IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 5, October 2025, pp. 3887~3896
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i5.pp3887-3896

Accuracy of long short-term memory model in predicting YoY inflation of cities in Indonesia

Harfely Leipary, Adi Setiawan
Master of Data Science Program, Faculty of Science and Mathematics, Satya Wacana Christian University, Salatiga, Indonesia

Article Info

Article history:

Received Jan 7, 2025
Revised Jul 1, 2025
Accepted Jul 13, 2025

Keywords:

Coefficient of determination
Inflation
Long short-term memory
Mean absolute percentage error
Root mean square error

ABSTRACT

Our research evaluates the effectiveness of the long short-term memory (LSTM) model in forecasting annual year-on-year (YoY) inflation across 82 cities in Indonesia based on time series data from BPS economic reports for 2014-2024. This study tests the accuracy of the model in reconstructing past inflation patterns, then evaluates the capabilities and limitations of the model in various urban area contexts with the root mean square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2) metrics. The findings show that LSTM performs well in metropolitan areas such as Jakarta, Bandung, and Surabaya with R^2 values >0.8 and the lowest MAPE of 10.91% in Jakarta. However, in small cities with higher economic volatility such as Tanjung Pandan, the model shows significant prediction errors (R^2 <0.50 and MAPE up to 283.11%). Moderate performance (0.50≤ R^2 ≤0.80) was found in cities such as Palembang, Semarang, and Makassar, reflecting the model's adaptive ability to moderate inflation patterns. These results emphasize the important role of structured economic data in improving the reliability of predictions, so that the policy implications of this study include the use of the LSTM model as an early warning system by fiscal and monetary authorities, as well as the need for a data-based inflation control strategy to strengthen regional and national economic resilience in supporting sustainable development towards Indonesia Emas 2045.

Corresponding Author:

Adi Setiawan
Master of Data Science Program, Faculty of Science and Mathematics, Satya Wacana Christian University Salatiga, Indonesia
Email: adi.setiawan@uksw.edu

1. INTRODUCTION

Inflation is one of the fundamental macroeconomic indicators to assess the economic stability of a country or region. Inflation reflects the increase in the price of goods and services over a certain period of time, which directly affects purchasing power, financial stability, and economic growth. In Indonesia, inflation is an important issue, especially at the city level, due to economic decentralization, which causes structural heterogeneity in the regional economy, fiscal policy, and responses to global dynamics such as the influence of the rupiah exchange rate against foreign currencies [1]–[4].

Accurate inflation forecasting is essential for effective fiscal and monetary policy management. Reliable predictions enable policymakers to implement strategic measures to maintain price stability and support sustainable economic growth. Advances in artificial intelligence certainly play an important role [5]. One method that shows great potential in time series data analysis such as this study is long short-term memory (LSTM). LSTM is one of the artificial neural network architectures designed to capture complex patterns and long-term dependencies in sequential data. Research by Sumarjaya and Susilawati [6] shows the success of LSTM in predicting monthly inflation in Denpasar, with accurate prediction results. However, this study only covers one city and has not considered the aspect of structural heterogeneity in other regions. Other studies [4] emphasize the relationship between inflation and economic growth, but do not explore the role of AI-based methods in inflation prediction. Research by Vargas [7] compared several machine learning methods, including LSTM, to predict inflation in Costa Rica. LSTM proved to be one of the best performing methods. This finding indicates the great potential of LSTM in inflation prediction, although this study is limited to the Costa Rican context and has not explored heterogeneity across regions. Further research [2] applied various machine learning algorithms to predict inflation during the economic crisis in Sri Lanka.

Therefore, there is still a research gap in previous studies. Gaps such as geographical differences, have not discussed the inflation specifications of cities in Indonesia, model assessment indicators are still rare using the coefficient of determination (R^2). So, it is clear that the novelty offered is to evaluate the accuracy of the LSTM model in predicting monthly inflation for 82 cities in Indonesia, which has never been done comprehensively. By integrating a data-driven approach and contextual analysis at the city level, this study makes a significant contribution to improving the quality of predictions. The results are expected to provide an in-depth understanding of the inflation patterns in regions of Indonesia, support data-driven policy making, and promote price stability and sustainable economic growth.

2. METHOD

2.1. Long short-term memory

The LSTM model as a variant of the recurrent neural network (RNN), as shown in Figure 1, begins with data collection and cleaning [8]. Inflation data from various cities is organized, then the cleaning process is carried out by removing irrelevant columns and removing not a number (NaN). Furthermore, a lag feature is added to include inflation values in previous months as a predictor, which helps LSTM understand temporal patterns [9]. The processed data is then normalized using standard methods to ensure that all features are on the same scale. This speeds up the training process. After the data is ready, the data is divided into two parts, namely training and testing data. Then the data dimensions are converted into a three-dimensional format to meet the input needs of the LSTM: number of samples, amount of time, and number of features. The LSTM model is designed with several layers, starting from the input layer, followed by several layers equipped with dropout layers to prevent overfitting, and ending with a dense layer that produces output in the form of inflation prediction values as output. The model is trained using Adam optimization and mean squared error (MSE) loss function to minimize the difference between predicted and actual values. After training is complete, the model is evaluated using metrics such as mean absolute error (MAE), mean absolute percentage error (MAPE), and R^2 [10]–[13]. If the validation results show that the model performs well, then the model is applied to the test data, and the predicted results are compared with the actual data. However, if the model does not perform well based on the evaluation metrics, hyperparameter tuning is performed to improve accuracy. After tuning, the model is evaluated again. If there is an improvement in performance, the model goes back to the validation stage to ensure the accuracy of the results. On the other hand, if the model has not improved significantly, the next step is to make further improvements, such as changing the preprocessing method or updating the number of epochs. Once the best model is obtained, the process can proceed to the implementation and policy analysis stage.

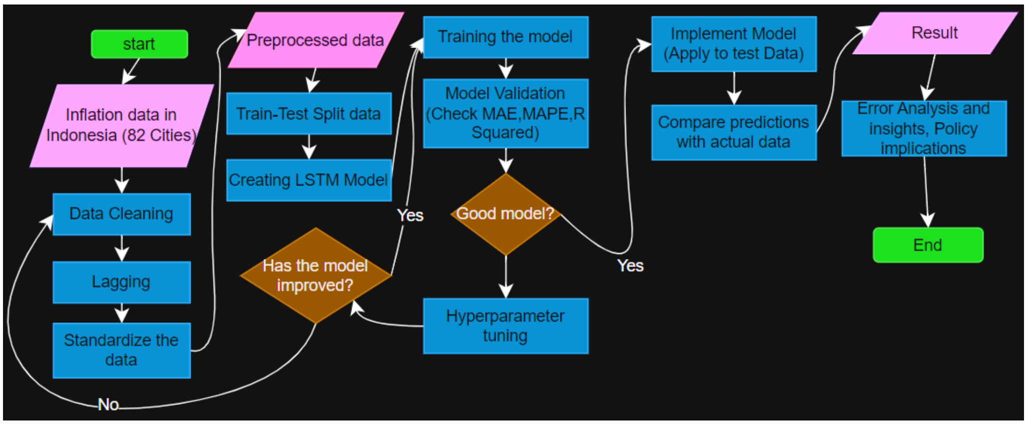

Figure 1. Research flowchart LSTM

2.2. Model validation

2.2.1. Mean absolute error

MAE is a metric used to evaluate the accuracy of a forecasting model. A smaller MAE value indicates a higher level of accuracy and smaller average prediction errors [14]–[16]. MAE is formulated in (1):

$$MAE = \frac{1}{n}\sum_{i}^{n}|f_i - y_i| \tag{1}$$

in (1), f_i represents the predicted value for the i-th data point where i=1,2,…,n. The variable y_i denotes the actual value corresponding to the i-th data point, with i taking the same range. Lastly, n refers to the total sample size used in the calculation.

2.2.2. Root mean squared error

Root mean square error (RMSE) is the average sum of the squares of the errors, and can also be defined as a measure of the error produced in a forecast or prediction model [17]. A lower value indicates a better RMSE value. The RMSE value can be found using (2) [15], [18], [19]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(f_i - y_i)^2}{n}} \tag{2}$$

in (2) for f_i represents the forecasted result value (prediction) for the i-th data point, where i ranges from 1 to n. The variable y_i indicates the actual observed value corresponding to the i-th data point within the same range. Lastly, n denotes the total size of the sample used in the analysis.

2.2.3. Mean absolute percent error

MAPE is a measure that shows the level of relative error by presenting it in percentage form. It states the percentage error of the forecast results with actual demand in a certain period of time. This value indicates the percentage of error and provides an overview of predictions that are made too high or low compared to the actual data [15], [17], [20]. The MAPE value is found using (3). In (3) described f_i as the predicted value for the i-th data point, y_i represents the corresponding actual value for the i-th data point, and n is the total number of data points in the sample [19], [21].

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - f_i|}{|y_i|} \tag{3}$$

2.2.4. Coefficient of determination

The R^2 is a statistical metric that shows how well a regression model predicts or explains the variability of the target data formulated according to (4). R^2 formula in (4) is the square of the correlation coefficient (r) between the actual value (y_i) and the predicted value ($\hat{y}_i$) used to measure the extent to which the predicted value can explain the variance of the actual value. After that, the r value is squared to obtain R^2, which shows the proportion of the actual value variance that can be explained by the prediction model. The r value (correlation coefficient) ranging from -1 to 1 in this study was not used considering that the definition of the value (negative) is difficult to interpret while R^2 which ranges from 0 to 1, where a value of 1 means that the prediction model is fully able to explain the variance of the actual value, while a value of 0 indicates that the model does not explain the variance in the actual value [15], [21]–[24].

$$R^2 = \left(\frac{\sum(y_i - \bar{y})(y_i - \bar{\hat{y}})}{\sqrt{\sum(y_i - \bar{y})^2 . \sum(\hat{y}_i - \bar{\hat{y}})^2}}\right)^2 \tag{4}$$

2.3. Data and scoring

This study uses annual inflation rate data consisting of 115 rows, i.e start from December 2014 until June 2024. In this case, 82 cities analyzed in the dataset includes Meulaboh, Banda Aceh, Lhokseumawe, Sibolga, Pematangsiantar, Medan, Padangsidimpuan, Padang, Bukittinggi, Tembilahan, Pekanbaru, Dumai, Bungo, Jambi, Palembang, Lubuklinggau, Bengkulu, Bandar Lampung, Metro, Tanjung Pandan, Pangkalpinang, Batam, Tanjung Pinang, Jakarta, Bogor, Sukabumi, Bandung, Cirebon, Bekasi, Depok, Tasikmalaya, Cilacap, Purwokerto, Kudus, Surakarta, Semarang, Tegal, Yogyakarta, Jember, Banyuwangi, Sumenep, Kediri, Malang, Probolinggo, Madiun, Surabaya, Tangerang, Cilegon, Serang, Singaraja, Denpasar, Mataram, Bima, Maumere, Kupang, Pontianak, Singkawang, Sampit, Palangka Raya, Tanjung, Banjarmasin, Balikpapan, Samarinda, Tarakan, Manado, Palu, Bulukumba, Watampone, Makassar, Pare-Pare, Palopo, Kendari, Baubau, Gorontalo, Mamuju, Ambon, Tual, Ternate, Manokwari, Sorong, Merauke, and Jayapura. After the data is cleaned, normalization is carried out using the standardization method to ensure a uniform data scale, with the formula in (5) where x is the actual value, and is the average value and is the standard deviation (σ):

$$z = \frac{x-\mu}{\sigma} \quad (5)$$

The dataset was then divided into a training set (80%) and a test set (20%), and its dimensions were converted to a three-dimensional format ($n \times t \times f$), where n is the number of samples, t is the number of lag features (lag time), and f is the number of features. In the model building stage, an LSTM architecture was designed with an input layer to receive the three-dimensional data, followed by multiple LSTM layers to capture temporal patterns with gradually decreasing units (256 → 128 → 64). A dropout layer was added to prevent overfitting, and a dense layer was used to generate monthly inflation predictions. The model was optimized using the MSE loss function and the Adam algorithm.

The model training process was carried out with parameters such as 100 epochs, a batch size of 16, and using an early stopping callback to stop training if the validation performance does not improve after 15 epochs, and ReduceLROnPlateau to adaptively reduce the learning rate. After training, the model was evaluated using metrics such as MAE, MAPE, and R^2. The evaluation results were compared for each city, and a prediction visualization in the form of an actual vs predicted graph was created to understand the model performance.

3. RESULTS AND DISCUSSION

Figure 2 shows the Annual Inflation Graph for the period 2015 to 2024 for five cities in Indonesia, which shows quite dynamic inflation rate fluctuations in the areas of Meulaboh, Banda Aceh, Lhokseumawe, Sibolga, and Pematang-Siantar. The x-axis represents the year, while the y-axis shows the inflation rate. Each city is represented by a different color and marker, indicating inflation fluctuations over time. The graph shows marked volatility, especially in the early years, with sharp peaks and troughs, indicating periods of high and low inflation. From 2020 onwards, there has been a marked decline followed by an upward trend that peaked around 2022 before stabilizing in recent years. This pattern suggests that external economic factors, such as global disruptions and domestic policies, have affected inflation dynamics differently across cities.

Figure 3 describes the comparison of inflation rates between Merauke City and Jayapura City, there are the last 2 cities in data. Both cities showed quite significant fluctuations in inflation rates during the period. Merauke tends to have a higher inflation rate than Jayapura, especially at the beginning of the observation period. However, over time, the difference in inflation rates between the two cities tends to narrow. There are several peaks and valleys of inflation that occur in both cities, indicating seasonal factors or certain economic events that affect price levels in both areas.

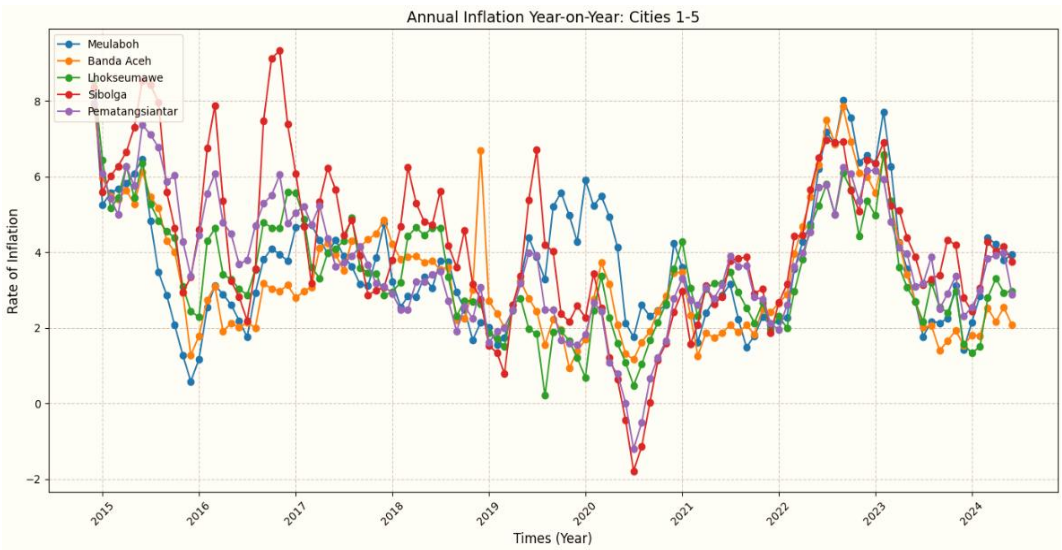

Figure 2. Annual inflation of 5 early cities

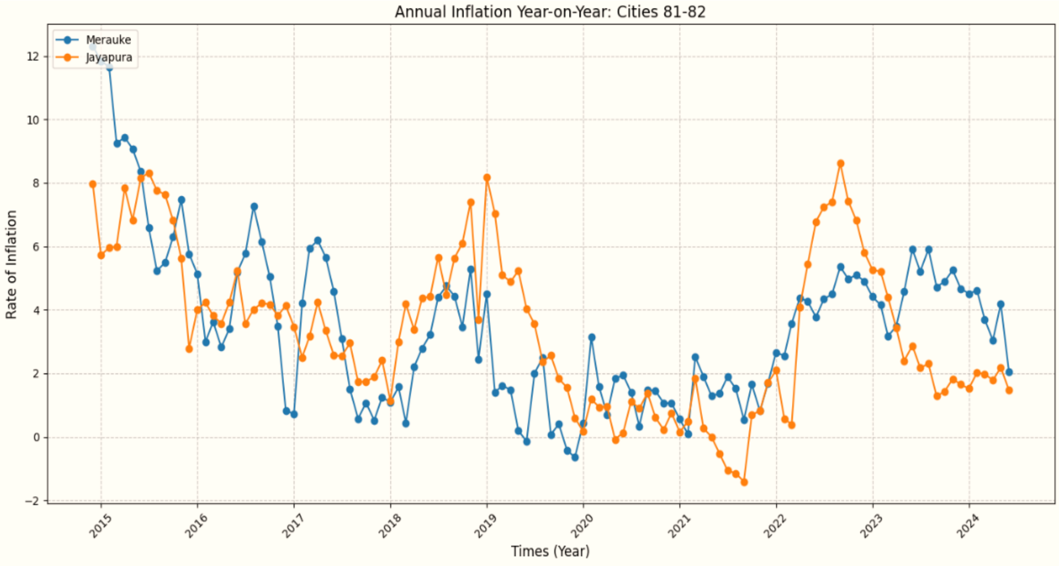

Figure 3. Annual inflation of the last 2 cities

Table 1 reflects the yearly inflation rates in five cities are Meulaboh, Banda Aceh, Lhokseumawe, Sibolga, and Pematangsiantar. Each row in the table represents the inflation rate for a certain period that is consistent across the five cities. In general, inflation in Lhokseumawe tends to be higher than other cities, with a maximum value reaching 8.53 (12/01/2014), indicating that the city may experience greater price pressures or more significant economic volatility than other cities.

Table 1. First eleven rows of data for the first five cities and statistical results of the cleaned dataset

YoY	Meulaboh	Banda Aceh	Lhokseumawe	Sibolga	Pematangsiantar
12/01/2014	8.2	7.83	8.53	8.36	7.94
12/01/2015	0.58	1.27	2.44	3.34	3.36
12/01/2016	3.77	3.13	5.6	7.39	4.76
12/01/2017	4.76	4.86	2.87	3.08	3.1
12/01/2018	2.15	6.7	2.63	2.76	3.08
12/01/2019	4.28	1.38	1.2	2.58	1.54
12/01/2020	4.24	3.46	3.55	2.42	2.78
12/01/2021	2.07	2.41	1.97	1.86	2.12
12/01/2022	6.56	6	5.37	6.43	6.16
12/01/2023	1.42	1.53	1.56	2.8	2.3
06/01/2024	3.94	2.07	2.97	3.75	2.89

In line with Tables 1 and 2 shows yearly inflation data in five cities in eastern Indonesia, namely Ternate, Manokwari, Sorong, Merauke, and Jayapura. Each row represents the inflation rate for the same time period among the five cities. Merauke City shows a consistently higher inflation rate than other cities, with a peak value reaching 12.31, indicating the potential for significant economic pressure in the region. In contrast, Manokwari tends to have the lowest and most stable inflation rate, with values mostly ranging around 5–6. Other cities, such as Ternate and Sorong, have fluctuating inflation rates but tend to be in the middle range. Meanwhile, Jayapura shows a more varied pattern, with inflation rates sometimes approaching the highest values in this dataset.

Table 2. Seven rows of the last 5 cities of the cleaned data set statistics results

Ternate	Manokwari	Sorong	Merauke	Jayapura
9.34	5.7	6.83	12.31	7.98
8.26	5.66	7.2	11.84	5.72
8.11	5.39	6.81	11.65	5.96
7.92	6.64	7.11	9.25	5.99
7.83	6	7.06	9.43	7.83
8.64	5.33	6.62	9.07	6.82
8.22	6.15	8.93	8.35	8.15

The plot in Figure 4 illustrates the comparison between the actual value and the predicted value of inflation in Tanjung Pandan City, which is the city with the lowest R^2 value among other cities with a value of 0.48 based on the data used. The horizontal axis (x-axis) represents the time index or a certain period, while the vertical axis (y-axis) shows the inflation value. The blue line depicts the actual value, and the orange line shows the predicted results from the model. From this visualization, it can be seen that the prediction model is quite good at capturing the general pattern of the actual data, although there are some striking differences in certain periods. The actual data shows larger fluctuations, especially in the period around the 5th and 13th index, where there is a sharp increase in the actual data that is not fully followed by the prediction model. In contrast, the prediction tends to be smoother with smaller fluctuations. In addition, in some parts, the predicted value is higher than the actual value, for example in the period around the 10th to 12th index. Conversely, in the final period, the actual value drops more drastically than the prediction. This difference indicates that the model has limitations in capturing extreme variations or spikes that occur in the actual data.

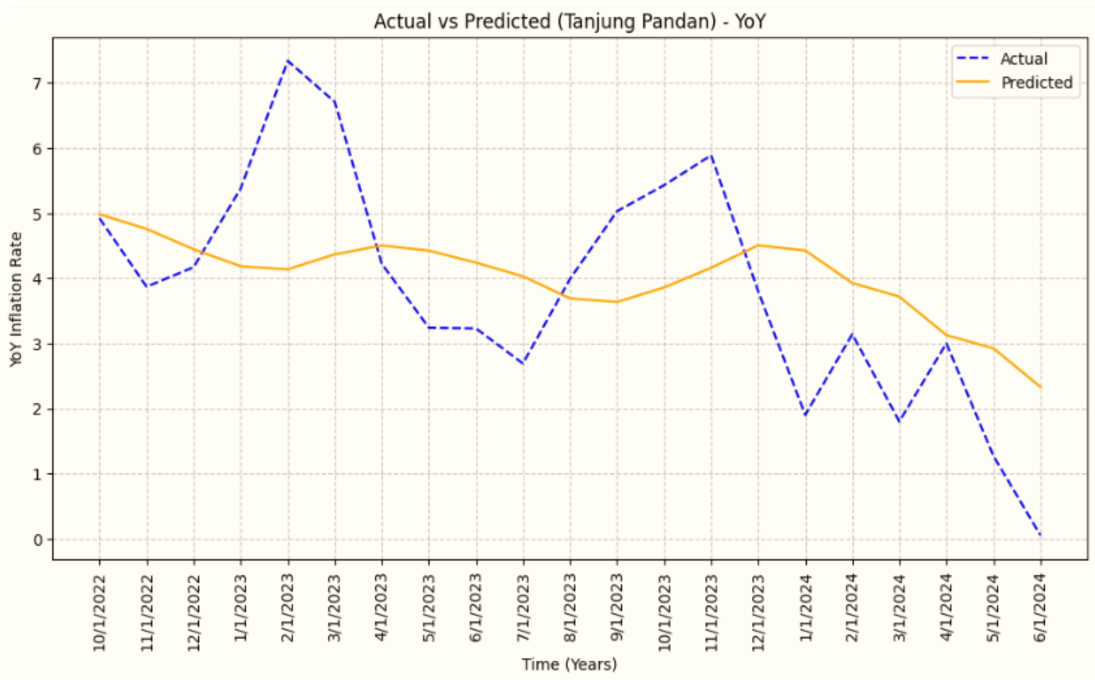

Figure 4. Results of actual data analysis and predictions for Tanjung Pandan City

Figure 5 represents a comparison between the actual and the predicted inflation values for Malang City based on the given data. According to the R^2 result, Malang City has the highest value of all the cities, at 0.92. The horizontal axis (x-axis) represents the time index or a certain period, while the vertical axis (y-axis) shows the inflation value. The blue line depicts the actual data, while the orange line depicts the predicted results from the model used. From this visualization, it can be observed that the trend of the predicted data follows the general pattern of the actual data, although there are significant differences at certain points. The actual data shows sharper fluctuations, especially around the 10th time index, where there is a drastic decrease in the actual value, while the prediction remains in a smoother decreasing pattern. In the next period, the prediction tends to approach the actual value, but still shows a small deviation. This plot reflects that the prediction model is quite capable of capturing the general pattern of the inflation trend in Malang City, but is less sensitive to sudden fluctuations or extreme changes that occur in the actual data.

Table 3 describes the distribution of cities with the lowest R^2 values. The analysis results show model's performance in predicting monthly inflation varies across these four cities, with accuracy levels ranging from low to quite good. Tanjung Pandan City recorded a value R^2 of 0.489, indicating low performance, with a very large error rate such as MAPE reaching 292.78%. This indicates that the inflation pattern in the city is difficult to capture by the model, possibly due to fluctuating data or significant anomalies [25].

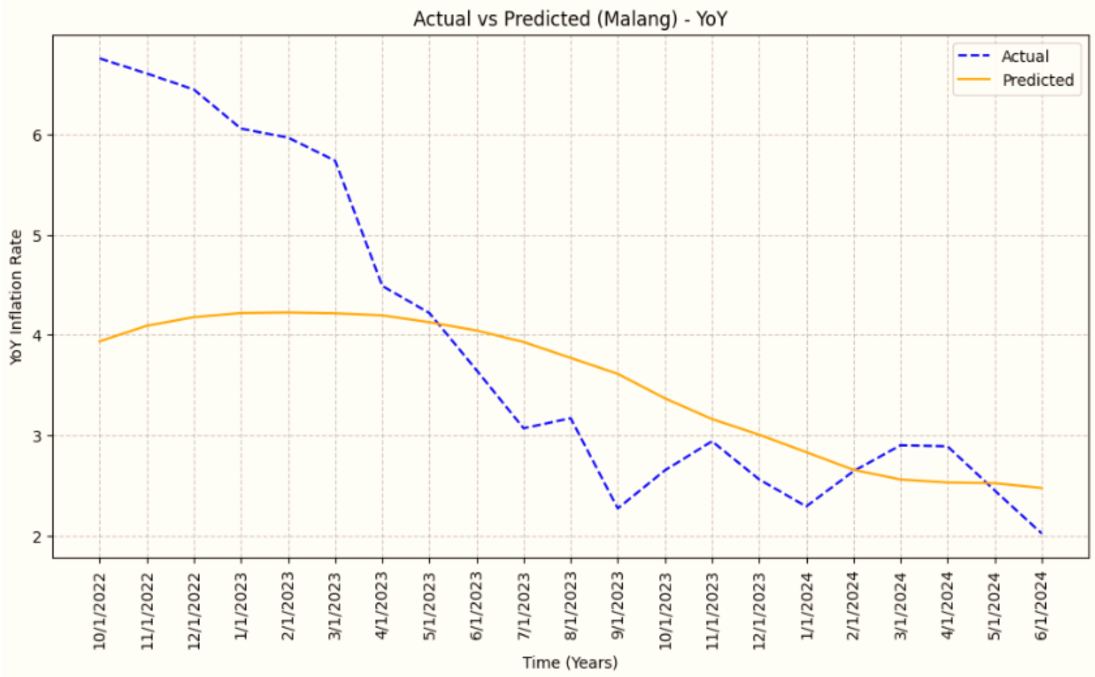

Figure 5. Results of actual data analysis and predictions for Malang City

Table 3. Distribution of cities with the lowest R^2 values ($R^2 \leq 0.60$)

City	MAPE	R^2
Tanjung Pandan	283.11	0.49
Manado	29.30	0.56
Mamuju	37.86	0.58
Ternate	17.41	0.58
Manokwari	19.10	0.60

Table 4 shows the performance of LSTM in predicting monthly inflation rates of 82 cities in Indonesia using four metrics MSE, MAE, MAPE, and R^2. The model performs well in cities such as Tasikmalaya, Malang, and Palangka Raya, with R^2 values ranging from 0.89 to 0.90 and MAPE below 30%, highlighting its ability to capture stable inflation patterns. For example, Malang achieves an MSE of 0.83 and MAE of 0.71 reflecting accurate predictions. In contrast, cities such as Tanjung Pandan (R^2=0.48, MAPE=292.78%) and Manado (R^2=0.56, MAPE=21.22%) show poor performance, possibly due to data volatility or anomalies. Metropolitan cities such as Jakarta (R^2=0.75), Bandung (R^2=0.78), and Surabaya (R^2=0.89) performed better, benefiting from structured economic systems and consistent data. However, smaller cities such as Sorong (R^2=0.68, MAPE=38.60%) and Gorontalo (R^2=0.67) faced challenges due to irregular patterns or unstable economic conditions. This study emphasizes the importance of adaptive models for more accurate inflation predictions across regions, supporting better economic policy planning [23], [26]–[28]. The results show that the LSTM model produces accurate predictions in large cities such as Jakarta, Bandung, and Surabaya with a R^2 above 0.8 and a low error rate. For example, Jakarta has a MAPE of 10.91%, reflecting an accurate prediction rate. However, in small cities such as Tanjung Pandan, the performance of this model is lower, with R^2 below 0.5 and MAPE of 283.11%. Several cities, including Tasikmalaya (R^2=0.90), Malang (R^2=0.90), and Palangka-Raya (R^2=0.893), recorded good performance with MAPE below 30%, indicating the model's ability to capture structured inflation patterns effectively. This variation is caused by differences in economic stability, data consistency, and other external factors. Large cities benefit from stable economic conditions and structured inflation trends, which allow the model to capture patterns with high accuracy. In contrast, cities such as Tanjung Pandan experience higher volatility due to irregular economic activities, making inflation trends more difficult to predict. The LSTM model is effective for areas with stable data patterns but requires additional approaches in areas with fluctuating data. Recommendations for development include the integration of more specific exogenous variables such as commodity prices, hyperparameter adjustments so that it is expected to support responsive economic policies based on local characteristics.

Table 4. Final results of model evaluation for each city

City	MSE	MAE	MAPE	R^2	City	MSE	MAE	MAPE	R^2
Meulaboh	3.17	1.38	33.22	0.76	Kediri	1.95	1.11	25.94	0.83
Banda Aceh	6.31	1.69	43.29	0.73	Malang	1.50	0.88	19.39	0.90
Lhokseumawe	1.18	0.92	30.77	0.68	Probolinggo	1.53	0.98	22.16	0.82
Sibolga	0.83	0.74	19.24	0.72	Madiun	0.26	0.42	14.61	0.83
Pematangsiantar	1.15	0.92	26.28	0.80	Surabaya	2.36	1.21	23.75	0.89
Medan	1.84	1.07	34.14	0.72	Tangerang	0.39	0.56	18.13	0.73
Padangsidimpuan	1.27	0.91	20.16	0.75	Cilegon	0.33	0.47	13.37	0.83
Padang	2.82	1.36	31.94	0.82	Serang	1.13	0.94	26.39	0.84
Bukittinggi	2.97	1.46	45.46	0.81	Singaraja	1.56	1.09	31.49	0.67
Tembilahan	0.57	0.61	26.53	0.63	Denpasar	2.84	1.20	24.26	0.89
Pekanbaru	0.63	0.63	21.50	0.82	Mataram	2.25	1.11	25.66	0.86
Dumai	1.09	0.82	19.44	0.79	Bima	0.85	0.69	21.56	0.78
Bungo	2.69	1.36	48.09	0.80	Maumere	10.07	2.68	67.25	0.75
Jambi	2.40	1.23	38.93	0.66	Kupang	2.54	1.18	30.52	0.85
Palembang	1.12	0.82	22.22	0.85	Pontianak	1.37	0.96	30.96	0.78
Lubuklinggau	0.84	0.75	25.20	0.80	Singkawang	0.82	0.75	25.68	0.80
Bengkulu	0.96	0.83	21.19	0.73	Sampit	0.46	0.53	16.94	0.81
Bandar Lampung	0.68	0.64	17.39	0.81	Palangka Raya	1.39	0.88	19.28	0.89
Metro	1.33	0.90	24.32	0.75	Tanjung	0.64	0.65	23.31	0.73
Tanjung Pandan	3.04	1.43	283.11	0.49	Banjarmasin	1.25	0.91	23.27	0.78
Pangkalpinang	2.63	1.21	39.49	0.86	Balikpapan	0.53	0.52	11.54	0.74
Batam	0.99	0.82	24.12	0.82	Samarinda	0.43	0.56	14.67	0.85
Tanjung Pinang	1.23	0.90	29.96	0.58	Tarakan	0.45	0.55	21.16	0.75
Jakarta	0.14	0.27	10.91	0.76	Manado	0.89	0.70	29.30	0.56
Bogor	0.36	0.49	11.58	0.78	Palu	0.23	0.40	12.43	0.61
Sukabumi	0.47	0.60	15.29	0.83	Bulukumba	0.29	0.45	14.76	0.76
Bandung	2.60	1.22	46.14	0.78	Watampone	0.34	0.43	16.35	0.80
Cirebon	0.92	0.83	22.31	0.77	Makassar	0.89	0.78	20.90	0.83
Bekasi	1.01	0.88	20.69	0.73	Pare-Pare	1.43	0.97	26.05	0.74
Depok	2.12	1.05	23.32	0.78	Palopo	0.92	0.77	27.69	0.77
Tasikmalaya	1.19	0.83	19.08	0.90	Kendari	2.55	1.31	28.58	0.73
Cilacap	3.00	1.20	23.73	0.86	Baubau	3.00	1.17	21.04	0.70
Purwokerto	2.50	1.08	21.94	0.82	Gorontalo	1.14	0.94	34.12	0.67
Kudus	1.52	0.88	17.97	0.84	Mamuju	0.89	0.77	37.86	0.58
Surakarta	2.43	1.13	21.75	0.89	Ambon	1.64	1.03	20.37	0.74
Semarang	0.70	0.67	18.07	0.77	Tual	1.05	0.74	16.99	0.72
Tegal	0.98	0.82	20.04	0.86	Ternate	1.24	0.78	17.41	0.58
Yogyakarta	2.44	1.15	22.10	0.82	Manokwari	0.91	0.78	19.10	0.60
Jember	5.93	1.77	32.48	0.81	Sorong	0.69	0.71	37.64	0.68
Banyuwangi	2.86	1.23	25.11	0.67	Merauke	0.76	0.65	18.17	0.66
Sumenep	3.80	1.66	29.46	0.85	Jayapura	0.33	0.45	16.56	0.77

4. CONCLUSION

This study describes how effective the LSTM model is at predicting inflation in various Indonesian cities, especially in areas with stable inflation trends, such as Jakarta, Bandung, Surabaya, Malang, and Tasikmalaya (R^2>0.80). However, the model accuracy decreases significantly in cities experiencing high economic fluctuations such as Tanjung Pandan, Sorong, and Gorontalo (R^2<0.50) which is mostly caused by variations in local economic conditions and seasonal influences. Cities with medium prediction performance ($0.50 \leq R^2 \leq 0.80$) such as Palembang, Semarang, Makasar, and several other areas show that the LSTM model is adaptive in capturing extreme inflation patterns although improvements are needed for prediction accuracy. The difference in performance shows the importance of improving the quality of input data, especially through more relevant feature selection, outlier detection and handling, and more comprehensive integration of external economic variables such as commodity prices, unemployment rates, interest rates, and fiscal policies. This model can be used as an early warning system to assist Bank Indonesia and local governments in directing price stabilization and inflation control policies. Cities with high volatility require rapid intervention, such as strengthening logistics distribution and stabilizing commodity prices. In contrast, cities with stable inflation can focus policies on infrastructure development and long-term investment planning. With the results of this study, it is hoped that it can be integrated into Indonesia's national inflation monitoring system as a strong and accurate database for state institutions in designing targeted monetary and fiscal policies so that they can have an impact.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Harfely Leipary		✓	✓			✓		✓	✓	✓	✓		✓	
Adi Setiawan	✓	✓		✓	✓	✓	✓	✓	✓	✓		✓		✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors declare that there are no known conflicts of interest regarding the publication of this paper. All author activities were conducted independently for academic and scientific purposes.

DATA AVAILABILITY

The data supports this study is publicly accessible in GitHub at https://github.com/harfelyleipary/Indonesia-Inflation-Dataset/blob/main/yoy_Inflation_dataset_Indonesia.csv.

REFERENCES

[1] H. Aginta, “Spatiotemporal analysis of regional inflation in an emerging country: the case of Indonesia,” *Regional Science Policy and Practice*, vol. 14, no. 3, pp. 667–688, 2022, doi: 10.1111/rsp3.12539.

[2] A. S. A. Raj, L. Kumarasankaralingam, M. Balamurugan, B. Maheswari, J. Gowri, and A. Dutta, “Forecasting the economic crisis of Sri Lanka: application of machine learning algorithms for time series data,” *Procedia Computer Science*, vol. 235, pp. 1087–1096, 2024, doi: 10.1016/j.procs.2024.04.103.

[3] V. Fibriyani, N. Chamidah, and T. Saifudin, “Estimating time series semiparametric regression model using local polynomial estimator for predicting inflation rate in Indonesia,” *Journal of King Saud University - Science*, vol. 36, no. 11, 2024, doi: 10.1016/j.jksus.2024.103549.

[4] N. Septrila and E. P. Kurniasih, “The effect of inflation and economic growth on unemployment rate in Asean countries,” *International Journal of Business, Technology and Organizational Behavior (IJBTOB)*, vol. 2, no. 6, pp. 580–594, 2022, doi: 10.52218/ijbtob.v2i6.226.

[5] M. Tao, “Digital brains, green gains: artificial intelligence’s path to sustainable transformation,” *Journal of Environmental Management*, vol. 370, 2024, doi: 10.1016/j.jenvman.2024.122679.

[6] I. W. Sumarjaya and M. Susilawati, “Forecasting monthly inflation rate in Denpasar using long short-term memory,” *Jurnal Matematika*, vol. 13, no. 1, pp. 11–24, 2023.

[7] A. R.-vargas, “Forecasting Costa Rican inflation with machine learning methods,” *Latin American Journal of Central Banking*, vol. 1, 2020, doi: 10.1016/j.latcb.2020.100012.

[8] A. Elhalwagy and T. Kalganova, “Multi-channel LSTM-capsule autoencoder network for anomaly detection on multivariate data,” *Applied Sciences*, vol. 12, no. 22, 2022, doi: 10.3390/app122211393.

[9] D. D. Pramesti, D. C. R. Novitasari, F. Setiawan, and H. Khaulasari, “Long-short term memory (LSTM) for predicting velocity and direction sea surface current on Bali strait,” *Barekeng: Journal of Mathematics and its applications*, vol. 16, no. 2, pp. 451–462, 2022, doi: 10.30598/barekengvol16iss2pp451-462.

[10] S. Zahara, Sugianto, and M. B. Ilmiddaviq, “Consumer price index prediction using long short term memory (LSTM) based cloud computing,” *Journal of Physics: Conference Series*, vol. 1456, no. 1, 2020, doi: 10.1088/1742-6596/1456/1/012022.

[11] B. M. Achmad, S. Sa’adah, and I. Kurniawan, “LSTM algorithm in predicting chronic kidney disease optimized using genetic algorithm,” *JUITA: Jurnal Informatika*, vol. 12, no. 2, 2024, doi: 10.30595/juita.v12i2.22965.

[12] A. Setiawan and R. W. N. Wijaya, “Comparison of ANN method and logistic regression method on single nucleotide polymorphism genetic data,” *Barekeng: Journal of Mathematics and its applications*, vol. 17, no. 1, pp. 197–210, 2023, doi: 10.30598/barekengvol17iss1pp0197-0210.

[13] A. Arista, “Comparison decision tree and logistic regression machine learning classification algorithms to determine covid-19,” *SinkrOn*, vol. 7, no. 1, pp. 59–65, 2022, doi: 10.33395/sinkron.v7i1.11243.

[14] J. Michańków, P. Sakowski, and R. Ślepaczuk, “Mean absolute directional loss as a new loss function for machine learning problems in algorithmic investment strategies,” *Journal of Computational Science*, vol. 81, 2024, doi: 10.1016/j.jocs.2024.102375.

[15] Z. Shahzad *et al.*, “A multi parameter forecasting for stock time series data using LSTM and deep learning model,” *Mathematics*, vol. 11, no. 3, 2023.

[16] S. Mukherjee, B. Sadhukhan, N. Sarkar, D. Roy, and S. De, “Stock market prediction using deep learning algorithms,” *CAAI Transactions on Intelligence Technology*, vol. 8, no. 1, pp. 82–94, 2021, doi: 10.1049/cit2.12059.

[17] Y. Liu, “Forecasting China’s consumer price index (CPI) based on combined ARIMA-LSTM models,” *Frontiers in Business, Economics and Management*, vol. 16, no. 2, pp. 264–270, 2024, doi: 10.54097/v49dwv67.

[18] S. M. Mostafa and H. Amano, “Effect of clustering data in improving machine learning model accuracy,” *Journal of Theoretical and Applied Information Technology*, vol. 97, no. 21, pp. 2973–2981, 2019.

[19] X. Tian *et al.*, "U + LSTM-F: a data-driven growth process model of rice seedlings," *Ecological Informatics*, vol. 84, 2024, doi: 10.1016/j.ecoinf.2024.102922.

[20] F. N. Sibai, A. El-Moursy, and A. Sibai, "Forecasting the consumer price index: a comparative study of machine learning methods," *International Journal of Computing and Digital Systems*, vol. 15, no. 1, pp. 487–497, 2024, doi: 10.12785/ijcds/150137.

[21] B. Mallala, A. I. U. Ahmed, S. V. Pamidi, M. O. Faruque, and R. R. M, "Forecasting global sustainable energy from renewable sources using random forest algorithm," *Results in Engineering*, vol. 25, 2025, doi: 10.1016/j.rineng.2024.103789.

[22] G. S. Mohammad, "Estimate accuracy of the sample determination coefficient R2 and R2 for the multiple regression models," *Journal of Southwest Jiaotong University*, vol. 55, no. 1, 2020, doi: 10.35741/issn.0258-2724.55.1.25.

[23] S. M. Al-Selwi *et al.*, "RNN-LSTM: from applications to modeling techniques and beyond—systematic review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 5, 2024, doi: 10.1016/j.jksuci.2024.102068.

[24] C. Mandang, D. P. E. Saerang, and W. Frederik, "The influence of price, service quality and physical environment on customer satisfaction case study markobar cafe mando," *Jurnal EMBA*, vol. 5, no. 2, pp. 678–686, 2017.

[25] H. C. Chen, W. J. Chen, and Y. Zhou, "Estimation of chromaticity coordinates for LEDs array by modulation of red or yellow LEDs with artificial neural network," *2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 88–91, 2013, doi: 10.1109/IIH-MSP.2013.31.

[26] Q. Q. He, C. Wu, and Y. W. Si, "LSTM with particle swam optimization for sales forecasting," *Electronic Commerce Research and Applications*, vol. 51, 2022, doi: 10.1016/j.elerap.2022.101118.

[27] S. Atef, K. Nakata, and A. B. Eltawil, "A deep bi-directional long-short term memory neural network-based methodology to enhance short-term electricity load forecasting for residential applications," *Computers and Industrial Engineering*, vol. 170, 2022, doi: 10.1016/j.cie.2022.108364.

[28] S. Kubheka, "South African inflation modelling using bootstrapped long short-term memory methods," *SN Business and Economics*, vol. 3, no. 7, 2023, doi: 10.1007/s43546-023-00490-9.

BIOGRAPHIES OF AUTHORS

Harfely Leipary is a graduate of Universitas Kristen Indonesia Maluku. He is currently advancing his expertise by pursuing a master's degree in Data Science at the Faculty of Science and Mathematics, Satya Wacana Christian University (UKSW). He can be contacted at email: leiparyharfely@gmail.com.

Adi Setiawan He is a distinguished academic with extensive expertise in mathematics, statistics, and data science. He served as the Dean of the Faculty of Science and Mathematics (FSM) at Satya Wacana Christian University (UKSW) from 2017 to 2022 and currently holds the position of Head of the Quality Assurance and Development Division (GPMF) for the 2022–2027 term. With a focus on the Master of Data Science program, he teaches various subjects, including statistics, abstract algebra, advanced statistics, mathematical statistics, probability theory, industrial statistics, and data mining. He earned his Bachelor's degree in Mathematics from Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, in 1991. He pursued advanced studies at Vrije Universiteit Amsterdam, obtaining his master's degree in Mathematics in 1997 and his doctorate in Statistics in 2007. His dedication to professional development is reflected in his participation in numerous workshops, such as "LaTeX for Researchers and Students" (UKSW, 2016), "Survey and Mapping Technology" (UGM, 2016), and workshops on writing high-quality journal articles (UKSW, 2017; UAD Yogyakarta, 2019). He profound knowledge and commitment to academic excellence make him a highly respected figure in the field of data science and statistical education. He can be contacted at email: adi.setiawan@uksw.edu.