US Unemployment Rate forecasting

CS501 Computing for Science and Engineering - Fall 2023

Russell Ho, Prathyush Kothari, Lucas Mori (team leader), Darshitaben Patel

October, 2023

1 Interim Report - Progress Status

In this section, we will provide comments on the progress made since the last update. A 'snapshot' of the final report as of the date of this submission will follow this section. This section will be removed upon final report submission.

Team's Shared Folder: `https://purdue0-my.sharepoint.com/:f:/g/personal/lkawano_purdue_edu/EqBwiRC1YzlEhV_k4UwwWVwBLwTh4t9EigzvL3dxj1IOKg?e=isAV9B`

Kaggle Dataset Link: `https://www.kaggle.com/datasets/calven22/usa-key-macroeconomic-indicators`

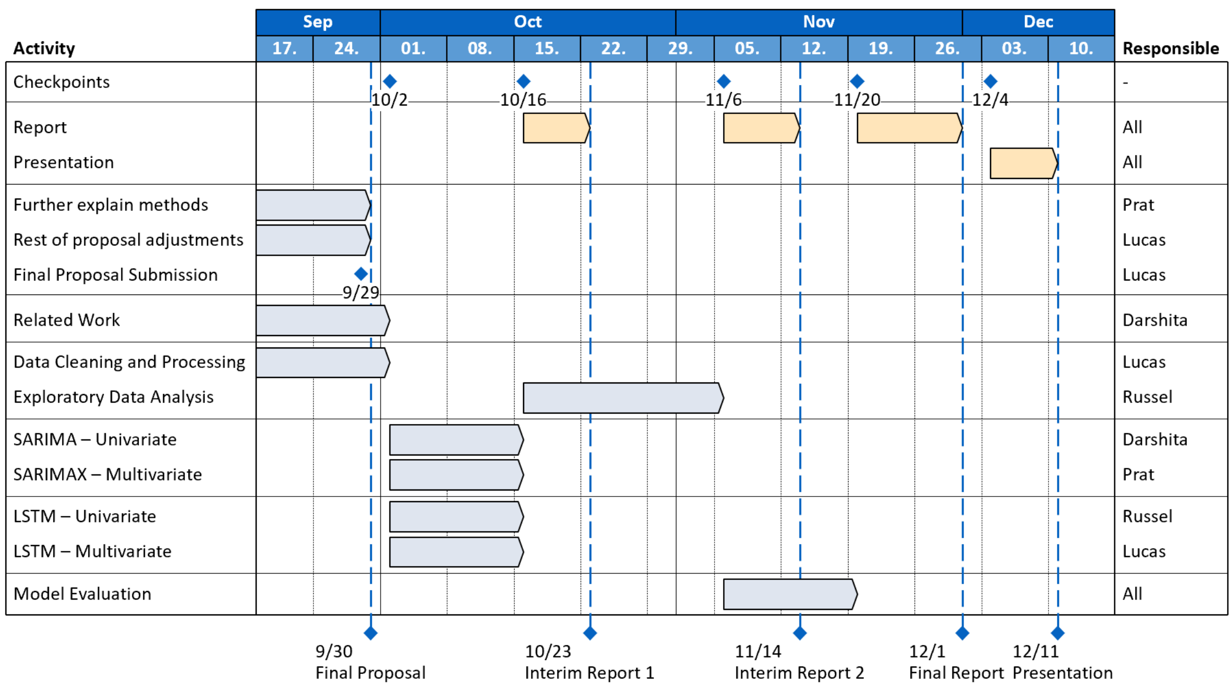

Figure 1: Plan of work

Figure 1 displays the team's schedule along with the members designated for each task. As of now, the project remains on course, with all tasks planned up to the 'Interim Report 1' being successfully completed.

Our primary strategy encompasses the following stages:

1. Developing initial 'vanilla' models for each approach (Note: 'vanilla' here refers to the basic, unmodified versions).
2. Refining and improving these baseline models.
3. Evaluating and comparing the top-performing models.
4. Drawing conclusions and providing recommendations.
5. Documenting the lessons we've learned.

We believe this structured approach enhances our prospects for the project's successful completion.

Progress:

- Implementation of the baseline LSTM Univariate model.
- Implementation of the baseline LSTM Multivariate model.
- Implementation of the baseline SARIMAX model.
- Completion of the literature review session.

Anticipated next steps:

1. Investigate avenues to optimize the baseline models.
2. Conduct Exploratory Data Analysis (EDA).
3. Compare the results obtained from the finalized models.
4. Draft and finalize the concluding report.
5. Design the final presentation.

Abstract

As part of the "CS501 - Computing for Science and Engineering" course, this research's primary objective is to develop an accurate unemployment rate forecasting model to address the challenges posed by dynamic labor markets and to facilitate informed decision-making for individuals, businesses, and policymakers. The study plans to develop two unemployment rate forecasting models using LSTM and ARIMA methodologies with python as the programming language of choice. The research will utilize USA Key Macroeconomic Indicators which contains data from 1981 to 2021. The project will follow a systematic approach, including data collection and cleaning, exploratory data analysis, data preprocessing, model implementation, and evaluation. The expected outcomes include a deep understanding of unemployment rate dynamics, two robust forecasting models (LSTM & SARIMAX), and valuable recommendations for stakeholders.

2 Introduction

The unemployment rate stands as a pivotal indicator of a nation's economic health, offering insights into labor market dynamics, economic cycles, and the overall well-being of its citizens. In the United States, variations in the unemployment rate over the years have reflected a confluence of factors, ranging from technological revolutions and policy shifts to broader global economic downturns. Historically, scholars and policymakers alike have dedicated significant effort towards understanding these fluctuations, as the ramifications of unemployment extend beyond the economic domain, influencing social stability, mental health, and general prosperity. With advancements in technology, the availability of richer datasets, and the emergence of sophisticated analytical tools, there is an unprecedented opportunity to dissect and forecast unemployment trends with enhanced precision and granularity.

This study endeavors to delve deep into the intricacies of the U.S. unemployment rate by leveraging a comprehensive dataset from the "USA Key Macroeconomic Indicators" available on Kaggle. Sourced from the Federal Reserve Bank of St. Louis and released by the U.S. Bureau of Economic Analysis, this dataset spans from 1981 to 2021 and encompasses a myriad of macroeconomic indicators. By intertwining traditional econometric models with cutting-edge machine learning techniques, our analysis seeks not only to chart the historical trajectory of unemployment but also to proactively forecast its future trends. Through a systematic exploration, we aim to shed light on the multifaceted interplay of economic factors influencing unemployment, thereby providing valuable insights for policymakers, economists, and the general public.

3 Literature Review

Developing an accurate unemployment rate forecasting model is of paramount importance in addressing the challenges posed by dynamic labor markets and facilitating informed decision-making for individuals, businesses, and policymakers. Numerous studies and research efforts have contributed to this field, aiming to enhance our understanding of unemployment rate dynamics. In this related work section, we will explore some key contributions in the area of unemployment rate forecasting. The forecasting of the U.S. unemployment rate is a critical and perennial challenge in the field of economics, with far-reaching implications for policy formulation, financial markets, and societal well-being. Over the years, extensive research has been conducted to develop effective models and methodologies for predicting this crucial economic indicator. A review of the literature reveals a rich landscape of studies that employ a diverse range of econometric techniques, time series models, and data sources. This section provides an overview of key developments in the field, highlighting seminal works, recent advancements, and the predominant themes in U.S. unemployment rate forecasting. It aims to underscore the importance of accurate unemployment rate forecasts and lay the foundation for the methodology employed in this research.

Challenges in Unemployment Rate Forecasting

Conventional forecasting models, while valuable in elucidating historical trends and patterns, often grapple with accommodating the myriad of factors influencing contemporary unemployment rates. Current forecasting models frequently encounter difficulties when trying to capture the intricate dynamics of unemployment rates. These limitations substantially impede their capacity to effectively predict shifts in labor markets. In today's swiftly evolving economic terrain, characterized by globalization, technological progress, the rapid evolution of job markets, changing skill requirements, demographic transitions, and external shocks, such as financial crises, public health emergencies, and pandemics, render the development of models that can deliver dependable and up-to-the-minute forecasts a challenging endeavor. Unemployment rates are marked by intricate and multifaceted dynamics.

This research project aspires to confront these hurdles by conceiving an advanced forecasting model for unemployment rates. It recognizes that the intricacies of today's labor markets demand a more all-encompassing and adaptable approach to forecasting. The primary strategy in this undertaking revolves around the amalgamation of historical labor market data with a diverse array of pertinent economic, demographic, and social indicators. Through this approach, the research seeks to create a forecasting model capable of furnishing forecasts that are not only more precise but also timelier.

Key Aspects of the Research Approach

Here are some key aspects of this research approach:

1. **Data Integration**: This research emphasizes the integration of various data sources, including historical labor market data, economic indicators (such as GDP growth, inflation rates, and industrial production), demographic information (age distribution, education levels), and social indicators (public sentiment, job search behavior). By amalgamating these data sources, the model aims to capture a more holistic view of the labor market.

2. **Adaptability**: The research acknowledges that labor markets are not static and are subject to rapid changes. The model developed in this study will be designed to adapt to evolving labor market conditions, ensuring that forecasts remain relevant and reliable in dynamic environments.

3. **Timeliness**: With the inclusion of real-time and high-frequency data, this research aims to provide timely forecasts. Timely information is crucial for businesses, policy-makers, and job seekers to respond effectively to labor market changes.

4. **Policy and Business Decision Support**: The primary beneficiaries of this research are policymakers, businesses, and individuals. By offering more accurate and actionable forecasts, the model will enable policymakers to design targeted interventions, businesses to make informed hiring and expansion decisions, and individuals to plan their careers and job searches effectively.

Model Development

In the quest for the noble objective of refining the precision of unemployment rate forecasting, this scholarly research endeavor meticulously delineates a systematic method to attain this lofty aim. More precisely, the study is dedicated to the creation of two distinct models for forecasting unemployment rates. These models undertake an exhaustive consideration of a plethora of determinants that exert influence over the labor market. The evaluation phase of this study involves the rigorous assessment of the predictive capabilities of both the SARIMAX and LSTM models. To accomplish this, the research employs a dedicated test set, which serves as the litmus test for the models' forecasting prowess. In the ensuing discourse, we shall expound upon the intricacies of this research approach:

Comprehensive Factor Consideration

The first key aspect of this research is the recognition of the multitude of factors that influence unemployment rates. These factors go beyond historical trends and involve a wide range of economic, demographic, and social indicators. By considering a comprehensive set of variables, the research aims to build models that are better equipped to capture the intricate dynamics of the job market.

LSTM Model

The utilization of LSTM, a type of recurrent neural network, represents a departure from traditional time series forecasting methods. LSTM is well-suited to modeling sequences of data and has demonstrated significant promise in capturing long-term dependencies in time series data. In this research, the LSTM model will be employed to analyze and forecast unemployment rates. LSTM's ability to learn from historical data and adapt to changing patterns makes it particularly valuable in the context of dynamic labor markets.

ARIMA Model

ARIMA, on the other hand, is a well-established time series forecasting method. Its strength lies in its ability to model and forecast time series data by considering autoregressive, differencing, and moving average components. The ARIMA model, which has been widely used in economics and finance, provides a reliable benchmark against which the LSTM model's performance can be compared.

Methodological Comparison

A crucial aspect of this research is the side-by-side comparison of the LSTM and ARIMA models. This comparison will allow for an assessment of the effectiveness and performance of each approach in forecasting unemployment rates. The aim is to identify the strengths and weaknesses of both models and potentially leverage the best features of each to create a more robust forecasting tool.

Model Adaptability

The adaptability of the developed models is of paramount importance. Given the ever-evolving nature of job markets, the research emphasizes the need for models that can adjust to changing conditions. LSTM's adaptability to shifting data patterns and SARIMA's ability to model different components of time series data will be closely examined in this context.

Practical Implications

The ultimate goal of this research is to provide practical, actionable insights for policymakers, businesses, and individuals. A more accurate unemployment rate forecasting model can empower these stakeholders to make informed decisions in a timely manner, navigate labor market fluctuations, and implement effective strategies.

Performance Metrics Utilization

The assessment is not limited to mere visual comparisons; rather, it employs quantitative performance metrics to furnish a comprehensive understanding of each model's accuracy. Key metrics, such as the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), are computed to gauge the disparities between the model's forecasts and the observed values. These metrics serve as a basis for numerical evaluation, enabling a precise assessment of the predictive strengths and weaknesses of each model.

Additional Information

The United States' national unemployment rate (UR), which has been seasonally adjusted, is disseminated on a monthly basis by the Bureau of Labor Statistics (BLS), accessible through their website at `www.bls.gov`. The existing seasonal adjustment approach employed by the BLS is denoted as "X-13 ARIMA," and more comprehensive insights into this method can be gleaned at `https://www.bls.gov/cps/seasonal-adjustment-methodology.htm`.

The UR index for the month denoted as 't' specifically pertains to individuals who find themselves without employment but remain available for work during the week inclusive of the 12th day of month 't'. Additionally, these individuals have actively sought job opportunities in the preceding four weeks, which culminate with the reference week.

This vital statistical information, at the federal level, spans from January 1948 and extends up to the most recent data.

In figure 2, it is evident that the data representing the unemployment rate (UR) exhibits a pronounced asymmetry in its cyclical patterns, especially during the periods of severe economic downturns, which predominantly shape the temporal evolution of the UR data. We observe that rapid, steep ascents culminating in sharp peaks typify the nature of general economic contractions, while the UR data feature prolonged and gradual descents, characteristic

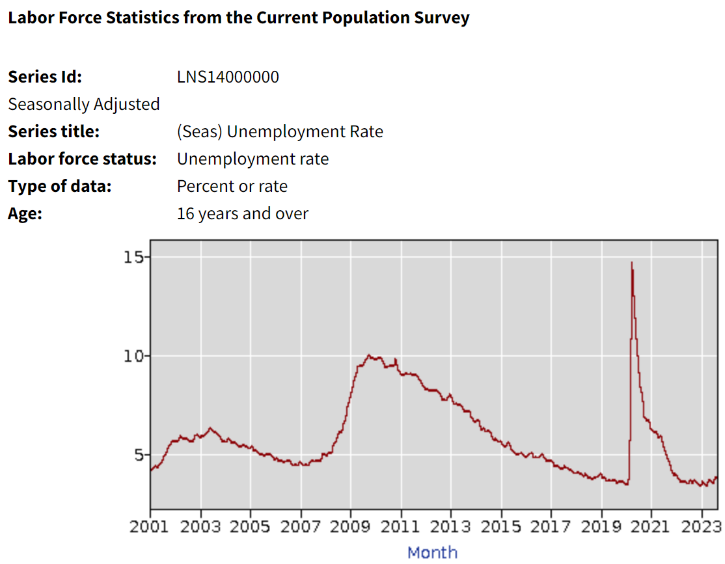

Figure 2: Unemployment Rate Patterns

of business expansions.

The endeavor of forecasting the unemployment rate in the United States necessitates the application of statistical models to anticipate forthcoming values, drawing insights from a rich tapestry of historical data. In assessing the precision of these predictions, we employ a set of widely recognized metrics, which encompass the Mean Absolute Error (MAE), Mean Squared Error (MSE), and the Mean Absolute Percentage Error (MAPE).

Quantitative Measures for Evaluation

To evaluate the precision of our forecasts, we employ the following quantitative measures:

1. Mean Absolute Error (MAE): MAE quantifies the average absolute disparity between the projected values and the real values. It serves as a yardstick for the model's predictive accuracy.

2. Mean Squared Error (MSE): MSE determines the average of the squared deviations between the projected and actual values, offering greater emphasis on larger discrepancies.

3. Mean Absolute Percentage Error (MAPE): MAPE evaluates the percentage variance between the predicted and actual values, thereby revealing the relative accuracy of the forecasts.

The mathematical expressions for these metrics are elucidated as follows:

$$\mathrm{MAE} = \frac{\sum |\mathrm{actual} - \mathrm{predicted}|}{n}$$

$$\mathrm{MSE} = \frac{\sum (\mathrm{actual} - \mathrm{predicted})^2}{n}$$

$$\mathrm{MAPE} = \left(\frac{\sum |\mathrm{actual} - \mathrm{predicted}| / \mathrm{actual}}{n} \right) \times 100$$

Here, "actual" pertains to the observed or historical values, "predicted" denotes the values forecasted, and "n" signifies the count of data points.

In the phase of interpretation, the outcomes of MAE, MSE, and MAPE should be scrutinized to gauge the efficacy of the model. A diminishment in MAE and MSE signifies heightened accuracy, while a decreased MAPE implies superior relative precision.

A commitment to continual enhancement is indispensable. This entails the refinement of the forecasting model, adjustments to model parameters, and the integration of fresh data as it becomes available. These measures serve to amplify the precision of forthcoming forecasts.

In summation, the delineated framework provides a top-tier perspective on the process of forecasting the unemployment rate in the United States and the concurrent assessment through MAE, MSE, and MAPE. The selection of the forecasting model and the intricacy of the analysis shall be contingent upon the specifics of the dataset, the temporal scope, and the requisite degree of sophistication.

4 Data and Methodology

4.1 About the Data

For the purpose of this project, we will be utilizing the dataset available on Kaggle titled "USA Key Macroeconomic Indicators" that contains data retrieved from FRED (Federal Reserve Bank of St. Louis) that was released by the U.S. Bureau of Economic Analysis. This dataset provides a comprehensive collection of key macroeconomic indicators for the United States, spanning from 1981 to 2021. One of the primary indicators present in this dataset is the unemployment rate. The unemployment rate is a vital economic metric that represents the percentage of the labor force that is jobless but actively seeking employment. Historically, fluctuations in the unemployment rate have been influenced by various factors, including economic downturns, technological advancements, and policy changes. By analyzing this dataset, we aim to gain insights into the patterns and factors affecting the unemployment rate over the years, which will be instrumental in building and refining our predictive models.

Features of the Dataset

1. **DATE**: Represents the date for the recorded data. This can be on a monthly or quarterly basis.
2. **unrate (Unemployment Rate)**: Indicates the percentage of the total labor force that is unemployed but actively seeking employment and willing to work.
3. **psr (Personal Saving Rate)**: Represents the percentage of people's income that they save rather than spend.
4. **m2 (M2)**: A measure of the money supply that includes cash, checking deposits, and easily convertible near money.
5. **dspic (Real Disposable Personal Income)**: The amount of money that households have available for spending and saving after income taxes have been accounted for.
6. **pce (Personal Consumption Expenditures)**: Represents the value of the goods and services purchased by, or on the behalf of, "households".
7. **reer (Real Broad Effective Exchange Rate)**: An index that describes the strength of a currency relative to a basket of other currencies, adjusted for inflation.
8. **ir (Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity)**: Indicates the return on investment for the U.S. government's 10-year treasury note.
9. **ffer (Federal Funds Effective Rate)**: The interest rate at which depository institutions lend reserve balances to other depository institutions overnight on an uncollateralized basis.
10. **tcs (Total Construction Spending)**: Represents the total amount spent on construction projects.

4.2 Methodology

The following step-by-step approach will be adopted:

Figure 3: Steps of our approach

1. **Data Collection and Cleaning**:
   - Obtain historical data for US unemployment rates. This data will preferably span several decades to account for various economic cycles.
   - Integrate external factors (e.g., economic indicators, policy changes) that could provide additional context and potentially improve the model's accuracy.
   - Clean the data to remove any outliers or missing values, ensuring that it is consistent and ready for analysis.
2. **Exploratory Data Analysis (EDA)**:
   - Visualize the data to identify patterns, trends, seasonality, and cyclic behaviors.
   - Determine statistical properties of the data like mean, variance, and autocorrelations.
   - Evaluate and choose potential external factors, such as GDP growth, inflation rates, and policy changes, based on their historical influence on the unemployment rate.
3. **Data Preprocessing**:
   - Split the data into training and test sets. The training set will be used to train the models, while the test set will be reserved for model evaluation.
   - Normalize or standardize data if necessary, especially for LSTM, which is sensitive to the scale of input data.
4. **Model Implementation**:
   - *SARIMAX Model*:
     - Use the Seasonal Autoregressive Integrated Moving Average (SARIMAX) model, which is well-suited for time series data with trends and seasonality.
     - Identify the best parameters (p, d, q)(P, D, Q)s for the SARIMAX model using grid search or similar techniques.
     - Train the SARIMAX model on the training dataset.
   - *LSTM Model*:
     - Use the Long Short-Term Memory (LSTM) model, a type of recurrent neural network (RNN), to model the unemployment rate.
     - Structure the data into suitable time steps for LSTM input.
     - Design the LSTM architecture, choosing the number of layers and units.

    - Train the LSTM model on the training dataset.

5. **Model Evaluation**:

    - Use the test set to evaluate the predictive performance of both the SARIMAX and LSTM models.
    - Compare the forecasts of both models with the actual values from the test set.
    - Compute performance metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and others to quantify the accuracy of each model.

5 Results

5.1 Exploratory Analysis

*UPCOMING TASKS

5.2 SARIMA

5.2.1 SARIMA - Univariate

Model Architecture and Hyperparameters for Baseline SARIMA models:

- Model Orders:

  order: p = 1; d = 1; q = 1.

  seasonal order: P = 1; D = 1; Q = 1; S = 12.

- Train set = 1981 - 2020
- Validation set = 2021

Upon evaluation, the SARIMAX univariate model achieved a Mean Absolute Error (MAE) of 1.7223 and a Mean Squared Error (MSE) of 4.5667.

The univariate approach to modeling the unemployment rate leverages the temporal structure inherent in the data, building a model that projects future values based solely on past observations. Although SARIMAX captures potential seasonality, trend, and autocorrelation patterns present in the unemployment data, its singular reliance on past unemployment rates can be a limitation.

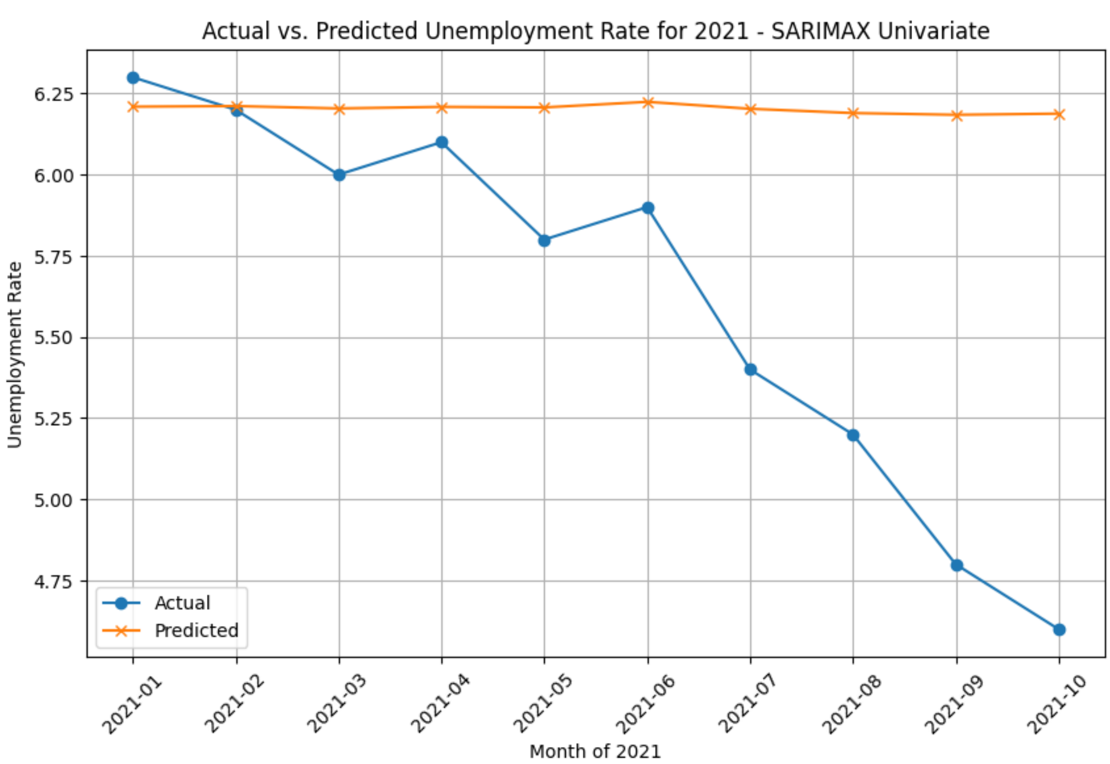

Figure 4: Actual vs Predicted Values - SARIMAX Univariate Model

In Figure 4, the SARIMAX univariate model showcases its capability to approximate the true trend of the unemployment rate. While the model exhibits a strong general understanding of the underlying patterns, discrepancies between predicted and actual values highlight the challenges in exclusively using past unemployment rates to forecast future values.

5.2.2 SARIMAX - Multivariate

Model Architecture and Hyperparameters for Multivariate SARIMA models:

- Exogenous Variables: Oil Prices, and US GDP
- Model Orders:

  order: p = 1; d = 1; q = 1.

  seasonal order: P = 1; D = 1; Q = 1; S = 12.
- Train set = 1981 - 2020
- Validation set = 2021

Upon evaluation, the multivariate model posted a Mean Absolute Error (MAE) of 1.6093 and a Mean Squared Error (MSE) of 3.5311. When compared to the univariate model, the multivariate version demonstrates improved accuracy, as highlighted by its reduced error metrics.

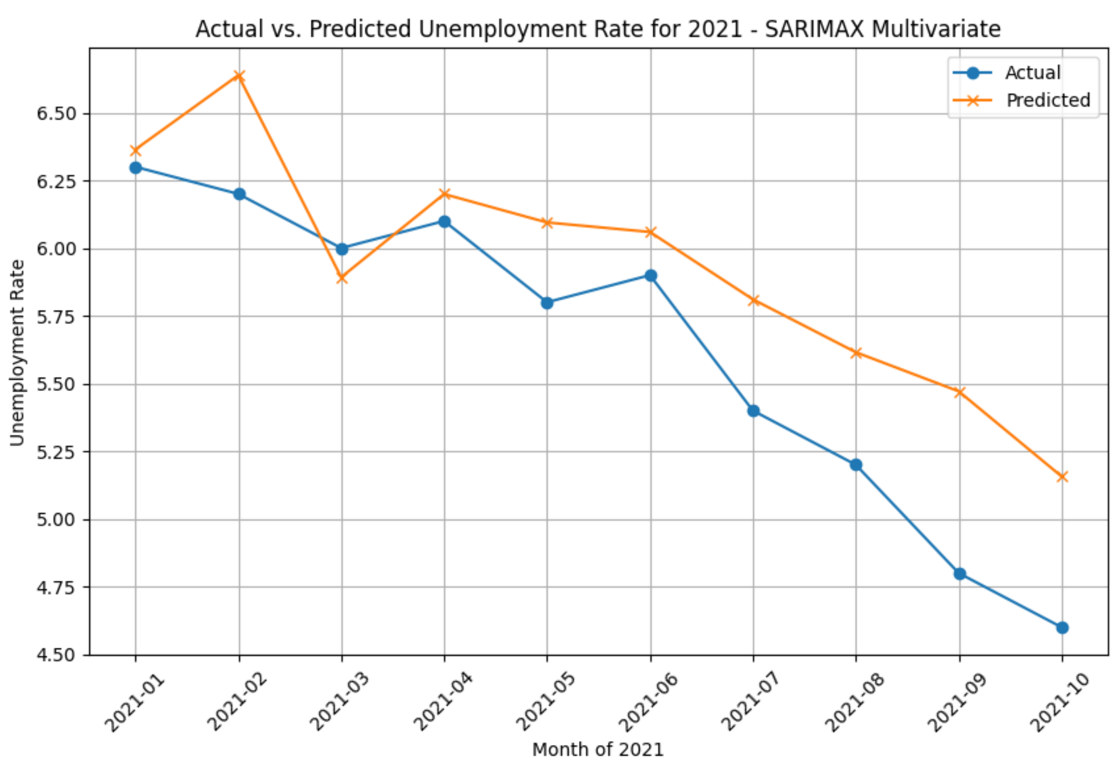

Figure 5: Actual vs Predicted Values - SARIMAX Multivariate Model

By examining 5, one observes that the multivariate model does capture the trend in the unemployment rate with greater fidelity than its univariate counterpart. The introduction of additional exogenous variables, such as GDP and oil prices, provides the model with a broader context, allowing it to draw upon relationships between unemployment and other economic indicators.

However, the efficacy of a multivariate approach is contingent on the relevance and quality of the exogenous variables. Irrelevant or noisy predictors can degrade the model's predictive power, making variable selection critical. Further analysis and potential feature engineering could yield even more accurate SARIMAX multivariate models.

5.3 LSTM

Model Architecture and Hyperparameters for Baseline LSTM models:

- Hidden Layers = single hidden layer of 64 neurons; ReLu activation
- Learning Rate = 0.0001
- Train set = 1981 - 1999
- Validation set = 2000 - 2020
- Test set = 2021

5.3.1 LSTM - Univariate

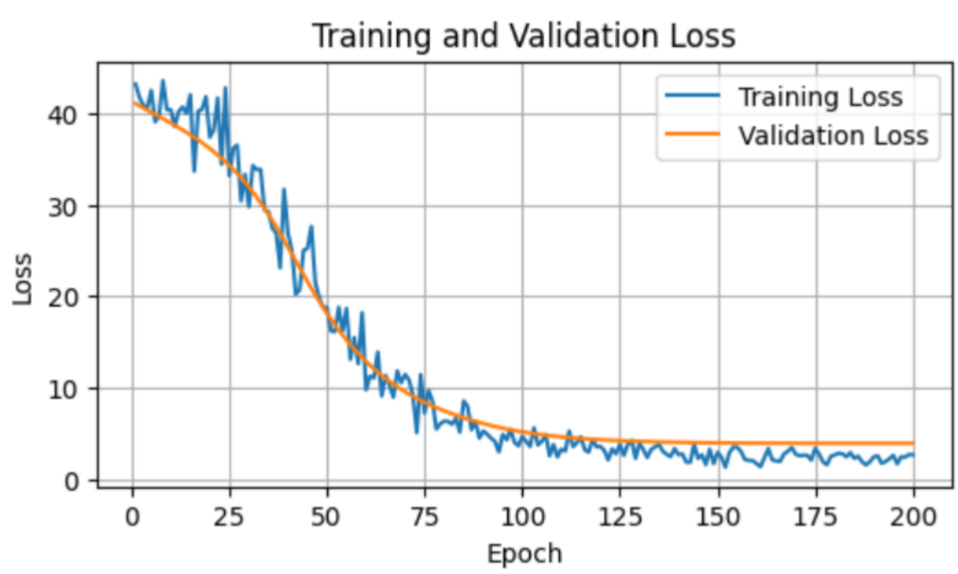

Figure 6: LSTM Univariate Model Training-Validation Loss

The results obtained from the training of the LSTM univariate model suggests that given sufficient training epochs, the model is able to reduce both training and validation loss despite the limited number of data points available.

In addition, evaluating variations of the trained model based on benchmarks such as the average of the root mean square error, mean absolute error, and mean absolute percentage error suggested that lowering the learning rate improved the model's overall accuracy in predicting the unemployment rate on the test data. Given a learning rate of 0.0001 with a hidden size of 64 while keeping the number of layers minimal, the model was able to achieve an average root mean square and mean absolute error of 0.56, with the average mean absolute percentage error being 10.97%. Although the accuracy on the test data showed promise, examining the prediction values next to the actual values revealed the shortcomings of working with limited data when utilizing a recurrent neural network.

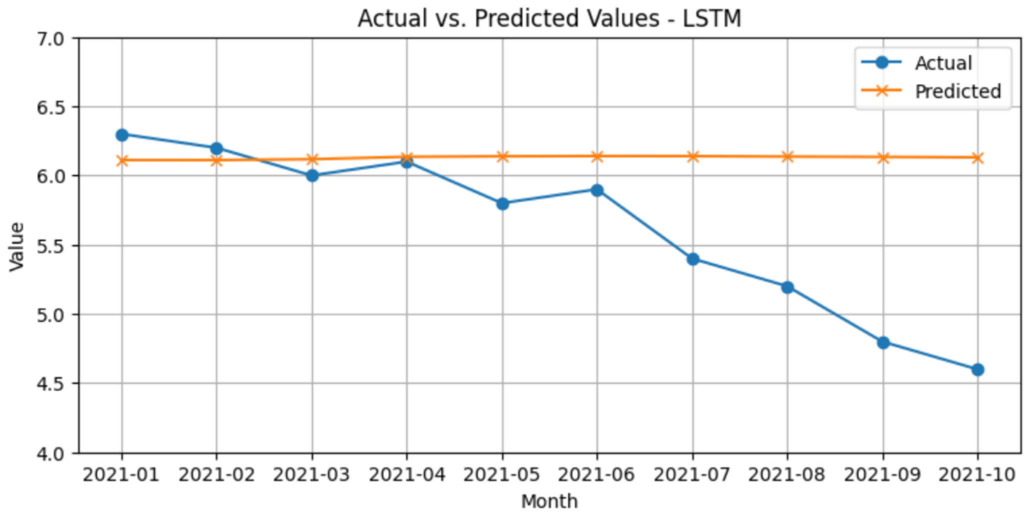

Figure 7: Actual vs Predicted Values - LSTM Univariate Model

Given that the LSTM model was making predictions based on time-series data, the model was trained based on the unemployment rates of consecutive years grouped by months. This was done in order to prevent the model from making predictions based on previous predictions, potentially causing issues such as autoregressive forecasting, which could be detrimental to the model's accuracy in making predictions especially in cases where the total amount of data available is limited. Hence the warranted approach in predicting unemployment rates based on different monthly sets of previous year's monthly data. This resulted in the model making similar predictions across different months resembling that of the average unemployment rate of the actual values on the test data in order to achieve a higher accuracy.

Despite the limited amount of data available for training, validation and testing, the greatest hindrance preventing further increases in model accuracy is the univariate nature of the model. Recurrent neural networks excel at the prediction of sequential time-series data. However, while Long Short-Term Memory networks have proven to be effective tools for time-series forecasting, the use of univariate LSTM models in scenarios characterized by limited data points reveals inherent shortcomings. One primary concern is the potential for overfitting, whereby the model may excessively adapt to the idiosyncrasies of the limited dataset, resulting in poor generalization to unseen data. Limited data also restricts the model's ability to capture complex temporal dependencies such as that of predicting unemployment rates which can be dependent on a myraid of different factors. LSTMs inherently require a sufficient volume of historical data to uncover meaningful patterns. Furthermore, the lack of diversity in a small dataset can limit the model's ability to adapt to changing patterns or seasonality. In such cases, the model may struggle to provide reliable predictions.

5.3.2 LSTM - Multivariate

The baseline multivariate LSTM model performs noticeably worse than the univariate LSTM model, with a MAPE of 48.56%, MSE of 7.41, and MAE of 2.69. This suggests that the additional input features provided to the baseline model might not be contributing positively to its predictive power. One possible explanation for this underperformance could be issues such as feature redundancy, noise, or irrelevant information in the additional input variables, which could be obstructing the model's ability to extract meaningful patterns and relationships from the data.

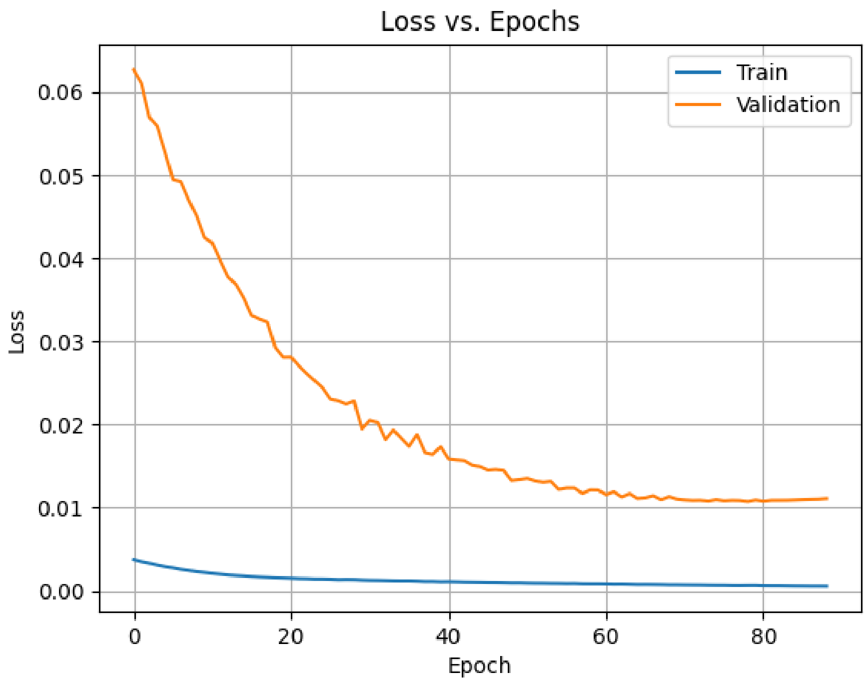

Figure 8: LSTM Multivariate Model Training-Validation Loss

In Figure 9, it is evident that the model correctly captures the downward trend of the test data, albeit with a consistent overestimation throughout the entire period. When compared to the graph of the univariate LSTM, the results appear more promising, despite the higher error.

To delve deeper into this issue, several follow-up steps can be taken. First and foremost, it's crucial to do a comprehensive feature selection or engineering process to pinpoint which features, if any, truly hold value for the task at hand. Techniques such as correlation analysis, mutual information, or recursive feature elimination can be instrumental in this regard. Additionally, exploring data preprocessing steps, such as scaling, normalization, or transformation, may have the potential to enhance the model's overall performance.

Moreover, it's advisable to experiment with various model architectures and hyperparameters for both the baseline and univariate LSTM models. This exploration can help determine whether alternative configurations could yield more favorable results. The inclusion of ensemble methods, cross-validation, and the application of rigorous evaluation metrics is essential to ensure a robust and equitable comparison between the two models.

The loss curve, which exhibits a gradual and smooth decrease in both training and validation losses, indicates that the model is effectively learning from the data without overfitting to the training dataset.

However, the significant size disparity between the test set and the training/validation sets can significantly influence model evaluation. This imbalance can lead to unreliable or biased performance estimates.

For instance, a small test set may not be representative of the broader distribution of real-world data. Consequently, if the model happens to perform well on this limited test set, it can mistakenly instill unwarranted confidence in the model's capabilities. This becomes especially problematic if the model has overfit the training data and the test set coincidentally contains examples that the model handles well but isn't indicative of the overall data

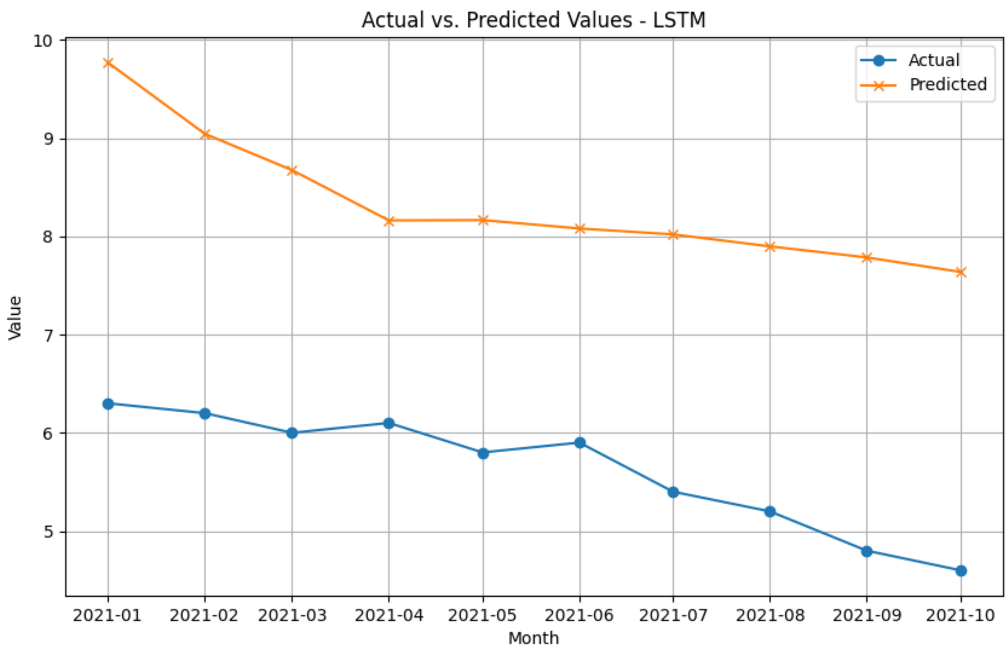

Figure 9: Actual vs Predicted Values - LSTM Multivariate Mode

distribution.

6 Discussion

6.1 Baseline models

In using Mean Absolute Error (MAE) as our primary metric for comparison of the baseline models, the univariate LSTM demonstrated the best performance, while the univariate SARIMA had the least favorable results.

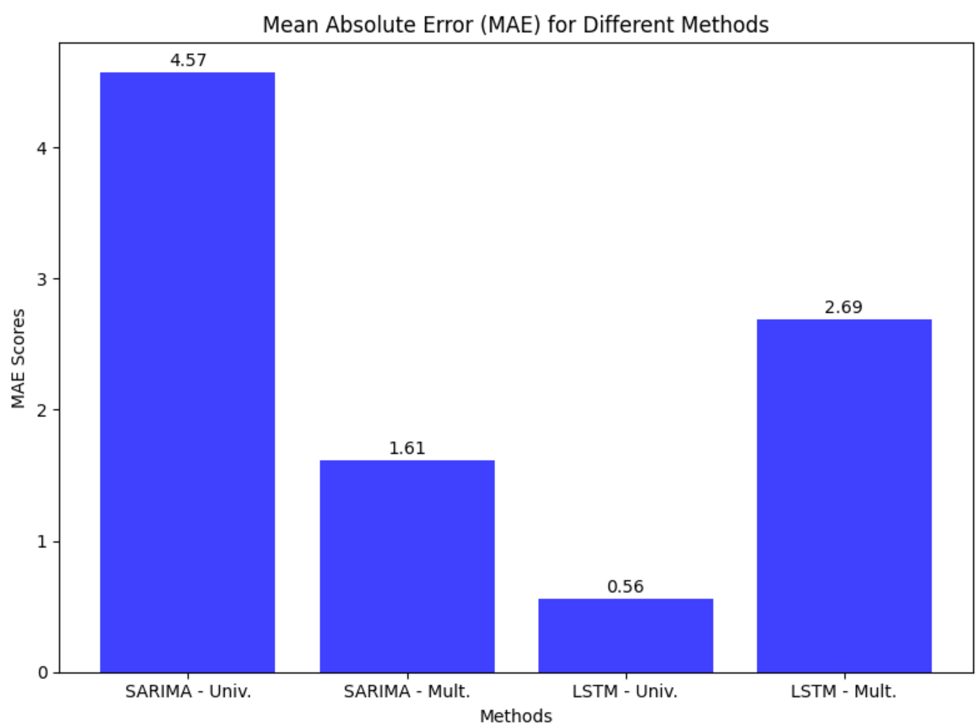

Figure 10: Mean Absolute Error (MAE) for Different Methods

6.1.1 LSTM

The exploration of the baseline Long Short-Term Memory (LSTM) network models yielded various insights and observations:

1. **Baseline Performance**:
    - The univariate and multivariate LSTM models displayed a reduction in training and validation losses over sufficient training epochs. However, the limited data size posed challenges in achieving high accuracy.
    - With specific hyperparameters: a learning rate of 0.0001 and a single hidden layer of 64 neurons, the univariate model predicted with an average root mean square error of 0.56 and a mean absolute percentage error of 10.97%.
    - In contrast, the multivariate LSTM model, using all available features as input, did not outperform as expected. With a significantly higher mean absolute percentage error (MAPE) of 48.56%, this suggests potential issues such as redundant features, noisy data, or irrelevant input variables, underscoring the importance of feature selection and engineering.
    - Interestingly, the multivariate model was able to capture the data trend, albeit with consistent overestimation.

2. **Data Limitations**:
    - The LSTM's performance indicates that it is paramount to have ample data, especially for recurrent neural networks, to capture intricate temporal patterns effectively.
    - Despite the design choice to prevent autoregressive forecasting, the univariate model still displayed tendencies to predict averages due to limited data, pointing towards challenges in predicting with sparse datasets.
    - The primary concern for the univariate model was overfitting given the limited data, which might make the model susceptible to poor generalizations for unseen data.
3. **Recommendations for Improvement**:
    - Feature selection methods, such as correlation analysis and recursive feature elimination, could help in refining the input data for the multivariate model.
    - Experimenting with different architectures, hyperparameters, and preprocessing steps might enhance performance.
    - Ensemble methods and cross-validation techniques would further ensure robust and comprehensive model evaluations.
4. **Evaluation Challenges**:
    - The disparity in size between the training/validation sets and the test set could skew model performance evaluation.
    - A smaller test set might not truly represent the broader real-world data distribution, which can lead to overconfidence in the model's capabilities, especially if the model overfits to the training data.

In summary, while LSTMs have shown promise in time-series forecasting, the current project underscores the importance of having sufficient and well-curated data, a robust feature selection process, and the need to carefully assess model evaluation metrics and procedures.

6.1.2 SARIMA

1. **Baseline Performance**
    - The univariate SARIMAX model demonstrated its capability to predict the unemployment rate based on historical data. The model was trained using the unemployment rate as the target variable without any exogenous inputs.
    - For the univariate SARIMAX model with specific orders (p, d, q, P, D, Q, S) of (1, 1, 1, 1, 1, 1, 12), the model achieved a Mean Absolute Error (MAE) of 1.7223 and Mean Squared Error (MSE) of 4.5667 which were calculated based on the validation set.
    - The multivariate SARIMAX model incorporated additional exogenous variables such as GDP and oil prices with a Mean Absolute Error (MAE) of 1.6093 and Mean Squared Error (MSE) of 3.5311. However, the inclusion of these variables did not necessarily guarantee a significant improvement in prediction accuracy.

The model's performance metrics, MAE and MSE, were indicative of its predictive capabilities.

2. **Data Limitations:**
   - The interpolation method used for GDP data might introduce biases or inaccuracies, especially if there are large gaps in the data.
   - The reliance on external data sources like FRED means that any inconsistencies or errors in their data will directly impact the model's training and predictions.

3. **Recommendations for Improvement:**
   - Feature engineering and selection are crucial. While GDP and oil prices were used as exogenous variables, other economic indicators might provide more relevant insights into the unemployment rate.
   - Hyperparameter tuning for the SARIMAX model, such as experimenting with different orders, can potentially improve model performance.

4. **Evaluation Challenges:**:
   - The split between training and validation data is crucial. If the validation set does not adequately represent the broader data distribution, the model's evaluation metrics might not be indicative of its real-world performance.
   - Overfitting remains a concern, especially if the model becomes too attuned to the training data's nuances and fails to generalize well to new, unseen data.

7 Conclusion

*UPCOMING TASKS

8 References

- Han Xiao, Rongda Chen, J. Guerard. (Year). *Forecasting the U.S. Unemployment Rate: Another Look*. Retrieved from `https://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1716&context=theses`

- Peter Ďurka, & Silvia Pastoreková. (2012). ARIMA vs. ARIMAX – which approach is better to analyze and forecast macroeconomic time series? In *Proceedings of 30th International Conference Mathematical Methods in Economics* (pp. 136-140). University of Economics, Department of Statistics, Dolnozemská cesta 1, 852 35 Bratislava 5, Slovakia.

- Stevens, J. P. (2012). *Applied Multivariate Statistics for the Social Sciences*. Routledge.

- Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Almquist and Wicksell.

- Whittle, P. (1963). *Prediction and Regulation*. English Universities Press. ISBN 0-8166-1147-5.

9 Appendix

9.1 Division of Work

We structured our task delegation according to the project plan below, with each team member assigned responsibility for specific activities as indicated in the right-most column.

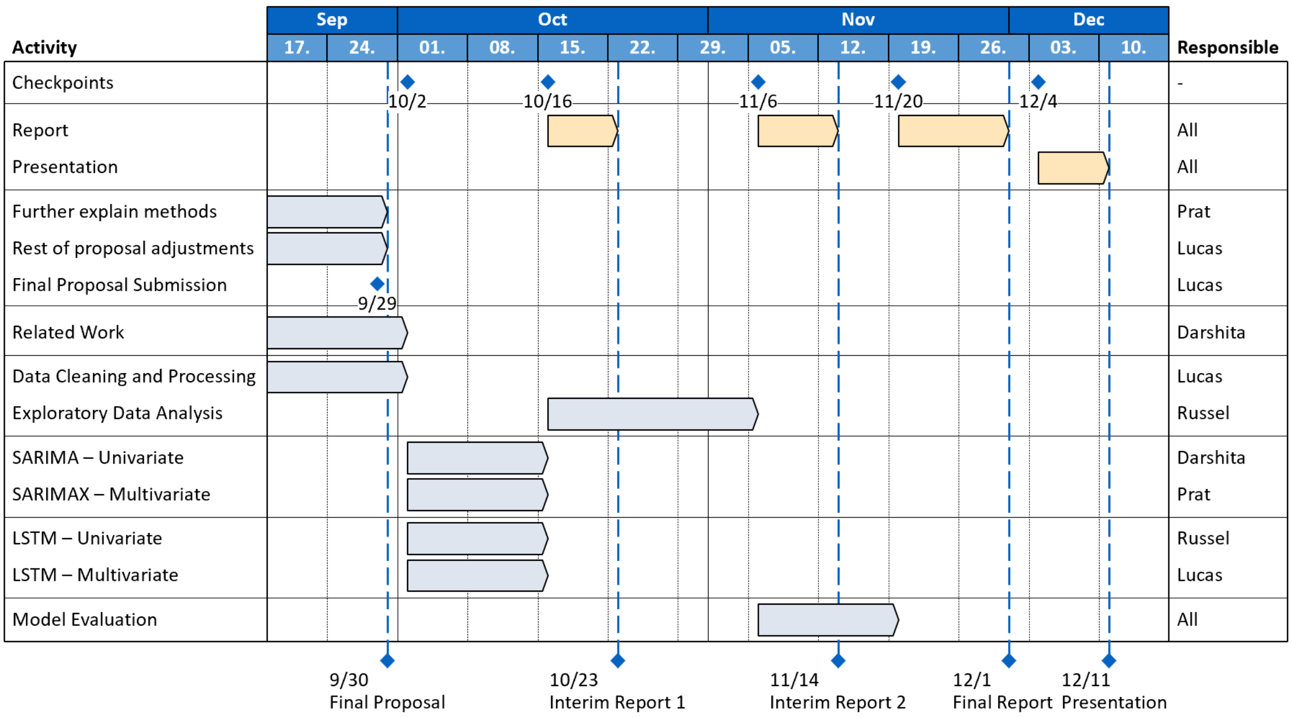

Figure 11: Plan of work

9.2 Program Architecture

The flowcharts below describe the program architecture of each of the models.

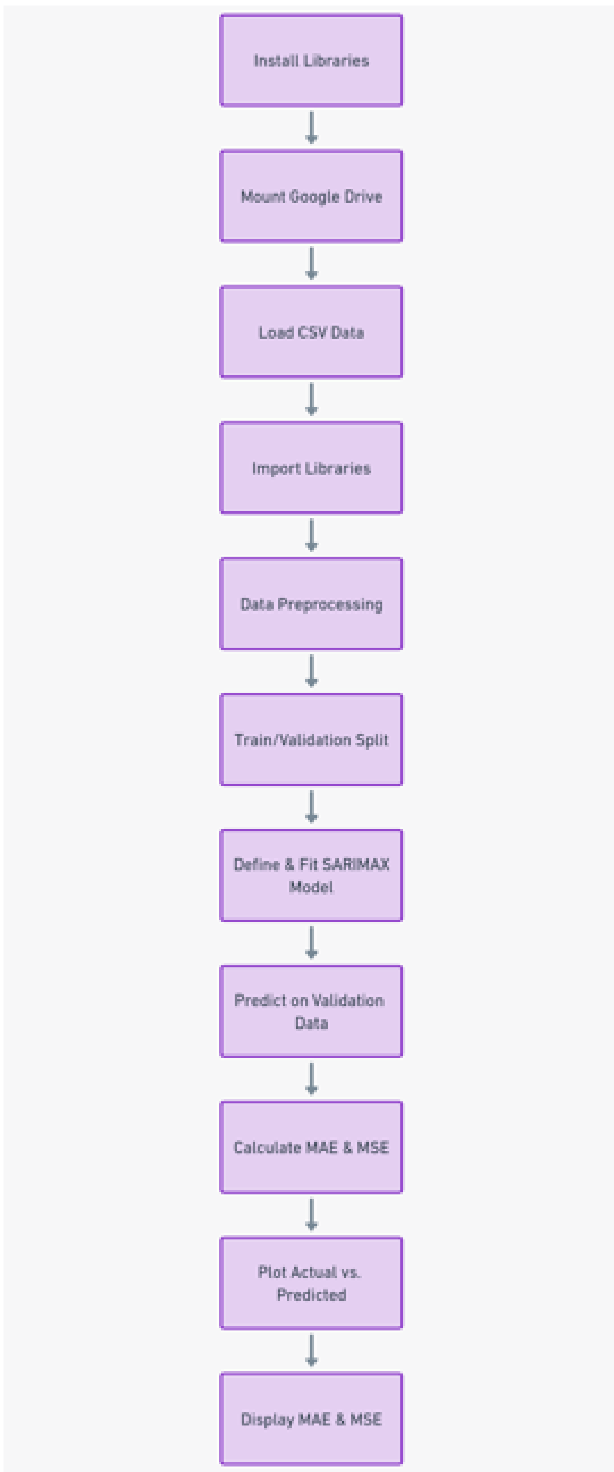

Figure 12: SARIMA Multivariate Architecture

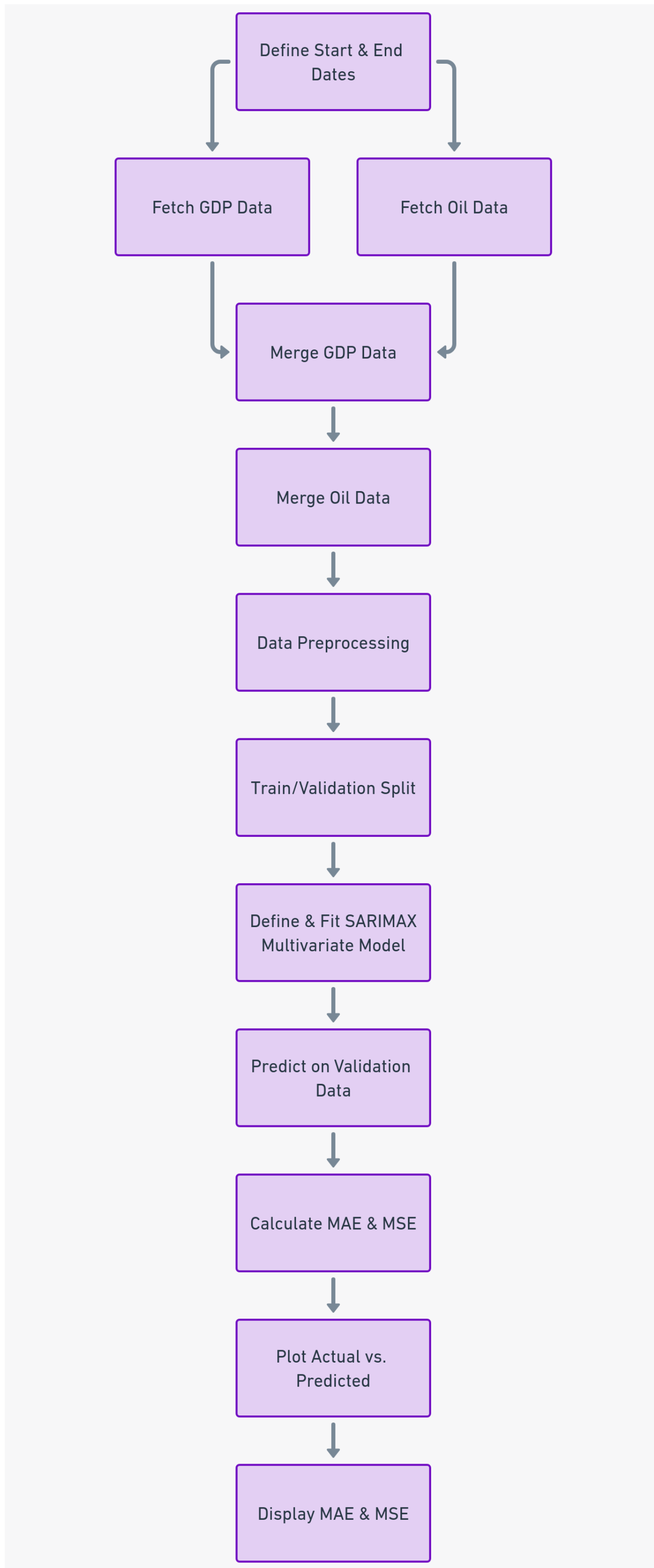

Figure 13: SARIMA Univariate Architecture

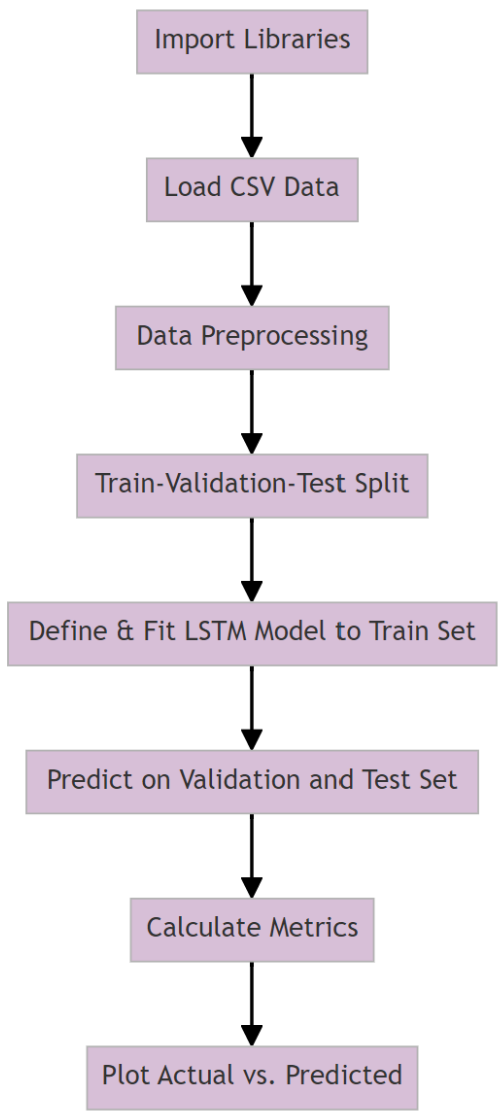

Figure 14: LSTM General Program Architecture