Forecasting inflation rates in Turkey with Linear Regression, SARIMA, and LSTM

Supervisor: Heikki Laiho

Can Ayan - 10622664@mydbs.ie – Dublin Business School

I declare that this dissertation that I have submitted to Dublin Business School for the award of Financial Analytics is the result of my own investigations, except where otherwise stated, where it is clearly acknowledged by references. Furthermore, this work has not been submitted for any other degree.

Signed: Can Ayan
Student Number: 10622664
Date:28/08/2023

Abstract

Turkey, categorized under emerging economies, has seen fluctuations in its inflation rates, which have recently been notably high. This prompts the need for robust estimation models that can accurately predict inflationary trends. This study aims to contribute to the existing literature by testing and comparing three distinct inflation forecasting models; multilinear regression, SARIMA, and Long Short-Term Memory (LSTM) in the context of Turkey between 2004 and 2023. By contrasting these models with the CBRT's median market participant survey, and using Root Mean Square Error (RMSE) for model evaluation, the study seeks to identify the most accurate model for predicting inflation in Turkey. The outputs of these study shows that usage of SARIMA and LSTM models together outperforms than the individual models and the benchmark survey. Individual level, SARIMA were performed better to capture extreme fluctuations in time series than others.

Keywords

Consumer Price Index (CPI), Forecasting, Regression, Machine Learning in Finance, Econometrics, Deep Learning, SARIMA, LSTM, Time Series

Contents

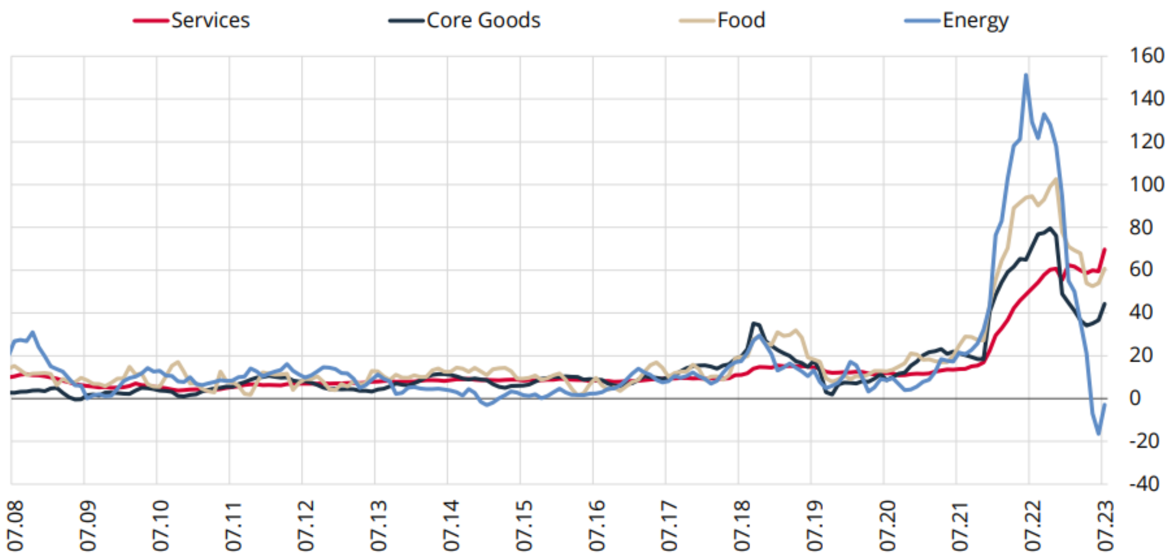
Introduction

Inflation, which is the basis of monetary policy, is one of the important macroeconomic indicators taken into consideration in a country. The level of inflation carries insights about the health of the economy and monetary policy is designed accordingly. The relationship of inflation with other economic indicators makes it important to both framing and analysis. Basically Inflation is the increase in the prices of goods and services over time (FED). In cases where inflation is higher than the desired level, purchasing power decreases. Hyperinflation refers to a situation where the prices of goods and services rise uncontrollably over a defined period of time (WEF). Hyperinflation can be associated with an economic crisis, and therefore monetary policy designs measures to counter rising inflation. This situation has also manifested itself in the example of Turkey, where inflation rates have increased recently, and the path of disinflation was announced by the CBRT (Inflation Report, July 2023). This situation increases the importance of estimating the data for the correct design of the implemented policies. In this direction, the CBRT also shares its recent inflation forecasts with the public.

Turkey, which is included in the category of emerging economies, has been experiencing higher inflation rates in recent years compared to the developments since January 2003, when the current data were prepared. When the literature on Turkish inflation is examined, the 1951-2015 range can be divided into three subgroups (TATLIYER, 2016). Accordingly, the average inflation rate in the block between 1051 and 1970 was 7.4% annually, shifting to higher levels after 1970. Especially in the 1990-1999 period, much higher inflation rates were observed with an average of 77.2% per year compared to the previous period. While inflation was 8.3% on average between 2004-2015, it increased to an average of 27.1% between 2018-2023.

Figure 1 : Turkey Inflation YoY for CPI and Subgroups (CBRT)

(Annual % Change)



As can be seen in the chart above, especially after 2021, annual inflation reached extreme values, including sub-groups, and showed high volatility. This situation necessitated a re-examination of the effects in time series estimation and created a unique situation for testing the success of estimation models. With the capture of these effect, there is room in the literature for retesting prediction models in the future. The necessity of a good estimation model, the necessity of re-adjusting the hyperparameters for the extraordinary period, possible changes in the factors affecting the inflation form the importance of estimating the inflation in Turkey with different models in this study. Accordingly, three estimation models that can be classified in different clusters in line with the literature search were applied on the data set. In order to evaluate the success of the forecasts produced by these models, the current inflation forecast data in the public are also used.

Rationale

The fact that more accurate predictions are made than the predictions produced by classical methodologies with the techniques based on machine learning that have been developed recently, has led to this study apply for renewed dataset by using Turkey inflation rates. However, Turkish inflation differs from the time series used in other studies in terms of its recent dynamics. In this way, it is aimed to contribute to the literature by testing the success of the techniques used in the newly developing inflationary environment. Accordingly, multilinear regression, which is a linear model, SARIMA, which is a non-linear model, and Long Short Term Memory (LSTM), which is a deep learning algorithm, were tested to compare the accuracy rates and the ability to capture shifts. Studies conducted in the literature review have shown that the success of the models varies according to many parameters, and this study aims to suggest additional results and parameter settings to the literature for future studies.

Literature Review

There are several articles and studies that focussed on forecasting inflation rates by using different methodologies and datasets. To conduct a literature review purposes, some articles with different aspects covered such as region, independent variables, frequency of data, modelling methodologies and evaluation metric's selection were examined. In addition, Central Bank of Turkey (CBRT) publishes inflation report quarterly and shares the Bank's official estimations based on a given econometric model for the following 12 quarters by using quarterly data. Also, after evaluate several methods, CBRT's recent inflation forecasts may be used as another benchmark. This literature review aims to gather decent information for building an econometric forecast model and a deep learning model. In this domain, this section will be extended by using articles that focus forecasting inflation rates with different datasets and methods. Because of the nature of this macroeconomic variable, It is possible to get benefit from other researches which focus on inflation rate for other countries.

The inflation report published by the CBRT consists of the general assessment economic outlook and medium-term projections. The section on the determinants of inflation in the report also provides information on the latest drivers of inflation in Turkey. This section begins with bold text, referring to the significant loss of value experienced in the Turkish lira. This situation indicates that the exchange rate should be used to make an inflation forecast based on the literature review. However, another point to be touched upon was inflation expectations. Another factor examined is commodity prices, while natural gas and oil prices affect inflation in Turkey, where the country is a net energy importer. Another issue included in the study is the results obtained by examining inflation in two separate groups, services and goods. Accordingly, while services inflation is more sensitive to internal factors, especially wages, goods inflation is more sensitive to exchange rate and external factors with the effect of import prices. On the other hand, while both groups are affected by demand, it is stated that the sensitivity to credit in the core goods group is higher. From this point of view, it can be deduced that the interest rates of consumer loans have a demand-side effect on inflation. In inflation indicators that are more sensitive to wages, it is seen that inertia is more prominent. As a result of a study conducted with the Bayesian VAR model in this report, a nominal increase of 1% in the minimum wage has an effect between 0.08% and 0.12% on inflation. Due to the fact that minimum wage increases are made in the same period of the year and are determined mostly according to past inflation data, both seasonality and auto regression analysis were the important component of inflation analysis. According to the result obtained from the CBRT's inflation forecast update and sources table in the inflation report, the output gap, food prices, exchange rate and oil prices are used as explanatory variables in the inflation forecast.

Literature on Turkey's Inflation Forecasting

One of the most recent studies on this subject published by Akbulut in 2022 to forecast Turkey's inflation rate. In this study, Vector Auto-Regression (VAR) model is used as a benchmark while some machine learning technics are conducted to compare results in a perspective of accuracy with root mean square errors (RMSE) and R-square metrics. Lasso (Least Absolute Shrinkage and Selection Operator) and Ridge regression models are used as a linear machine learning models where the neural network model is used as a nonlinear machine learning model. Also, It is seen that Augmented Dickey-Fuller (ADF) hypothesis test is used to track stationary condition of the time series. The chosen dataset for benchmark model includes Turkish Lira currency, oil prices, money supply and BIST100 index. Because of the feature selection nature of nonlinear algorithms, several other data is used for machine learning models. According to key findings, nonlinear methods produce more accurate prediction than the linear benchmark model and other linear machine learning models. Thus, usage of nonlinear machine learning algorithms for forecasting Turkish CPI is suggested as a complementary tool.

In 2018, another article published by Kucukefe for aiming to produce more accurate inflation forecasts for Turkey than the Central Bank of Turkey's official forecast. In this study, several machine learning methods are conducted such as Bayesian Ridge Regression, Kernel Ridge Regression, Random Forests Regression, and Support Vector Machines. Similar to Akbulut, RMSE is selected as a evaluation metric criteria. The key findings shows that machine learning technics produced more accurate forecasts than the CBRT's official forecast.

To extend data understanding on Turkey's inflation forecasting, It is useful to get benefit from different structured articles which focuses on an efficacy of a forecasting model rather then making a comparison. Saz (2011) published an article to examine efficacy of SARIMA models to forecast Turkey's inflation. This study also illuminates on information for the data such as stationary and seasonality effects. The timespan of study contains the period from 2003 to 2009 and results suggest for a single best SARIMA model. According to evaluation metrics such as mean error (ME), mean absolute error (MAE), RMSE and Theil's U, SARIMA (0,0,0)(1,1,1) which is a seasonal autoregressive moving average model is suggested as a conclusive.

Aras and Lisboa (2022), published an article about explainable inflation forecast models by machine learning. This study uses tree-based models to analyse data for Turkey. It is preferred that Autoregression (AR) model as a benchmark while using wide set of machine learning technics such as Lasso, Ridge, Lars, Elastic Net Regression, Orthogonal Matching Pursuit, Recursive Feature Elimination with Random Forest, Boruta, Random Forest, Extremely randomized trees, Adaboost, GBDT, XGBoost. Also for the feature selection, It is seen that wide selection of economic data is used as independent variables such as Economic Confidence Index, Consumer Confidence Index, Reel Sector Confidence Index, Industrial Production Index, Unemployment Rate, Core Inflation, BIST-100 Index, Capacity Utilisation Rate, Commercial Credit Interest Rates for TL, Commercial Credit Interest Rates for USD Dollar, Consumer Credit Interest Rates for TL, Euro/Dollar Parity, Dollar Exchange Rate, Interest Rates for Deposits in Turkish Lira, Interest Rates for USD Dollar Deposits, Bullion Gold Selling Price (TRY/Gr), Current Account, Currency in circulation, CPI Based Real Effective Exchange Rate, 12-months-ahead CPI survey of expectation (mean), Commodity Price, Consumer Confidence Indicator, Industrial Confidence Indicator, Production in Industry, S&P 500 Index, VIX Index. Some of these methods are also significantly successful to track non-linear relationships between variable which deliver an advantage for feature selection by using wide range of dataset. The variables include monthly observations from January 2007 to August 2021. To compare result, RMSE and mean absolute error (MAE) is preferred. In this study sum of 58 different forecasting models has been evaluated. After running these models, It has been seen that for 1 month ahead forecasting horizon CPI based real effective exchange rate without lag, dollar exchange rate with one lag, logarithm of the core inflation price index with 1 lag are the top contributor or the most significant explainers of headline inflation. Their finding shows that tree-based ensemble models can be advantageous by providing better accuracy together with explainable predictions.

Literature on Inflation Forecasting

Comparison of different frameworks to make a macroeconomic forecast has a literature background decades ago while reaching time series data is another challenge to implement these frameworks. Stock and Watson (1998) made a comparison between 49 univariate forecasting methods with 215 monthly US macroeconomic time series data for a wide time horizon between 1959 and 1996. In this research, four methods are used where these models are autoregressions, exponential smoothing, artificial neural networks, and smooth transition autoregressions. The study uses mean square errors as a evaluation criteria and the findings show that the overall best performance is produced from autoregression with unit root pretests but the performance can be increased by combining other models.

Seeking for the most accurate model for inflation forecast will always be a challenge across many different methodologies. Theoharidis, Guillen and Lopes suggest an hybrid model which is merged version of Variational Autoencoders and Convolutional Long-Short Term Memory Networks (VAE-ConvLSTM) to forecast US inflation. They use US inflation data as a monthly time series from 1978 to 2019 and compare the proposed model with several benchmark models which includes some econometric and machine learning based models. This study stands out from a practical point of view in terms of comparing 25 different methods.

One of the recent studies for univariate inflation forecasts, shows mixed signals for accuracy. In 2022, Anna Almosova conducted a study to forecast monthly US CPI inflation in univariate form. The outputs of the study shows that LSTM model outperforming when It compared to autoregressive model, neural network and Markov-switcing models. On the other hand LSTM’s performance is similar to seasonal autoregressive model SARIMA. Another important output from this study is about how difficult it is to set up a proper neural network model. Across of many hyperparameters for neural network models, It seems only a few of them meaningfully affect the performance on inflation data. Study shows that the neural network model performs better when Bayesian information criterion (BIC) is used for the lag length selection. Thus, to forecast inflation the researcher should consider hidden units number and sufficient amount of time for training.

Another recent study from Yang and Guo (2021) aims to apply a method based on GRU-RNN for inflation forecasting and uses AR and VAR models as benchmark. In this study 11 independent variables is chosen to forecast China’s consumer prices which are year-on-year growth of narrow (M1) and broad (M2) money supply, weekly interbank lending rate, industrial and retail sales of consumer goods growth rates in year-on-year basis, national housing boom index (HINDEX), the cumulative growth rate of real estate development investment and commercial housing index in year-on-year basis, Shanghai and Shenzhen 300 index, exchange rate of RMB against USD and CPI index. The output of the study shows that GRU-RNN method has significant advantages in sequence prediction for China’s inflation rate.

The research paper by Livia Paranhos introduces a compelling approach to inflation forecasting through neural network models, with a specific emphasis on the LSTM models. The LSTM model is praised for its ability to capture long-term trends where the model enables to make better predictions for an extended sequence of past observations into its forecasts. In comparison to traditional feed-forward neural networks and standard linear benchmarks, LSTMs were found to perform substantially better, particularly for long-horizon forecasts. The study presents an empirical analysis using monthly US data and compares different neural network specifications. This comparison extends to models that consider solely inflation data, those that consider a pool of economic predictors excluding CPI data, and finally, models that combine both. This stratified approach helps to isolate the effect of various economic indicators on the forecast, thereby enriching the robustness of the study's findings. The work also were get benefit from using deep learning algorithms to gather valuable insights into the behaviour of inflation during periods of economic uncertainty, particularly during and after the Great Recession. The neural network models that included macroeconomic variables were found to accurately capture the fluctuations of inflation during these tumultuous periods, suggesting a possible role for nonlinearities and the importance of macroeconomic information. This study is important by analysing LSTM model in sample which were confirmed a good performance at capturing the business cycle dynamics. Furthermore, the study found that output, income, and consumption data were significant predictors of inflation. Additionally, the study goes beyond merely presenting a predictive model; it also rationalizes the advantages of using LSTM networks in terms of both performance and efficiency in parameter estimation.

Similar to Turkey, Brazil market has also experience on higher inflation rates and currency volatility. Another research which is conducted for Brazil is also important in view of literature review. Machine learning methods for inflation forecasting in Brazil: new contenders versus classical models is the name of the article which is published by Araujo and Gaglianone in 2020. At this study, they are comparing traditional econometric models with machine learning models. Their findings show that nonlinear automated procedures especially random forest model is outperforming comparing to traditional forecasting methods. Their measure for performance is mean-squared error. As a result the finding of this study suggests the use of machine learning methods to gather better accuracy for forecasting inflation in Brazil.

Haider and Hanif (2009) published an article on forecasting Pakistan inflation with neural network methods. Using RMSE as an evaluation metrics, the findings of this study were shown that using artificial neural network (ANN) methods outperforms when It is compared with autoregressive (AR) or autoregressive integrated moving average (ARIMA) models. In this study, univariate forecast neural network model was used. In this study, the number of layers was determined as 12. RMSE was used as forecast evaluation metric. Annual headline inflation data from July 1993 to June 2006 was used with monthly frequency.

Another study were conducted by Farah Fauziah which contributes an important extension to the existing literature on inflation forecasting by focusing on the Indonesian inflation data. This research is particularly relevant given Indonesia's adoption of an inflation targeting system and the socio-economic implications of inflation on poverty alleviation in developing countries. Long Short-Term Memory (LSTM) model were utilized by addressing the limitations of RNNs such as the vanishing and exploding gradient problems in this article. 27 different models to identify the optimal LSTM architecture were tested to decide how to tune hyperparameters such as learning rate, number of nodes in the hidden layer, and the optimizer based on the smallest root mean square error value. The selected model had 5 nodes in the hidden layer, used the Adam optimizer, and had a learning rate of 0.01, achieving a minimal RMSE of 0.24 after training for 55 epochs which gave insight on application of LSTM model. The researchers argue that the selection of the LSTM model is well-suited to capture these long-term trends in the data, a feature particularly crucial for understanding the persistent nature of inflation in developing countries like Indonesia. Moreover, the study offers timely insights into inflation trends amid the ongoing COVID-19 pandemic. It observes that Indonesia's inflation rates have remained stable and below 2% since the pandemic's onset in early 2020. Forecasting results predict a continuation of this trend until at least September 2022, categorizing it as "mild inflation." However, the research also acknowledges limitations in extreme data situations, as the LSTM model couldn't precisely predict the inflation spikes, such as the one in November 2005. This speaks to the need for continued refinement of machine learning models for economic forecasting, especially in contexts with high volatility and exogenous shocks like pandemics which is a relevant case when Turkey’s CPI data deviation in recent years considered. The study shows the LSTM model's ability to capture long-term dependencies and adapt to variable economic conditions, making it a suitable tool for modelling and forecasting inflation.

For Indonesia inflation data another study published in 2023 by using LSTM model as well (Diksa, Kuswanto,Fithriasari). Recognizing inflation's nonlinear and dynamic nature, the study demonstrates that LSTMs offer more accurate predictions than traditional models. Study suggests that the number of hidden neurons and epochs used in the LSTM model must be optimized to prevent overfitting. Specifically, using too many hidden neurons or epochs leads to poor generalization to new data and poor accuracy score for the test data. The study suggests that LSTM-based forecasting models could be invaluable tools for economic policy-making in Indonesia, particularly in guiding monetary policy to manage inflation effectively.

Another research on inflation forecasting is published by Adolfo Rodríguez-Vargas which get leverage from machine learning methods for inflation forecasting in Costa Rica. This study uses evaluation of five machine learning techniques, which are two variants of k-nearest neighbors (KNN), random forests, extreme gradient boosting, and LSTM to assess their predictive capabilities. The results outperform when It compared with the univariate methods traditionally employed by the Central Bank of Costa Rica. Study applies a selection of statistical tests and criteria, including those outlined by Diebold and López (1996) and Cumby and Huizinga, to evaluate forecast errors. Notably, the LSTM model emerges as the best performer across the board in terms of accuracy for varying forecast horizons. LSTM model also meets the desirable properties of optimal forecasts, including a lack of undesirable correlation patterns in its forecast errors. This makes LSTM a potent tool for long-term policy planning, especially given its superiority over traditional methods. Univariate KNN also performed well, particularly in the short-term forecasting horizon. Random forests, while not outperforming LSTM, were notable for their ability to correctly predict the direction of inflation changes in at least 60% of the cases. The worst performer among the tested methods was KNN with exogenous variables, indicating that simplicity might sometimes be more effective than complexity in forecasting models. This paper gives an evidence to support for the adoption of machine learning methods in inflation forecasting.

Literature on Internal and External Factors of Inflation

Before deciding a model to forecast inflation It is important to understand data and the components in line with theory. There are several layers to focus on which factors affect inflation and how. It is also may be examined in Turkey level or in general terms of inflation. This review plays a crucial role in developing business understanding. The affect may be divided by two as internal and external while it is also possible in a form of demand-pull or cost-push (McKinsey, 2022).

Early literature shows a significant relationship between money supply and inflation (Gordon, 2021) as a strongest component of US inflation. Aksi Granger causality test shows that lagged exchange rate influence inflation but lagged inflation does not cause exchange rate changes. Also labour market tightness affects inflation by wages where the minimum wage is an important aspect of expected inflation. These components may be categorized as internal factors for US inflation data. According to conducted research for eight non-eurozone EU member states' inflation drivers, short run inflation dynamics are mainly influenced by domestic shocks (Globan, 2014).

For external shock oil and global food prices are most common drivers. There is a linkage between inflation and oil prices. Direct inflationary effects of oil price shocks are more sizeable for developed countries than emerging economies meanwhile the food prices increases have more significant impacts on emerging economy inflations (Galesi and Lombardi, 2009).

In an article published by the CBRT in 2016, which factors have been decisive in Turkey's inflation dynamics in the last 10 years were examined. Accordingly, between 2006 and 2015, CPI remained at a relatively high level, averaging 8.2% on an annual basis. It is desired to analyze how much inflation is affected by which factors. Accordingly, rigidity in inflation, unprocessed food, exchange rate, import prices, output gap, real unit wages, taxes and other factors are analyzed according to how much they affect inflation. Accordingly, 3% of inflation in this process is due to rigidity, in other word inertia in inflation. This shows the importance of examining both the expectation channel and autoregressive dynamics in inflation. However, another important effect comes from the prices of unprocessed food. This situation reveals the importance of detecting seasonal effects on the data. Another important variable is the exchange rate (USD/TRY) effect. Accordingly, approximately 2% of inflation stemmed from the exchange rate. Another striking point in the study is that although there was no significant change in average inflation in two separate 5-year periods, there were changes in the factors affecting inflation. This makes the need for update visible for the feature selection process.

In 2017, another article was published in the same journal under the leadership of Ali Hakan Kara. This article focused on the exchange rate transition to inflation. Accordingly, the transition effect from exchange rate to inflation varies according to seasonal factors. This study also creates a good literature on which level we can expect the dimensions of the relationship between exchange rate and inflation in a model to be established. While the exchange rate passivity can reach 25 percent in overheating periods, it may fall below 10 percent in cooling periods.

Literature on Time Series Analysis and Forecasting

Forecasting has a wide range of methodologies and classifications when you consider many factors of your data. Different type of information requires different approaches to reach proper outcomes within line practitioner’s objective. Beginning to explain notions from general to more specific would be a good starting point to understand whole picture and why this study aims to perform aforementioned practices to solve problematic.

First of all, It is essential to understand data which is also the first part of CRISP DM methodology to decide correct forecast type. Data may be divided by two main groups as quantitative and qualitative. In this case quantitative data refers to statistical concept meanwhile qualitative data refers to judgmental. Qualitative data may be found in a type of unstructured or semi-structured and It may contain non-numerical information. Data may be collected as a form of survey and to deal with this kind of data, labelling becomes more important. On the other hand, quantitative data has numerical information that means it is suitable to make comparison and calculation with variables. This kind of data may be divided by two subgroups as causal or time series according to Zietlow (2018). Causal data also known as regression refers a causal relationship in observed data. The practitioner may try to find an answer for this relationship to capture the way of causality, existence of causality or how strong It is. While dealing with a causal dataset, It is important to understand the difference between correlation and causation. The other type of quantitative data is called as time series which includes advances methodologies to gather insight from data. Moving average, exponential smoothing and classical decomposition maybe used for this purpose. Time series often includes four components and It is an important starting point to analyse each of these component’s presence and level at the data understanding process. These components are trend, seasonal, cycle and noise impacts on the time series. The trend indicates a long-term increase or decrease behaviour observed at the data. It may be linear or non-linear. Seasonal patterns can be captured when a time series is affected by seasonal factors for examples the seasons of the year or the special weeks like quarter ends. Also, a cycle impact may be observed in a time series data in respected of fluctuations of variable. Finally, noise is the last component of time series that refers a random variation in the data.

Analyse Technics

Methodologies may be divided by two main group to analyse time series data. The first one is called as time series analysis and It is important to build a data understanding and data preparation. Without a decent understanding of data may reduce the success of the modelling process. It is similar for data preparation phase. The time series data may be suitable with the aimed forecasting model. Thus, the other part is time series forecasting which includes several methodologies based on different approaches such as decomposition, smoothing, regression or machine learning. Deep learning may be located as a subfield of machine learning technics. In another perspective forecasting can be analysed univariate or multivariate.

After estimating inflation with a methodology, analysing the margin of error in the test set with the test-train split method can be seen as the first step in evaluating the results we have obtained. The total value, directions, and deviations of these errors will give us some insight into the success of our model. But other methods are also applied to gain deeper insights into the evaluation of results. The literature research shows us that; In most studies for inflation forecasting, the forecasting success of the chosen methodology is compared with the results of a different benchmark model. However, estimates shared by government agencies and market experts can also be useful as benchmarks.

Data & Business Understanding

Both of monetary and fiscal policies directly monitor price level of the goods and services which is also known as inflation in economic theory. To track changes on prices of a basket of goods and services, consumer price index (CPI) commonly used as a measure. It is one of the most widely used indicators of inflation and It is crucial for making economic policy decisions. This index is also used as a benchmark for discussion on wages at country level. Thus, It is possible to consider pre and post inflation effects on overall economy with various of stakeholders.

Monetary policy is governed by central banks of the countries which prioritize price stability. It is clearly stated on Turkey's central banks website as "The Central Bank of the Republic of Türkiye is primarily responsible for steering the monetary and exchange rate policies in Türkiye. The primary objective of the Bank is to achieve and maintain price stability.". To monitor price level, CBRT uses headline CPI data which announced monthly by Turkish Statistical Institute (TURKSTAT). TURKSTAT is one of the key data sources in Turkey which publishes regularly official statistics on population, resources, economy, society, and culture. TURKSTAT publishes 457 different statistics on inflation and prices and 72 of them is related to CPI. CBRT also collects combination of some of TURKSTAT's, CBRT's itself and other official statistics in electronic data delivery system.

CPI measures the change in the prices of goods and services for household consumption over time. The main objective of the CPI with the base year 2003 is; to calculate the inflation rate by measuring the change in the prices of goods and services that are subject to consumption in the market. For this purpose, all final monetary consumption expenditures of households, foreign visitors and corporate population in the country are considered. This concept left out the consumption expenditures of households for their own consumption and the relative rents applicable to the households. Individual Consumption Classification by Purpose (COICOP) is used by TURKSTAT in determining the weights and calculating the index, and these expenditures are gathered under 12 main groups and 43 subgroups. 404 items are included in the index. In the index, prices are compiled from a total of 228 districts, including all 81 provincial centres. Within the scope of the CPI, 564 710 prices are compiled from 27 411 workplaces per month and 5 246 tenants are followed within the scope of the index. The number of workplaces and prices may vary during the year according to the seasonal structure. The item baskets and weights are updated at the end of each year and the series is continued with the chain Laspeyres formula. As of December each year, new items are included in the index or items that lose their importance are removed from the index and new weights are used in the index calculation. The index is calculated by dividing the current prices by the prices of the previous December, which is the "new price reference period ", and multiplying by the December index, chaining is performed. The price index results are announced to the public on the 3rd of each month at 10:00 AM.

Benchmark Models and Methodology

Analysis

This study aims to predict inflation rates in Turkey with techniques based on machine learning and to evaluate the results of these estimations. At the same time, this analysis will be conducted for subgroups of inflation. These estimates will be based on monthly data. For the headline inflation, which will be analysed in the first part, the forecasts in the CBRT's quarterly inflation report will be used as a first comparison criteria. However, these forecasts are published quarterly and include forecasts for the next 12 quarters. For this reason, another public survey data will be used as a benchmark in order to better compare the results of the study. The market participants survey by CBRT has been implemented since August 2001, and statistics on the results are available as of that date. This indicator, which is produced in order to monitor the short and long-term expectations regarding basic macroeconomic variables such as consumer inflation, exchange rate, current account balance, GDP growth rate and interest rates, is created by monitoring the expectations of decision makers and experts in the financial and real sector regarding various macroeconomic variables. For example, the July 2023 Market Participants Survey was answered by 41 participants, consisting of representatives of the real sector, financial sector, and professionals, and the results were evaluated by aggregating the answers of the participants. The data indicate monthly inflation expectations for the next 3 months, year-end inflation expectations, and 12-month and 24-month future inflation expectations. Median expectations of market participants are used in the data and are available through the CBRT's electronic data distribution system. It is considered that this data can be used to create a benchmark in this regard. However, this data does not provide an indication of subsets of inflation. In addition, it does not provide information on which method set the market experts base their expectations on. At this point, the use of different methods for benchmark creation continues, as witnessed in the literature review. Deploying a benchmark model should be competitive and efficient to give a better comparison for objective technics' output.

Literature review process shows a wide usage of auto regression based linear models as a benchmark econometric model. There are multiple goals when a macroeconomist dealing with a macro economic data such as describing and summarizing data, forecasting data, advising to policy makers by digging insights from the data. In these aspects, every methodology has its own strength and weakness so It is still a challenge to find an overall dominant strategy to handle a time series data. Meanwhile some technics have a old history and wide literature, and some of machine learning technics proportionally new and have less academic research. Christopher Sims (1980) provided a new framework to capture dynamics in time series which is called vector autoregression (VAR). After two decades of literature, Stock and Watson (2001), analysed the efficiency of this framework for these aspects. At this study, VAR model is evaluated for four criteria which are data description, forecasting, structural inference, and policy analysis. According to this study VAR model is useful in general for these tasks but there are some limitations for this model. State of the art of this study indicates a weakness when the system is small such as three variables. Before deploying a benchmark model It is useful to understand the logic behind regression analysis to determine which benchmark model suits to Turkey CPI data.

This study aims to contribute to the literature by examining the success of linear, non-linear and deep learning methods in inflation forecasting through the example of Turkey inflation data. The fact that the data has been above the average of the data from 2003, when the new index started to be recorded, especially in recent years, shows the importance of retesting these models. At this point, multivariable linear regression and SARIMA will be used as benchmark models for data estimation, and LSTM model will be considered as an advanced deep learning model.

Regression Analysis

Time series analysis can be computed by linear or non linear models. Linear models are statistical models which try to explain or forecast dependent variables with a linear combination of parameters. Linear regression is a model which is a famous concept in statistics and econometrics. All independent variables multiply by assigned coefficients to find best linear line which minimize the errors. To to that, ordinary least square (OLS) method is commonly used. The stationary of time series play an important role for the success of the model. For an objective of estimating coefficients which minimize the sum of squared residuals between the observed and estimated values, some statistical practices should be followed. The basic of regression analysis may be started from classical linear regression model which indicates a single equation between two variables in a linear aspect. The equation of simple linear regression can be shown as following:

$$y_t = \alpha + \beta x_t + \mu_t$$

where the subscript t indicates the observation number (for time series case It is date). \`y\` is dependent (endogenous) variable where the \`x\` joined to the model as independent (exogeneous) variable. α and β stand for coefficients of the model. The aim of the model is to find best α and β to forecast y when It considered for all of the observations. \`μ\` indicates randomness in the model, a random disturbance term. The decide best parameters OLS approach commonly adopted where the other alternative is maximum likelihood.

BLUE Model

To build a good model BLUE (Best Linear Unbiased Estimator), it is a set of properties that an estimator should possess to be considered optimal in the context of linear regression analysis. An estimator is considered BLUE if it satisfies the following conditions:

Linear: The estimator is a linear function of the data.

Unbiased: The expected value of the estimator is equal to the true parameter value.

Efficient: The estimator has the smallest variance among all linear unbiased estimators. This can be achieved by selecting exogenous variables is in line with economic theory.

OLS method is the most common technique used to obtain the BLUE estimates in multiple linear regression. It aims to minimize the sum of squared differences between the actual and predicted values, thus providing the best possible estimates of the coefficients.

To determine the selected model is a good or appropriate model there some reference points. According to the essentials of econometric book (Gujarati and Porter) these reference criteria include parsimony, identifiability, goodness of fit, theoretical consistency and predictive power.

1. Parsimony: This principle emphasizes that a model should be as simple as possible, with the fewest number of parameters required to explain the data adequately. By employing a parsimonious model, researchers can avoid overfitting, which occurs when a model captures the noise in the data instead of the underlying pattern.
2. Identifiability: Identifiability refers to the ability to uniquely estimate a model's parameters. An identifiable model allows for consistent and unbiased estimation of its parameters.
3. Goodness of Fit: The goodness of fit measures the degree to which a model accurately represents the data. Commonly used goodness of fit measures includes the R-squared, adjusted R-squared, and mean squared error. These metrics help researchers assess the proportion of the total variation in the dependent variable explained by the model.
4. Theoretical Consistency: Theoretical consistency implies that the chosen model aligns with the underlying economic theory or established relationships. A model with strong theoretical foundations is more likely to provide meaningful insights and generalize well to new data. Researchers should consider both the theoretical consistency and empirical performance of a model when making their selection.
5. Predictive Power: A good model should have strong predictive power, meaning it can accurately forecast future outcomes based on the given data. Evaluating predictive power typically involves using techniques like cross-validation or out-of-sample testing. Models with strong predictive power are valuable for decision-making and policy evaluation, as they enable researchers and policymakers to make informed predictions about future trends and potential interventions.

Heteroscedasticity

While deploying a classical linear regression model OLS widely used as an estimation technique and It comes with Its own assumptions. If these assumptions do not meet then It may cause the selection of coefficient process or associated standard errors are wrong. Heteroscedasticity stands for the lack of homoscedasticity where a regression model assumes a consistent variance in data which also means homoscedasticity. If It requires a more technically explanation than heteroscedasticity refers that data with unequal variability across data, standard errors of a variable are not constant over time. If heteroscedasticity exists across data then It cause weakness for the model. First of all, to test model overall statistical significance which may be measured by applying F test may turn unreliable.

There are couple ways to detect heteroscedasticity and the first one is getting benefit from scatter plot of residuals. It may be track by visual analysis of distribution of error terms after deployed a regression model. It is an important point before evaluating result after regression.

Stationarity

Before deploying a model, checking stationary aspect of time series is another important part of analysis. The notion is stationary is related to observe time dependence. In other words, if there is a trend affect in data, that means this time series is non-stationary. A situation of non-stationary again does not meet an OLS model assumption. It may be seen visually and also It is possible to apply hypothesis test such as Augmented Dickey-Fuller (ADF) test to capture any trend effect.

As a part of logic test, It is expected to see a trend some macroeconomic data such as nominal gross domestic product or consumer price index in an inflationist economy. In case of emerging markets, it is usual to see an absence of deflation that means prices usually go up but the dimension of these movements fluctuates. Thus, even though it is possible to see different inflation rates usage of index data cause a non-stationarity. On the other hand, some financial data like commodity prices may not be show signals of a dominant trend for a period. At this stage, demand and supply relationship may have a dominant impact on prices, thus erases any trend effect by fluctuating to different ways. This logic test may not be enough to track any trend effect alone. To build a decent forecasting model, It is important to apply an econometric test to check stationarity for all series in data. This may be classified under data preparation process according to CRISP DM methodology.

With an objective of capturing trend effect under statistically significance, ADF test is applied all of the column of raw data. In line with logic test, CPI data, USD/TRY exchange rate and total consumer loan amount shown a sign of trend meanwhile Brent oil prices in USD does not according to tests output. To arise this effect also monthly percentage changes of each time series involved to the raw data. Thus, Brent oil prices in USD , USD/TRY exchange rate and total consumer loan showed a strong significance in favour of non-stationarity. Surprisingly monthly percentage change of CPI showed stationarity, according to output of state of art. It may be indicating existence of structural break down and under this condition It is possible to avoid from stationarity impact by splitting time series.

Seasonality

Literature review shows that auto regression based forecasting models are used widely for inflation forecast. As It is mentioned before, time series mainly have four components and auto regression based model significantly useful to capture these components' impacts. It also may be developed in linear perspective by adding seasonality which is an important component of CPI. In line with theoretical consistency, the fact that %25.43 weight of Turkey's CPI is deriving from food and non-alcoholic beverages which is highly correlated by seasons of the year because of the weather conditions. This weather conditions are correlated with household's heating needs which can be found in housing index which is a subset of CPI with a weight of %16.62. It may be seen another seasonality impact for education, recreation, culture, and hotels group. Thus, It is highly important to capture seasonal effects on CPI when building a benchmark linear model to reach a better model in domain of mentioned criteria.

Auto Regression

Autoregressive model takes leverage from past patterns of time series to make better prediction for future behaviour. It requires a significant relationship between values statistically. An autoregressive model captures the relationship between a variable and its lagged values. Thus, under an assigned lag order the model tries to measure significance of auto regression. The auto regressive model may be presented as a linear regression of last observation by explaining from lagged observations. Auto regression may be shown as the following equation:

$$x_t = \alpha + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \beta_p x_{t-p} + \mu_t$$

where x_t is variable's value at time t, in other words the last observation. Where the α is constant β represents coefficient of each lagged observations. μ_t is the noise error term of the model. Determining p is the crucial in this model, It requires to examine each lagged significance individually.

Multilinear Regression Model

As mentioned above, regression analysis, which is a linear method, works with the assumption that a variable is affected by other variables with a constant coefficient over time and calculates this constant coefficient by optimizing with the OLS method. It is important to ensure stability in the data set for a better regression method. Although monthly changes are used in this data, the undetectable stagnation in the inflation data may also adversely affect the overall success of the model. In addition, monthly inflation change data for the multilinear regression model has been tried to be calculated by using monthly changes of Brent oil, consumer loan interest and exchange rate data. However, it is aimed to capture the autoregressive effects by adding the monthly inflation change to the linear regression method with a 1-month lag. Accordingly, the mathematical representation of the two models can be given as follows.

$$CPI_t = \alpha + Oil_t\beta + \text{Exchange Rate}_t ¥ + \text{Consumer Loan Interest Rate}_t € + \mu_t$$

$$CPI_t = \alpha + Oil_t\beta + \text{Exchange Rate}_t ¥ + \text{Consumer Loan Interest Rate}_t € + CPI_{t-1}\Omega_+ \mu_t$$

As can be seen from the equations, a constant term, an error term and the multiplication of the explanatory variables by a certain constant are used to estimate the CPI data. The linear regression method can be considered as a mathematically simple and useful model. The coefficient assigned for each independent variable and the overall significance of the model will decrease if there is a relationship between the preferred independent variables. Therefore, using linear regression on very large datasets brings with it the need for better handling of time series analysis.

SARIMA Model

SARIMA stands for Seasonal Autoregressive Integrated Moving Average. It's a non-linear time series forecasting model that extends the standard ARIMA model to handle seasonality in the data. ARIMA models are widely used in time series analysis for predicting future values based on past observations. SARIMA get benefit from both autoregressive (AR) and moving average (MA) components along with integration (I) and seasonality (S) components to capture various patterns present in the data.

The autoregressive component considers the relationship between a current observation and a specified number of previous observations as It is mentioned in previous passage. It quantifies the impact of past values on the current value. The AR component is denoted by "p" and represents the order of the autoregressive component. A higher "p" indicates that the current value depends on more previous observations.

The moving average component represents the relationship between the current observation and a specified number of past forecast errors (the difference between predicted and actual values). A moving average term in a time series model is a past error. The MA component is denoted by "q" and indicates the order of the moving average component.

The integrated component deals with differencing the time series data to make it stationary. Stationarity is important because most time series models assume that the data has constant statistical properties over time. The order of differencing is represented by "d". If the data is already stationary, d=0. If differencing is performed once to achieve stationarity, d=1, and so on. Align with state of art outputs, the non-stationarity of CPI data was handled by a positive number value of integrated component.

The seasonal autoregressive component captures the relationship between the current observation and its corresponding observation from the same season in the previous terms. It models the seasonal pattern. The order of the seasonal autoregressive component is denoted by "P". In this study, monthly data was used thus the seasonality component was chosen as 12 to capture monthly seasonality effects.

Similar to the integrated component, the seasonal integrated component involves differencing the data but over the seasonal period. It's denoted by "D" and indicates the order of seasonal differencing.

Lastly, the seasonal moving average component is like the regular moving average component but applied to the seasonal differences. It's denoted by "Q" and represents the order of the seasonal moving average component.

To sum up SARIMA model is generally represented as SARIMA(p, d, q)(P, D, Q, s), where (p, d, q) are the non-seasonal AR, I, and MA components while (P, D, Q, s) are the seasonal AR, SI, SMA, and the length of the seasonality.

Machine Learning

The words artificial intelligence, machine learning, deep learning and algorithm were often confused with each other and misunderstood. However, artificial intelligence, machine learning, deep learning, and algorithm are different terms that have been around long enough to be assigned a fixed meaning by literature.

Algorithm

Algorithm is a list of rules to follow to solve a problem. The dictionary meaning of algorithm is defined as a set of mathematical instructions or rules that, especially if given to a computer, will help to calculate an answer to a problem. Algorithm is the way in which the solution of a problem or how to reach the determined goal is explained thus a fundamental of other terms. As It can be understood from explanations, the algorithm is not a result, but a path leading to the result. Algorithms may be shown in every field; It has gained a central importance in our lives by affecting our lives even more with technological developments. For the algorithm to be successful, it is not only sufficient to reach the result, but it also needs to be simple and fast. Besides consisting of a strictly defined set of instructions, the algorithm must always have an end and be proven to work in any case. Everything from the sorting of numbers, the face detection of the photo camera to the word search, briefly machine learning, artificial intelligence or deep learning systems were built on an algorithmic logic. Algorithms are divided into several types according to their usage areas, complexity, design methods and application methods.

Machine Learning

In this period when technology is advancing rapidly, the concepts of machine learning and deep learning have become one of the popular sub-titles of the field of artificial intelligence.

Machine learning refers to the use of algorithms and statistical models by computer systems. In this way, automation can be provided in certain tasks. In other words, a machine learning model learns from data to get better at a particular task. For example, an email service can use machine learning to classify emails that arrive in your inbox as "important" or "junk". In simple terms, it uses the algorithm it learns for new data by detecting which variables in the data. Machine learning usually may be examined by three categories.

1. Supervised Learning:

Supervised learning is a machine learning approach in which learning algorithms are trained under "supervision" with training data. This data comes in pairs as inputs and expected outputs. In other words, properties and tags are identified as an input. The purpose of the algorithm is to create a function by extracting the relationship between the inputs so that it can make accurate predictions for new and previously unseen data. Thus, It may be useful for certain task like classification or regression. Categorizing emails as spam or not is an example of this featuring.

2. Unsupervised Learning:

Unsupervised learning works with unlabelled data. The algorithm analyses this data and tries to discover structures or patterns within it but is not guided by a true or false output. Clustering is widely use as a technic for unsupervised learning. It is used to separate data into groups with similar characteristics. For example, customer segmentation. Also, dimension reduction may be used to reduce complexity of the data. In this case, there were no tagging for expected output and algorithm tries to find best structure of reduced dimensions.

3. Reinforcement Learning:

In reinforcement learning, an agent performs actions in each environment and receives rewards or punishments for those actions. Its purpose is to learn action strategies that will maximize total reward over time. A reinforcement learning problem is usually defined by the components such as agent, environment, actions, and reward. For example, in a computer game, the agent's goal might be to get the highest score. The agent gains or loses points by performing certain actions. Reinforcement learning can help the agent learn how to play this game. Each of these approaches is suitable for specific problems and scenarios. In addition, each of these methods has a wide field of research and application and is the cornerstone of artificial intelligence and deep learning.

Artificial Intelligence

According to IBM, artificial intelligence (AI) is a field, which combines computer science and robust datasets, to enable problem solving. It also contains some sub-groups of machine learning and deep learning. There are two types of AI which are weak and strong. Weak AI—also known as Artificial Narrow Intelligence (ANI) is AI trained and focused to perform specific tasks. Strong AI is made up of Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI). Super Intelligence (ASI) also known as superintelligence would surpass the intelligence and ability of the human brain. Strong AI is still entirely theoretical with no practical examples in use today. both deep learning and machine learning are sub-fields of artificial intelligence, and deep learning is actually a sub-field of machine learning.

Deep Learning

Meanwhile, deep learning has a long history and many aspirations. Modern deep learning provides a powerful framework for supervised learning. By adding more layers and more units within a layer, a deep network can represent functions of increasing complexity.

Deep learning is a subclass of machine learning and artificial intelligence and consists of a set of algorithms that attempt to learn at multiple levels. The deep expression in deep learning refers to successive layers. How many layers are in a model, the depth of the model is expressed by this number of layers. In deep learning, layers structured one after another perform learning through structures called artificial neural networks.

In 1944, the concept of neural networks was initially introduced by Warren McCullough and Walter Pitts, two researchers from the University of Chicago. Later, in 1952, they joined MIT and became founding members of a department that is often referred to as the pioneer in cognitive science. Neural networks, also known as artificial neural networks (ANNs) are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another. Deep learning methods have been developed on artificial neural networks (ANN) studies in general. However, unlike these studies, they are based on more hidden neurons and layers.

Deep learning encompasses computation models with multiple processing layers, allowing the representation of the available data at multiple levels of abstraction. In deep learning methods, also known as deep networks, there are different layers stacked on top of each other to create a representation of the data. Deep learning methods efficiently perform effective higher-level abstractions from raw data, can form automatic sets of features, thus enabling the automatic extraction and utilization of features that are typically determined by humans.

The learning phase of some deep learning algorithms can be quite lengthy, and in various studies, semi-supervised learning approaches have been proposed to shorten the learning processes of deep neural networks. In some works, approaches have also been presented to imbue the capabilities of deep learning into shallow but widely used machine learning methods like support vector machines (Cho and Saul, 2009).

Deep learning methods have yielded successful results in processing various types of data, including video, audio, and text. Often, one method can excel in processing textual data, such as natural language processing while another method might achieve better results in processing video and audio data. In some studies, the proposed deep learning approaches can be successfully used for multimodal learning.

Deep learning methods can be applied to various domains in more detail, including (Deng and Yu, 2014):

1. Language modeling and natural language processing

2. Speech and audio processing
3. Information retrieval
4. Object recognition and computer vision
5. Multimodal and multitask learning

In the landscape of artificial intelligence, a variety of deep learning models showcase unique strengths. While Deep Neural Networks (DNNs) excel in many areas, their training might occasionally be time-consuming. Meanwhile, Convolutional Neural Networks (CNNs), albeit data-intensive, reign supreme in image analysis, offering noteworthy achievements and even broadening their capabilities to linguistic challenges. On the other hand, Deep Auto-encoders, being unsupervised, optimize the learning journey by negating the necessity for annotated datasets. Deep Belief Networks (DBNs) provide adaptability, serving both labelled and unlabelled scenarios, but their training can be as lengthy as that of DNNs. The Deep Boltzmann Machine differentiates itself from DBNs mainly due to its bi-directional ties across most tiers, which amplifies its computational load. In terms of linear data, Recurrent Neural Networks (RNNs) stand out, with versions like LSTM becoming indispensable in linguistic tasks.

Deep Neural Networks (DNNs)

In the literature, one of the first algorithms about neural networks is the perceptron algorithm (Rosenblatt, 1958). In this network, there is an input layer that is directly connected to the output. With this algorithm, classifiers can be developed to separate patterns that can be linearly separated. For more complex problems, multiple hidden layers have been added to this algorithm, and the weights of each layer have been adjusted with the learning method called the delta rule.

With the addition of more hidden layers (more than two) to neural networks of this type, it has become possible to detect non-linear complex relationships, and the resulting neural networks are called deep neural networks (DNNs). DNNs can be used for both supervised (if there are labeled data) and unsupervised learning problems, such as clustering. DNNs are widely used for classification and regression purposes and achieve successful results.

Convolutional Neural Networks (CNNs)

CNNs, are a specialized kind of neural network for processing data that has a known grid-like topology. Convolutional is a kind of linear operation. In convolutional network terminology, the first argument to the convolution is often referred to as the input, and the second argument as the kernel. The output is sometimes referred to as the feature map. CNN is a deep learning algorithm that is generally used in image processing and takes images as input. This algorithm, which captures and classifies the features in the visuals with different operations, consists of different layers. The visual, which passes through these layers, which are Convolutional Layer, Pooling and Fully Connected, is subjected to different processes, and reaches the consistency to enter the deep learning model. When creating CNN models, data pre-processing were less time consuming when It was compared to classical machine learning algorithms.

Recurrent Neural Networks (RNNs)

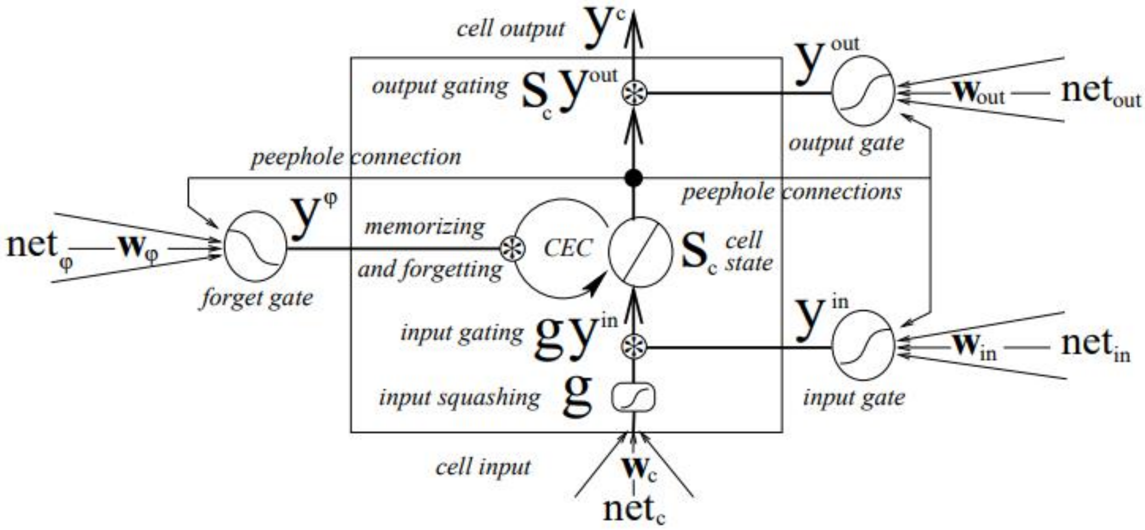
RNNs are a family of neural networks for processing sequential data. It reuses hidden layer properties at one point as input for another. Therefore, it is a deep learning model that is suitable when the input data has directionality. Thus, It may be well performed on time series data. RNNs are a kind of deep learning constructs that are generally used to predict the next step. The biggest difference from other deep learning structures may be classified as that they were able to remember relationships across data. RNNs make associations between inputs to follow the next step and remember all their associations while they are being trained. RNNs use a loop-like structure that rotates within themselves so that the relationships they have established are permanent. They have some advantages such as possibility of processing input of any length, shared weights across time, accounting historical information into computation meanwhile some disadvantages such as possible time inefficiency for computation, difficulty of accessing information from a long time ago and lack of consideration any future input for the current state (Amidi). There are four types of gates in order to remedy the vanishing gradient problem.

Long Short Term Memory (LSTMs)

RNN models help predict time series by carrying past information into the future. However, in long series, weakness may occur in the information carrying level of RNNs. Long short-term memory (LSTM), which is a sub-type of the RNN structure, emerges as a more convenient model for long sequences by adding a carrier strip that enables the transport of many time slots. The prediction is computed and updated sequentially after seeing each of the data lags. Intermediate output, which is called the “state” of the network, is used as an additional input at the next time step (Almosova, 2022).

The basic unit of an LSTM network is the memory block which contains one or more memory cells and three adaptive, multiplicative gating units shared by all cells in the block. Each memory cell carries its own recurrently self-connected linear unit, which is called the cell state. In this way, the problem of going back in time in RNNs is solved.

Figure 2: LSTM memory block with one cell (cse-lab.seas.harvard.edu)



Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) architecture that was introduced by Hochreiter & Schmidhuber in 1997. The LSTM is designed to alleviate the vanishing gradient problem, which makes it difficult for standard RNNs to learn from data where longer-term dependencies exist.

Forger gate may be presented with following equation where ft is the forget gate's activation vector, σ is the sigmoid function, Wf is the matrix, bf is the bias, h_{t-1} is the hidden state from the previous time-step, and xt is the input at the current time-step (Sen, 2018).

$ft=\sigma(Wf\cdot[h_{t-1},xt]+bf)$ (1)

Input gate (i_t) controls values in the cell state to update can be shown as following equation.

$it=\sigma(Wi\cdot[ht-1,xt]+bi)$ (2)

Equation 3 and 4 shows the cell state which is a combination of previous cell state and the candidate cell state.

$C\sim t=\tanh(WC\cdot[ht-1,xt]+bC)$ (3)

$Ct=ft*Ct-1+it*C\sim t$ (4)
Output gate controls the parts of the cell state make it to the output.

$ot=\sigma(Wo\cdot[ht-1,xt]+bo)$ (5)

Lastly, hidden state is a function of the cell state and output gate.

$ht=ot*\tanh(Ct)$ (6)

At the data preparation stage, all the values rescaled with minimum maximum scaler algorithm to get better results from LSTM model. While the parameters were tuning for LSTM model, had get benefit from literature review and the default algorithms. LSTM units which is also called neurons were tested between 100 and the 2000 by iteration. Activation algorithm were chosen as ‘relu’ which was default in Keras library. Similarly, ‘adam’ is used as a model optimizer. The loss function of model was set up for mean squared error which is align with this study’s evaluation metric. The number of epoch was chosen between 50 and 500 by iteration where the one epoch refers one forward pass and one backward pass of all the training examples. The number of samples that will be used to update the model weights in each iteration, also called as batch size was tuned to 32 while verbose argument was set equal to 1.

Optimization for Training Deep Learning Models

According to deep learning book from MIT, there are several points to consider for optimization of the deep learning model. In machine learning, optimization of the loss function is crucial for model training. However, the true loss function is often computationally intractable to minimize directly. Practitioners typically optimize a surrogate loss function, like negative log-likelihood, which serves as an efficient and effective proxy. This approach can sometimes result in better performance, even allowing the model to learn more than it would by directly minimizing the original loss function. For example, using a log-likelihood can make a classifier more robust by pushing classes further apart, even when the training set 0-1 loss has already reached zero. Machine learning algorithms commonly employ early stopping criteria based on the true underlying loss function to avoid overfitting which is used in this study for LSTM forecasting with multiple exogenous variables model. This results in the algorithm stopping before the surrogate loss function has fully converged. Thus, It may create a key difference for machine learning optimization and classical optimization approaches.

In machine learning, the objective function often decomposes as a sum over individual training examples, distinguishing it from general optimization algorithms. This enables the use of approximate methods like stochastic or minibatch gradient descent, where parameter updates are based on a subset of the training data. For example, the Maximum Likelihood Estimation problem often decomposes into a sum of log-probabilities over individual examples. Computing the exact gradient for optimization would require evaluating the model on the entire dataset, which is computationally expensive especially when It is a case dealing with wide dataset. However, the law of large numbers allows us to approximate this gradient by averaging over a small, randomly sampled subset of data. This approach is computationally efficient and also mitigates the issue of data redundancy. But this process may add more stochasticity to the model and should be evaluated carefully. The reduction in computation time often outweighs the loss in gradient estimate accuracy, leading to faster overall convergence. In this context, "batch" or "deterministic" methods refer to algorithms that use the entire training set for each update, while "stochastic" or "online" methods update parameters based on individual or small sets of examples. In this study deterministic approach is preferred.

Results and Discussion

After reviewing the literature, a range of information on economic forecasting is explained, with a review of related concepts. In this way, inflation in Turkey was examined conceptually and emerging concepts on how it could be predicted were investigated. In this study, the methods that can be used for macroeconomic forecasting and which benchmarks to use to measure the success of these methods were analysed. Accordingly, the first metric that came to the fore was the CBRT's quarterly publication, which included recent inflation forecasts. Although monthly inflation forecasts can be revealed indicatively by a study on the shared graph in these publications, they are not officially shared by the CBRT. However, in these publications, headline inflation three months later and its deviation paths in the 70% confidence interval, with the lower and upper bands, are included. As mentioned before, current inflation data is used by equating it to the index value of 100 for January 2003. In the forecasting models included in this study, the period from January 2004 to January 2023 is used. With the mainly 75-25 train test split approach, the period of this data before April 2018 is reserved for machine learning. With the models created from the data between January 2004 and April 2018, the realizations between April 2018 and January 2023 were tried to be estimated and the margins of error were calculated accordingly. Both the official estimates, the estimates of the market participants survey, and the statistical methods used have recently become more suitable to obtain a higher margin of error. The reason for this is the high fluctuations observed in the factors contributing to inflation in recent years. As this data increases the estimation uncertainty, it also reveals the necessity of questioning the validity of model parameters used in the past. While the average of monthly headline inflation changes from January 2004 to April 2018 was 0.69, this average increased to 2.26 in the period from April 2018 to January 2023. This change has manifested itself with the observation of trend effects in the data despite monthly changes and indicates the possibility of a structural break down. This situation leaves the door open for new studies to be made for Turkish inflation, using more than one train set of different periods.

Benchmark estimations

Mentioned changes in the pace of inflation rates are correlated with the high deviations of the estimators in line with the expected. As explained in detail in previous chapters, market professionals survey was conducted monthly with several questions on macro-economic with a purpose of gathering market expectations on next future' economic data. Thus, inflation expectations of survey participants were used as a primary benchmark in this study. Outlining this major change in structure of Turkey's inflation path in recent years, the variance of net errors of these estimates increased significantly. For benchmark purposes, the RMSE is calculated for the median of this survey for the current month inflation data. The calculated RMSE shows a score of 1.90 for the same period which were conducted as a test set in this study which refers to last %25 of all data. Meanwhile, the success rate of median expectations of survey participants were proportionally better in the past. As an example, RMSE were calculated for the period between 2014 January and 2017 November which contains same length of dataset. During this period the calculated RMSE of the median of the same survey were 0.48. This dramatic change is correlated with the size change of observed data but also indicates a further need to recalibrate current models hyperparameters. Also, one of the other outputs of survey was one month further inflation expectations. When the same dataset used to calculate another RMSE value by using these median expectations, 0.79 value was reached for the old period. For the recent term, the RMSE value deteriorated to 2.28.

The drastic behavior differences that the data set has shown in recent years caused a train test split ratio, which is not normally used very often, to be tested in this study. Accordingly, the data set was further divided into two parts, 95% train and 5% test set. Looking at the current month median expectations of the market participants survey, the RMSE calculated for the 12-month period from February 2022 to January 2023 is 1.75. Considering the median of the estimates shared for the next 1 month, this ratio is calculated as 2.14. While evaluating all these results, it should also be taken into account that the participants who participated in this survey have the opportunity to revise their forecast models every month and have the opportunity to manually add one-off effects to the model. At this point, all calculated RMSE values can be considered as a very challenging benchmark before evaluating model outputs.

Models

Feature Selection

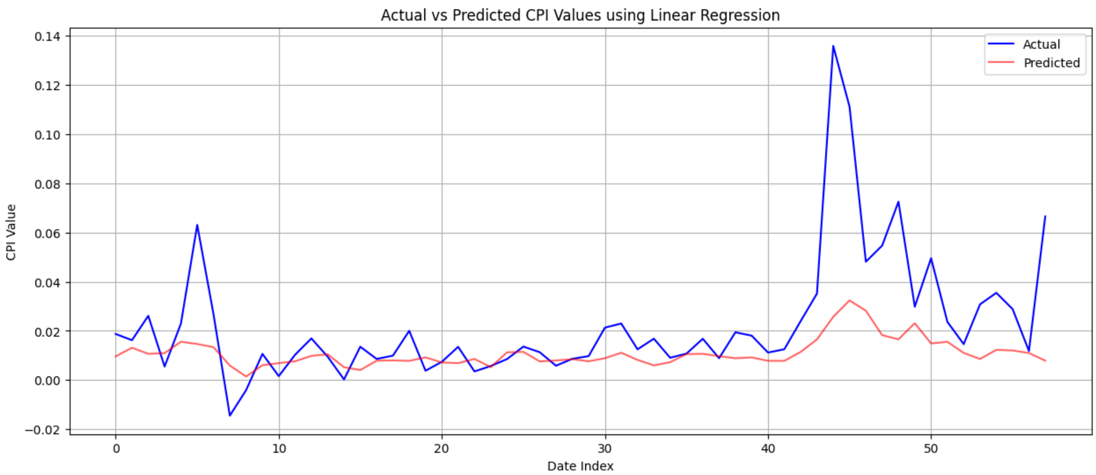
While constructing all the models, three independent variables that were thought to be significant according to the literature review were added to the model. LSTM is run for univariate form and with these independent variables as well. To capture better autoregressive affects in multilinear regression model, the model is run also with lagged CPI data for month. Also a mixture of SARIMA and LSTM is tested to reach better accuracy. At this point SARIMA's outputs are added to LSTM model as fourth exogenous variable.

In order to estimate the data more accurately, it was preferred to use data compatible with the theory and known to have a high relationship with inflation. Among the higher-frequency data that can be obtained more functionally in the real world scenario, before the data is announced, the exchange rate, consumer loan interest rate and Brent type oil prices are included in the model according to the inflation relationship. Since the exchange rate affects inflation as an external factor through import prices, it is theoretically expected to show a positive correlation with inflation. Accordingly, it can be thought that an increase in the exchange rate will also cause an increase in inflation. In the Central Bank study, which was mentioned in the literature review, the relationship between inflation and the exchange rate was determined, and empirical studies were carried out to calculate the strength of this relationship. USD/TRY daily buying prices are used for the exchange rate. The interest rate of another variable consumer loan is added to the model with the assumption that it can affect inflation on the demand side. Accordingly, the interest rate can be considered as an internal factor. The last independent variable was added to the model as the import price of Brent oil in USD. As a net energy importer, Turkey might be shown low demand elasticity to changes in global energy prices. Therefore, increases in energy prices may affect producer prices and therefore consumer prices in terms of cost. Exchange rate and oil prices can be obtained instantly as market data, and daily from the data source of this study. Interest rates are published on a weekly basis. The literature on how to set up the model used in the benchmark section shows that the variables must be free from trend effects, that is, stationary, to create a good linear model. Data were collected from the CBRT's electronic data distribution system via API connection, both as monthly changes and nominal values. The data were tested for stationarity with the ADF Dickey Fuller test. After the test applied to the entire data set, it was observed that the nominal values of inflation and exchange rate were clearly under the trend effect, that is, they were not stationary. By using all data as monthly changes, they were brought to the same scale and recovered from non-stationary effects. On the other hand, non-stationnary was observed in the monthly changes in inflation.

Multilinear Regression Outputs

As the first benchmark model, the multi linear regression model was applied to the data set with the OLS method. Brent oil price, consumer loan interest rate and USDTRY exchange rate were set as exogenous variables with their monthly changes. While there are many factors that are affected by the dynamics of the inflation data, the model should be established by considering both time series analysis (trend, seasonality, cycle and noise effects) and explanatory variables. Multilinear regression analysis alone may be insufficient to capture all these effects. At this point, the first model created to predict inflation was made using the three variables mentioned above and a RMSE value of 2.82 was obtained with 75-25% train test split approach. When the train set was extended to 95% level, a poor accuracy was obtained with a RMSE value of 3.35 compared to all other models. As a result, in order to improve the model a little more, monthly inflation change was added as an explanatory variable with a delay of 1 month. This model has reached a more successful prediction power with two different train sets. Although the model operated with 75% train split has a lower predictive power than the market participant survey median with an RMSE value of 2.43, it is a good benchmark for other models. In the chart below, there is a view of the realizations and predictions for the developed linear regression model. Although the deviation values are low in the months that are close to the mean, it is seen that it is difficult to determine the accelerations in the months when the standard deviation is high.

Figure 3



SARIMA Outputs

As explained in the previous sections of the study, the SARIMA model is an advanced model used in time series analysis and is a derivative of ARIMA that includes seasonality. The first three of the hyperparameters are set as autocorrelation (p), integrated (d) and moving average (q). The 'p' parameter specifies the number of delays of the autoregressive terms, while the 'd' parameter specifies how many times the time series should be differentiated. The 'q' parameter specifies the lag numbers of the moving average terms. It is seen that the SARIMA model, which was defined as a benchmark for estimating Turkish inflation in the literature review in 2011, were recommended to set these first three parameters as equal to 0. The following three parameters consist of capital letters of the same notation. Accordingly, while the 'P' parameter shows the number of lags for seasonal autocorrelation terms, the 'D' parameter indicates how many times the seasonal difference should be taken to provide seasonal stability. The 'Q' parameter specifies the number of lags for seasonal moving average terms. The last parameter used in the SARIMA model is denoted by 'S' and represents the seasonality period. In monthly time series usage, it is set as 12 if seasonal effects from month to month are desired. According to the same study, it was preferred to use 12 lag terms for the seasonal period and 1 for the other three seasonal ARIMA parameters. The first applied SARIMA (0,0,0,1,1,1,12) model was run for the first 75% of the data and its predictive power was checked with the remaining 25% test data. Accordingly, it generated an RMSE of 2.76 and underperformed when compared to the hit of the poll's median expectations. Afterwards, the hyperparameters were optimized according to the model dataset to find the combination that produced the most accurate result between 0 to 3. As a result of the optimization, it is seen that the SARIMA(0,2,0,2,2,1,12) model reached the RMSE score of 1.98. Although the results are accurate, the significance of the model in terms of econometrics seems to be open to improvement due to the increase in the average of the monthly change observed in the inflation data in recent years. The assigned coefficients of exogenous variables are not aligned with logical expectation for consumer loan and oil prices. Indeed, the Z values show statistically insignificance individually for these 2 variables.

Table 1

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                               CPI-1   No. Observations:                 171
Model:             SARIMAX(0, 2, 0)x(2, 2, [1], 12)   Log Likelihood                -270.994
Date:                             Sat, 26 Aug 2023   AIC                            555.988
Time:                                     14:32:26   BIC                            576.825
Sample:                                          0   HQIC                           564.454
                                             - 171
Covariance Type:                               opg
===================================================================================
                        coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
usdtry_currency-1     0.0551      0.014      4.068      0.000       0.029       0.082
consumer_loan-1       0.0178      0.010      1.713      0.087      -0.003       0.038
oil_prices-1         -0.0016      0.004     -0.379      0.705      -0.010       0.007
ar.S.L12             -0.7036      0.090     -7.796      0.000      -0.881      -0.527
ar.S.L24             -0.3337      0.097     -3.458      0.001      -0.523      -0.145
ma.S.L12             -0.9959      6.494     -0.153      0.878     -13.723      11.732
sigma2                1.7088     10.981      0.156      0.876     -19.814      23.232
===================================================================================
Ljung-Box (L1) (Q):                  50.79   Jarque-Bera (JB):                 0.47
Prob(Q):                              0.00   Prob(JB):                         0.79
Heteroskedasticity (H):               0.48   Skew:                            -0.14
Prob(H) (two-sided):                  0.01   Kurtosis:                         2.94
===================================================================================
```

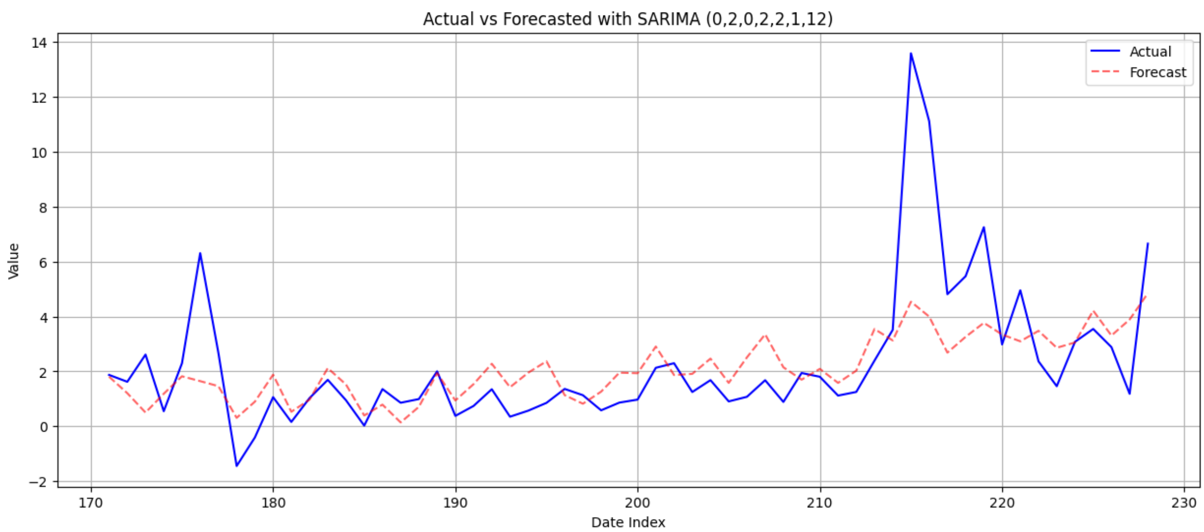
For this reason, the data set was distributed at different rates and the SARIMA model was reworked. The results found with the previous optimization in the 95%-05% split data shows better accuracy in favor of SARIMA (0,2,0,2,2,1,12) with an RMSE of 3.01 when It compared to SARIMA (0,0,0,1,1,1,12) where the RMSE is equal to 2.57. But, the summary table of the model's statistics remained an open door to model significancy. For consumer loan interest rate change and Brent oil prices monthly changes the individually significance tests failed again. Established benchmark model has shown that an inflation model applied with SARIMA has the power to make predictions to compete with expert expectations but . On the other hand, it is seen that it is insufficient to detect non-linear effects in the time series and leaves room for support.

Table 2

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                                CPI-1   No. Observations:                217
Model:             SARIMAX(0, 2, 0)x(2, 2, [1], 12)   Log Likelihood                -399.055
Date:                              Sat, 26 Aug 2023   AIC                            812.110
Time:                                      00:29:39   BIC                            834.876
Sample:                                           0   HQIC                           821.332
                                              - 217
Covariance Type:                                opg
=====================================================================================
                        coef    std err          z      P>|z|      [0.025      0.975]
-------------------------------------------------------------------------------------
usdtry_currency-1     0.0959      0.011      8.934      0.000       0.075       0.117
consumer_loan-1       0.0228      0.012      1.824      0.068      -0.002       0.047
oil_prices-1          0.0061      0.005      1.342      0.180      -0.003       0.015
ar.S.L12             -0.7446      0.088     -8.481      0.000      -0.917      -0.573
ar.S.L24             -0.3366      0.092     -3.646      0.000      -0.518      -0.156
ma.S.L12             -0.9954      3.077     -0.324      0.746      -7.026       5.035
sigma2                2.8332      8.436      0.336      0.737     -13.700      19.367
===================================================================================
Ljung-Box (L1) (Q):                  37.76   Jarque-Bera (JB):               189.66
Prob(Q):                              0.00   Prob(JB):                         0.00
Heteroskedasticity (H):               2.12   Skew:                            -0.87
Prob(H) (two-sided):                  0.00   Kurtosis:                         7.56
===================================================================================
```

As can be seen in the graph below, the estimations of the optimized SARIMA model detect deviations and perform more successfully than the linear regression model. However, it performs better than the market average forecast for inflation 1 month later. Two established econometric models have shown that the use of non-linear models in inflation estimation is important practice.

Figure 4



LSTM outputs

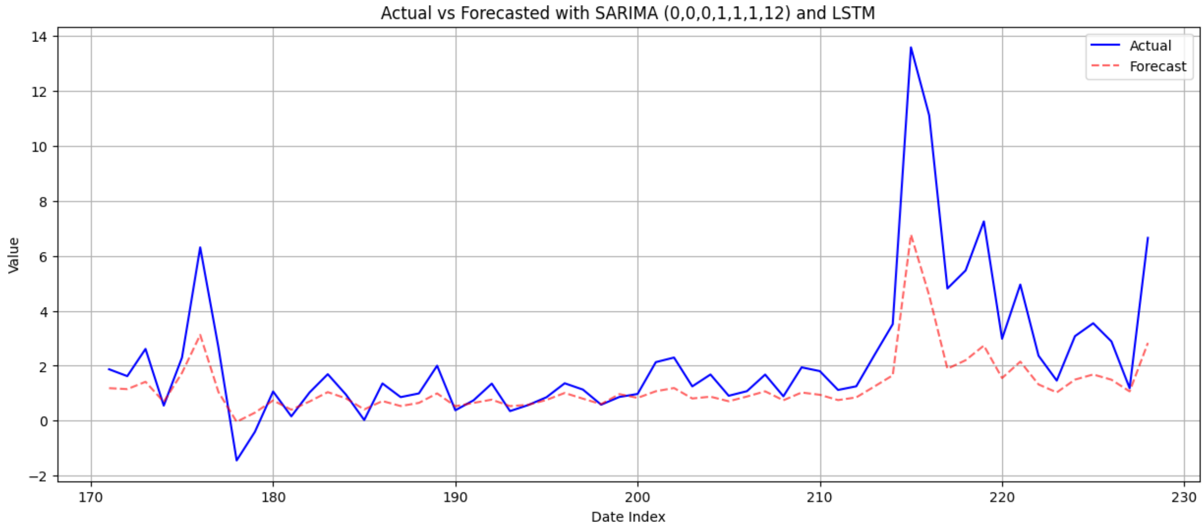
Long Short-Term Memory (LSTM), a variant of RNN, which is one of the types of deep learning, was preferred in the last part of the study. The high ability of this model, which was seen to be successful in literature studies, to handle long time series, brought this model to the scope. It is known to have a special structure called memory cell within the RNN model.

Input gate, forget gate, output gate, new memory cell, cell state update and hidden state update form the mathematical background of the model. Keras library on Phyton were used to deploy LSTM model and 22 different parameters were existed to tune in the function. That clearly indicates, there is a higher risk for deploying a decent model then regular linear approach or SARIMA model in the case of parameter tuning. The choice of sequence length, batch size, number of layers, number of neurons in each layer, learning rate, dropout rate, etc., can affect the performance. It's often a challenge to get the hyperparameters right.

While running the LSTM model, the stochastic process included in the algorithm caused different results to be produced in each run of the algorithm. This randomness seed parameter was eliminated by setting it to 100. When the model was run with only three explanatory variables, it produced the worst RMSE values in the train test split approach of 75-25% with 2.88. At this point, it was seen that using only independent variables while constructing the model with LSTM was not effective. To validate this output, the model has been optimized for parameters such as dropout rate, batch size, epoch in order to produce the most accurate result in order to check whether there is a hyperparameter set in which the model performs better. As suggested in the literature review, the use of 500 epochs for inflation estimation was tested in all LSTM models, but no significant positive difference could be reached for this data set.

As seen in the literature review, LTSM models were run with datasets with a large feature selection and positive results were obtained. However, in this study, 3 explanatory variables with easily obtainable forward prices and the data itself were set as explanatory variables to scale them to the next periods. At this point, the outputs of SARIMA models, which are more successful than linear regression and can detect non-linear relationships, are added to the LSTM model as an explanatory variable, apart from 3 independent variables. First, the LSTM model, which was created by adding the outputs of the SARIMA (0,2,0,2,2,1,12) model, which gave better results, was run and the RMSE value of 2.13 was obtained. While this result was more accurate than the results of other established models, the same model was reconstructed in a way that included the outputs of the SARIMA (0,0,0,1,1,1,12) model, which was also recommended in the literature review. In this model, the RMSE value of 1.83 was obtained. This level is more successful than other established models as well as more successful than the median expectations of market experts.

Figure 5



As seen in the graph above, this model predicts particularly hard deviations more accurately than other models. The combination of SARIMA and LSTM models could not show similar performance when they were run with a 95% train set, and they showed evidence to learned more from the recent upward trending data and produced very high estimates. Finally, the LSTM model is run univariate without any independent variables. Optimal lags was found with the Bayesian Information Criterion (BIC) approach, as suggested in the literature review. While applying this method, the optimization tools mentioned in the Deep Learning Book were also applied. With the 2.63 RMSE, a weaker accuracy rate was obtained compared to other models.

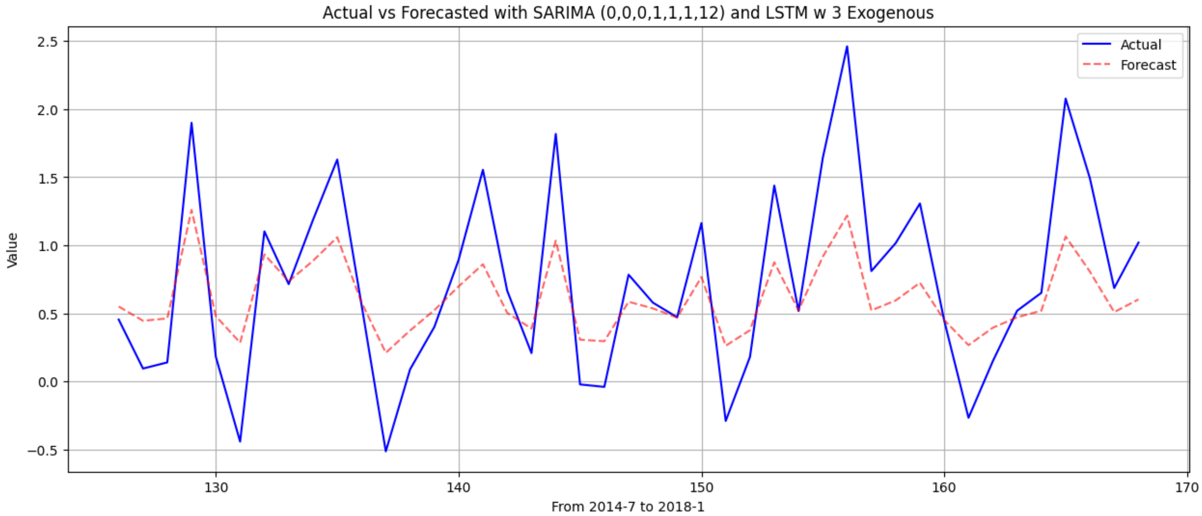
SARIMA and LSTM Outputs for subgroups of CPI

SARIMA and LSTM models were reworked for the subgroups of Turkey's inflation. These subgroups are food and non-alcoholic beverages, housing, and transportation, respectively. These three subgroups make up 57.12 percent of the total inflation index. Although these three groups are sub-groups that make up inflation, they are more exposed to different effects within themselves. The rates of being affected by the selected variables will also differ. However, it is seen that the standard deviations of these subgroups are higher than inflation. In the period used for the test set, the standard deviation of the headline inflation was 2.54%, while it was calculated as 3.4% for the food group, 3.08% for the housing group and 4.84% for the transportation group. Accordingly, it can be expected that the RMSE values of the prediction models will be higher. However, it should be noted that the hyperparameters of the SARIMA model are optimized according to the main inflation. Accordingly, 2 hyperparameters designed for SARIMA were applied to these three subgroups. While it was observed that the SARIMA(0,0,0,1,1,1,12) model suggested in the literature review produced significant results in all groups, the other model produced very unsuccessful results. When the SARIMA and LTSM models were combined in all groups, it was observed that the hit rate obtained from only the SARIMA model was improved. In models trained only with LTSM (with 3 exogenous variables), slightly less success than the SARIMA(0,0,0,1,1,1,12) model was observed for all 3 subgroups. It was observed that the LTSM model improved the outputs of the SARIMA model, which was trained for headline inflation, for each subgroup.

SARIMA and LSTM Outputs for different time spans

Under methodology section, the ongoing stationarity even monthly change for CPI data was mentioned. It's called a structural break when a time series abruptly changes at a point in time (Stata). Although this study concentrates on estimation methodologies rather than analysis of Turkey's inflation data on its own, this situation was also taken into account in order to construct a good model. At this point, SARIMA and LSTM tests were applied using the same parameters from January 2004 to January 2018. When the ADF test was performed, it was observed that the monthly CPI data was stable. Similarly, 75% train 25% test split approach was used. In order to develop a benchmark for this period, RMSE values were calculated for the same time period by using the median results of the expectation survey. Accordingly, the RMSE value of the forecast for the current month was 0.48, and the RMSE value of the forecast for the next month was calculated as 0.79. While the SARIMA(0,0,0,1,1,1,12) model showed a very successful performance with the RMSE value of 0.553, it was observed that the SARIMA parameters optimized for the main data set could not produce significant results. In the model in which the outputs of the SARIMA(0,0,0,1,1,1,12) model were given as input to the LSTM model with 3 explanatory variables, the RMSE value improved a little bit and was calculated at the level of 0.547. Thus, it was seen once again that the SARIMA and LSTM models exhibited a very competitive success together. When the number of LSTM units (neurons) increased from 100 to 2000 It has been observed an improvement to 0.466. Individual LSTM model with explanatory variables reached 0.667 RMSE value with 100 neurons while It has been seen 0.688.

Figure 6



When the second part of the data, that is, between January 2018 and January 2023, was scanned with the ADF test, it was observed that results did not identify non-stationarity in the monthly CPI data. Thus, the tests were repeated with two stationary time series, which are more significant econometrically. However, when the second time interval was reserved for the 25% test set, it was insufficient for training the data. Therefore, SARIMA and LSTM models were run once again with a 90-10% train test split set approach to gather an indicative result, although which is not the optimal scenario. The output of this structure validated the previous results. Best RMSE was reached with combined model of SARIMA(0,0,0,1,1,1,12) and LSTM with 100 neurons to1.52 while it was 2.42 with 2000 neurons. It may be an interesting point for further research to examine relationship between neurons and the pace (fluctuation level) of data. Meanwhile, the survey based benchmark results were calculated 1.70 and 1.74 for the current month and one month ahead expectations respectively. Thus, again SARIMA and LSTM combined model outperformed. With this timespan the negative correlation between accuracy rate and the LSTM units were seen for individual LSTM data as well. But this time LSTM data performed more accurate than SARIMA model. Overall, splitting data to 2 time spans, handled the stationarity problem of dependent variable and validate obtained result at previous sections.

Conclusion

In this study, inflation in Turkey between 2004 and 2023 was analysed and it was desired to estimate recent data. At this point, the use of classical econometric methods and advanced methods has been tested. For this purpose, studies that predict both Turkish inflation and other inflation indices such as main subgroups of headline inflation were examined. Methodologies based on classical econometric methods with a clustering approach of linearity and machine learning were conducted. Different models were constructed according to the results of the literature review, and the median results of the market participant survey conducted by the CBRT were used as a benchmark to compare the accuracy success of these models. RMSE values were taken as basis to examine the accuracy of the models.

According to outputs of state of art, combined model of SARIMA and LSTM outperformed than individual models and the market median expectations. High volatile pace of recent data was captured better with nonlinear models where the optimized hyperparameters SARIMA model worked better than all the individual models. According to findings of this study, SARIMA and LSTM combined models may be useful for CPI forecasting even under extreme volatility conditions. There is a room for analysing the relationship between deep learning parameters and forecast accuracy.

Bibliography

1. Akbulut, H. 2022. "Forecasting inflation in Turkey: A comparison of time-series and machine learning models," Economic Journal of Emerging Markets. Available at: https://journal.uii.ac.id/JEP/article/view/20760/13556 (Accessed: April 26, 2023).
2. Almosova, A. 2022. "Nonlinear inflation forecasting with recurrent neural networks," DOI: 10.1002/for.2901.
3. Araujo, G. S. & Gaglianone, W. P. 2020. "Machine learning methods for inflation forecasting in Brazil: new contenders versus classical models," Available at: https://www.sciencedirect.com/science/article/pii/S2666143823000042 (Accessed: April 26, 2023).
4. Aras, S. & Lisboa, P. 2022. "Explainable inflation forecasts by machine learning models," Expert Systems with Applications, 207, p. 117982. Available at: https://www.sciencedirect.com/science/article/pii/S0957417422012106 (Accessed: April 26, 2023).
5. Central Bank Of Turkey "Merkezin Güncesi," Available at: https://tcmbblog.org/.

6. Cho, Y. & Saul, L. 2009. "Kernel Methods for Deep Learning," Advances in Neural Information Processing Systems 22 (NIPS 2009).
7. Costa, J. & Cabral, M. 2022. "Statistical Methods with Applications in Data Mining: A Review of the Most Recent Works" ," Available at: https://doi.org/10.3390/math10060993 (Accessed: July 26, 2023)
8. Deng, L. & Yu, D. 2014. "Deep Learning: Methods and Applications," Foundations and Trends® in Signal Processing, 7(3-4), pp. 197-387.
9. Diksa, G. B. N., Kuswanto, H., & Fithriasari, K. (2023). Forecasting Indonesia inflation using long short term memory method.
10. Galesi, A. & Lombardi, M. J. 2009. "External Shocks and International Inflation Linkages: A Global VAR Analysis," ECB Working Paper No. 1062.
11. Globan, T., Arcabic, V. & Soric, P. 2014. EFZG Working Paper Series, No. 05.
12. Goodfellow, I., Bengio, Y. & Courville, A. 2016. Deep Learning. MIT Press. Available at: https://www.deeplearningbook.org/ (Accessed: April 26, 2023).
13. Gordon, R. J. 1981. "Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment," DOI 10.3386/w0708.
14. Gujarati, D. N. & Porter, D. C. 2010. Essentials of Econometrics, Fourth Edition, McGraw-Hill International Edition.
15. Haider, A. & Hanif, M. N. 2009. "Inflation Forecasting in Pakistan Using Artificial Neural Networks," Available at: https://www.jstor.org/stable/25825345 (Accessed: April 26, 2023).
16. IBM. "Neural Networks," Available at: https://www.ibm.com/topics/neural-networks (Accessed: April 26, 2023).
17. IBM. "What is Artificial Intelligence (AI)?," Available at: https://www.ibm.com/topics/artificial-intelligence.
18. Kara, A. H., Sarikaya, C. & Ogunc, F. 2016. "WHICH FACTORS HAVE BEEN DETERMINANT IN INFLATION DYNAMICS OVER THE PAST DECADE?"
19. Kara, H. & Öğünç, F. 2008. "Inflation Targeting and Exchange Rate Pass-Through: The Turkish Experience," Emerging Markets Finance and Trade, 44(6), pp. 52-66.
20. Klapper-Rybicka, M, & Schraudolph, N. & Schmidhuber, J. Unsupervised Learning in LSM Recurrenct Neural Networks
21. Kucukefe, B. 2018. "Forecasting Inflation Using Summary Statistics of Survey Expectations: A Machine-Learning Approach," Available at: https://dergipark.org.tr/en/pub/endustrimuhendisligi/issue/53786/669930 (Accessed: April 26, 2023).
22. McKinsey. 2022. "What is inflation: The causes and impact," Available at: https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-inflation (Accessed: August 17, 2022).
23. MIT News. 2017. "Explained: Neural networks and deep learning," Available at: https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414 (Accessed: April 26, 2023).

24. Paranhos, L. 2021. "Predicting Inflation with Neural Networks," University of Warwick.

25. Rodríguez-Vargas. 2020. "Forecasting Costa Rican Inflation Using Neural Networks: The Role of Macroeconomic Variables," UCB.

26. Rosenblatt, F. 1958. "The perceptron: A probabilistic model for information storage and organization in the brain," Psychological Review, 65(6), pp. 386–408.

27. Savitriab, F. F., Siregarbc, R. F., Hariantobd, F. Y. & Napitupulube, H. 2023. "Forecasting Inflation in Indonesia using Long Short Term Memory," Universitas Padjadjaran Sumedang, Indonesia.

28. Saz, G. 2011. "The Efficacy of SARIMA Models for Forecasting Inflation Rates in Developing Countries: The Case for Turkey," International Research Journal of Finance and Economics, Issue 62.

29. Sen, S. & Raghunathan, A. 2018. "Approximate Computing for Long Short Term Memory (LSTM) Neural Networks," IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS, 37(11).

30. Stanford University. "CS 230 - Recurrent Neural Networks Cheatsheet," Available at: https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks.

31. Stock, J. H. & Watson, M. W. 1998. "A Comparison of linear and nonlinear univariate models for forecasting macroeconomic time series," NBER Working Paper Series.

32. Stock, J. H. & Watson, M. W. 2001. "Vector Autoregressions," Journal of Economic Perspectives, 15(4), pp. 101-115.

33. Tests for structural breaks in time-series data | Stata (Accessed: August 28, 2022).

34. Theoharidis, A. F., Guillén, D. A. & Lopes, H. "Deep Learning Models For Inflation Forecasting," Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/asmb.2757 (Accessed: April 26, 2023).

35. Yang, C. & Guo, S. 2021. "Inflation Prediction Method Based on Deep Learning," Hindawi Computational Intelligence and Neuroscience, 2021, Article ID 1071145. Available at: https://www.hindawi.com/journals/cin/2021/1071145/ (Accessed: April 26, 2023).