinear sentiment-economy relationships with high accuracy, while Random Forest averages predictions from multiple decision trees. Together, these approaches provide a multi-dimensional understanding of how sentiment influences economic dynamics over time.

Heatmap analysis reveals correlation patterns, the Granger causality test identifies the direction of influence, and time-series visualizations track how sentiment and economic trends evolve over time.

4 Methodology

This study integrates time series forecasting, sentiment analysis, and correlation analysis to examine the economic effects of the Russia-Ukraine war. The methodology consists of three main phases: economic data preparation, sentiment analysis, and correlation analysis, as shown in Fig 1.

Phase 1 : Economic Data Collection and Forecasting for Correlation Analysis

Economic data for Russia and Ukraine were sourced from the World Bank and IMF. Missing values were imputed using the most accurate time series forecasting models, selected based on error analysis.

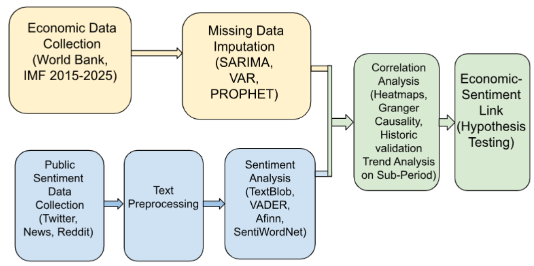


Fig. 1: Proposed methodology flowchart

Data Collection and Missing Value Imputation Using Time Series Models Yearly economic data for Russia and Ukraine (2015–2025) was sourced from the World Bank and IMF, covering CPI, inflation, GDP growth, unemployment, government debt, trade, FDI net inflows, and military expenditure. Data was stored in CSV files for each country. However, Trade, FDI, and military expenditure data were missing for 2024–2025 in both countries, while Russia's CPI data was unavailable for 2022–2023. To address these gaps, SARIMA, VAR, and Prophet time series models were tested and the model with the lowest prediction error was selected for each series.

Phase 2 : Sentiment Analysis of Public Reactions to the War

Tweets, Reddit comments, and news articles were collected, preprocessed, and analyzed for sentiment using TextBlob, VADER, AFINN, and SentiWordNet.

Data Collection and Sentiment Scoring A total of 503,696 text records were collected from Twitter, Reddit, and news articles (via Kaggle and Google NewsAPI), covering the years 2015–2025. This enabled sentiment trend analysis before, during, and after the peak Russia-Ukraine war period. Duplicates were removed, and all records were merged into a dated text file. The raw text was cleaned using standard NLP techniques, including lowercasing, removal of URLs, hashtags, usernames, punctuation, special characters, stopwords, emojis, and repeated characters, along with lemmatization. Each entry was then analyzed using four sentiment models: TextBlob, VADER, Afinn, and SentiWordNet. Sentiment polarity scores were computed and aggregated annually to track sentiment trends over time.

Phase 3 : Correlation Analysis Between Sentiment & Economic Indicators

This section outlines the methods used to examine the relationship between sentiment and economic indicators through statistical analysis and machine learning-based predictive validation.

Visualization & Statistical Testing To examine sentiment's link to economic indicators, methods included Spearman heatmaps (showing direction and strength of associations), Granger causality tests (assessing sentiment's predictive power), and time-series plots (revealing parallel trends). Pearson correlations with one-year lagged sentiment further highlighted its potential as an early signal for changes in GDP, inflation, and unemployment, emphasizing sentiment's forecasting value.

XGBoost and Random Forest for Historical Validation Using Sentiments Historical validation in machine learning involves training models on past data (2015–2023) to predict future outcomes (2024–2025). In this case, sentiment analysis models are used to predict key economic indicators for Ukraine and Russia. The goal is to determine which sentiment model most accurately captures economic trends using machine learning models: XGBoost and Random Forest.

Hypothesis Testing

This study tests five key hypotheses:

1. Sentiment-Driven Economic Shocks Are Stronger in Crisis Periods.
2. Sentiment-Based Economic Predictions Vary Across Economic Structures.
3. Military Expenditure Has a Strong Trade-Off with Economic Growth and Trade.
4. Economic Recovery Patterns Differ Based on Conflict Impact and Policy Response.
5. The Predictive Power of Sentiment Analysis Models Evolves Over Time.

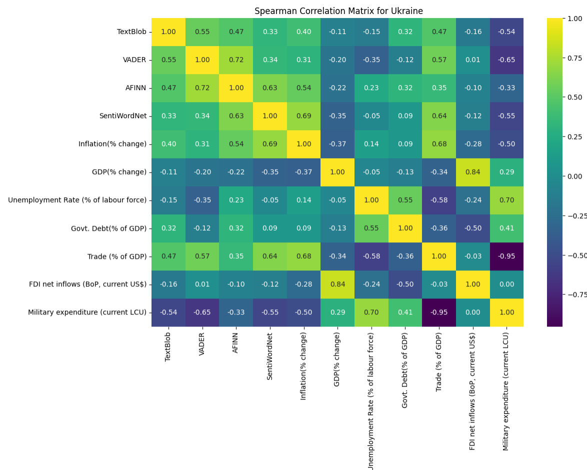
By integrating these methodologies, this study provides a rigorous, data-driven approach to understanding how sentiment influences economic trends during geopolitical conflicts.

5 Results and Discussion

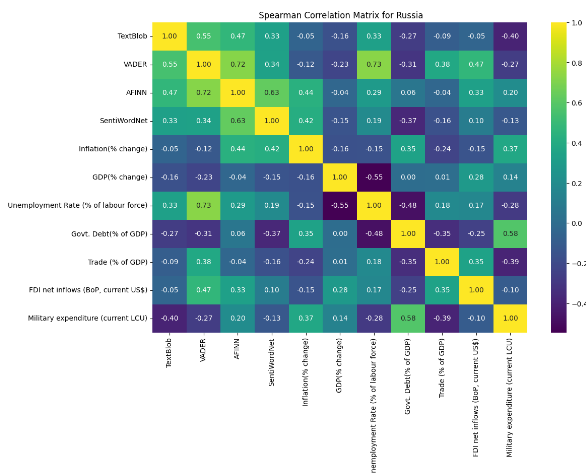
To fill gaps in the economic data (2015–2025), SARIMA, VAR, and Prophet models were evaluated using MAE, RMSE, and MAPE. SARIMA showed the lowest errors for Ukraine and was used entirely. For Russia, SARIMA worked best for CPI and trade, while VAR outperformed for FDI and military spending, resulting in complete datasets for both countries. Now, to investigate the relationship between sentiment and economic indicators during the Ukraine-Russia conflict, we used correlation heatmaps, Granger causality tests, trend analysis, and machine learning models to uncover patterns and evaluate predictive performance.

5.1 Correlation Analysis: Spearman Heatmaps

Fig 2 shows Spearman correlation heatmaps for Ukraine and Russia, we can draw insights into how economic factors interact with sentiment scores in both countries.



(a) Spearman heatmap for Ukraine



(b) Spearman heatmap for Russia

Fig. 2: Spearman correlation heatmaps for Ukraine and Russia economic factors with sentiment score

Spearman heatmaps reveal differences between Ukraine and Russia in sentiment-economic links. Ukraine shows a strong negative correlation between military spending and trade (-0.95) and high inflation correlation with SentiWordNet sentiment (0.69), indicating public sensitivity to economic instability. Unemployment and government debt similarly align with negative sentiment. In Russia, unemployment correlates strongly with VADER sentiment (0.73), but other links are weaker. Government debt has modest negative correlations, and mili-

Table 1: Granger Causality Test Results for Ukraine

Lag	TextBlob	VADER	AFINN	SentiWordNet
1	0.0390	0.7027	0.9906	0.5549
2	0.1279	0.7699	0.2058	0.1105

Table 2: Granger Causality Test Results for Russia

Lag	TextBlob	VADER	AFINN	SentiWordNet
1	0.5270	0.9202	0.6476	0.1679
2	0.0014	0.9541	0.6543	0.5647

tary spending's weaker impact on trade (-0.39) suggests Russia's trade is more resilient to defense expenditure.

5.2 Granger Causality Test

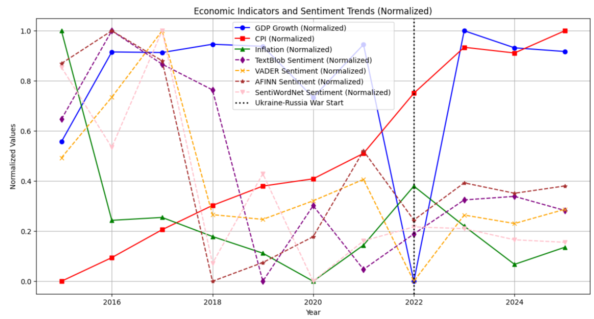
The Granger Causality test assesses whether past sentiment scores can predict economic factors, with p-values below 0.05 indicating significant causal links. Results show differing impacts in Ukraine and Russia. In Ukraine, TextBlob sentiment at lag 1 ($p = 0.0390$) significantly predicts economic indicators, while SentiWordNet ($p = 0.1105$) and TextBlob ($p = 0.1279$) at lag 2 suggest a possible delayed effect. In Russia, only TextBlob at lag 2 ($p = 0.0014$) shows significance. This indicates sentiment influences Ukraine's economy more immediately, while in Russia, effects appear more gradually reflecting Ukraine's higher volatility and Russia's slower, more stable economic response.

5.3 Comparing XGBoost and Random Forest for Historical Validation Using Sentiments

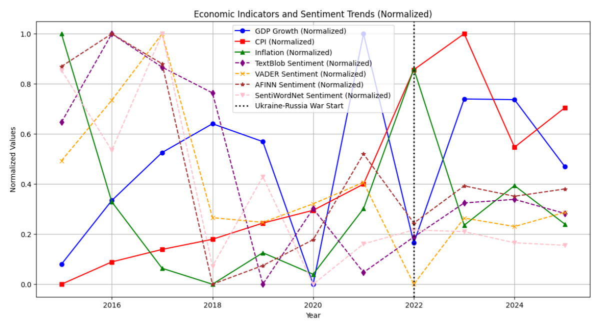
For Russia, XGBoost outperforms Random Forest (MAE: 75.95B vs. 134.30B), showing better predictive power. For Ukraine, Random Forest (MAE: 212.80B) slightly outperforms XGBoost (MAE: 229.61B); however, the difference is marginal. Overall, XGBoost is superior, offering more accurate forecasts for Russia and comparable performance for Ukraine. This suggests XGBoost's strength in capturing complex sentiment-economic relationships makes it more effective for economic forecasting than Random Forest.

5.4 Trend Analysis

This section analyzes the temporal relationships between sentiment and economic indicators, highlighting how their interactions shifted before, during, and after the Ukraine-Russia war.



(a) Line plot for Ukraine



(b) Line plot for Russia

Fig. 3: Line plots for Ukraine-Russia economic indicator vs sentiment trends

Line Plot and Pearson Correlation Analysis with Lagged Sentiment The line plots in Fig. 3 show the evolution of GDP growth, CPI, inflation, and sentiment scores for Ukraine and Russia. From 2015 to 2021, Ukraine's economy was volatile, with fluctuating GDP and unstable inflation. Sentiment mirrored this, with TextBlob strongly correlating with GDP (0.71) and inflation (0.78), indicating tight sentiment-economy links. Russia showed steadier growth, lower inflation, but gradually declining sentiment and weaker correlations. The 2022 war triggered GDP drops, inflation spikes, and sentiment crashes, with correlations peaking at 1.0, signaling synchronized economic and psychological shocks. Line plots confirm these disruptions, especially for Ukraine. Sentiment models diverged: TextBlob and SentiWordNet tracked economic shifts, while VADER and AFINN showed weaker or opposing trends, particularly in Russia. Post-war (2023–2024), Ukraine rebounded faster despite persistent inflation, while Russia's recovery was slower. Sentiment partly recovered, but reversed correlations (e.g., TextBlob-GDP: –1.0) reveal the model-dependent and dynamic nature of sentiment during recovery phases.

Performance Metric Table

Stage	Method	Country	Metric(s)	Value(s) / Key Findings
Missing Data Imputation	SARIMA	Ukraine	MAE, RMSE, MAPE	Lowest error across all indicators – used entirely for Ukraine.
	SARIMA, VAR	Russia	MAE, RMSE, MAPE	SARIMA best for CPI and Trade; VAR best for FDI and Military Spending.
Correlation Analysis	Spearman Heatmap	Ukraine	Spearman's ρ	Trade–Military Spending: -0.95, Inflation–SentiWordNet: 0.69, Unemployment and Debt correlate negatively with sentiment.
		Russia	Spearman's ρ	Unemployment–VADER: 0.73, Military Spending–Trade: -0.39; overall weaker links.
Granger Causality Test	Granger Causality	Ukraine	p-value (< 0.05)	TextBlob (lag 1): 0.0390 (significant), lag 2 results suggest delayed influence.
		Russia	p-value (< 0.05)	TextBlob (lag 2): 0.0014 - shows delayed sentiment impact.
Machine Learning Models	XGBoost vs. Random Forest	Russia	MAE	XGBoost: 75.95B, Random Forest: 134.30B – XGBoost superior.
		Ukraine	MAE	Random Forest: 212.80B, XGBoost: 229.61B – small difference; both effective.
Trend Analysis	Line Plots, Pearson Correlation	Ukraine	Pearson's r	Pre-war: GDP–TextBlob:0.71, Inflation–TextBlob:0.78. During War: GDP–TextBlob 1.0. Post-war: GDP–TextBlob: -1.0.
		Russia	Pearson's r	Pre-war: stable growth. During War: GDP–TextBlob: 1.0. Post-war: mixed recovery and correlations.

5.5 Hypothesis Testing results

Based on the findings, the following hypotheses were assessed:

1. **Sentiment-Driven Economic Shocks Are Stronger in Crisis Periods :** During peak war years, Pearson correlations between sentiment and GDP/inflation reached ±1.0, showing extreme reactivity to crisis conditions. This highlights the importance of monitoring public sentiment during global shocks like wars, pandemics, or financial crises to anticipate economic instability.
2. **Sentiment-Based Economic Predictions Vary Across Economic Structures :** XGBoost outperforms for Russia, while Random Forest performs slightly better for Ukraine. In stable, centrally controlled economies (e.g., Russia), sentiment influences unfold gradually. In volatile economies (e.g., Ukraine), effects are immediate. This suggests model choice should align with the economic structure.
3. **Military Expenditure Has a Strong Trade-Off with Economic Growth and Trade :** Ukraine shows a strong negative correlation between military expenditure and trade (-0.95), compared to Russia's weaker link (-0.39). This implies high defense budgets may hinder trade and growth in some economies. Nations with large military spending (e.g., U.S., China, India) must balance security with long-term economic health.
4. **Economic Recovery Patterns Differ Based on Conflict Impact and Policy Response:** Line plots show Ukraine's faster post-war recovery, while Russia's was slower, marked by persistent inflation. This suggests recovery speed depends on initial shock depth and policy response- vital insight for post-crisis economic planning.
5. **The Predictive Power of Sentiment Analysis Models Evolves Over Time:** Granger causality shows immediate sentiment impact in Ukraine, but delayed effects in Russia. Model effectiveness also varies over time, stressing the need for regular recalibration in economic forecasting.

These tests show sentiment is a leading economic indicator during crises, though its predictive strength varies by method and context.

6 Conclusion and Future Work

The study highlights distinct sentiment-economy dynamics in Ukraine and Russia. Ukraine shows stronger, more volatile sentiment links to inflation, trade, and military spending, while Russia's are steadier. Granger causality reveals immediate sentiment impact in Ukraine but delayed effects in Russia. XGBoost outperforms Random Forest overall, particularly for Russia. Line plots show Ukraine's faster post-war recovery. Lagged Pearson correlations confirm sentiment's predictive power, especially during crises. Overall, sentiment plays a key role in forecasting, with Ukraine's economy more sentiment-driven and reactive, while Russia's responds more gradually. However, this study is limited by the yearly resolution of economic data, potentially masking short-term sentiment effects. Future research could improve accuracy by integrating deep learning NLP models and applying sentiment analysis in policymaking to better manage economic disruptions during conflicts.

Bibliography

[1] Emmanuel Joel Aikins Abakah, David Adeabah, Aviral Kumar Tiwari, and Mohammad Abdullah. Effect of russia–ukraine war sentiment on blockchain and fintech stocks. *International Review of Financial Analysis*, 90:102948, 2023.

[2] Andres Algaba, David Ardia, Keven Bluteau, Samuel Borms, and Kris Boudt. Econometrics meets sentiment: An overview of methodology and applications. *Journal of Economic Surveys*, 34(3):512–547, 2020.

[3] Osman Aygun, Irfan Aygun, and Mehmet Kaya. Using aspect-based sentiment analysis to evaluate the global effects of the food security crisis during the russia-ukraine war. *Global Food Security*, 44:100828, 2025.

[4] Scott R. Baker, Nicholas Bloom, and Steven J. Davis. Measuring economic policy uncertainty*. *The Quarterly Journal of Economics*, 131(4):1593–1636, 07 2016.

[5] Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, 2011.

[6] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[7] Jonas M. Bruhin, Rolf Scheufele, and Yannic Stucki. The economic impact of russia's invasion of ukraine on european economies. Working Paper 2023-04, Swiss National Bank, 2023.

[8] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery.

[9] Eilyn Chong, Chiung Ching Ho, Zhong Fei Ong, and Hong H. Ong. Using news sentiment for economic forecasting: A malaysian case study. In *Machine Learning in Central Banking*, number 57 in IFC Bulletin. Bank for International Settlements, 2022.

[10] Moritz Grebe, Sinem Kandemir, and Peter Tillmann. Uncertainty about the war in ukraine: Measurement and effects on the german economy. *Journal of Economic Behavior & Organization*, 217:506, 2024.

[11] Marwan Izzeldin, Yaz Gülnur Muradoğlu, Vasileios Pappas, Athina Petropoulou, and Sheeja Sivaprasad. The impact of the russian-ukrainian war on global financial markets. *International Review of Financial Analysis*, 87:102598, 2023.

[12] Tianyu Ray Lia, Anup S. Chamrajnagara, Xander R. Fonga, Nicholas R. Rizika, and Feng Fua. Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model. *arXiv preprint*, 2018.

[13] Mantas Lukauskas, Vaida Pilinkienė, Jurgita Bruneckienė, Alina Stundžienė, Andrius Grybauskas, and Tomas Ruzgas. Economic activity forecasting based on the sentiment analysis of news. *Mathematics*, 10(19), 2022.

[14] Brahami Menaouer, Safa Fairouz, Mohammed Boulekbachi Meriem, Mohammed Sabri, and Matta Nada. A sentiment analysis of the ukraine-russia war tweets using knowledge graph convolutional networks. *International Journal of Information Technology*, 2025.

[15] Efstathios Polyzos. Inflation and the war in ukraine: Evidence using impulse response functions on economic indicators and twitter sentiment. *Research in International Business and Finance*, 66:102044, 2023.

[16] Geetanjali Sahi, Aditi Khandelwal, and Shubheshwar Gupta. War of tweets: Sentiment analysis on ukraine russia conflict. In Manish Gupta, Shikha Agrawal, Kamlesh Gupta, Jitendra Agrawal, and Korhan Cengis, editors, *Machine Intelligence and Smart Systems*, pages 62–72, Cham, 2025. Springer Nature Switzerland.

[17] Kazuhiro Seki, Yusuke Ikuta, and Yoichi Matsubayashi. News-based business sentiment and its properties as an economic index. *Information Processing & Management*, 59(2):102827, 2022.

[18] Aaryan Sinha, Bijayalaxmi Rout, Sushree Mohanty, Soumya Ranjan Mishra, Hitesh Mohapatra, and Samik Dey. Exploring sentiments in the russia-ukraine conflict: A comparative analysis of knn, decision tree and logistic regression machine learning classifiers. *Procedia Computer Science*, 235:1068–1076, 2024. International Conference on Machine Learning and Data Engineering (ICMLDE 2023).

[19] Zunaidah Sulong, Mohammad Abdullah, Emmanuel J. A. Abakah, and Simplice A. Asongu. Russia-ukraine war and g7 debt markets: Evidence from public sentiment towards economic sanctions during the conflict. *International Journal of Finance & Economics*, 2023.

[20] Paul C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 2007.