PIONEER
ACADEMIC PUBLISHING LIMITED

Journal of World Economy
ISSN 2709-3999
www.pioneerpublisher.com/jwe
Volume 3 Number 4 December 2024

Predictive Modeling and Analysis of Socio-Economic Indicators: A Data-Driven Approach to Forecasting and Decision-Making for Mongolia

Beiduo Teng[1]

[1] Guangdong Country Garden School, Guangdong, China
Correspondence: Beiduo Teng, Guangdong Country Garden School, Guangdong, China.

doi:10.56397/JWE.2024.12.08

1. Introduction

Most recently, socio-economic dynamics, human rights, and development challenges have gained wide attention, specifically in emerging nations like Mongolia. Over the last few decades, concern with the disciplinary interface of socio-economic systems, human rights, and development problems has escalated, particularly in the context of emergent Mongolian nations. The country's socio-political environment, which consists of the traditional and post-modern ideals of development, makes the country an exciting object of analysis for the significance of socio-economic factors for the country's development. This project uses advanced machine learning methods to analyze these dynamics and identify fundamental patterns that shape Mongolia's developmental path.

This research uses broader elements like GDP growth, unemployment, poverty, and more specific ones like voice and accountability, disaster displacement, and so forth to explain the current socio-economic situation in Mongolia. The project uses time series forecasting, clustering, and causal inference to draw insights and present actionable recommendations that may help policymakers, researchers, and development agencies make better decisions.

This report will describe the method used when analyzing the data collected, the result obtained from the data analysis, and the potential implications of the result for the future development of Mongolia in terms of both social and economic development.

2. Background

Mongolia, a center between Russia and China, has experienced economic changes in recent decades. Mongolia has rich mineral resources and has become more integrated into the global economy; the country has seen high growth, development, and instability. Yet, despite the availability of resources in Mongolia, the government still has many problems. This is coupled with unemployment, persistent poverty, social inequality, environmental challenges like harsh winters, and displacement due to natural disasters.

Governance and human rights are also significant concerns. Maintaining Democracy and human rights and improving governance through voice and accountability are substantial and essential to the country's sustainability and continuous improvement of its performance. However, their effectiveness in boosting economic growth and stabilizing society is still a concern.

This report examines these complex linkages through a historical analysis of data from the World Bank. The work employs the most advanced machine learning models to analyze the socio-economic trends and the relationships between various socio-economic variables and to predict the most critical indicators that characterize Mongolia's development. The outcomes of this research can help prepare policy strategies, disseminate practical information for economic development planning, and identify conditions where social issues can be addressed by targeted measures to reduce social problems.

3. Data Description

3.1 Data Source

The dataset applied in this project is obtained from the World Bank data, a reliable international organization with numerous databases on global economic and social statistics for countries. The information included in this dataset covers key socio-economic indicators of Mongolia concerning economic development, citizenship, politics, social security, and population changes. It contains financial data (GDP growth, unemployment rates) and non-financial data (voice and accountability ratings, etc.), which allow for the detailed study of the country's evolution over time.

Data from several decades encompasses historical context and recent trends. This allows an understanding of how outgoing patterns of socio-economic changes occurred in Mongolia and multi-causal dependencies and interconnections. The data set is highly integrated, including economic, social, environmental, and governance indicators, which creates an opportunity to examine the development problems in Mongolia from several angles.

3.2 Data Structure

The dataset's structure is tabular, where each row corresponds to a particular socio-economic characteristic of Mongolia, and each column groups attribute and value measures of the indicators in several years. The dataset is many-faceted and has numerous socio-economic characteristics denoted by the variables used in the research.

3.2.1 Data Dictionary

The following table describes the key variables (indicators) within the dataset:

Variable Name	Description	Type
Country Name	The name of the country to which the data belongs (Mongolia in this case).	Categorical
Country Code	The 3-letter country code for Mongolia (MNG).	Categorical
Indicator Name	The name of the socio-economic indicator being measured (e.g., GDP growth, Unemployment).	Categorical
Indicator Code	The code is associated with each indicator in the World Bank database.	Categorical
Year	The year for which the data is recorded.	Numeric
Value	The value of the respective socio-economic indicator for the given year.	Numeric
GDP growth (annual %)	The country's Gross Domestic Product (GDP) is the annual percentage change.	Numeric
Unemployment, total (% of total labor force)	The percentage of the total labor force that is unemployed.	Numeric
Poverty headcount ratio at national poverty lines	Percentage of the population living below the national poverty line.	Numeric
Voice and Accountability (Percentile Rank)	The percentile rank reflects perceptions of a country's governance, specifically its political freedom.	Numeric
Intentional homicides (per 100,000 people)	The number of intentional homicides per 100,000 population.	Numeric

Internally displaced persons (new displacement)	The number of people newly displaced due to disasters within the country.	Numeric

3.3 Rationale for Using This Dataset

The chosen dataset is directly aligned with the core objectives of this project, which is to explore and predict the socio-economic trends of **Mongolia** while identifying significant relationships between various development indicators. The rationale behind selecting this dataset includes the following key factors: economic performance, social stability, and governance variables, such as GDP growth rate, unemployment, poverty rates, and voice and accountability, respectively. This enables a comprehensive evaluation of Mongolia's development path. The dataset contains historical data that spans several decades, allowing one to make long-range forecasting and use machine learning techniques like time series analysis and LSTM neural networks. This is important to determine how historic movements could influence future movements. The dataset is composed of economic, social, and political variables, which enables us to analyze how two perhaps unrelated concepts might affect one another (such as the influence of governance on growth or the role of natural disasters on poverty). This diversity is helpful for the comprehensive approach to the analysis, which is necessary because of the project's interdisciplinary character. The data set for this study complies with World Bank guidelines for data collection and definitions of the indicators. This makes the data suitable for other complex analytical works such as clustering, causality, and machine learning. This information is critical for policy analysis. Other numeric variables such as the unemployment rate, poverty rate, and governance index are convenient for policy recommendation. As such, by predicting these indicators and discovering the underlying causal mechanisms, the project can offer specific suggestions to guide the allocation of interventions.

This dataset is perfect for the project objectives of finding and predicting future opportunities for social injustice, human rights abuse, and economic development. Such possibilities make it feasible to employ various advanced machine learning models to accomplish these goals and to gain essential insights into the socio-economic problems of Mongolia.

3.4 Data Cleaning and Preprocessing

To make it appropriate for analysis, the data needed extensive preprocessing. In this, it was necessary to remove the cases where values were missing, scale the numerical variables, and transform the data to be appropriate for forecasting and regression modeling. To achieve manageability and intense focus in the dataset, only country-specific information relevant to Mongolia was included, which was subsequently refined to concentrate on key socio-economic indicators.

4. Methodology

This section provides the methods adopted in this study, the techniques used, and the models applied to study and forecast major sociometric indicators in Mongolia. EDA follows data preprocessing and combines machine learning and statistical modeling to extract information and predict future developments. The approach was chosen to maximize the findings' credibility, reliability, and generalisability to policy and development practice.

4.1 Data Preprocessing

Therefore, pre-processing helps clean the data before it is processed through Analytical or Machine learning algorithms. Firstly, because of the nature of the data collected from different sources over several decades, where several socio-economic indicators are measured, data cleaning and preparation were obligatory to prepare the data for analysis.

4.1.1 Filtering and Data Cleaning

The raw dataset contained several global socio-economic features; however, only features corresponding to Mongolia were selected for this project. This involved screening out other records that are not related and only selecting the Mongolian data alone. Most indicators included missing observations, especially in the first years, where the information was probably not well recorded. Some columns with missing data in all the years were omitted to ease the analysis of the dataset. As for the key indicators, such as GDP growth, unemployment, and poverty rate, the rows for the critical years, including 2020 and 2021, were excluded in this case as well to avoid noise in the analysis.

4.1.2 Reshaping Data for Time Series Analysis

To facilitate time series forecasting, the data was transformed from a wide format where years were used as columns to a long format where each record had a year, an indicator, and its associated value. This structure allowed for the application of complex algorithms for temporal analysis, including the LSTM (Long Short-Term Memory) networks and the ARIMA (AutoRegressive Integrated Moving Average). This transformation was essential to facilitate the ability of the models to use history and make future values.

4.1.3 Handling Missing Values

In time series analysis, it becomes essential to deal with missing data to ensure that bias does not influence the forecast. Rather than filling in missing values that could be completely misleading, only records containing missing ones in one or several vital variables have been omitted. This was due to the realization that imputing missing socio-economic data in essential attributes such as the GDP could distort actual conditions. These records were excluded to improve the dataset's quality for predicting future outcomes and to make it a cleaner data set.

4.1.4 Normalization and Scaling

To get the data ready for machine learning models, especially neural networks such as LSTM, the numeric values of the indicators were normalized with MinMaxScaler. This change converted values to a scale from 0 to 1, a necessity for neural networks to execute correctly. To guarantee that all indicators, no matter their unit of measurement (e.g., percentages, population counts), get equal treatment by the model, normalization helps to avoid dominance by a single feature because of scale differences.

4.2 Exploratory Data Analysis (EDA)

Exploratory Data Analysis was done to ascertain the distribution of the data, discover trends, and uncover relationships connected to multiple socio-economic indicators. Hence, line plots, correlation heat maps, and bar charts were used to analyze the data and identify the pattern.

4.2.1 Time Series Analysis

The line plots were applied to show the time series data of the selected essential variables, including GDP growth, unemployment, and homicide rates. This supported the recognition of cyclical patterns, trends, and every anomaly or significant change that might have occurred in response to societal-political events. The time-series forecasting models used in the study were based on these visualizations.

4.2.2 Correlation Analysis

A correlation matrix was computed to analyze the association between various socioeconomic factors. This matrix also enabled the identification of both positive and negative connections between variables. It provided information on how variables, e.g., unemployment and poverty, might be connected or impacted by other variables, such as GDP growth. Knowledge of these relationships was essential for causal inference and policy analysis conducted in the later stages of the project.

4.3 Machine Learning Techniques

The socio-economic factors were analyzed and predicted using traditional machine learning Models and advanced neural networks.

4.3.1 ARIMA for Time Series Forecasting

The ARIMA model was used to forecast future values of the primary variable — GDP growth. ARIMA is a robust method of analyzing a data set regarding the dependence of observation on past observations, making it relevant for time series data. This was done by implementing the Auto ARIMA technique whereby the model checks through different options for the best fitting components for the ARIMA structure to reduce error magnitudes. Therefore, the selected model was applied to forecast future GDP growth rates for the following five years, which helped to identify the trends in the Mongolian economy.

4.3.2 LSTM (Long Short-Term Memory Networks) for Forecasting

To boost the accuracy of time series predictions, LSTM neural networks were put into action. LSTM represents a recurrent neural network (RNN) that works with sequential data and captures long-term dependencies, making it a perfect fit for forecasting socio-economic indicators. With the help of historical data, the LSTM model achieved the ability to predict upcoming GDP growth values. The selection of LSTMs was motivated by their talent for controlling the vanishing gradient problem and capturing both long-term and short-term movements in the data.

To stop overfitting, the LSTM architecture integrated Dropout layers. Metrics like Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R^2, and Mean Absolute Percentage Error (MAPE) were applied to analyze the performance of the model. These measurements thoroughly reviewed the model's ability to forecast future results.

4.4 Clustering and Anomaly Detection

4.4.1 K-Means Clustering

K-means clustering was used to cluster similar socio-economic trends. This unsupervised learning method was used to partition the socioeconomic data of Mongolia into clusters and get groups of years where indicators were identical. This was particularly useful for identifying trends in inequality, economic or poor performance, or the reverse, and determining which year was above or below average regarding socio-economic performance.

4.4.2 Isolation Forest for Anomaly Detection

For detecting outliers in the most critical variables, for instance, a sharp increase in unemployment or an increase in GDP growth rates by several points, the algorithm of the Isolation Forest was used. This algorithm helps find outliers in high-dimensional data; in this case, that means years where the socio-economic values substantially differed from the norm. In pointing out anomalies, the model offered suggestions on years that should be studied in more detail because of changes in policy, shocks, or governance problems.

4.5 Data Analysis and Results

4.5.1 Exploratory Data Analysis

The EDA investigation of Mongolia's socio-economic indicators revealed significant findings concerning trends and the relationships among critical variables such as GDP growth, unemployment, poverty, homicides, and voice and accountability. This investigation aids in the understanding of the essential socio-economic dynamics and guides likely policy recommendations. A summary of the findings is in line with the project objectives below.

(1) Time Series of Intentional Homicides (per 100,000 people)

The graph of intentional homicides as a time series reveals a dramatic fall from 2005 to 2010, with homicide rates going from about 15 per 100,000 people to less than 8. The trend suggests that Mongolia observed essential improvements in public safety during this period. The homicide rate stayed around 6 to 8 after 2010, indicating that the decrease in violent crime maintained that trend.

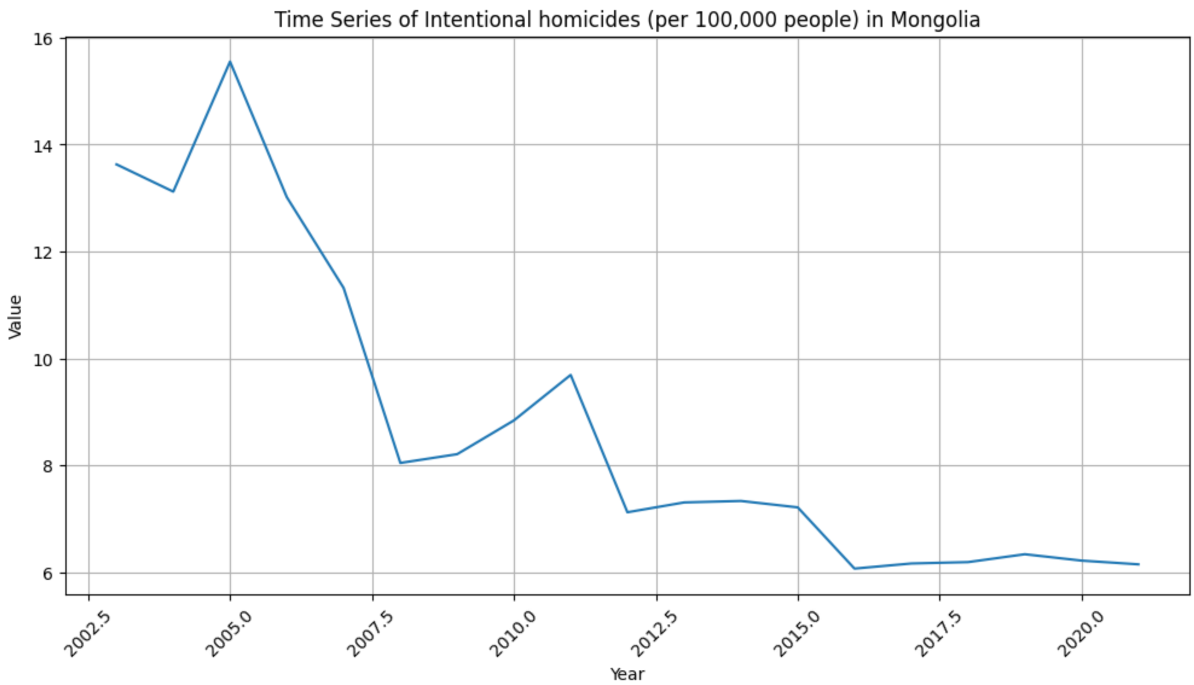

Figure 1. Time Series of Intentional Homicides (per 100,000 people)

The decline in homicide rates might be caused by better law enforcement or improvements in social factors. Analyzing the causes of this decline can provide an understanding that can benefit developing nations facing similar challenges. This also indicates human rights concerns since lower violent crime rates signify public safety and security rise.

(2) Time Series of Internally Displaced Persons Due to Disasters

The data for internally displaced persons (IDPs) shows that in 2010, a peak displacement of over 20,000 cases occurred, with a substantial drop in the years following. Still, from 2015 onwards, the numbers show fluctuations in displacement, with smaller spikes occurring in 2018 and 2020.

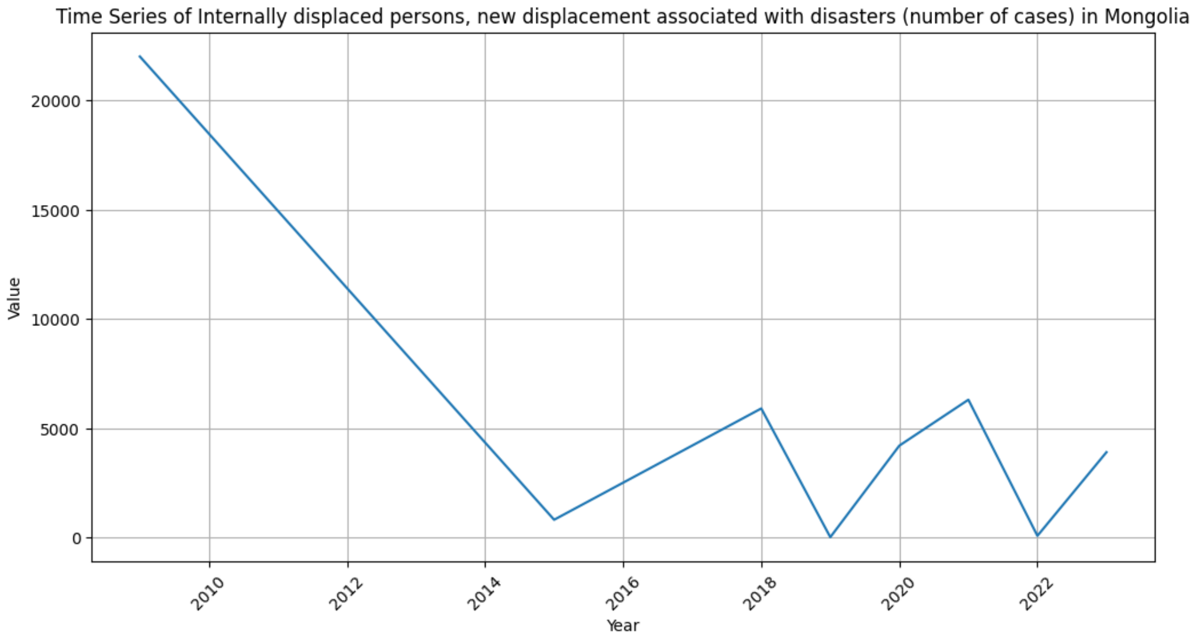

Figure 2. Time Series of Internally Displaced Persons Due to Disasters

Heavy displacement in 2010 could have emerged from severe environmental factors, including Mongolia's harsh winters (Dzuds) or other natural disasters. The changes in later years represent the persistent exposure of specific populations to displacement, requiring improved disaster preparedness and mitigation strategies to be developed. The insight pertains directly to development challenges because displacement intensifies poverty and social inequality.

(3) Time Series of Voice and Accountability

The voice and accountability percentile score declines from the late 1990s to 2021. The data reached its most significant point in 1998 but has steadily fallen to nearly 60 percentiles in recent years, with some oscillations in the 2010s.

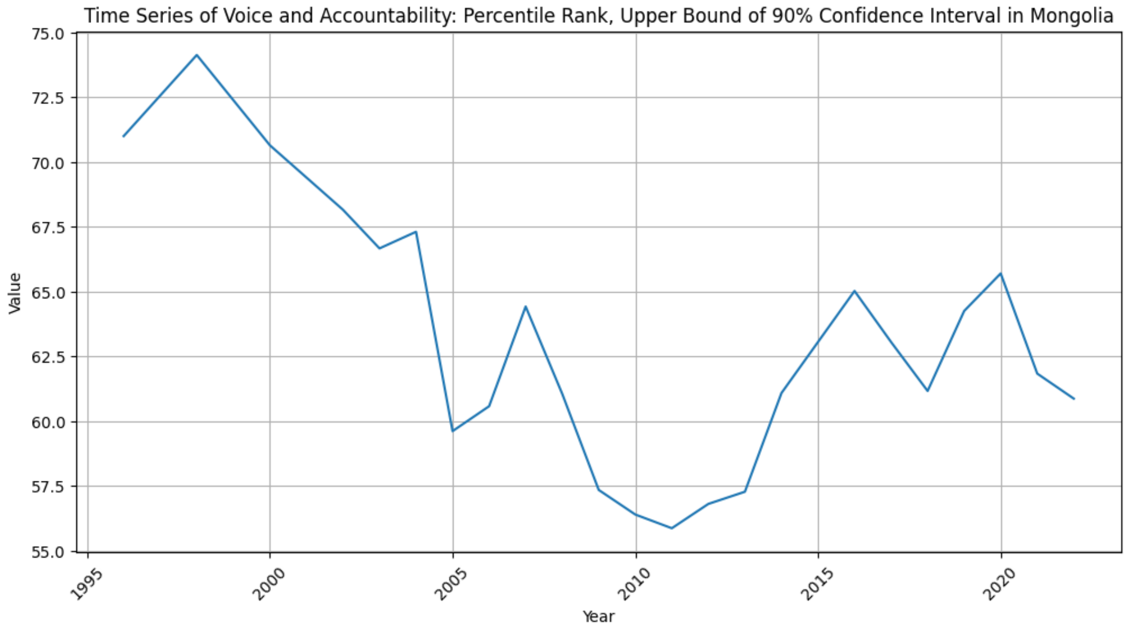

Figure 3. Time Series of Voice and Accountability

A reduction in voice and accountability suggests that Mongolians may have faced reduced political freedoms or a weakening in democratic governance as time goes on. Given that governance is essential for nurturing inclusive growth, this trend has the potential for extensive effects on social development and economic stability. It also fits with the project's objective of recognizing governance challenges and their more extensive consequences for socio-economic development.

(4) Time Series of GDP Growth (Annual %)

Mongolia's economic growth has shown remarkable volatility through the years, with steep drops during the late 1990s and early 2000s, followed by accelerated growth that reached approximately 15% in 2011. The growth rate fell sharply once more after 2012, accompanied by more minor recoveries in the years since.

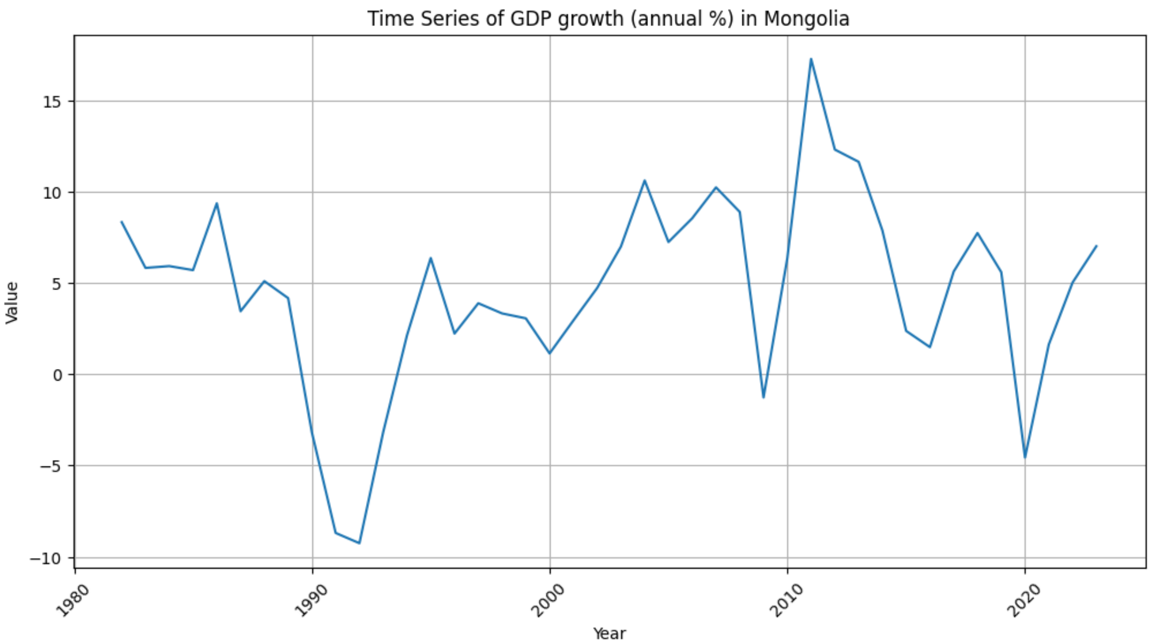

Figure 4. Time Series of GDP Growth (Annual %)

The significant variations in GDP growth indicate that Mongolia's economy is quite unstable, probably because of its dependence on natural resource exports. Such volatility challenges the progress of sustainable economic development and efficient long-term planning. Calculating future GDP trends is essential for creating policies to protect the economy from such shocks.

(5) Correlation Matrix for Key Indicators

The correlation matrix presented reveals significant correlations between the indicators. For example, annual data shows that GDP has a negative correlation with homicides and displacement, suggesting that a higher rate of economic growth might be connected to lower rates of crime and displacement.

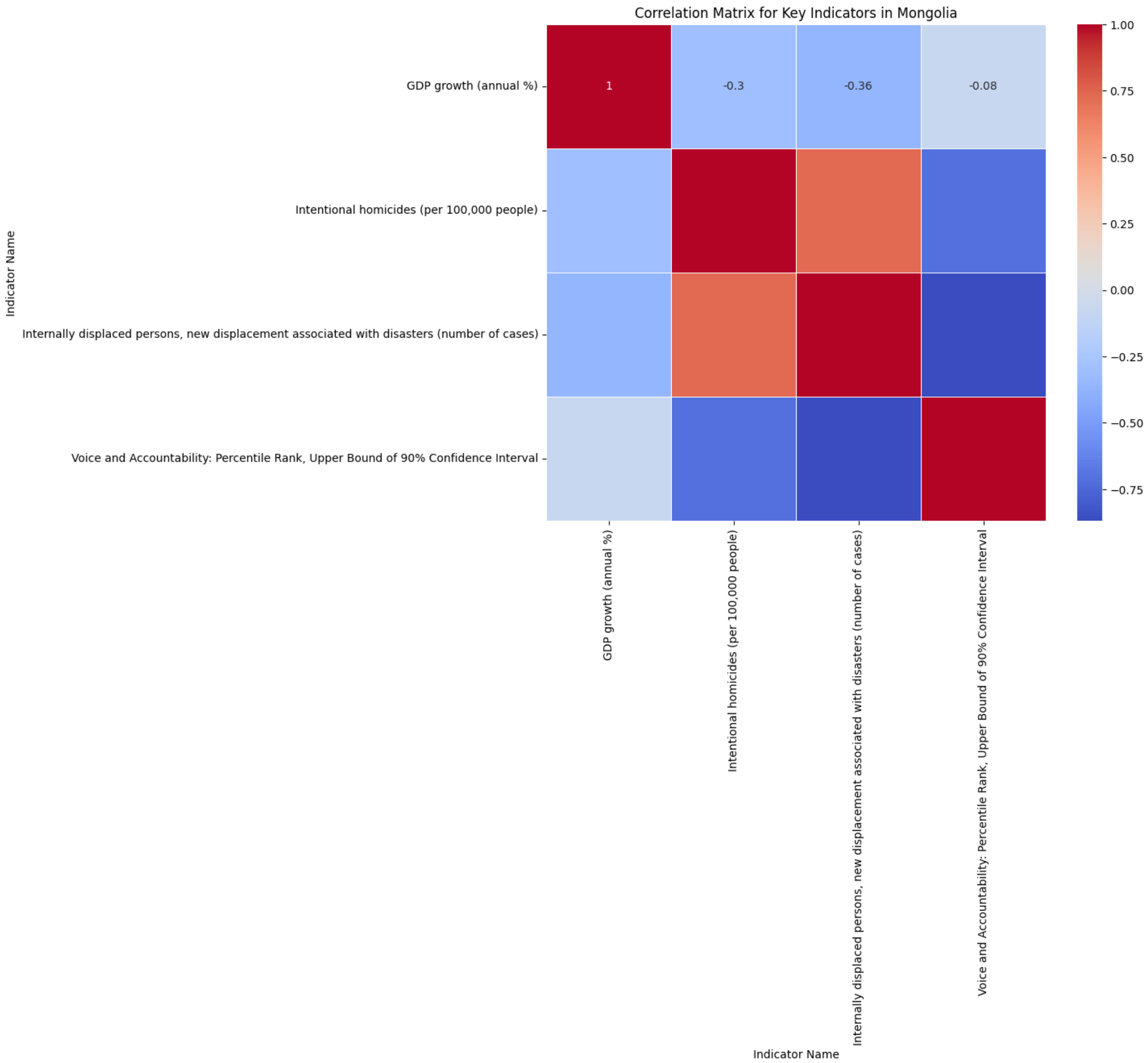

Figure 5. Correlation Matrix for Key Indicators

Understanding these correlations can help determine the directions in which interventions may be most effective. For instance, enhancing GDP could help lead to a decline in the levels of crime and displacement. This insight justifies a positive relationship and interdependency of economies and social aspects in policy-making processes.

(6) Comparison Between Homicides and GDP Growth

The time series plot of homicides and GDP growth indicates that economic growth was primarily associated with low homicide rates. The inverse is visible when the GDP grows very high, which could be observed in the early 2010s when the homicide rate declined sharply.

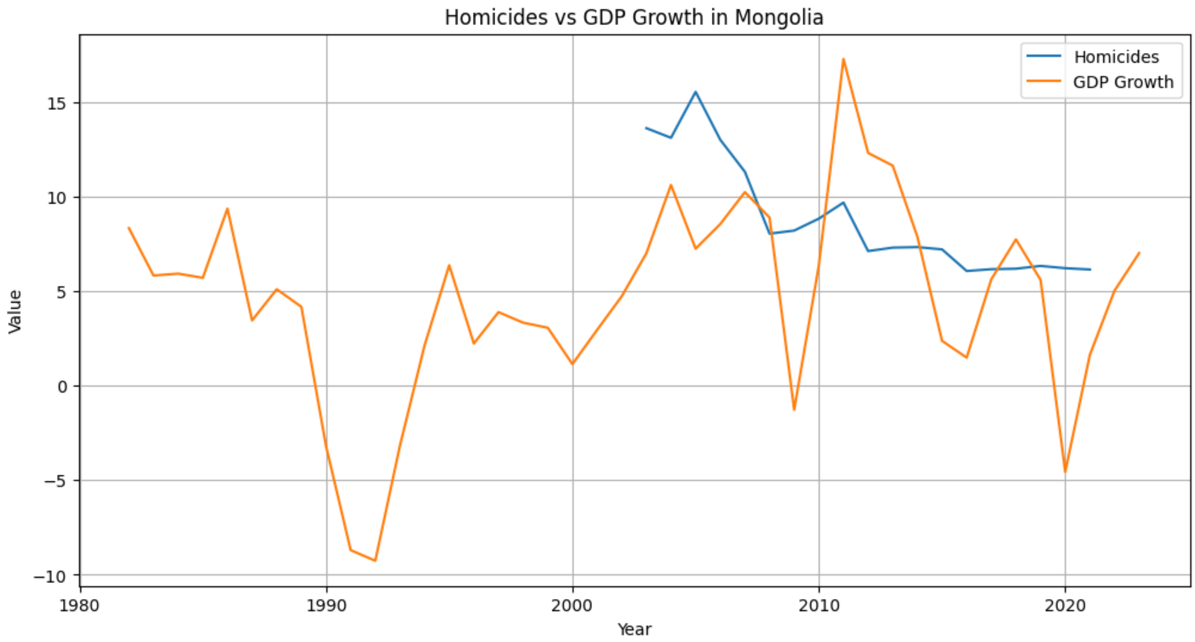

Figure 6. Comparison Between Homicides and GDP Growth

This negative correlation means that economic development can help maintain a stable society and decrease violent crime rates. However, the relationship is probably reciprocal, and enhanced safety facilitates a good business climate and economic development. This finding aligns with the project's objective of seeking relationships between social factors and financial consequences.

(7) Analysis of New Displacement Due to Disasters

The bar chart displaying new displacement due to disasters reveals significant variation, with a pronounced peak in **2010**, followed by smaller spikes in recent years.

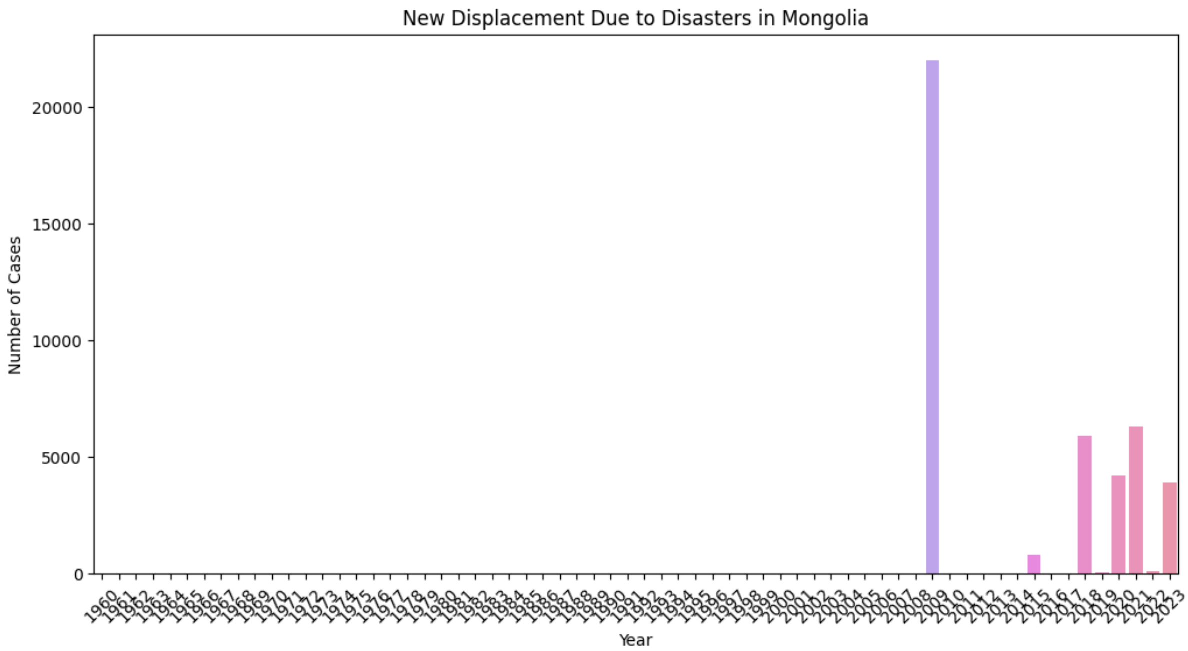

Figure 7. Analysis of New Displacement Due to Disasters

This high displacement in a few years indicates more environmental threats, which should be related to the disasters prone within the country, such as the severe winter and droughts. These areas must be addressed through more effective infrastructure and social investment, protection and reduction of vulnerability, and preparedness for disaster. This idea corresponds to the project's objective of assessing how environmental and social factors affect development results.

4.5.2 Summary of Insights

The EDA explores key patterns and associations with the socio-economic environment of Mongolia. The following are the significant findings:

- **Declining homicide rates** and improved public safety over time, linked to social stability and human rights improvements.
- **Economic volatility** and the challenges associated with sustaining long-term growth, emphasizing the need for economic diversification.
- **Governance challenges**, indicated by declining voice and accountability, may hinder inclusive development and exacerbate inequality.
- The role of **environmental disasters** in displacing vulnerable populations, highlighting the need for more robust disaster management policies.

5. Arima Forecasting

The analysis derived from the ARIMA model is useful in predicting Mongolia's GDP growth. Finally, applying a stepwise search to reduce AIC to its lowest, the model chosen was ARIMA (0,1,0) with no seasonality. This model configuration gave an AIC of 241.559, the least of all the tested configurations. Notably, other potential models, which include ARIMA (1,1,1), ARIMA (1,1,0), and ARIMA (0,1,1), had higher AIC or even infinite AIC; therefore, they were not so suitable for this dataset.

The ARIMA (0,1,0) model recommends that the data represent a random walk without autoregressive or moving average elements. In other words, the model implies that the GDP growth rate change is dependent only on the previous period's data and is not multileveled. The sigma^2 value of 20.1874 is for variance and the standard error of 4.391, which is statistically significant to infer that the model captures variability in the data.

The Ljung-Box test (Q) gave a p-value of 0.86; therefore, there is no autocorrelation in the residuals, indicating that the model was adequate. Also, from the model, it means that the Jarque-Bera test for the normality of residuals was at 0.92, which further confirms that its residuals are normally distributed. The model's residuals also do not have much skewness (-0.16) and kurtosis (3.06), which are ideal values for a good time series model.

Conclusively, it can be commented that the ARIMA (0,1,0) model seems to adequately capture the random fluctuations in the GDP growth trends in Mongolia. This model is particularly suitable for short-term projections due to its inefficiency for tasks requiring autoregressive or moving averages.

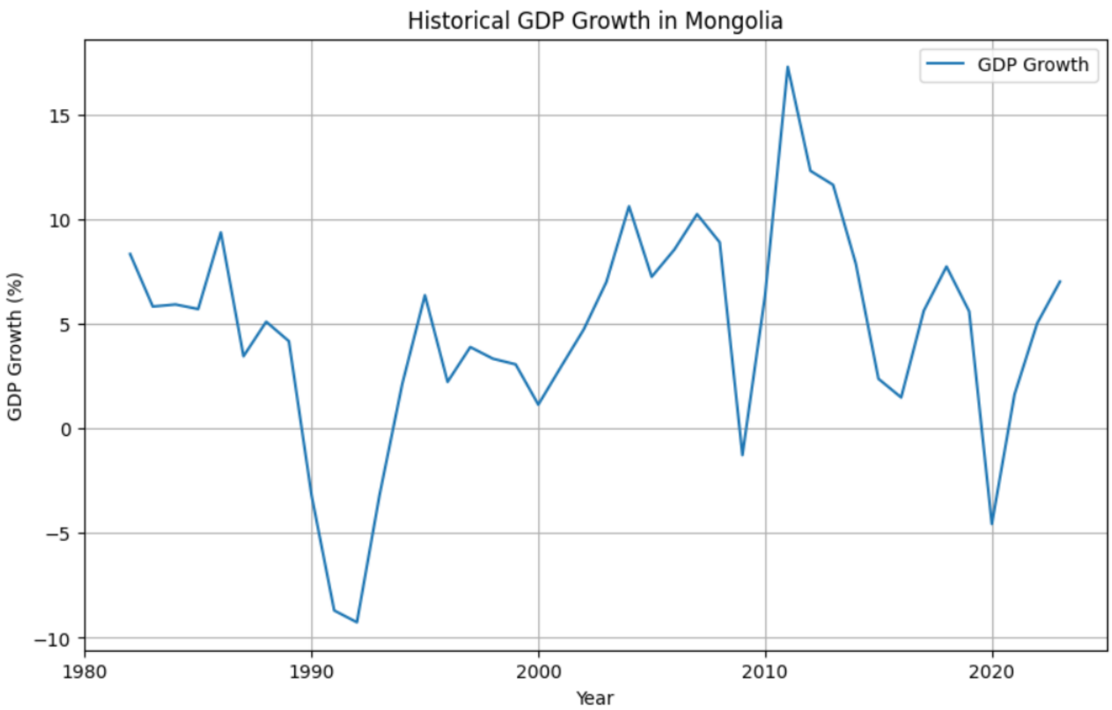

Figure 8. Historical GDP Growth

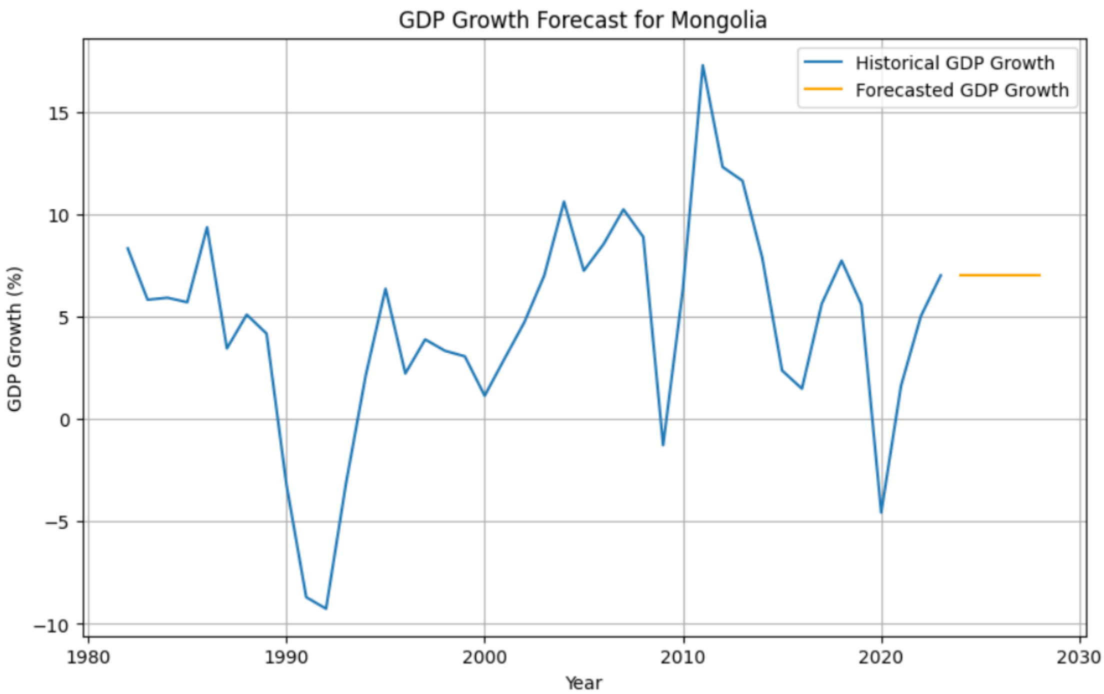

Figure 9. Forecasting of GDP growth for Mongolia

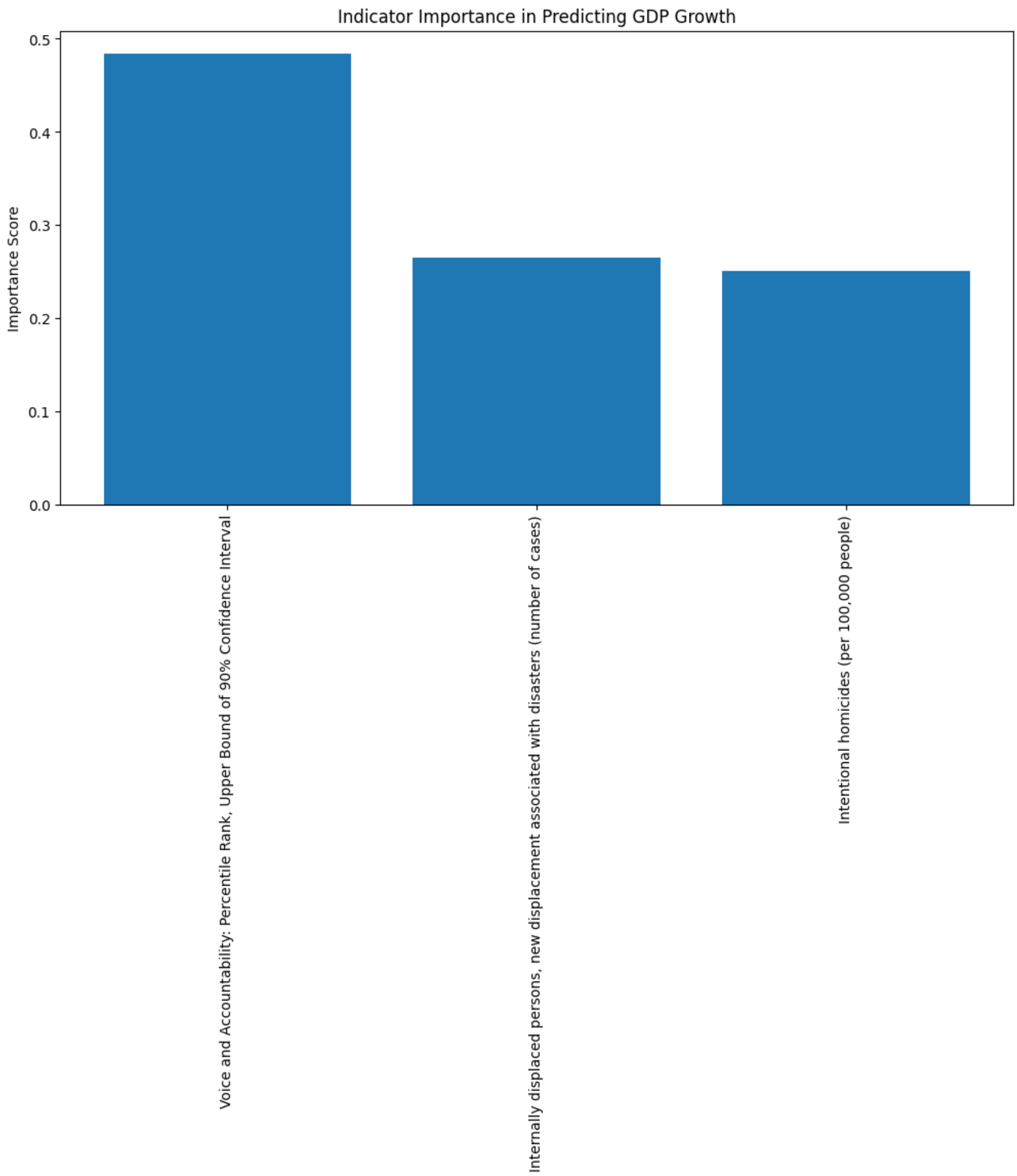

Figure 10. Indicator Importance

6. Anomaly Detection

Anomaly detection of Mongolia's GDP growth involves detecting any unusual event associated with the dataset, and the anomalies bring out key points that are almost out of the normal range. The graph above shows this anomaly detection where the GDP growth rate is shown against various years of reference, and any anomalies are marked in red. These are shock points through which the GDP growth deviates from the usual pattern of behavior.

The anomalies can be seen during periods of significant economic events, such as:

Early 1990s: There exists a sharp economic downturn, with GDP growth below -5%; it may be assumed that the period in question was marked by significant financial problems that Mongolia experienced from a centrally planned economy to a market one.

Around 2010, Another significant skewness was observed when Mongolia shortly possessed high GDP growth rates and then lost them, which may be associated with fluctuations in the prices of international commodities, mining operations, or the country's domestic economic measures.

Around 2020: A slight decline in the current GDP growth after suffering the impacts of the global pandemic and the disruptions in the economies of different countries.

These anomalies help identify the specific years that experienced economic fluctuations or shocks in Mongolia's growth; this detection may help policymakers or analysts look at these years more closely to understand what happened. They can also be used to anticipate and respond to unusual changes in the selected economic indicators.

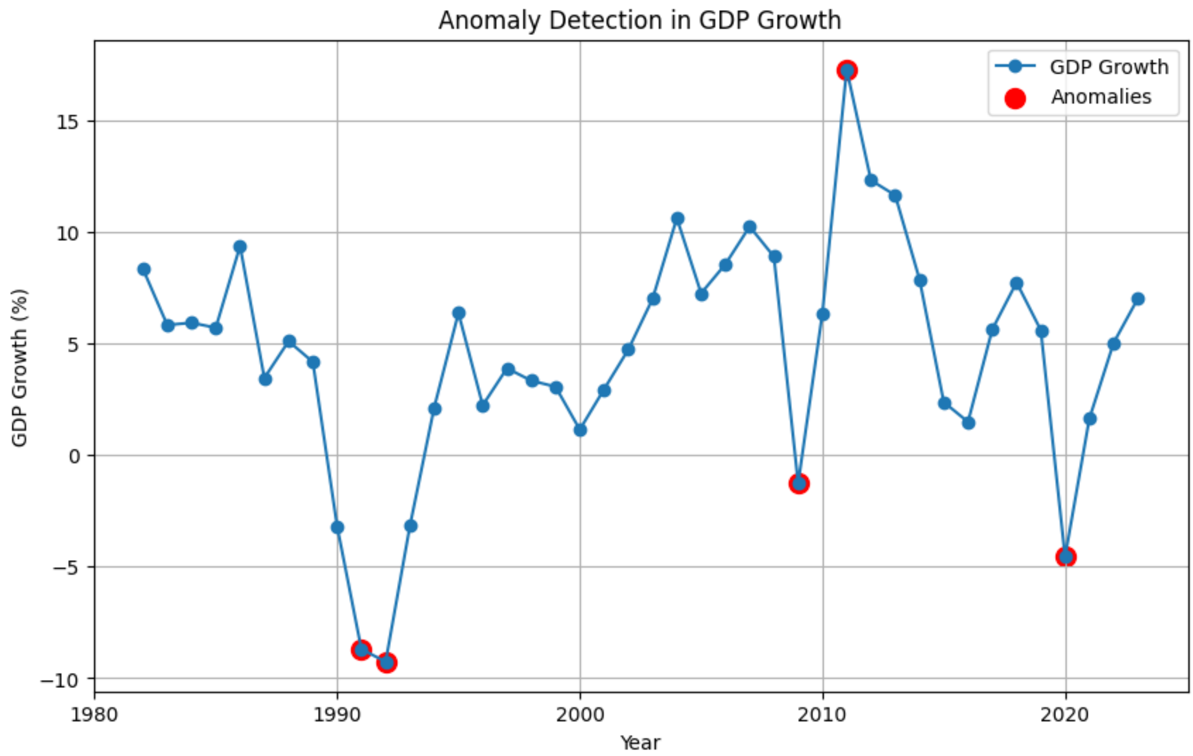

Figure 11. Anomaly Detection

7. LSTM Modelling

In this analysis, the LSTM model helped us understand the time-series forecasting of GDP growth in Mongolia. Because of the success of deep learning approaches, we were able to identify more detailed characteristics in the GDP time series data, which may not have been possible using conventional statistical methods.

1) **Training Performance**:

- The training RMSE of 2.80 indicates that the model performed well on the training set, effectively learning the underlying trends from historical GDP data.
- The training MAE of 2.26 enhances the understanding of the accuracy of the predictions made by the model of actual GDP values to attest that the predictions made during the training phase were near actual values.
- With an R^2 value of 0.77 on the training set, the model explained 77% of the volatility in the GDP growth; thus, the model highlighted the significant factors driving the Mongolian economic growth.

2) **Test Performance**:

- The test RMSE of 5.81 also indicates that the model performed well when faced with new data it hadn't encountered before. It shows that the model was not overly complex and was capable of being generalized in the future periods to come.
- The MAE of 4.84 in the test phase established that the model could present

accurate predictions of GDP growth in the validation period, thus supporting the usefulness of its predictions on actual future data.

These outcomes suggest that the LSTM model accurately tracked the GDP growth, identifying patterns from one series and predicting subsequent patterns. Also, employing historical data allows the stakeholders to recognize how the GDP growth might change, which is essential for economic forecasting.

Visualization Insights:

- The GDP Growth Prediction plot demonstrates that the model tracks the GDP growth line. The training forecasts are similar to the actual GDP increase, while the test forecasts are consistent with the overall trend.
- The Residual Plot confirms that the deviations between actual and predicted values are acceptable and small, and the model used in the training and testing phase is stable and reliable.

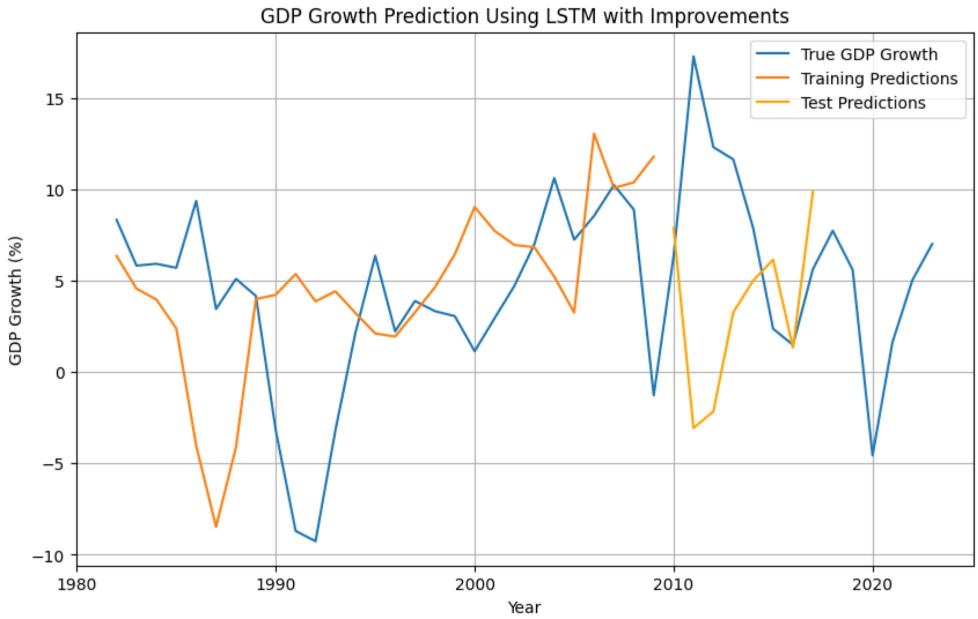

Figure 12. Prediction of GDP Growth

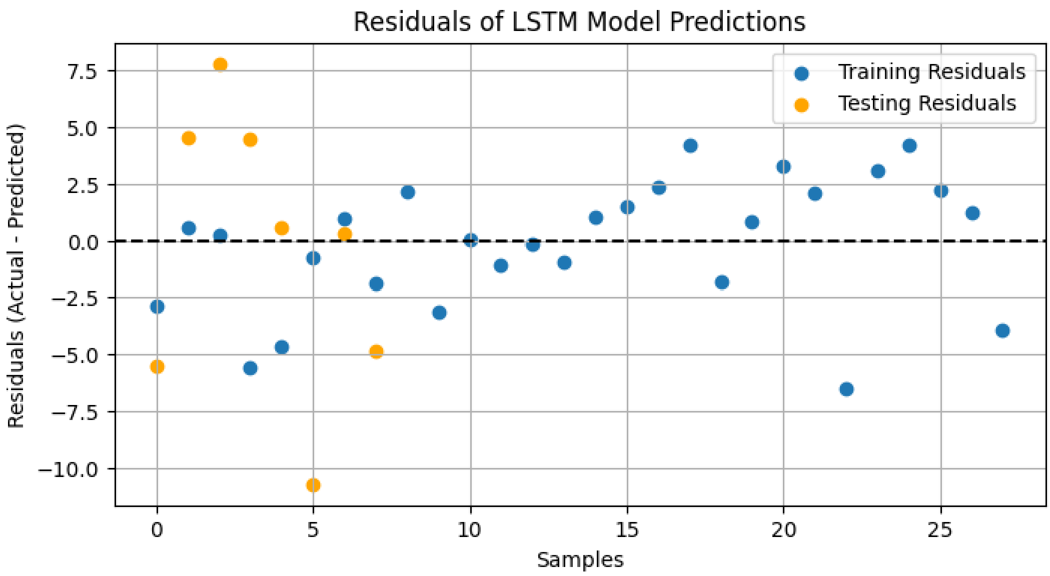

Figure 13. LSTM Residuals

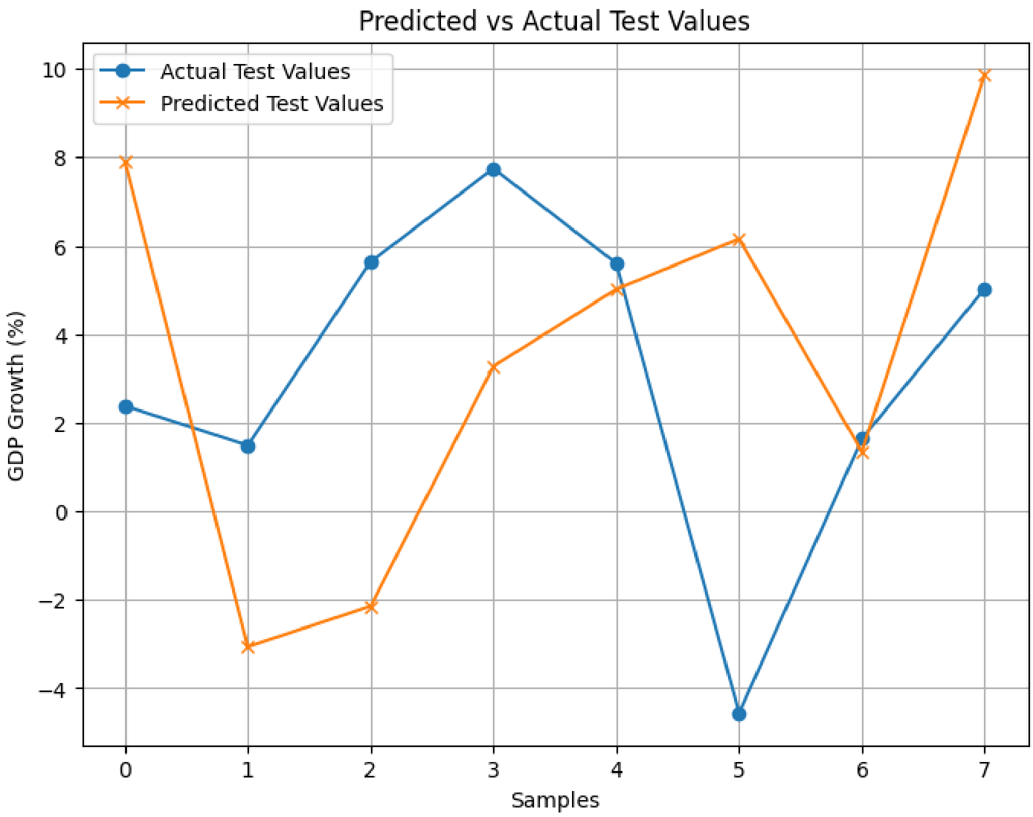

Figure 14. Predicted and Actual values

8. Discussion and Findings

This project applied a multi-method approach to explore, analyze, and predict socio-economic trends in Mongolia, focusing on indicators such as GDP growth, unemployment, and other key variables. The results shed light on essential patterns and relationships within the dataset using various statistical, machine learning, and deep learning techniques. Below is a detailed discussion of the methodologies and findings from each analysis phase.

The initial EDA provided crucial insights into the socio-economic landscape of Mongolia over time. Time series visualizations for several key indicators revealed trends aligned with Mongolia's economic and social transitions. For instance, the time series plots for GDP growth and unemployment highlighted rapid economic expansion and contraction periods, often coinciding with global financial conditions.

Additionally, the correlation matrix between selected indicators uncovered significant relationships, such as the inverse relationship between GDP growth and intentional homicides, reflecting how social stability and economic performance can be interlinked. The heatmap visualizations further underscored these connections, demonstrating that economic growth is not isolated but influenced by a complex web of socioeconomic factors.

The EDA also spotlighted other critical indicators, such as internally displaced persons due to disasters, where significant variations were observed over time. These insights offer a starting point for deeper analysis and policy-making considerations for Mongolia's socio-economic development.

8.1 Forecasting and Predictive Modelling

The project employed advanced forecasting techniques such as ARIMA and LSTM to predict future trends in Mongolia's GDP growth. The ARIMA model was selected based on its performance in minimizing the AIC value during model selection, providing a simple yet effective method to forecast GDP growth. The model's forecast provided short-term predictions aligned with recent historical trends, adding value to policymakers and economists looking for near-future economic forecasting.

The LSTM model, which leverages deep learning for time series forecasting, was instrumental in capturing the more complex patterns in the GDP growth data. Despite the intricacies of forecasting economic variables, the LSTM model demonstrated predictive solid performance. As reflected in the test RMSE and MAE, the model effectively generalized and maintained consistency across the validation period. The visualizations of actual GDP growth versus the predicted values reinforced this point, illustrating that the model could align its predictions closely with actual data. The application of LSTM in this context proves that advanced neural network models can capture time-dependent relationships, providing a solid tool for forecasting highly variable socio-economic data.

8.2 Anomaly Detection

The anomaly detection using Isolation Forest provided valuable insights into outlier years where socio-economic performance deviated sharply from the norm. The anomalies identified in the GDP growth data, such as the significant economic dip around 1990 and the rapid spike in 2010, reflect periods of historical significance. Detecting these outliers can assist policymakers in identifying past crises or opportunities, which can be leveraged to inform future decision-making strategies.

8.3 Feature Importance

The Random Forest feature importance analysis offered a deeper understanding of the socio-economic indicators most predictive of Mongolia's GDP growth. Among all the indicators, "Voice and Accountability" ranked as the most influential in predicting GDP, followed by "Internally Displaced Persons" and "Intentional Homicides." This finding highlights the importance of good governance and social stability as crucial drivers for economic prosperity. These indicators emphasize the multifaceted nature of economic growth, where non-economic factors such as governance and social factors play pivotal roles.

8.4 Findings and Implications

The correlation between different socio-economic variables, including GDP, unemployment, and governance structure, shows that economic growth is multifaceted. GDP is a result of the policies conducted in the economy and social and governance factors. ARIMA and LSTM were able to make the GDP growth forecast. However, the LSTM was more detailed by considering the historical data. The models developed here can be used in future forecasting exercises, making this an effective method for economic planning. The feature importance analysis showed that governance factors, which include "Voice and Accountability," are the significant determinants for GDP growth, therefore pointing towards institutional quality as the critical determinant of economic development. This finding implies that enhancing governance standards might be as essential for Mongolia's economic growth as a continuous economic reform. The GDP data anomalies correlate with historical events like the early 1990s recession. These findings show how machine learning helps determine the stages where policy concerns should be inclined.

9. Conclusion

In conclusion, this project successfully used statistical methods, machine learning, and deep learning approaches to give a holistic view of Mongolia's socio-economy. The findings thereby point to the fact that economic, social, and governance indicators are intertwined in determining the country's development path. These tools and models can be used in future research to gain more insights and predict the direction of socio-economic activities with policy and strategic relevance in Mongolia.