Explainable AI: Investigating Transformer, NBeats, and LSTM models for Inflation Forecasting Economic Indicators

MSc Research Project
Data Analytics

Pradeep Kumar Reddy Bathula
Student ID: X22165801

School of Computing
National College of Ireland

Supervisor: Aaloka Anant

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Pradeep Kumar Reddy Bathula
Student ID:	X22165801
Programme:	Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Aaloka Anant
Submission Due Date:	14/12/2023
Project Title:	Explainable AI: Investigating Transformer, NBeats, and LSTM models for Inflation Forecasting Economic Indicators
Word Count:	5726
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	31st January 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Explainable AI: Investigating Transformer, NBeats, and LSTM models for Inflation Forecasting Economic Indicators

Pradeep Kumar Reddy Bathula
X22165801

Abstract

Our research examines the application of Explainable Artificial Intelligence (XAI) principles to understand how sophisticated machine learning models, such as Trans- form, NBeats, along with LSTM, anticipate rising prices. It utilizes key economic indicators such as Personal Consumption Expenditures (PCE), Producer Price Index (PPI), Gross Domestic Product (GDP), and Consume r Price Index (CPI). Main objective is understanding how algorithms interpret and incorporate economic data into their predictions. This will increase our understanding of complex eco- nomic forecasting processes. Research uses XAI for determining how each economic indicator improves algorithms' forecast accuracy. This technique lets us compare model computing capabilities to determine relative importance of PCE, PPI, GDP, CPI into inflation estimates. Our results underscore the relevance of components to improving prediction accuracy through explaining underlying model concepts. This investigation expands to explore AI models in economic research. Research using XAI improves economic decision- making by providing clear and understandable AI-driven forecasting results. Our study stresses explaining the ability in using AI models for complicated economic projections to contribute to the fast- growing area of AI in economics. It also stresses openness and honesty whereas using AI for complex economic forecasting and research.

1 Introduction

The general trend toward higher prices and higher expense of living is known as inflation(Oner; 2019). It causes consumer-oriented goods and services to have greater prices. Popular metrics used to track and evaluate inflation include CPI, PPI, PCE, and GDP (Zimmermann; 2015).We included XAI in our research to examine the model's predictions. Particularly, we made use of the well-known XAI approach SHAP. These methods help us understand how the model makes its predictions through drawing attention to the most important attributes and the roles they play.

1.1 Background and Motivation

According to Salman et al. (2018),Sridhar and Sanagavarapu (2021),Tuominen et al. (2023),These research's show that, in one article, the Transformer multi-head is regarded as the best; in another, the LSTM multi-layer; and in a third, the N-BEATS multi-layer. Using Explainable Artificial Intelligence (XAI) methodologies, we compare the performance of three models and improve their interpretability in this study.

1.2 Research Question

1.2.1 Question

In the context of Explainable AI (XAI), how can Transformer, NBeats, and LSTM models understand and use these metrics? and which of these metrics—PCE, CPI, PPI, and GDP— has the biggest influence on the projection produced by the models?

The latter phase of our study involves assessing many neural network modules for forecasting inflation by using XAI. The models included in this set are Transformer, LSTM, and N-BEATS. Such models use XAI as a primary tool for analyzing comprehend several economic variables, including GDP, Consumer Price Index (CPI), Producer Price Index (PPI), PCE. Experts meticulously analyze historical data upon inflation and other crucial economic variables to train assess models. By using Explainable Artificial Intelligence (XAI), we could comprehensively examine interrelationships and mutual influences of Gross Domestic Product (GDP), Consumer Price Index (CPI), Producer Price Index (PPI), and Personal Consumption Expenditures (PCE) into predicting models' frameworks of. To make models' internal decision-making processes more transparent and understandable, study employs XAI.

2 Related Work

Related Work section, we take close look at previous research that utilized similar approaches to analyze their strengths and weaknesses. Its purpose is to survey the current literature extensively to find areas where research is lacking to provide a basis for future methods. Finding weaknesses in earlier research requires thorough reliable technique, as this study shows. Our method incorporates results from literature review to tackle these challenges. Our primary goal is to make substantial advancements over previous approaches, leading to findings which may help both academic research and real-world applications.

2.1 The Past Study

2.1.1 Phillips curve forecasts

Phillips curve predictions consist of a wide variety of methodologies for projecting hyperinflation and its variants. To foretell the future, these techniques analyze economic indicators such as growth, rate of unemployment, and output gap. Policymakers' economists could gain better grasp of dynamics of inflation through studying connections amongst these factors inflation(Stock and Watson; 2008).Instead of seeing both unemployment and inflation as both sides of the same coin, economists should begin to examine the distributional and misallocationary effects of active fiscal policy. In cases whenever negative output gaps are significant, increasing aggregate demand (AD) might lower unemployment while just slightly raising inflation. In deep recessions, the phenomenon of crowding out does not occur, hence this decrease in unemployment might endure for some time.

Consequently, private investment should stay stable regardless of monetary policy or government expenditure, allowing the former to continue its long-term effects of lowering unemployment and increasing economic growth (Pettinger et al.; 2023)

2.1.2 Naive approach

Predicting future inflation rates using a naive method is very risky since it assumes the rate will remain constant relative to its earlier pattern. Potential inflation rates might be affected by fundamental economic drivers, seasonality, current trends; however, those variables are not included in this method. Ignoring the intricacies of economic development, Naive Forecast takes simplified approach that depends only on past outcomes predicts that certain trends as predictive of future actions(Thomas; 1999). Many times, multiple regression models surpass naive prediction methods when it comes to MAE (Mean Absolute Error) and MAPP since the latter considers more factors simultaneously. A multiple regression model produces more accurate and trustworthy predictions than naive approach since it considers all relevant elements rather than just past trends(Dhakal; 2018).

2.1.3 Fisher's model

For calculating nominal interest rate, one may use Fisher's interest rate model for adding the projected real interest rate and inflation rate. The predicted rate of inflation may be calculated by subtracting the current nominal interest rate from the expected rate of real interest. According to Thomas (1999), this method accounts for the real and nominal components of interest rates to predict future inflation patterns. Although Fisher equation provides a means of predicting hyperinflation by connecting real and nominal interest rates, it does so with limitations, especially whenever dealing with lower nominal rates. Spending may not be stimulated by lower nominal rates into liquidity trap. Because demand is unlikely to decrease regardless of a rise in real interest rates, consumer confidence may remain high. Efficient monetary policy should consider these factors. It is possible as the Fisher equation isn't an ideal tool for forecasting future inflation due to these limitations(; 2021).

2.2 Multiple linear regression

The goal of Multiple Linear Regression technique is to find the best linear equation that links independent variables with dependent ones. Forecasts of GDP growth, unemployment, interest rates, and money supply are just a few economic variables this method considers when calculating inflation. Inflation forecasts using estimated coefficients from regression analysis of historical data upon these variables are possible (Elsiddig and Mohamed; 2015). Assuming some degree of relationship amongst independent variables is crucial to efficiency of multiple regression analysis model. Basic idea of classic multiple regression analysis is to find out how those independent variables, as economic indicators, affect dependent variable, such as inflation, but if they don't have any link, model won't work(Jeon; 2015).

2.3 Time Series Model

2.3.1 ARIMA

One way to predict and analyze time series data is via ARIMA, which signifies Auto-Regressive Integrated Moving Average. In this model, concepts of autoregression (AR) moving average (MA) are fused. Differences between autoregression and moving average lie in their respective consideration of data series lags and prediction error delays. This synergy allows ARIMA to effectively characterize and predict time-varying data having serially dependent values(Banerjee; 2014). P, d, q are the 3most important parameters in the AR- IMA model that define the model's parameters. According toKelikume and Salami (2014),The order Moving Average (MA), degree of differencing (I), and Autoregressive (AR) component are determined, respectively, by these parameters. Even though ARIMA models are widely used, there are certain restrictions on how well they can predict time series data. When attempting to predict the outcomes of various economic or policy scenarios, this is particularly true(Kenny et al.; 1998).

2.3.2 VAR

Two critical decisions need to be taken when assessing the VAR algorithm. The selection of variables is the first stage in the VAR model. Because there are no exact theoretical models for the dynamics of inflation, this method is based on statistical concerns. Finding the best delay condition is also crucial.The decisions made in this context impact the way the model represents the interaction and time-based relationships among the selected variables. For VAR analysis to provide precise dependable results, it's essential to meticulously assess these parameters(Moser et al.; 2006). VAR models are sometimes referred to as" black boxes" since they operate in- dependently of economic theory frameworks. This phrase arises from their inability to clearly explain present economic explanations for dynamic interactions. This quality may make it more challenging to understand economic processes or causal relationships that cause VAR model's outputs to show apparent correlation(Bentour; 2015).

2.4 Macroeconomic Variables

The goal of economic forecasting is to foretell future economic conditions by examining key factors. Research into this area often utilizes macroeconomic data including GDP growth, private consumption expansion, private investment growth, and CPI inflation. In addition, forecasting approach incorporates financial predictors like private sector credit growth, stock prices, property prices, bank prime lending rates, deposit rates to offer complete picture of economic trends expected future changes (Chen and Ranciere; 2017). It may come as a surprise, but studies have shown adding more variables to financial forecasting models doesn't always give rise to more accurate predictions. Factors formed from a selection of variables, instead of all of them, might often provide more accurate projections, according to recent research. Economic estimate accuracy is best improved by focusing upon quality relevance of variables instead of their number(Aras and Lisboa; 2022)

2.5 Machine Learning

2.5.1 Random Forest

When it comes to classification regression, many turn to Random Forest (RF), robust versatile machine learning technique. Being able to handle several types of data is where its strength lies. A popular option in many machine learning applications, Random Forest is a nonlinear non-parametric approach which can handle difficult datasets which don't fit traditional linear models(Aras and Lisboa; 2022). An assortment of tree predictors is produced using Random Forest, where every tree is built using randomly selected subset of dataset. The more trees there are in the forest, the less generalization error will be the efficiency of Random Forest model is greatly affected by two primary variables: robustness of every single tree degree to which they are correlated. A better Random Forest model as a whole benefits from stronger individual trees less correlation between them(Breiman; 2001). When working with numerical traits, Random Forest model proves to be quite effective. It can produce rules depending on precsie numerical value of feature and whether it's present. Model can consider features' quantatitive qualities and their ocurrence across the dataset because of this unique capcity.As a result, Random Forest could be enhanced to make it a more accurate and practical tool for a variety of analytical forecasting jobs(Afonso et al.; 2019).

2.6 Deep Learning

Use Long Short-Term Memory (LSTM) neural networks to enhance weather prediction. The LSTM model was extended with extra features like pressure and meteorological variables.The dependability of climate predictions in Indonesian airport zones was later improved by using the multi-layer LSTM architecture to capture complicated patterns(Salman et al.; 2018).

As Claimed by Sridhar and Sanagavarapu (2021),state that the price of Doge Coin will be determined by a transformer machine learning model with a number of heads of attention. It goes beyond current models to take into account both short- and long-term patterns for improved forecasting.Real-time examination of the characteristics and market capitalization of cryptocurrency is the main area of interest. Our study focuses on the implementation of the Multi Head Self Attention Model.

N-BEATS has been shown to be useful for many different time-series prediction applications. It performs better than normal statistical methods and can compete with more complicated machine learning models. The architecture is a useful tool for accurate and understandable time series forecasting because of its capacity to handle single-variate time series data and detect temporal trends (Tuominen et al.; 2023).

2.7 Explainable AI

In high-stakes industries like healthcare, Explainable Artificial Intelligence (XAI) plays an important part in improving transparency and trust in advanced machine learning algorithms. Two popular XAI techniques are Shapley Additive explanations (SHAP) and Local Interpretable Model Agnostic Explanation (LIME). SHAP reviews characteristics and explains how they affect model outcomes using game theory, but the conclusions might change depending on the model. It is better for users to consider feature order instead of exact scores. Due to an expectation of feature independence, SHAP can cause problems when working with related features. can make the models easier to understand and make it easier for them to be used properly(Salih et al.; 2023).

3 Methodology

3.1 Data Collection:

The dataset, which was acquired from FRED (Federal Reserve Economic Data), includes 764 months (about 63 and a half years) of economic data for study and covers PCE, GDP, PPI, and CPI from January 1960 to August 2023. Since we only have monthly data, we can only use a certain number of rows. Gathering many datasets including economic indicators, every with a 'DATE' column serving as a date index, was the first step of the time series forecasting project. To creating single, consistent dataset, we used 'pd.merg e' function to combine two sets of data into single file named 'merged data.csv.' Data consistency is ensured by including all main economic indicators in one uniform dataset.

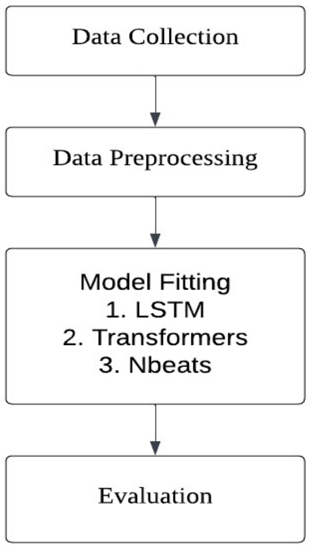

Figure 1: Methodology Flowchart

3.2 Data Preprocessing:

Time series data index was defined as the date. Data imputation was not necessary since no missing values were found. To ensure constant scaling, Min-Max was utilized for data standardization. The training of the model is at a crucial stage. The input and target sequences, usually consisting of 12-time steps, were generated using normalized data.

3.3 Model Fitting:

The LSTM, Transformer, and N-BEATS models were created specifically to predict future values into time series data. The LSTM model utilized thick output layer, dropout regularization, and many LSTM layers to effectively capture temporal correlations. The TensorFlow/Keras Transformer model captures complex temporal interactions. System uses multi-head self-attention forward-processing layers. To effectively teach N-BEATS model complex time series patterns, forecasting back cast sub- networks were used. Hyper parameters including input shape, dropout rates, and layer configurations have been adjusted for every model for optimizing architecture and training.

3.4 Evaluation:

After training, we tested every model against data calculated RMSE (Root Mean Squared Error) MAE to measure prediction accuracy. All economic indicator models were assessed for predictive power. Additionally, XAI approaches were employed to comprehend models' predictions decision-making procedures more fully. XAI methods like SHAP LIME were used to analyze the outcomes of models.

$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$

$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$

$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$

4 Design Specification

4.1 Techniques and Architecture

To gain a deeper understanding of future inflation patterns, it is crucial to analyze past economic trends. The first step acknowledges the importance of acting. The true power of time-series data can be unlocked by harnessing blend of state-of-the-art machine learning models: LSTM, NBeats, Transformer. Our predictions are more accurate as those models complement one another by addressing various aspects of time-series analysis.

4.1.1 LSTM

Recurrent neural networks like LSTM may predict data points using prior time steps MOGHAR and HAMICHE (2020). LSTMs (Long Short Term Memory) discover tiny correlations in long-term data. For storing retrieving temporal data, such networks include recurring hidden layers with different memory blocks. LSTMs are good in long-term forecasting because they can store and utilize information. Input, forget, output gates let them gather important data and remove less important aspects for accurate projections.

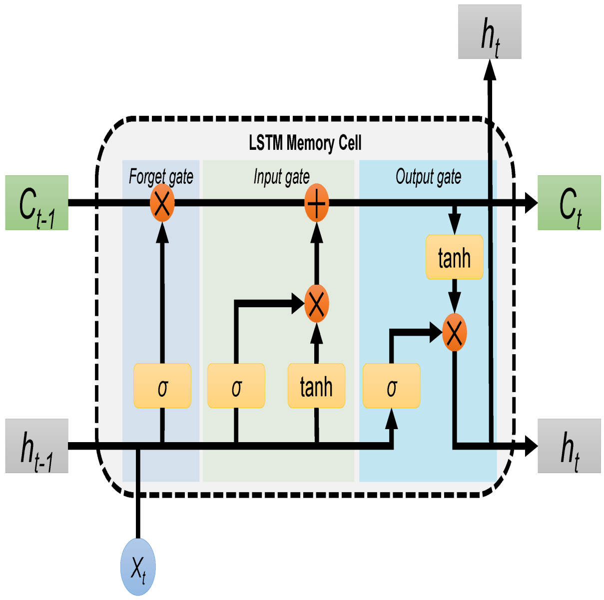

Figure 2: Architecture of LSTM Model (Alghofaili et al.; 2020)

Information is controlled by such gates (Salman et al.; 2018).

4.2 Transformer

As per Vaswani et al. (2017), Transformer model is an innovative approach which relies solely upon self-attention for representing input output. It differs from traditional sequence aligned RNNs in this way. The purpose of the study is to evaluate the model's capacity to produce predictions utilizing a variety of interrelated factors over extended periods of time. Transformers are capable of self-focus, which may include having several heads. Pay heed to identify trends in time series data, single-head attention is used to calculate weights. By using numerous weight matrices and carrying out simultaneous computations, multi-head attention improves the model's ability to identify subtle linkages within multivariate time series data. It increases overall productivity by concurrently collecting several dependencies and data points.

4.3 NBeats

A deep learning framework created specifically for time series forecasting is the N-BEATS architecture. The completely CNN uses a straightforward but effective method by include both forward and reverse residual links. N-BEATS is built using interconnected layers, or blocks, that use different trained models to generate accurate predictions. These blocks fall into two categories: trends and seasonality, which each captured distinct patterns in time series data. Because it eliminates the requirement for custom feature development tailored to time series data, the concept is incredibly flexible NBEATS's design makes it simple to comprehend and analyze learned patterns (Oreshkin et al.; 2019).

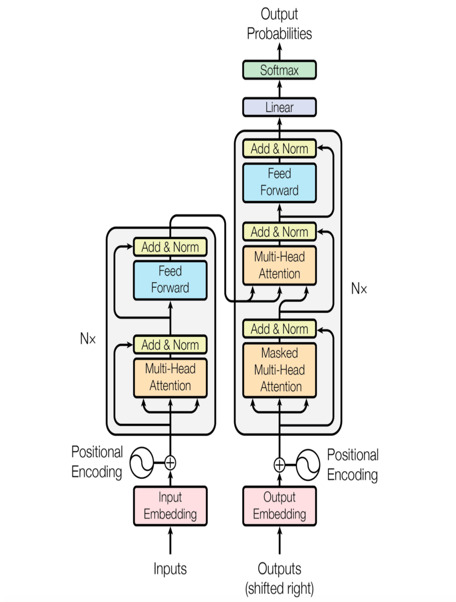

Figure 3: Transformer Architecture(Vaswani et al.; 2017)

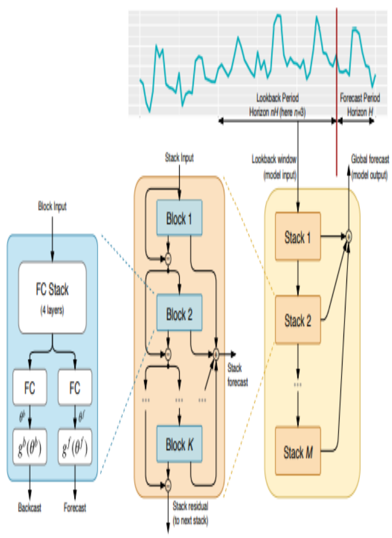

Figure 4: NBEATS Architecture(Oreshkin et al.; 2019)

4.3.1 XAI SHAP

The Robust values from gaming theory are the source of SHAP values, which are used by forecasting algorithms to assess the significance of distinctive traits. An essential part of our approach is the SHAP Kernel Explainer, which assigns these quantities in a model-agnostic manner. A thorough examination of every variable influencing a model's predictions might help one better understand how complex machine learning algorithms work inside. This procedure expands knowledge and enhances comprehension of their operation (Lundberg and Lee; 2017).

4.4 Frame Work

Using a precise and tested technique, we aggregate and provide the data for our whole inflation estimate. The panda library, a Python package that is widely used, makes it easier to combine data from multiple sources while preserving secrecy and integrity. This technique requires that the data be accurate, consistent, and complete. Using these preliminary methods lowers noise and incorrectness in the data, allowing valid predictive models.

The first step should be to develop model procedures for training and evaluation. These strategies must be effective for our estimate to be valid. As a result, we train and adjust machine learning models using previous data. Verifying the accuracy of our system's inflation prediction using actual data is crucial

Explainable AI (XAI) concepts improve system comprehension transparency. Because stakeholders in financial forecasting want to know how decisions are made, XAI is crucial in this field. By giving explanations that are both clear and easy to understand, it helps us to demystify complicated model predictions.

High-performance computing resources are crucial due to the computational demands of our sophisticated models. These models require a lot of processing power for training real- time predictions, but we make best use of their potential by using frameworks such as TensorFlow or PyTorch. That way, we know our system will be able to withstand computing economic forecasting demands remain stable

5 Implementation

5.1 Data Combining and Data Preprocessing

Federal Reserve Bank of St. Louis providedFRED (2022) database used into this research for PCE, PPI, CPI, GDP. At first, the CSV files with every dataset were imported into distinct Pandas Data Frames, with the names df1, df2, df3, df4 correspondingly. We used Pandas' merge function, using 'DATE' Askey for joining, to combine these datasets intone Data Frame. A new CSV file called" merged data.csv" was created to store combined data after merger. Added study and analysis shall be conducted using this material.

After reviewing the combined dataset, it was noted that none of the important columns—'PCE,' 'PPI,' 'CPI,' and 'GDP'—have missing values. This means that all economic indicators have comprehensive and continuous data in the collection. Data

Table 1: Summary Statistics

	PCE	**PPI**	**CPI**	**GDP**
count	764	764	764	764
mean	5604.07	117.42	136.99	99.98
std	4960.88	62.70	80.21	1.26
min	323.60	31.30	29.37	92.03
25%	1088.02	59.77	55.53	99.42
50%	3994.15	116.85	137.50	99.98
75%	9777.80	171.25	208.71	100.76
max	18726.90	280.25	306.27	102.98

handling and imputation procedures were not necessary since there were no missing values. Consequently, this dataset is flawless and well suitable for analysis and decision-making, without the usual issues caused by missing data.

5.2 Exploratory Data Analysis

For EDA, Python is the language of choice, namely the Pandas Matplotlib libraries. At outset, df.info () provides a higher-level description of Data Frame, including its structure, data types, missing values. By highlighting distributions and likely outliers, df.describe() offers statistical insights into numerical columns. In addition, line graphs are used for time series analysis to display major economic data including GDP, CPI, PCE, PPI. To show patterns trends into data to set stage for more research model creation, these visualizations using 'DATE' as index are crucial.Time series trends, distribution histograms, and comparative box plots are shown in Figures 5, 6, and 7, respectively. These visualizations offer insights into data distributions, temporal patterns, and comparative studies throughout time.

5.3 Model Architecture and Training

5.3.1 LSTM

LSTM models excel in time series forecasting. Building dataset comprising input sequences target values is integral part of LSTM technique. An important part of this method is the "look back" parameter, that regulates the number of time steps that are looked at to anticipate cur- rent one; a common value is 12. Complex temporal patterns into data may be more easily discovered using this method.

Two 50-unit layers make up a typical LSTM model for time series prediction. Time series analysis benefits from these layers' ability to process sequential information. To

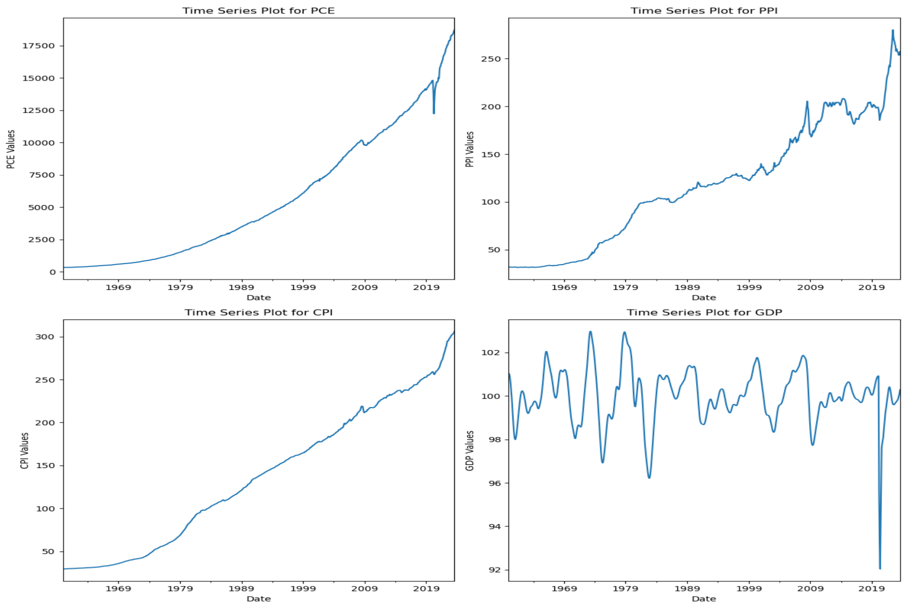

Figure 5: time series plot of PCE, PPI,GDP and CPI

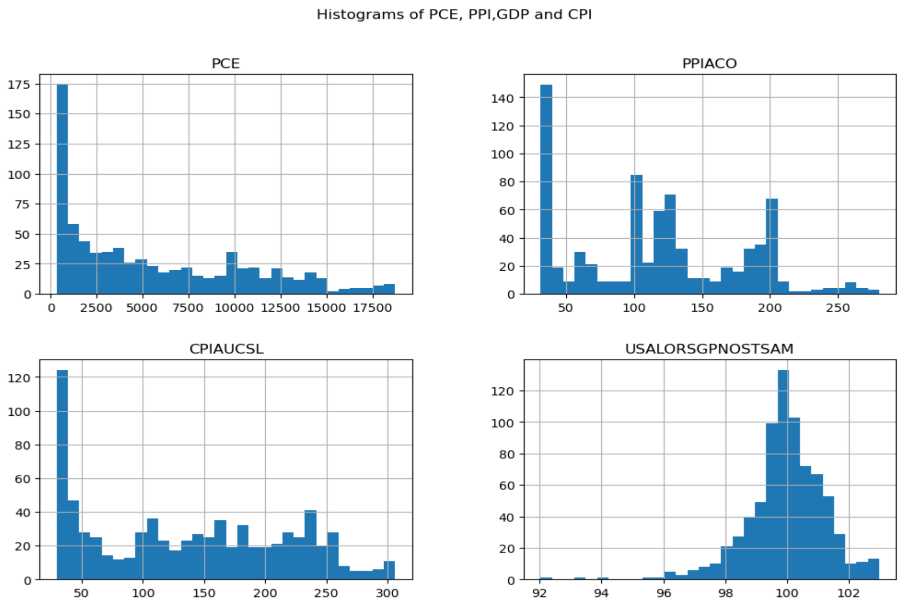

Figure 6: Histograms of PCE, PPI,GDP and CPI

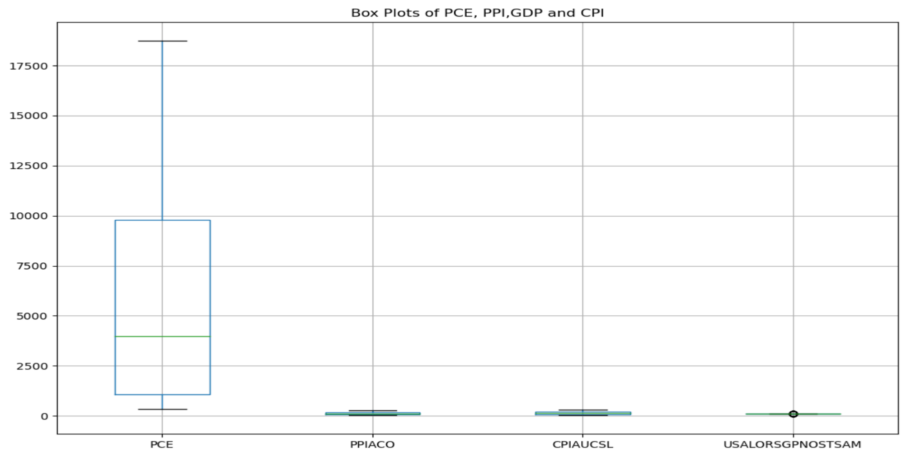

Figure 7: BOX of PCE, PPI,GDP and CPI

prevent model from overfitting and guarantee it applies well to fresh data, the model has- dropout layer having 20 dropout rates. Final predictions for output are generated by Dense layer that has been tuned to fit feature sizes of dataset. Parameters such as dropout rate, the number of units for every LSTM layer," look back" value are critical

" Adam" optimizer" mean squared error" loss function are often used by model during training. A batch size of 32 is used throughout 50 iterations of the training method. An essential aspect of training is using checkpoints for loading ideal model weights, guaranteeing peak performance. To make model's predictions on test data more precisely pertinently they are rescaled following training for matching scale of original time series data.

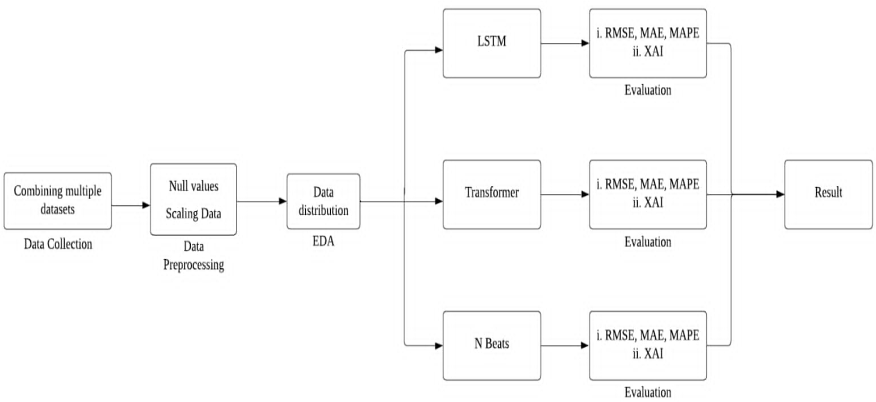

Figure 8: Implementation Architecture

5.3.2 Transformers

Sophisticated processing of sequential data by transformers has led to their increased prominence in time series forecasting. Transformers use their ability with sequential connections to build input sequences target values that go along with them using sliding window mechanism, as opposed to more conventional methods.

Time series forecasting using Transformer model involves complex and layered architecture. The process begins with an input layer and continues via many Transformer blocks. A model's ability to concentrate on many parts of a sequence simultaneously is enhanced by components of every block, like Layer Normalization, Multi-Head Self-Attention, residual connections. To avoid overfitting, which global average pooling layer is used to extract comprehensive insights by sequence, FFN with dropout is employed for training. To improve Learned characteristics, the last layer is often MLP. It ends with output layer that is intended to predict time series values.

Among the most crucial hyperparameters in Transformer models are attention heads, feed- forward dimension, dropout rate, number of blocks, MLP units. A model's ability to show and make use of temporal patterns is significantly affected by these settings. Training models take place over 50 epochs having batch size that is set utilizing Adam optimizer mean squared error loss function. While training a model, it's customary to use validation split for checking development. Time series analysis relies upon forecasted values of model, derived from the training stage then applied to the test dataset.

5.3.3 N-BEATS

N-BEATS proves superior effectiveness in predicting time series due to its unique structure approach. N-BEATS use a 12-step" look back" window, comparable with LSTM, for analyzing input sequences. Although, it does this via succession of densely linked layers, allowing it for capturing complex temporal trends.

The capability of N-BEATS to decode input sequences by tightly interconnected layers is essential to its architecture. Key outputs of model, known as 'Theta Back cast' and 'Theta Forecast,' are findings obtained from analysis of temporal data performed by each layer.For ensuring precise forecasts, it is crucial to understand data patterns in both historical future contexts. The modular design of N-BEATS makes it highly suitable for a wide range of forecasting activities and allows it to easily adjust to your specific requirements.

Distinguishing factors of N-BEATS are its hyperparameters, like input shape, Theta size, number of blocks, neurons, and layers. Hyperparameters play a critical role in enabling-model forrecognizing7 exploiting dataset trends over time. The model is trained utilizing Adam optimizer and the MSE (Mean Squared Error) loss function for 50 epochs. The batch sizing is set to 32 and there are 10 validations split. To make sure expected values are in line with the real data for the best analysis and interpretation, the model's predictions upon test data are changed back to original scale following training. With this rescaling in place, N-BEATS becomes a strong and flexible tool for time series forecasting, which is essential for model's forecasts to have any practical utility.

6 Evaluation

6.1 Observation 1 : Evaluation of LSTM

Accuracy of LSTM model's predictions is figured out by comparing its efficacy on training test datasets. To quantify prediction errors, assessment metrics like RMSE, MAE, MAPE are used; lower values indicate better performance. According totable," GDP" is the top feature. It has the lowest RMSE (1.02), MAE (0.43), MAPE (0.44), indicating that it makes the most precise forecasts out of all features that were assessed.

Table 2: Evaluation Metrics for Different Features

Feature	**RMSE**	**MAE**	**MAPE (%)**
PCE	652.53	465.22	3.06
PPI	21.71	16.99	7.61
CPI	6.12	5.28	2.04
GDP	1.02	0.43	0.44

Figure 9: LSTM

6.2 Observation 2: Evaluation of Transformers

Following the training transformer model, its prediction abilities were tested using a test dataset. Important evaluation metrics for determining accuracy of model are included in the table, including RMSE, MAE, MAPE, for number of attributes. Accurate consistent predictions were made by feature" GDP," which had greatest results in terms of RMSE (3.92), MAE (3.49), MAPE (3.51). This means the previous model generated very precise

Table 3: Evaluation Metrics for Different Features

Feature	**RMSE**	**MAE**	**MAPE (%)**
PCE	4156.92	3912.59	28.27
PPI	32.30	27.42	12.89
CPI	39.16	30.41	11.61
GDP	3.92	3.49	3.51

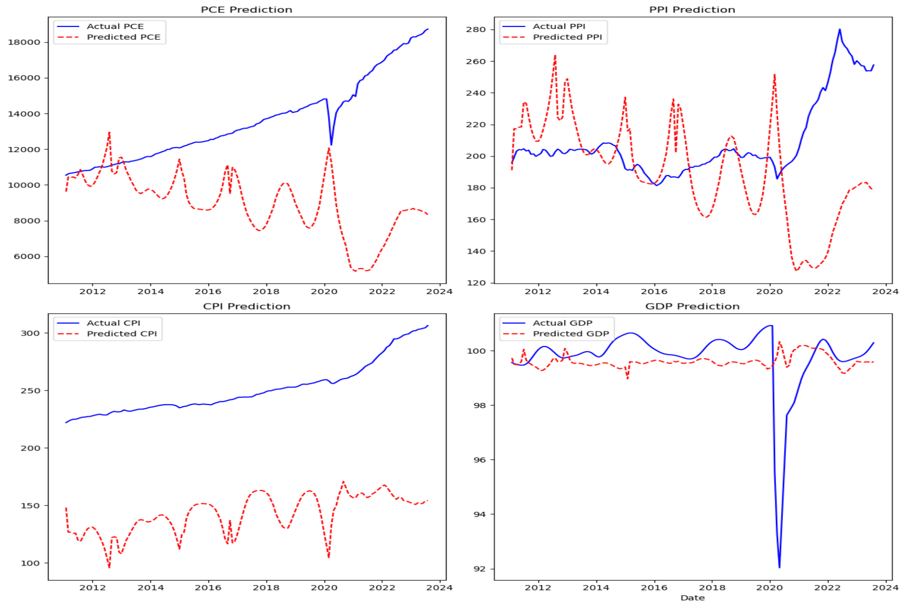

Figure 10: Transformers

predictions with little room for mistakes whenever this feature was utilized.

6.3 Observation 3: Evaluation of NBEATS

A deep learning network called an N-Beats model was utilized to assess accuracy of time series forecasts. RMSE, MAE, MAPE were metrics used for this purpose. down order to lower these indicators, study zeroed down on number of attributes. Having lowest RMSE (0.78), MAE (0.35), MAPE (0.36)," GDP" fared better than others. Consequently," GDP" proved to be most valuable characteristic for-Beats model when it came to making precise predictions which deviated least by actual values.

Table 4: Evaluation Metrics for Different Features (Including NBeats)

Feature	**RMSE**	**MAE**	**MAPE (%)**
PCE	545.16	417.72	3.12
PPI	12.00	10.30	5.02
CPI	13.27	12.11	4.75
GDP	0.78	0.35	0.36

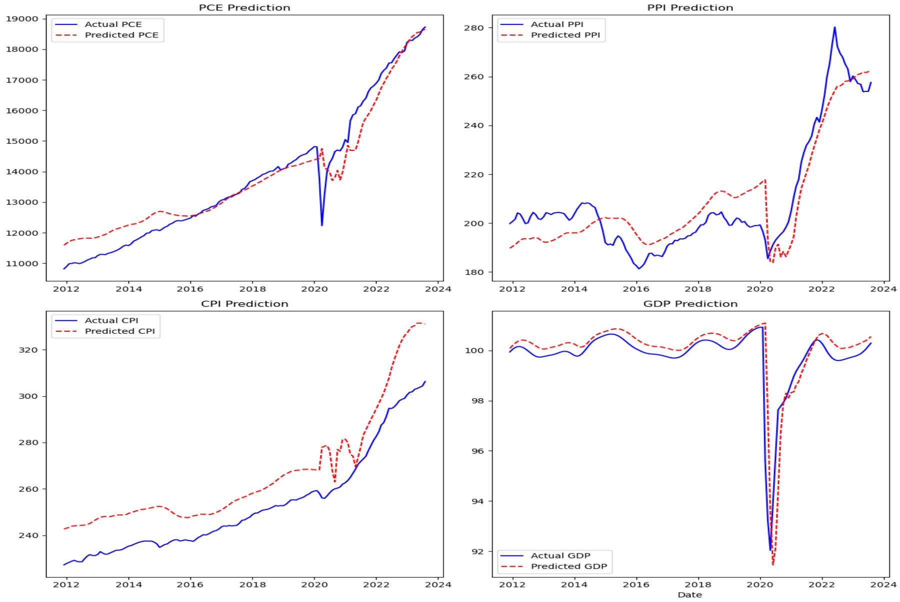

Figure 11: NBEATS Evaluation

6.4 Observation 4: Evaluation Using XAI

SHAP KernelExplainer provides explanations of model's results after its application to time series forecasting using machine learning model. First phase of process is too transform3Dtime series data in two-dimensional format that meets requirements of SHAP. A The custom model prediction function is used to restore the original 3D format of the re-shaped data to ensure compliance with the model. To measure the impact of every attribute upon predictions, SHAP's KernelExplainer generates SHAP values utilizing some of modified training data. By integrating SHAP's interpretability with deep learning's time series analytical complexity, this technique improves model transparency by exposing feature contributions combining two.

6.4.1 LSTM XAI

Feature GDP exhibits considerably higher impact upon model forecasts than others, according to SHAP analysis of LSTM model for time series forecasting. Mean SHAP value for this feature is 0.020396. Even if there are slight differences in impact among characteristics, GDP has the largest impact into this specific model. Table and figure results are displayed.

Feature	Mean SHAP Value
PCE	0.020270
PPI	0.020317
CPI	0.020353
GDP	0.020396

Table 5: Mean SHAP Values for LSTM

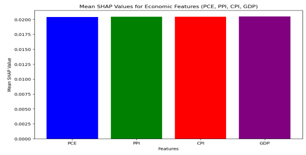

Figure 12: SHAP Values for LSTM

6.4.2 Transformer XAI

With an average SHAP value of 0.01433128713367212, feature GDP is having a larger impact on model's predictions compared to other features of the transformer model used in time series forecasting. Even though there is lot of rivalry amongst characteristics, GDP stands out as most important in this transformer model, even if changes in SHAP values are minimal. Table and figure results are displayed.

Feature	Mean SHAP Value
PCE	0.014183
PPI	0.014033
CPI	0.014247
GDP	0.014331

Table 6: Mean SHAP in Transformer Model

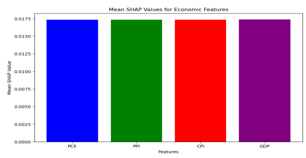

Figure 13: SHAP in Transformer

6.4.3 N-BEATS XAI

Having an average SHAP value of 0.07153067508567916, feature GDP is most predictive of model's outputs compared to other attributes. In this N-BEATS model arrangement, GDP has a much larger influence, but other features contribute equally due to tiny changes into SHAP values. Table and figure results are displayed.

7 Conclusion and Future Work

In this research, The N-BEATS model is the best performer among the three models (LSTM, Transformers, and N-BEATS) used in this examination for time series forecasting. It constantly obtained the lowest values for RMSE, MAE, and MAPE for all characteristics, with the" GDP" feature being the most significant. The powerful performance of the N-BEATS model was further validated by the XAI analysis utilizing SHAP values. Research into the future should aim to determine what exactly makes" GDP" such a powerful tool for prediction. More accurate robust forecasting tools for different time series applications might be result of investigating its relationships with other variables while employing these findings to improve model designs.

Feature	Mean SHAP Value
PCE	0.071159
PPI	0.071134
CPI	0.071372
GDP	0.071531

Table 7: Mean SHAP Values for N-BEATS Model

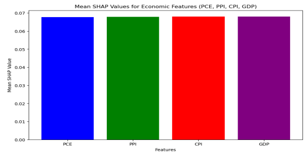

Figure 14: SHAP Values for N-BEATS

References

(2021).
URL: *https://www.studysmarter.co.uk/explanations/macroeconomics/financial-sector/fisher-effect/*

Afonso, B., Melo, L., Oliveira, W., Sousa, S. and Berton, L. (2019). Housing prices prediction with a deep learning and random forest ensemble, *Proceedings of 16th National Meeting on Artificial and Computational Intelligence*, SBC, Porto Alegre, RS, Brasil, pp. 389–400.
URL: *https://sol.sbc.org.br/index.php/eniac/article/view/9300*

Alghofaili, Y., Albattah, A. and Rassam, M. (2020). A financial fraud detection model based on lstm deep learning technique a financial fraud detection model based on lstm deep learning technique, *Journal of Applied Security Research* **15**.

Aras, S. and Lisboa, P. (2022). Explainable inflation forecasts by machine learning models, *Expert Systems with Applications* **207**: 117982.

Banerjee, D. (2014). Forecasting of indian stock market using time-series arima model, *2014 2nd International Conference on Business and Information Management (ICBIM)* pp. 131–135.
URL: *https://api.semanticscholar.org/CorpusID:2495102*

Bentour, E. M. (2015). A ranking of var and structural models in forecasting.
URL: *https://mpra.ub.uni-muenchen.de/61502/*

Breiman, L. (2001). Random forests, *Machine learning* **45**(1): 5–32.

Chen, S. and Ranciere, R. (2017). Financial information and macroeconomic forecasts.
URL: $https://papers.ssrn.com/sol3/papers.cfm?abstract_id = 2924356$

Dhakal, C. (2018). A naïve approach for comparing a forecast model.

Elsiddig, I. and Mohamed, E. (2015). Multiple linear regression model of inflation rate in sudan, *African Journal of Social Sciences* **5**: 81–91.

FRED, S. L. (2022). Welcome to fred, your trusted source for economic data since 1991.
URL: *https://fred.stlouisfed.org/*

Jeon, J. (2015). The strengths and limitations of the statistical modeling of complex social phenomenon: Focusing on sem, path analysis, or multiple regression models.
URL: *https://zenodo.org/record/1105869*

Kelikume, I. and Salami, A. (2014). Time series modeling and forecasting inflation: Evidence from nigeria.
URL: $https://papers.ssrn.com/sol3/papers.cfm?abstract_id = 2322918$

Kenny, G., Meyler, A. and Quinn, T. (1998). Forecasting Irish inflation using ARIMA models, *Research Technical Papers 3/RT/98*, Central Bank of Ireland.
URL: *https://ideas.repec.org/p/cbi/wpaper/3-rt-98.html*

Lundberg, S. M. and Lee, S. (2017). A unified approach to interpreting model predictions, *CoRR* **abs/1705.07874**.
URL: *http://arxiv.org/abs/1705.07874*

MOGHAR, A. and HAMICHE, M. (2020). Stock market prediction using lstm recurrent neural network.
URL: *https://www.sciencedirect.com/science/article/pii/S1877050920304865*

Moser, G., Rumler, F. and Scharler, J. (2006). Forecasting austrian inflation.
URL: *https://www.sciencedirect.com/science/article/pii/S0264999306001222*

Oner, C. (2019). Inflation: Prices on the rise.
URL: *https://www.imf.org/en/Publications/fandd/issues/Series/Back-to-Basics/Inflation*

Oreshkin, B. N., Carpov, D., Chapados, N. and Bengio, Y. (2019). N-BEATS: neural basis expansion analysis for interpretable time series forecasting, *CoRR* **abs/1905.10437**.
URL: *http://arxiv.org/abs/1905.10437*

Pettinger, T., yusuf, A., Tuyenikeloa, Milano, CES), a. g. M., Kutsy, Y., Anuar, A. I., Waqas, M., Mr.Git, Babayo, B. and et al. (2023). Phillips curve.
URL: *https://www.economicshelp.org/blog/1364/economics/phillips-curve-explained/*

Salih, A., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Menegaz, G. and Lekadir, K. (2023). Commentary on explainable artificial intelligence methods: Shap and lime.

Salman, A., Heryadi, Y., Abdurahman, E. and Suparta, W. (2018). Single layer multi-layer long short-term memory (lstm) model with intermediate variables for weather forecasting, *Procedia Computer Science* **135**: 89–98.

Sridhar, S. and Sanagavarapu, S. (2021). Multi-head self-attention transformer for dogecoin price prediction, *2021 14th International Conference on Human System Interaction (HSI)*, pp. 1–6.

Stock, J. H. and Watson, M. W. (2008). Phillips curve inflation forecasts.
URL: *https://www.nber.org/papers/w14322*

Thomas, L. B. (1999). Survey measures of expected u.s. inflation, *Journal of Economic Perspectives* **13**(4): 125–144.
URL: *https://www.aeaweb.org/articles?id=10.1257/jep.13.4.125*

Tuominen, J., Pulkkinen, E., Peltonen, J., Kanniainen, J., Oksala, N., Palomäki, A. and Roine, A. (2023). Forecasting emergency department crowding with advanced machine learning models and multivariable input.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. and Polosukhin, I. (2017). Attention is all you need, *CoRR* **abs/1706.03762**.
URL: *http://arxiv.org/abs/1706.03762*

Zimmermann, C. (2015). Federal reserve economic data: Your trusted data source since 1991.
URL: *https://fredblog.stlouisfed.org/2015/03/the-many-flavors-of-inflation/*