NATIONAL RESEARCH UNIVERSITY HIGHER SCHOOL OF ECONOMICS

Social Network Analysis

A predictive analysis of trade relations using LSTM

By Muhammad Asaad Khan

Introduction:

In recent years, our contemporary world has been interconnected through trade and technology at unprecedented levels. Despite the global trade network being affected through sanctions and trade wars, international trade is still ongoing at unprecedented levels. The primary aim of this project is to gain insights into the structural properties of the global trade network by leveraging network analysis techniques. Specifically, we aim to explore the dynamics of international trade relationships and understand the centrality and influence of different countries within the trade network. By examining various centrality measures, the project aims to uncover patterns and trends that will assist in effective decision-making for enhancing trade relations.

Analyzing international trade through the lens of network analysis provides a holistic view of trade relationships that traditional bilateral analysis may overlook. This approach enables us to consider the interdependencies between countries and the broader structure of the trade network. Understanding these dynamics is crucial for policymakers, economists, and businesses as it can reveal vulnerabilities, strengths, and opportunities within the global trade system. Furthermore, the emergence of disruptions in international trade in terms of trade wars and sanctions makes it vitally important to understand how international trade networks functions and to gain insights in terms of making international trade networks more resilient.

This project will employ the "World Export & Import Dataset (1989 - 2023)" dataset, taken from Kaggle, which includes detailed records of trade values, tariff rates, and policy indicators across multiple nations over several decades. By employing advanced algorithm capable of analyzing and interpreting complex networks, the project aims to explore patterns and trends that govern trade relationships.

Problem Statement:

Despite the extensive data available on international trade, the challenge remains to effectively predict which countries are likely to form stronger trade ties or partnerships in the future. Traditional econometric models, while useful for analyzing past trends, often ignore the nuanced interactions taking place in global trade networks. Furthermore, the dynamic nature of trade policies and economic indicators makes it necessary to adopt a more flexible and robust approach to forecasting.

Objectives:

The primary objectives of this study are to:

- Construct a comprehensive graph model of international trade relationships using the dataset at hand.
- To analyze the status of trade relations through centrality measures.
- Develop and evaluate a LSTM predictive model that can forecast the future of trade partnerships.
- Provide actionable insights that can guide economic policy and strategic trade decisions.

By addressing this problem, the project aims to contribute to a more nuanced approach in predictive modeling and designing international trade strategy, potentially leading to more stable and beneficial economic outcomes globally.

Data and methods:

The data for this analysis was sourced from the comprehensive trade database, which contains annual trade values between various countries. For this project, we featured engineered to focus on Russian Federation, China, India, and Brazil, that is, some of the countries within BRICS.

The features in the dataset and their description are provided below:

- ReporterISO3: ISO3 code of the reporting country.
- ReporterName: Name of the reporting country.
- PartnerISO3: ISO3 code of the partner country.
- PartnerName: Name of the partner country.
- Year: The year of the trade data.
- TradeFlowName: The type of trade flow (e.g., export, import).
- TradeValue in 1000 USD: The value of trade in thousands of USD

Long Short-Term Memory (LSTM) network was employed for predicting future trade trends. LSTM is a type of recurrent neural network (RNN) that is particularly well-suited for time series forecasting due to its ability to learn long-term dependencies.

The model was trained using the historical trade data. The dataset was split into training and validation sets with an 80-20 ratio. The training set was used to fit the model, while the validation set was used to monitor the model's performance and prevent overfitting.

Mean Squared Error (MSE) was used as the loss function, which is standard for regression problems. The model was trained for 100 epochs with a batch size of 32. Early stopping was implemented to terminate training when the validation loss ceased to decrease, which helps in preventing overfitting. The model was implemented using Python and the Keras library with a TensorFlow backend. This provided a flexible and efficient environment for developing and training the LSTM model. Data preprocessing and analysis were conducted using Pandas and NumPy, while visualization was carried out using Matplotlib.

Centrality Measures:

Centrality measures provide a quantitative framework to analyze and understand the structure and dynamics of international trade networks. In other words, they help identify which countries are most influential in the trade network. These are useful measures for helping stakeholders make informed decisions in trade policy, strategy, and risk management. At the same time, centrality measures assist in predicting trade patterns as well. For instance, a country with increasing centrality over time might become a significant future trading actor.

Figure 1: Betweenness Centrality and Out-degree Centrality

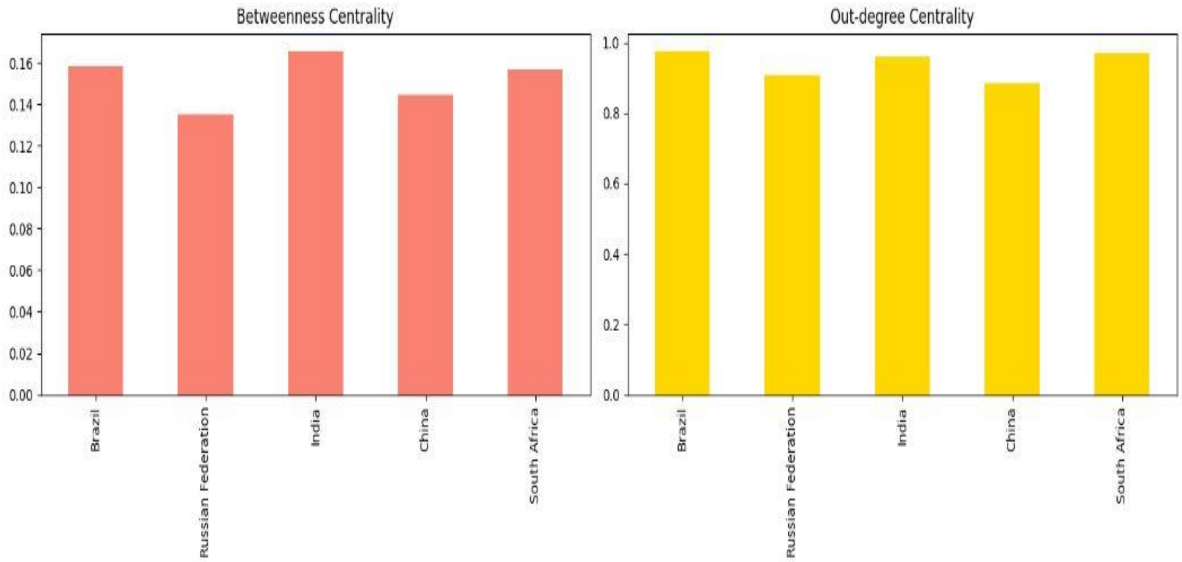


Figure 1 shows the between centrality and out-degree centrality of the relevant countries. Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. A node with high betweenness centrality has considerable influence over the flow of information or goods in the network because it connects different parts of the network. Brazil and India have the highest betweenness centrality, suggesting they are crucial intermediaries in the trade network, facilitating trade between other countries.

Out-degree centrality measures the number of direct connections a node initiates. In a trade network, it indicates how many countries a given country exports to. Brazil and South Africa have the highest out-degree centrality, indicating they initiate trade with many other countries, playing a key role in exporting goods and services.

Figure 2: Degree Centrality and Closeness Centrality

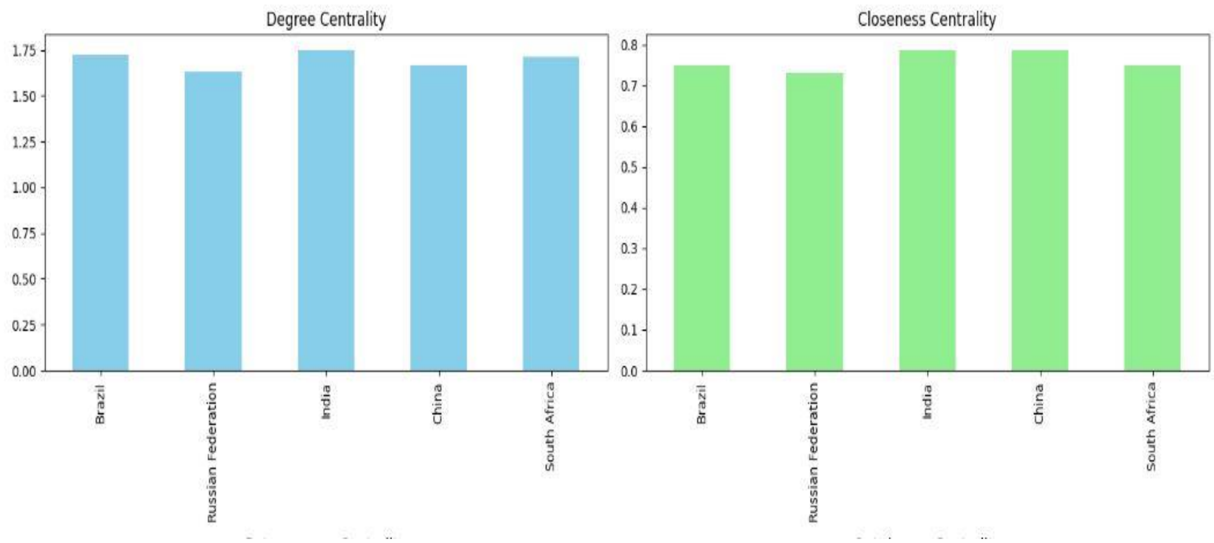


Figure 2 shows the degree centrality and closeness centrality. Degree centrality is the simplest centrality measure, counting the total number of direct connections a node has. It includes both incoming and outgoing connections. Closeness centrality measures how close a node is to all other nodes in the network. It is the reciprocal of the average shortest path distance from the node to all other nodes. Countries with high closeness centrality can quickly reach other countries within the trade network. This makes them efficient hubs for distributing goods and services, enhancing their strategic importance in global trade. Countries included in the analysis have similar closeness centrality, indicating that they are all relatively well-connected within the trade network and can quickly interact with other countries.

Data preprocessing and results:

LSTMs have been widely used and proven effective in various time series forecasting applications, including stock market prediction, weather forecasting, and financial modeling. Their success in these areas provides a strong justification for their application in trade trend prediction.

Trade data is inherently sequential, where past trade volumes, values, and trends influence future trade activities. Long Short-Term Memory (LSTM) networks are specifically designed to handle such sequential data.

For LSTM modelling, first, ‘time_series’ was feature engineered to convert year index to datetime format. Next, the data was normalized to a range of 0 to 1 using MinMaxScaler. Normalization is crucial for LSTM models as it helps in faster convergence during training. To further prepare the data for the LSTM model, sequences of a specified length (seq_length) were created. Each sequence consists of seq_length timesteps used to predict the next timestep. The dataset was split into training and testing sets, with 80% of the data used for training and 20% for testing.

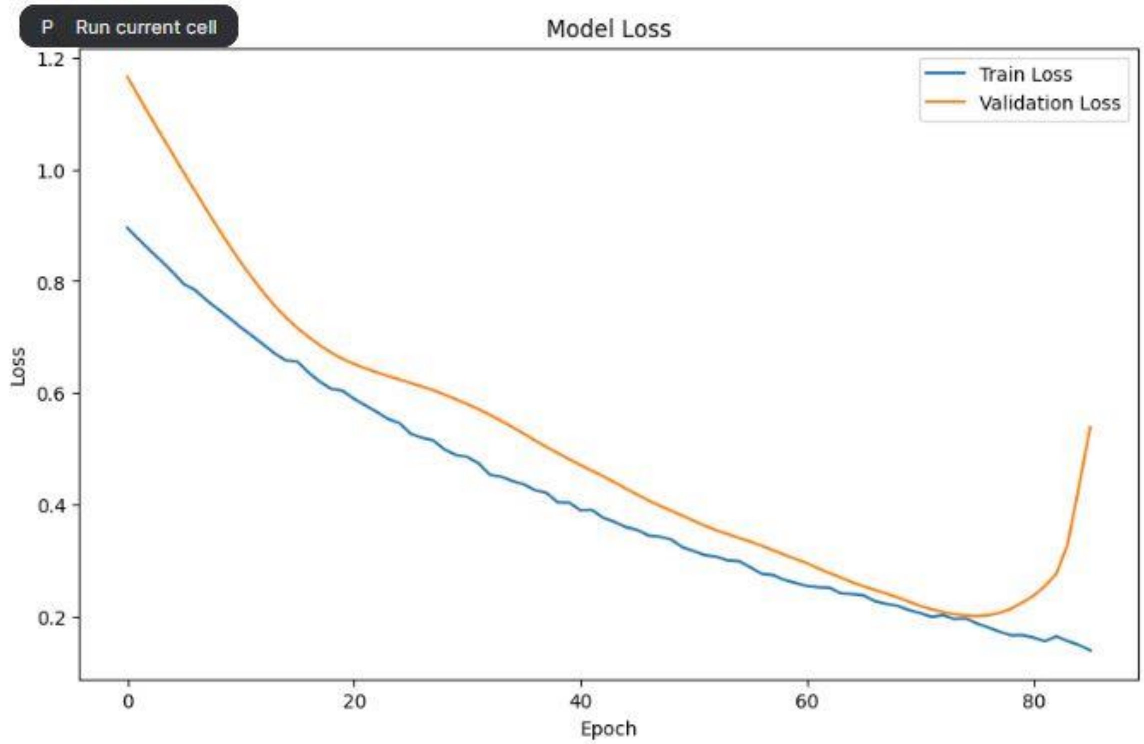
The LSTM model is built with the following layers:

- First LSTM Layer: 50 units, returns sequences, includes L2 regularization to prevent overfitting.
- Dropout Layer: 50% dropout rate to prevent overfitting.
- Second LSTM Layer: 50 units, does not return sequences, includes L2 regularization.
- Dropout Layer: 50% dropout rate.
- Dense Layer: Single unit for the output.

The model is trained for up to 100 epochs with early stopping based on validation loss to prevent overfitting. The patience parameter of 10 means training stops if the validation loss does not improve for 10 consecutive epochs.

Figure 3 shows the training and validation loss curves which were plotted to visualize the model's performance over time periods The plot shows the loss (mean squared error) on the y-axis and the number of epochs on the x-axis.

Figure 3: Brazil-China loss curve



Initially, both training and validation losses decrease, indicating that the model is learning. However, around epoch 70, the validation loss starts to increase while the training loss continues to decrease. This divergence indicates the model is overfitting.

Figure 4: Forecast of Brazil-China trade trend

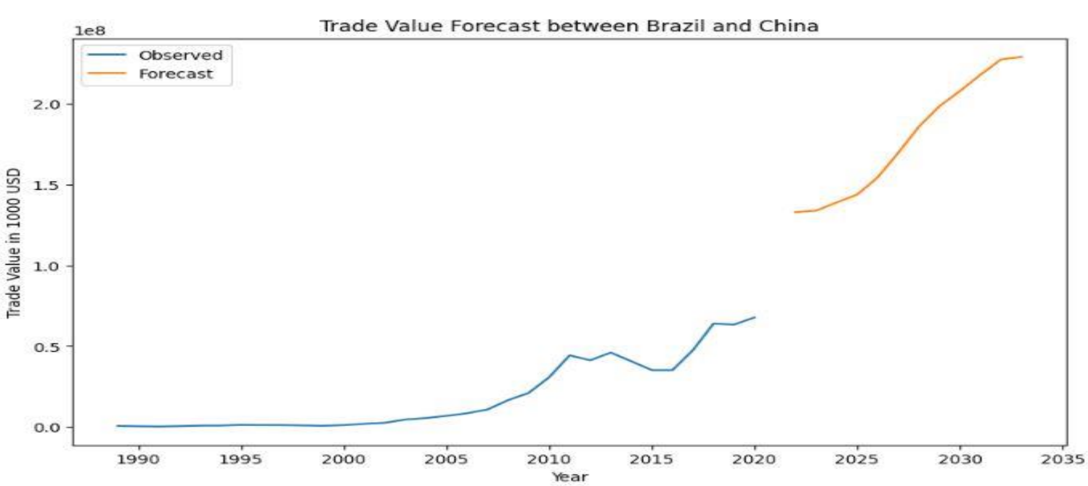


Figure 4 shows the past and the forecasted trade trend between Brazil and China. The trade value exhibits some fluctuations but generally continues to trend upwards. There are peaks around certain years. However, the forecasted trend suggests a significant upward trend. This suggests a continued and possibly accelerated growth in trade between Brazil and China. Factors contributing to this growth could include further economic integration, increased commodity exchanges, and stronger diplomatic ties.

Figure 5: Russia-China loss curve

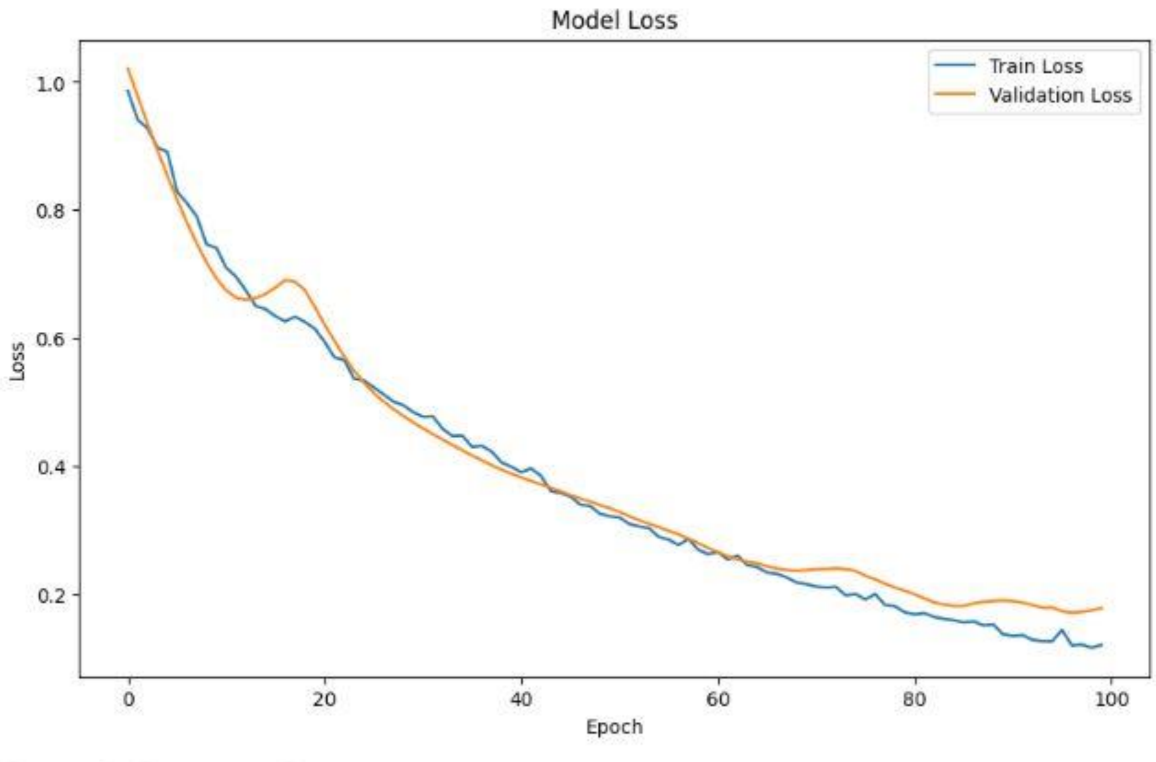


Figure 5 shows the loss curve of Russia-China model. Both the training loss and validation loss start at around 1.0 and decrease rapidly. This indicates that the model is learning effectively from the data, reducing the error significantly in the initial stages of training. The validation loss closely follows the training loss, suggesting that the model is generalizing well to the validation data at this stage.

There are some fluctuations in the validation loss, which is normal as the model tries to generalize better. These fluctuations indicate that the model is encountering different challenges in the validation data compared to the training data. The training and validation losses are still closely aligned, indicating good generalization without overfitting. The model demonstrates effective learning, as evidenced by the significant reduction in both training and validation losses over the epochs. The close alignment of training and validation losses throughout most of the training process indicates that the model generalizes well to unseen data, with minimal overfitting.

Figure 6: Forecast of Russia-China trade trend

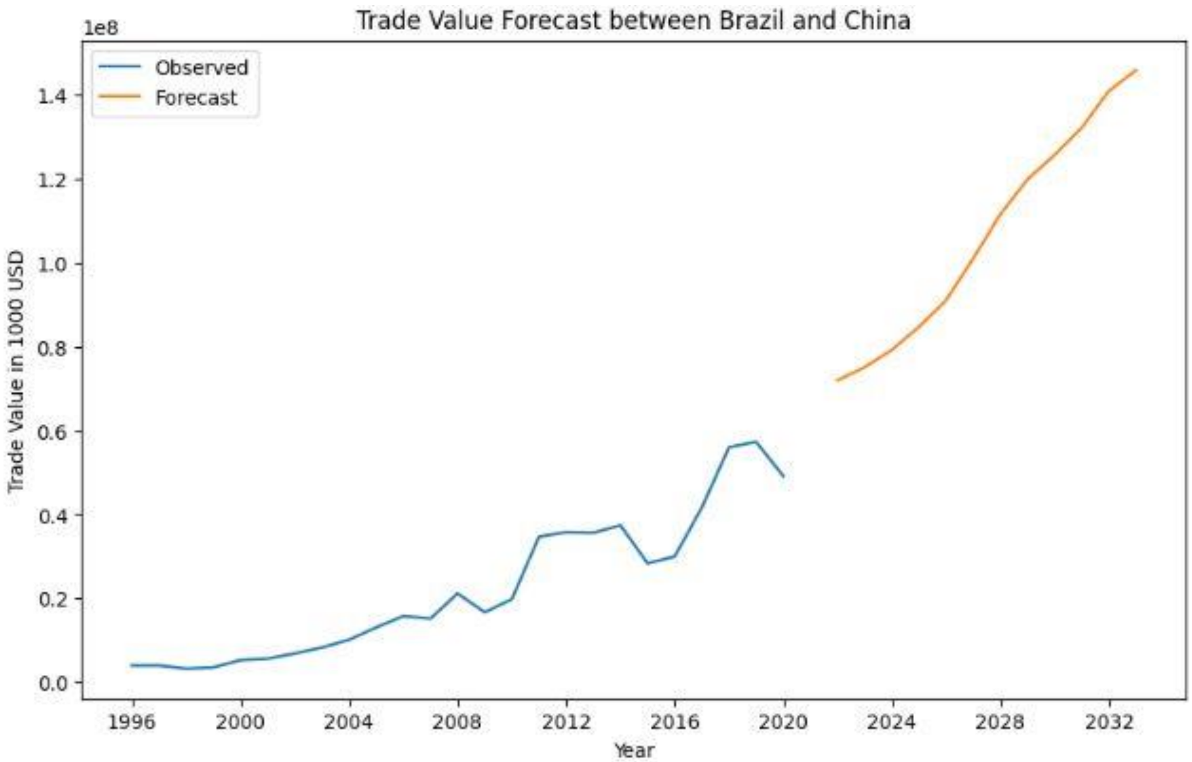
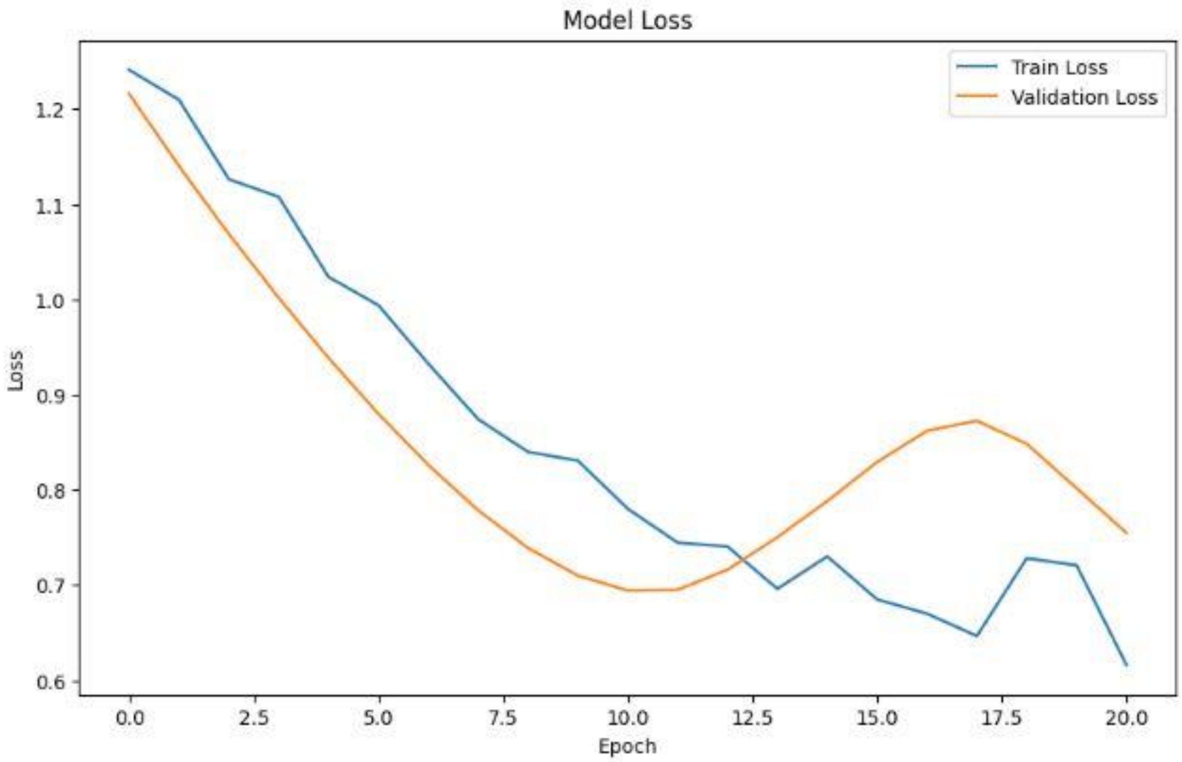


Figure 6 predicts an acceleration in trade ties in between Russia and China in the future. This is due to several factors including the rise of China and the contemporary geopolitical factors and the rise in regionalism.

Figure 7: Russia-India loss curve



Both the training loss and validation loss start at around 1.2 and decrease rapidly, indicating that the model is learning effectively and quickly reducing errors in the initial training stages. The training loss continues to decrease steadily, indicating that the model is progressively learning from the training data. The validation loss also decreases initially but starts to fluctuate around epoch 10, indicating some variability in model performance on the validation data. Around epoch 15, the validation loss starts to increase while the training loss continues to decrease, indicating the onset of overfitting.

The model is fitting the training data very well but is not generalizing as effectively to the validation data. The increase in validation loss after epoch 15, while the training loss continues to decrease, is a clear sign of overfitting. The model starts to memorize the training data rather than generalize well to new, unseen data.

Figure 8: Forecast of Russia-India trade trend

Figure 8 shows the historical and forecasted trade values between Russia and India. The forecast indicates a stable and sustained growth in trade between Russia and India, without the sharp fluctuations observed in the historical data. Historical data shows periods of rapid growth and fluctuations, but the forecast suggests a more consistent increase in trade value. This stability and growth highlight the strength of the economic relationship between the two countries and provide a promising outlook for future trade activities.

Conclusion:

The analysis of trade trends between Brazil, China, Russia, and India using Long Short-Term Memory (LSTM) networks provides significant insights into future trade patterns. The observed data, coupled with forecasted trends, highlights the evolving trade dynamics and the strengthening economic relationships between these major players in international trade. The use of LSTM networks has proven effective for forecasting trade values. LSTMs are well-suited for handling sequential data and capturing long-term dependencies, making them ideal for time series forecasting. The ability of LSTMs to manage non-linear relationships and robustness to noise further enhances their utility in economic and trade data analysis.

Despite measures like early stopping and regularization, the LSTM models here showed overfitting, particularly since the dataset had limited data features. Overfitting occurs when the model learns the training data too well, including its noise, which can negatively impact its performance on unseen data. This problem can be dealt with by using a more nuanced dataset.