Comparing the performance of various forecasting models upon the inflation data

Abstract

This paper compares three forecasting models: Long Short-Term Memory (LSTM), Autoregressive Integrated Moving Average (ARIMA), and random forest regressors, focusing on inflation prediction. We trained the model on 10-year monthly data of inflation values. We highlight the significance of inflation forecasting and the transformative impact of artificial intelligence on predictive analytics. Using established metrics like R-squared (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), we evaluate each model's performance. Our findings provide insights into their predictive accuracy and strengths of each model but also acknowledges the limitations of our research such as limited data to test for, non-stationary data set and unreliability of certain models on their own for time-series forecasting.

Introduction Background & motivation

In the ever-evolving landscape of forecasting models, the rapid advancements in artificial intelligence have ushered in a new era of innovation. With an array of techniques emerging, from traditional statistical methods to cutting-edge machine learning algorithms, the field of forecasting has witnessed a surge in diversity and efficacy.

This paper endeavors to delve into the realm of forecasting by conducting a comparative analysis of prominent models, including Long Short-Term Memory (LSTM), Autoregressive Integrated Moving Average (ARIMA), and random forest regressors. Our chosen arena for evaluation is the realm of inflation data—a cornerstone of economic analysis and decision-making. By subjecting these models to the rigors of inflation forecasting, we aim to discern their respective strengths and weaknesses.

The significance of inflation forecasting cannot be overstated. It serves as a compass for policymakers, guiding the formulation of monetary policies essential for economic stability. Likewise, businesses rely on accurate predictions to strategize pricing, production, and resource allocation. Investors, too, are keenly attuned to inflation trends, as they shape investment strategies and portfolio diversification.

To gauge the effectiveness of each forecasting algorithm, we employ a battery of metrics, including the widely recognized R-squared (R^2), Mean Absolute Error (MAE), and Root Mean

Squared Error (RMSE). These metrics serve as benchmarks, illuminating the predictive prowess of each model and facilitating informed comparisons.

Literature Review

Staudemeyer, R. C., & Morris, E. R. (2019)[1] break down the architecture and operations of LSTM networks in a step-by-step manner, making it easier for readers to grasp concepts behind this type of neural network. The paper explores previous well-known works apropos LSTM such as the paper by Sepp Hochreiter and Jürgen Schmidhuber in 1997, which introduced the LSTM, titled "Long Short-Term Memory" and later research upon this model, covering topics such as Basic Structure of an LSTM cell, Gates and their functions, Gradient flow and vanishing gradients, Backpropagation and few applications.In conclusion, this paper is a good start for someone willing to learn about LSTM but is also fairly complex so one must have prior knowledge of certain machine learning topics.

Kelikume, I., & Salami, A. (2014)[2] in their paper aim to provide insights into the factors driving inflation in the Nigerian economy. The paper utilizes ARIMA, SARIMA and also methods like LSTM and Vector AutoRegression. Data on key economic indicators such as Consumer Price Index, Retail Price Index.The research was able to produce better accuracy models that helped gain insights of future values of inflations in Nigeria.

Hafer, R. W., & Hein, S. E. (1990) [3] explore how certain forecasting methods have fared over. Particularly, this paper focuses on comparing the metrics of performance obtained by time series forecasting and interest-based models such as the Fisher equation or the expectations hypothesis.

The paper by Livia Paranhos[4] predicts inflation rate particularly using ordinary recurrent neural networks and LSTMs to obtain the best possible forecasting accuracy.

Research Questions

Compare a few of the popular time-series forecasting models.

Data and Methodology

Data

We've collected data for 10 years, monthly, for the inflation rates and below are the following parameters.

Table 1: Description of variables

Sl No	*Name*	*Description*	*Source*

1	INFR	Inflation rate	https://tradingeconomics.com/india/inflation-cpi
2	Time	Month corresponding to the inflation value recorded	

Method

Before training each model, the data is split into two parts: training set and testing set. The training set is used to train each model while the testing set is used to determine their performance. The first 80-85% data points are considered for training while the remaining are considered for testing. We used popular packages such as Scikit-learn, Prophet and Torch, which contained sort-of a base model of each machine learning model, upon which we built our required models. The following sections give a brief on how each model functions and key equations governing each one.

Time series algorithms such as ARIMA (AutoRegressive Integrated Moving Average) are widely used for forecasting economic indicators like inflation. ARIMA models leverage past observations to make predictions, making them suitable for capturing temporal patterns in inflation data. Below is a general equation for the ARIMA model:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + (1 - B)^d X_t + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

where: Y_t is the differenced series at time t
c is a constant, ϕ_1 , ϕ_2 … ϕ_p are the AR coefficients,
p is the order of the autoregressive component
d is the order of differencing, θs are the moving average coefficients, q is the order of the moving average, X_t is the original time series, εt is the error term at time t

We determined values of p, d and q using the auto_arima function from the pmdarima package available for python. This function automatically determines the optimal values for p, d and q via brute force on a small search space.

Next we utilized the arima module from sci-kit to train our data using the optimal p, d and q values and determined the error metrics.

LSTM Neural Networks :

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to handle long-term dependencies in sequential data. LSTMs use memory cells and gates to selectively store and retrieve information, making them effective for tasks like speech recognition, language translation, and time series forecasting. The following are the important equations pertaining to LSTM:

LSTM gate equations:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

Cell State update:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

Hiddes state:

$$h_t = o_t \odot \tanh(C_t)$$

Where:
f_t: Forget gate output
i_t: Input gate output
o_t: Output gate output
$\tilde{C}_t$: Candidate cell state
C_t: Updated cell state
h_t: Hidden state
W_f, W_i, W_o, W_c: Weight matrices
b_f, b_i, b_o, b_c: Bias vectors
σ: Sigmoid activation function
□: Element-wise multiplication

Random Forest Regressor :

Random forest regressors are ensemble learning algorithms that operate by constructing a multitude of decision trees during training. These trees then independently predict the inflation rate, and the final prediction is an average of all the individual tree predictions. Random forest regressors excel at handling large datasets with many features and can capture nonlinear relationships effectively.

For training using random forest regressor, we used the random forest regressor from the scikit-learn package.

Results and Discussion

The following is the plot for inflation rates plotted against dates. Also given are the ACF, PACF plots and test results.

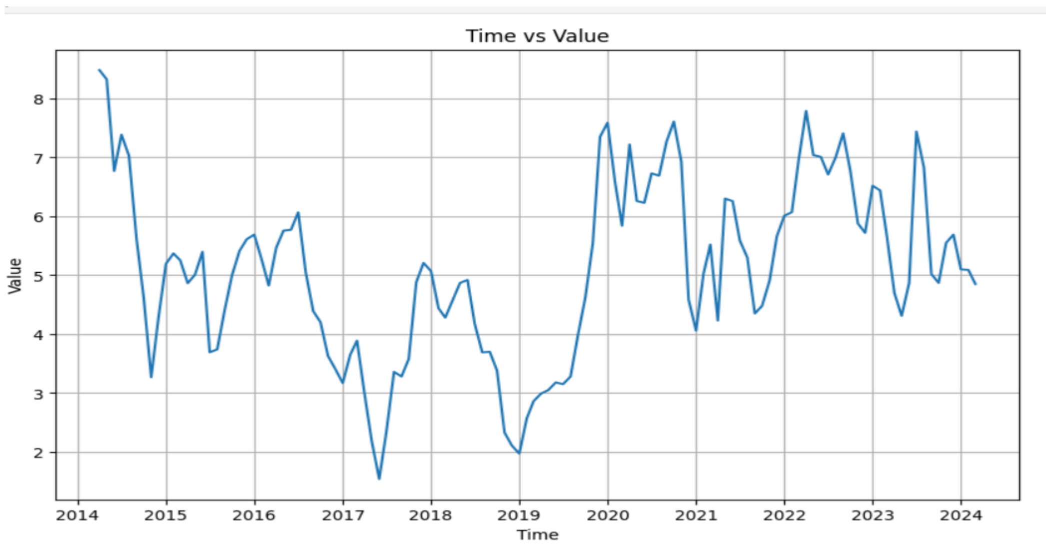

fig-1

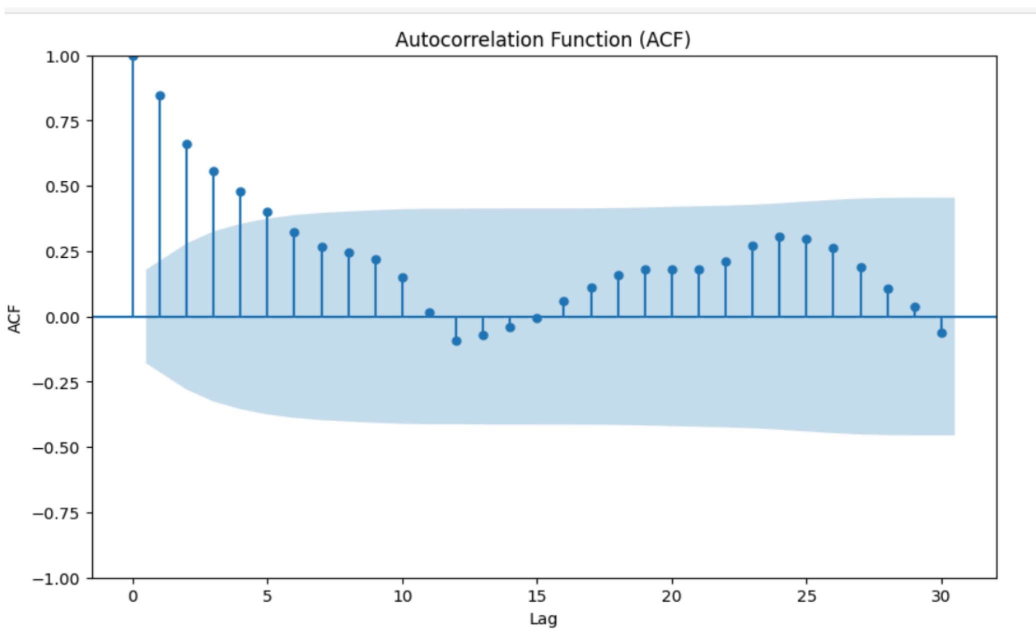

fig-2

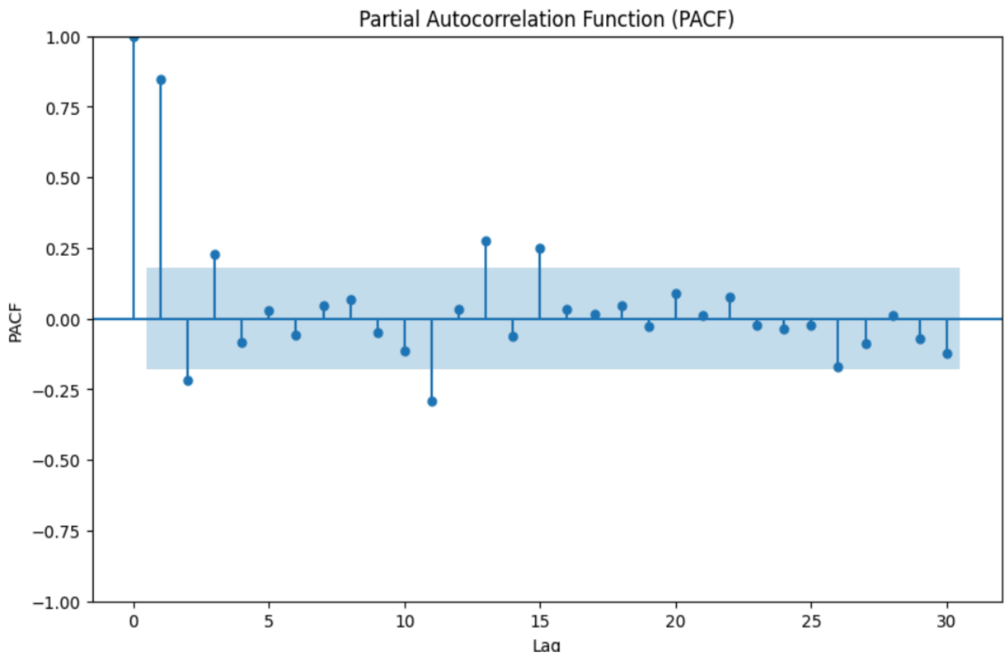

fig-3

```
ADF Statistic: -1.9776348492021176
p-value: 0.29646470451809726
Critical Values:
   1%: -3.492995948509562
   5%: -2.888954648057252
   10%: -2.58139291903223
```

fig-4

The table below depicts the error metrics for each of the models.. These values correspond to the optimal values of the hyperparameters.

Metrics of testing data:

Model	MAE	RMSE
LSTM	1.1572	2.010
ARIMA	1.7129	3.5997
Random-forest regressor	1.6326	1.8186

Metrics of training data:

Model	MAE	RMSE
LSTM	1.0285	1.1205
ARIMA	0.6126	1.5528
Random-forest regressor	0.1574	0.2206

Conclusion

Looking at the performances, it looks like the ARIMA model is the worst of them all. LSTM had a low MAE but a high RMSE as compared to Random forest regressor. This indicates that the random-forest regressor is good at capturing typical error size while LSTM consistently gives larger errors.

But the above conclusion regarding which model is better is not something we can take home. There are limitations to the research conducted; the major of the ones are listed below.

Firstly, we cannot assess the relative performance of models based solely upon their performance on a single dataset. We need to consider at least a good number of datasets to be able to come to a proper conclusion. A further improvement on this research would be to consider a couple more time-series datasets, which need to be reliable, and then find a model that consistently outperforms the others upon each dataset.

Regarding this dataset, from the p-value and comparison of t-statistic with critical values (fig 1-4), it is clear that the data is non-stationary. Hence, the ARMA models can be pointing towards misleading and inconclusive results. To overcome this issue, we might need to take much larger datasets as compared to just 10-year data or use any other familiar data, which we know beforehand to be stationary. Additionally, it would be advisable to contemplate additional dependent variables that may directly impact the variable being forecasted.

Random forest regressor shows a lot smaller errors as compared to the errors shown on the testing data. This usually means overfitting the training data and random forest regressors do tend to overfit data. We might need to create a hybrid model to ensure propitious results.

Another point to note is that there is constant research going on in each of these following models. Many research papers consider hybrid models as opposed to relying on a single model. One such hybrid model combines ARIMA model with Elman Artificial Neural Network(ANN) outperformed traditional ARIMA model [5]. Thus, we must also focus our attention on trying to optimize our forecasting models to achieve best results possible.

Bibliography

- Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks. arXiv preprint arXiv:1909.09586 [1]
- Hafer, R. W., & Hein, S. E. (1990). Forecasting Inflation Using Interest-Rate and Time-Series Models: Some International Evidence. *The Journal of Business, 63*(1), 1-17. https://www.jstor.org/stable/2353234 [3]
- Kelikume, I., & Salami, A. (2014). Time series modeling and forecasting inflation: Evidence from Nigeria. *The International Journal of Business and Finance Research, 8*(2), 41-51. Available at SSRN: https://ssrn.com/abstract=2322918 [2]
- Paranhos, L. (2021). Predicting inflation with recurrent neural networks. *arXiv preprint arXiv:2104.03757*. Retrieved from https://doi.org/10.48550/arXiv.2104.03757 [4]
- Xiao, Yi & Xiao, Jin & Wang, Shouyang. (2012). A hybrid model for time series forecasting. Human Systems Management. 31. 133-143. 10.3233/HSM-2012-0763.[5]