

CPI Prediction Model December 2022

By Ryan Peng, Lead Data Scientist

Table of Contents

03, 04

ABSTRACT
I . INTRODUCTION

15, 16

V . MODEL ASSUMPTIONS AND
VALIDATIONS
FIGURE 3

24, 25

REFERENCES

05, 06, 07, 09, 10

II . DATA DESCRIPTION AND
PREPARATION
FIGURE 1
FIGURE 2

17, 18, 19, 20

VI. MODEL TUNING AND
PERFORMANCE
VII. MODEL RESULT
FIGURE 4, FIGURE 5

11, 12, 13, 14

III. FEATURE ENGINEERING
IV. MODEL ALGORITHM

21, 22, 23

FIGURE 6, FIGURE 7
VIII. DATA SCIENTIST
REFLECTIONS AND FUTURE
POSSIBILITIES

ABSTRACT

The Core Consumer Price Index (CPI) measures the changes in the price of goods and services, excluding food and energy from the perspective of the consumer. It is widely used as a measure of inflation and as a key indicator of economic performance. A sustained increase in the CPI can signal a decline in purchasing power for consumers and can negatively impact the economy as a whole. As such, accurately predicting future changes in the CPI is of great importance for policymakers, businesses, and individuals.

For research, we aim to build a Long Short-Term Memory (LSTM) model that will be trained on historical CPI data and other features to predict future CPI trends. The value economists can derive from this is being able to make more informed decisions about economic policy and personal financial planning, using the predictions generated by the model. Furthermore, by gaining insights into the underlying factors that drive changes in the CPI, we can gain a better understanding of the economy as a whole and how it functions.

I . INTRODUCTION

Predicting the CPI is a challenging task due to the many factors that can influence changes in prices. These factors include changes in supply and demand, currency exchange rates, and government policies, among others. Additionally, demographic changes, such as population growth and aging, can affect the CPI by altering the demand for certain goods and services.

International trade and events, such as natural disasters, can also have a significant impact on the CPI by affecting the availability and cost of imported goods [1]. Furthermore, events specific to certain sectors, such as changes in technology or changes in regulations, can also affect the CPI by altering the costs and supply of certain goods and services. Additionally, socio-economic factors such as labor market conditions and consumer sentiment can also drive the movement of prices. Overall, predicting the CPI requires a thorough understanding of the economic and social factors that can affect prices and the ability to effectively incorporate these factors into the prediction model.

Predicting the CPI can be difficult for a number of reasons:

- **Data Quality:** Collecting accurate and timely data on prices can be challenging, especially for items that are infrequently purchased or are subject to rapid price changes.
- **Substitution Bias:** As prices change, consumers may substitute away from goods that have become more expensive and towards those that have become less expensive. This can make it difficult to accurately measure the overall change in prices.
- **Quality Change:** The prices of goods and services can change not only due to changes in price, but also due to changes in the quality of the goods and services. For example, a computer that costs the same as last year's model but is faster and has more memory will have a lower effective price than the previous model.
- **Exclusion of Non-market Goods and Services:** Prices of goods and services that are not traded in the market are not included in the basket of items used to calculate the

CPI. These goods and services such as housework, own production and imputed rent etc. This can be a limiting factor to accurately measure the overall change in prices.

- **Measurement Error:** Despite best efforts to measure prices accurately, errors in measurement can occur and can lead to an upward or downward bias in the calculated rate of inflation.
- **Unforeseen Events :** There can be unforeseen events like natural disasters or pandemics that can also cause massive shifts in the economy and prices, making it difficult to predict how they will affect the CPI.

Because of these factors and others, accurately predicting the CPI can be difficult [2], and the estimate of the future inflation is always uncertain.

In order to tackle this complexity, we will utilize the power of machine learning techniques, specifically LSTM networks and historical data, to gain insights into the underlying factors and predict future changes in the CPI.

II. DATA DESCRIPTION AND PREPARATION

LSTMs require sequential data in order to learn patterns in the data and generate predictions. The data should be structured in such a way that each observation is associated with a certain time step. Additionally, the data should be scaled and normalized appropriately so that all values lie within a certain range, and the features should be encoded in a way that the model can understand.

The dataset we used comes from St. Louis Fred Economic Research. The St. Louis FRED Economic Research provides a wealth of information on the economic situation of the United States and its individual states, including data on consumer spending, unemployment, GDP, inflation, and more [3]. The data we collected can be used to gain a comprehensive understanding of the current economic climate on both a national (U.S.) and state level, allowing for detailed analyses and predictions about the future of the economy. The dataset we utilized and compiled has a total of 12 features:

- Unemployment Rate
- Personal Saving Rate
- M2 (Money Supply (M1) + saving deposits)
- Real Disposable Personal Income
- Personal Consumption Expenditures
- Real Broad Effective Exchange Rate
- Market Yield for U.S. Treasury Securities at 10-Year Constant Maturity
- Federal Funds Effective Rate
- Total Construction Spending
- Industrial Production: Total Index
- Core Consumer Price Index

Data Exploration and Wrangling

We can plot our data and features to understand the patterns in economic activity over the past few decades. We used monthly data from 1994 onwards to 2022 after removing missing values.

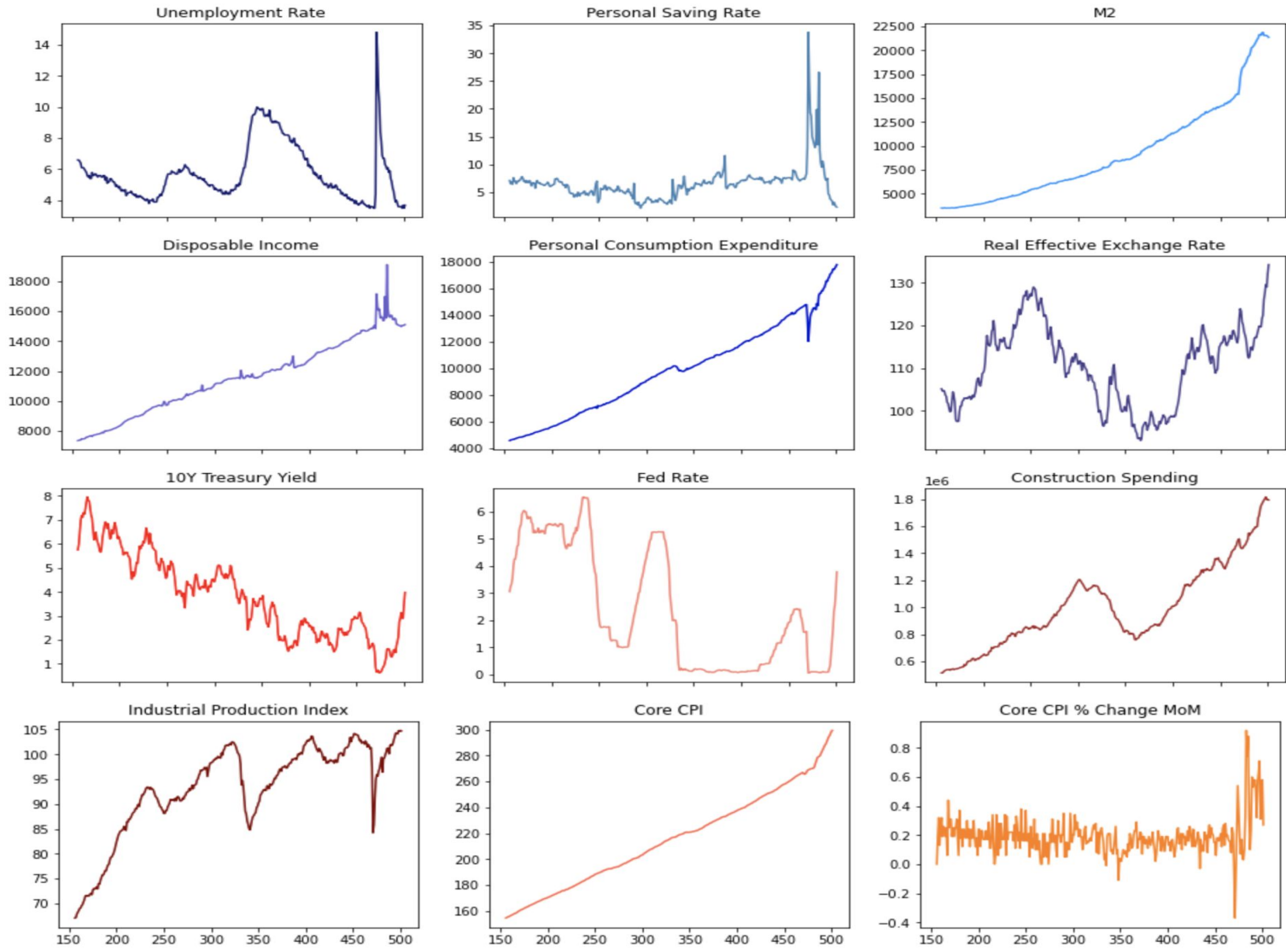


Figure 1. Distribution of Dataset Features

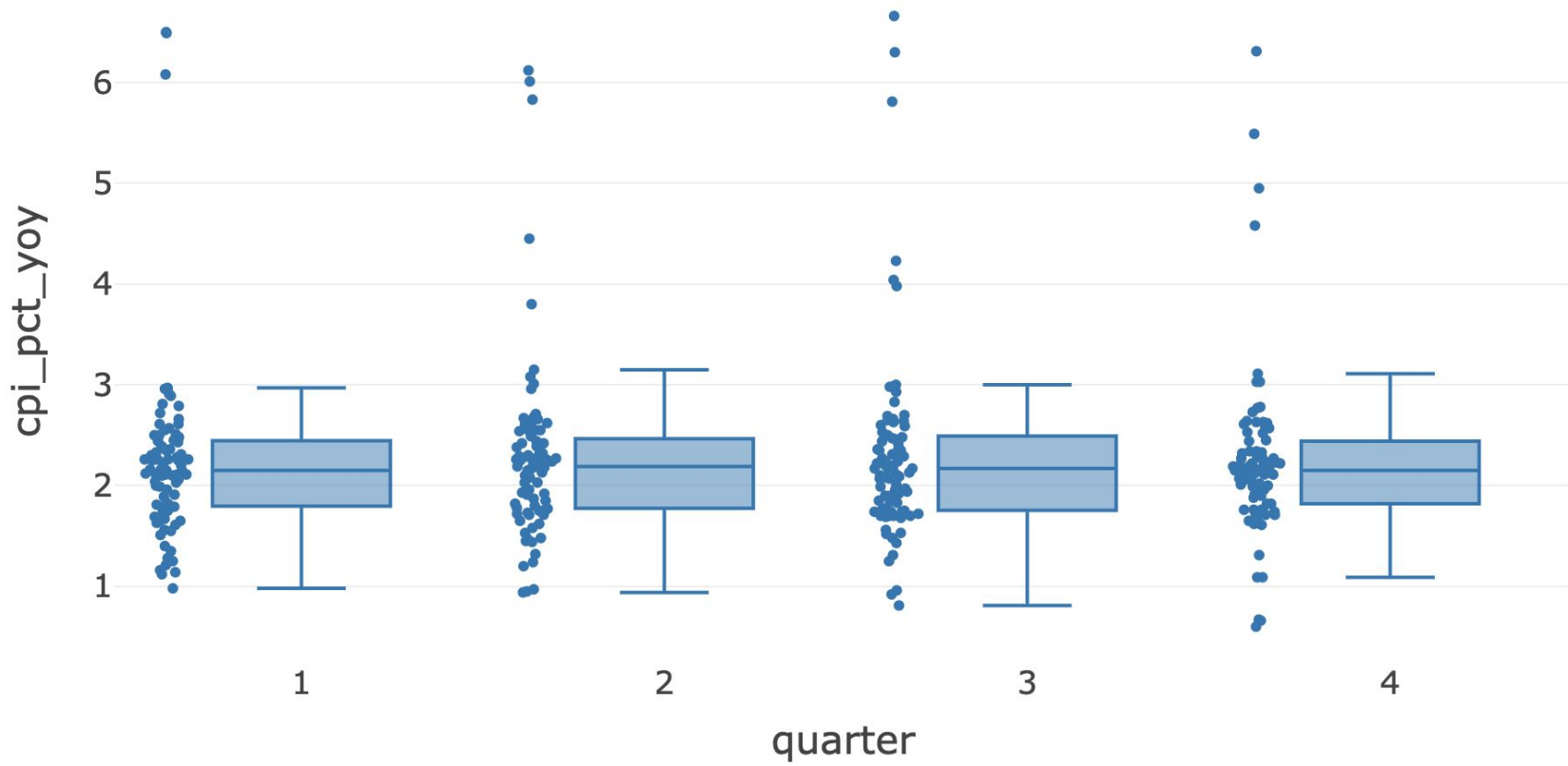


Figure 2. CPI Percent Change over Different Quarters

From Fig. 1, we can see that core CPI has a strong positive trend and even increased at a higher rate for the past few years. The main reason why core CPI is increasing is due to rising prices of goods and services in the economy. This can be attributed to a number of factors, such as increasing production costs, rising wages, and increasing demand. Additionally, inflationary pressures such as low central bank interest rates, increasing money supply, and currency devaluation can also contribute to an increase in core CPI [4]. We can see that in the past few years, possibly due to the COVID-19 pandemic, some features exhibit larger outliers, and core CPI has had a significant jump.

Looking at the core CPI change on an annualized basis in Fig. 2, we can see that the earlier parts of the years showed more significant increases in core CPI with outliers in the third quarter showing the most CPI percent change above four percent. In the last few months, there are more outliers observed (below the lower whiskers) meaning that they showed little change from the previous year.

The dataset contains many numerical features and we applied a smoothing technique to the data to reduce the effect of outliers. We also removed any historical data with missing values or features that had a significant amount of missing data. Since there are no categorical features, one-hot encoding was not necessary to convert any variables into numerical representations. Finally, we applied min-max normalization to the data to ensure that all the values lie within a certain range. This allowed us to effectively train the model on the data and make accurate predictions.

Feature Engineering

In order to see which features are useful for forecasting core CPI, we will be using the **Granger causality test**. One of the key assumptions of the **Granger causality test** is that the time-series being tested must be stationary, meaning that their statistical properties do not change over time. If the time-series values are not stationary, the results of the Granger causality test may be unreliable or misleading. If the time-series are found to be non-stationary, they may need to be transformed, such as by differencing or detrending, in order to make them stationary before the Granger causality test can be performed.

The Granger causality test typically involves running a stationarity test, such as the Augmented Dickey-Fuller (ADF) test [5], on the time-series data. The ADF test is a statistical test used to determine whether a time-series is stationary (i.e., does not have a unit root, meaning it does not have a trend or stochastic trend).

The null hypothesis of the test is that there is a unit root, and the alternative hypothesis is that there is no unit root. If the p-value from the ADF test is below a certain critical value, then the null hypothesis is rejected, and we can conclude that the time-series is stationary. The ADF test is commonly used in econometrics and finance to test for stationarity in financial time-series data such as stock prices, interest rates, and exchange rates.

After the data is statistically stationary, we used the Granger causality test to determine that the real effective exchange rate and 10Y treasury yield are not significant, thus we can first proceed to exclude them from our future models, dropping them from the dataframe.

The real effective exchange rate is an index that measures the value of a country's currency relative to a basket of other

currencies, adjusted for changes in the relative prices of goods and services in the country [6]. A 10-year Treasury yield is the interest rate at which the US government can borrow money for a period of 10 years [7].

Both the real effective exchange rate and the 10-year Treasury yield are important economic indicators in their own right, but they may not be significant in predicting core CPI, which measures changes in the prices of a basket of consumer goods and services. The real effective exchange rate measures the relative value of a country's currency, but it may not be directly related to the prices of goods and services consumed by households. The prices of these goods and services are affected by various factors, including domestic and international supply and demand, production costs, and monetary policy. Similarly, the 10-year treasury yield is a measure of the cost of borrowing for the US government and can be affected by various factors such as economic growth, inflation expectations and monetary policy. However, it may not have a direct correlation with the changes in the prices of goods and services consumed by households.

In general, there are many factors that contribute to changes in the prices of goods and services consumed by households and a lot of different indicators that can be used to predict inflation. The real effective exchange rate and treasury yield may not always be significant in predicting the core CPI, but it is important to consider other factors such as labor market conditions, economic activity, and monetary policy when making predictions about inflation.

By manually testing different features, we also have determined that including the federal funds effective rate increased our model error as well most likely due to spikes in values in recent years. The federal funds effective rate is the interest rate at which depository institutions lend and borrow overnight funds among themselves, usually on an uncollateralized basis [8]. The rate is determined by the market forces of supply and demand for overnight funds. Core CPI, on the other hand, measures the changes in the prices of a basket of goods and services consumed by households, excluding certain items such as food and energy.

The federal funds effective rate and core CPI are not closely related because they measure different things. The federal funds effective rate is a measure of the cost of overnight borrowing and lending, while core CPI measures changes in the prices of a basket of consumer goods and services. Furthermore, the rate is determined by the forces of supply and demand in the overnight lending market, while core CPI is determined by the changes in the prices of the goods and services included in the basket. While the Federal Reserve may use the federal funds rate as a tool to influence inflation, it is not a direct indicator of it and therefore might not be a useful predictor for core CPI.

III. TOOLS REQUIRED

We are using a number of software and packages commonly used by the data science community for machine learning and data visualization. For machine learning, we are using Keras/Tensorflow [9]. Other commonly used packages are also imported for data parsing and visualizations such as Numpy, Matplotlib, Pandas, and Scipy. Ipywidgets is used for demonstration and selection of targets and other parameters. BOTO3 is used for accessing and interacting with S3 data storage. Plotly [10] is used creating interactive figures to inferences with the APIs.

IV. MODEL ALGORITHM

LSTMs are a type of recurrent neural network (RNN) capable of learning long-term dependencies in data. Unlike traditional RNNs, which are limited by the short-term memory of the hidden layer neurons, LSTMs are able to remember information for long periods of time, allowing them to more effectively learn from data with temporal or time-series dependency. This makes LSTMs particularly well-suited for tasks like using historical data to predict future events as well as language translation and speech recognition, where understanding the context and meaning of words is crucial for achieving good performance.

LSTM networks can be used for univariate time-series forecasting, which involves using historical data for a single variable to predict future values of that variable. In this case, the input to the LSTM network would be a sequence of past values for the singular variable, and the output would be a predicted future value. The network can then be trained using supervised learning, where the

training data includes both the input and the known correct output. Once trained, the LSTM network can be used to make predictions on new data.

One advantage of using LSTM networks for univariate time-series forecasting is that they are able to learn and model long-term dependencies in the data, which can be useful for making accurate predictions. For example, if there is a seasonal pattern in the data, such as an increase in demand for a product during the holiday season, an LSTM network can learn and incorporate this information into its predictions. Additionally, LSTM networks are able to handle noisy or missing data, making them robust to real-world data imperfections.

In addition to univariate time-series forecasting, LSTM networks can also be used for multivariate time-series forecasting, which involves using historical data for multiple features to predict future values of those features. In this case, the input to the LSTM network would be a sequence of past values for each variable, and the output would be a predicted future value.

Supervised learning is also utilized, where the training data includes both the input and the known correct output.

An advantage of using LSTM networks for multivariate time-series forecasting is that they are able to learn and model the relationships between the different variables, which can be useful for making accurate predictions. For example, if there is a strong relationship between the demand for a product and the price of the product, an LSTM network can learn and incorporate this information into its predictions. Overall, LSTM networks can be a powerful tool for both univariate and multivariate time-series forecasting, and can provide accurate predictions even in the presence of long-term dependencies and noisy or missing data.

We have constructed two models using LSTM networks, one where it inputs a univariate dataset and another where it incorporates other economic metrics as inputs. The inputs of the models are set to utilize 12 data points (months) as the input and returns a single predicted output (core CPI).

Univariate Model

For the univariate model, we constructed a simple model with a single LSTM layer with 64 neurons. The number of layers is chosen because of the complexity of the problem. We want enough neurons to capture the high dimensionality of the data (so the model can capture the underlying trends). Most parameters are chosen through a fine-tuning and trial-and-error approach.

Dropout regularization and early stopping were not used due to the simple nature of the layers (there is only one layer). Implementing dropout would cause important information to be lost due to the neural network architecture and dataset size. Dropout regularization at optimal levels of 0.1/0.2 for the model were trialed and significantly decreased the performance of prediction on the test set. Dropout regularization erases important context information, especially in this time-series problem with limited data, timesteps, and layer height. Furthermore, the training and validation losses are carefully monitored for any potential overfitting. Similarly, a small learning rate of 0.001 is used in conjunction to the size of the small dataset. The overall batch size is set as to be 100 epochs. Batch size fine-tuning is done based on the observation of the model's performance.

Multivariate Model

The initial multivariate model was trialed with similar parameters as the univariate model but we observed poor performance. This can be due to the increased complexity and the added feature dimensions in the data. Thus, we added another layer to the LSTM with 100 neurons and also adjusted the original layer to 100 as well. The model was adjusted to run for a total of 250 epochs with default batch size of 32. A GridSearchCV attempt was done before to find the optimal epochs and batch size.

A model that is trained and tested on the same data is said to be overfitting, as it has memorized the training set. Overfitting is not desirable because it means that the model performs well on the training data but not on new, unseen data. The model showed significant overfitting, so regularization with early stopping and dropout were implemented. The final dropout layer was set at 20% and added between the second LSTM layer and dense layer. The early stopping patience was set at 50. The hyper-parameters were also tuned through manual experimental runs observing the performance and output result.

V. MODEL ASSUMPTIONS AND VALIDATIONS

Since we are interested in the monthly forecasting of core CPI, historical CPI data will serve as our baseline in the dataset. We employed a training/validation ratio of around 80/20%, respectively. The idea behind this ratio is to have a larger portion of the data set (80%) allocated for training the model and a smaller portion of the data set (20%) allocated for testing the model. The purpose of using this ratio is to ensure that the model has enough data to train on, while also reserving a sufficient amount of data to evaluate the model's performance. Having a larger training set allows the model to learn from more examples and potentially achieve better performance. The smaller test set is used to validate the generalization ability of the model, that means how well it performs on unseen data.

Additionally, using a ratio of 80/20 allows us to have a good balance between the amount of data used to train the model and the amount of data used to evaluate the model's performance [11]. It is important to note that this ratio can be adjusted depending on

the specific use-case and the amount of data available. If more data is available, a ratio of 90/10 or even higher can be used, if less data is available a ratio of 70/30 can be used. The key is to have a balance between training and testing, having enough data for both.

Since our predictions come from a time-series, we can utilize a walk-forward validation so that the training period is used to validate the results. In a walk-forward validation process, a number of periods is used as the training starting point. The basic idea behind walk-forward validation is to divide the time-series data into overlapping windows. The model is trained on one window, and then tested on the next window. This process is repeated, walking forward through the data, training on one window and testing on the next. This approach mimics the way that the model will be used in practice, where it will need to make predictions based on new data as it becomes available. The advantages of walk-forward validation include that it allows you to evaluate a model's ability to handle non-stationary time-series data, and it gives a better understanding of how the model will perform on new, unseen data. It's also useful in cases where it's hard to have large amounts of historical data and you have to work with a smaller training set.

The process of walk-forward validation can be a bit more time-consuming and computationally intensive than other forms of cross-validation, but it is often considered to be a more rigorous method for evaluating the performance of a time-series model. It's important to keep in mind that, like all validation methods, this too has its limitations. The choice of window size, rolling window or expanding window, how many periods to forecast, how many forecasting iterations to make all can affect the performance of the model. So, as always, it is important to experiment with different configurations and to use walk-forward validation in conjunction with other evaluation techniques to get a more complete picture of a model's performance.

Over a period of time, time-series predictions tend to become less accurate due to changes in stationarity and new trends and seasonality. As the forecast horizon increases, the uncertainty also increases, making it more difficult to make accurate predictions. Therefore, as new data becomes available on a monthly basis, the model is retrained to allow for more accurate predictions. An example of walk-forward validation is shown below in Fig. 3.

The amount of time to be included in the training set in regard to the amount of historical data is still to be determined and requires additional research in the future.



Figure 3. Example of Walk-Forward Validation

VI. MODEL TUNING AND PERFORMANCE

We utilize the ADAM [12] as the optimizer and binary cross entropy as the loss function. ADAM is computationally efficient compared to other optimization algorithms because it only requires first-order gradients. This is particularly useful in large-scale deep learning applications where computation time can be a bottleneck. ADAM has been widely used and tested in a variety of applications, including image recognition, natural language processing, and speech recognition. This widespread use and testing have led to a large body of research that supports its effectiveness.

To prevent overfitting, early termination was performed which monitors the validation loss. The learning rate was also automatically adjusted based on the changes due to plateauing. Cross validation was employed with four different splits with each split containing a different amount of training data. Due to the amount of training required, the parameters optimizations were performed using Amazon SageMaker SDK [13, 14], which utilizes a Bayesian related search technique to discover the optimal

hyper-parameters set. Sagemaker allows for various selections when dealing with hyperparameters optimizations, such as ContinuousParameters, IntegerParameters, and CategoricalParameters. Afterward, we utilize regex to extract the metrics from AWS Cloudwatch and compute the optimal hyper-parameters.

VII. MODEL RESULT

For the univariate model, the training and validation loss were observed and shown below in Fig. 4. The validation loss shows significant spikes and fluctuations due to the small batch size being passed in. This shows that the model was slightly overfitting to the small dataset. To address the overfitting issue in the univariate model, several techniques could be applied. One way to mitigate overfitting is to increase the size of the dataset used for training. Having more data allows the model to learn more general patterns and improves the model's ability to generalize to unseen data.

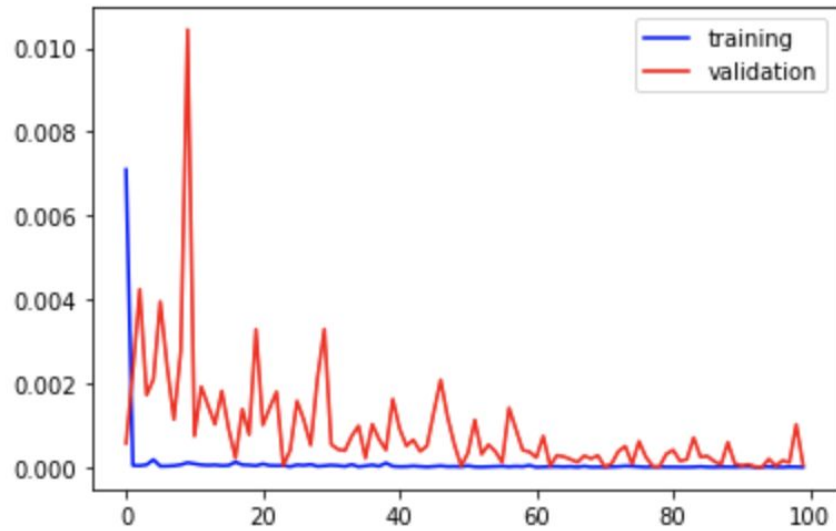


Figure 4. Univariate Training and Validation Loss

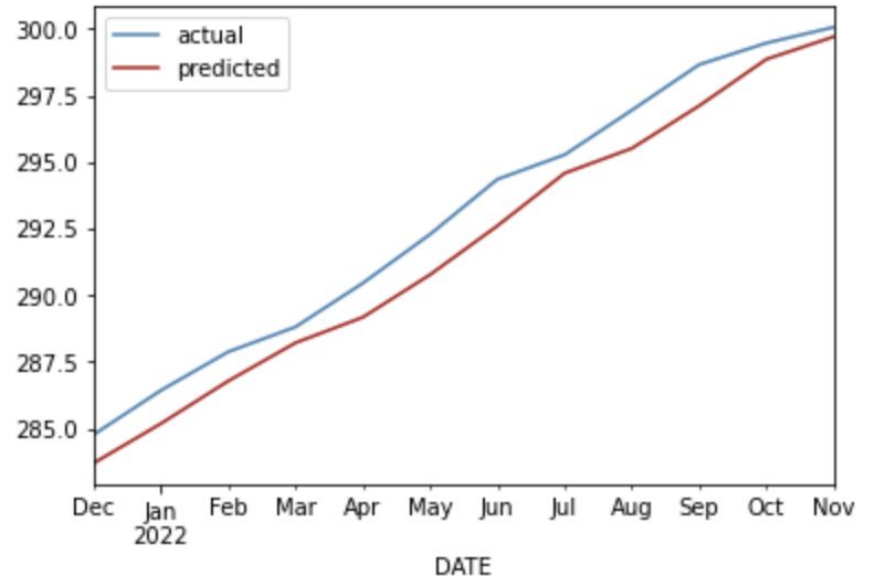


Figure 5. Univariate Actual vs. Predicted Core CPI

Using the model we fitted, we tested against our data and predicted the next year of core CPI. The result is shown below in Fig 5. As shown, we can observe that the univariate LSTM is able to approximate the shape of the CPI, but suffers from some under estimation in the CPI value. For December 2022, it predicts a CPI value of 300.32, which means the forecasted U.S. CPI changes year-over-year (YOY) is **5.46%**.

For the multivariate results, the training and validation loss were observed and shown below in Fig 6. Compared to the univariate model, the multivariate one experiences less spike fluctuations and can be attributed to an increased batch size and early stopping which prevents overfitting. The result of the multivariate model is shown in Fig 6. for loss during training. We can also see that from Fig. 7, our predictions overestimate at first, then underpredicts for the most recent months. For December 2022, the multivariate model predicts a CPI basis point of 300.06, which corresponds to a CPI change YOY of **5.37%**.

The mean model error is also lower in this model, however, the patterns and trends captured by the univariate model are not apparent in the multivariate model. This can be explained by several reasons:

- **Complex interactions:** In a multivariate model, the relationships between multiple variables can be complex and nonlinear, making it more challenging to identify patterns and trends.
- **High dimensionality:** When dealing with a high-dimensional dataset, it can be difficult to identify the most relevant variables and how they interact with each other.
- **Limited representativeness:** Some patterns and trends may not be captured in the multivariate model because some of the variables included in the model do not adequately represent the underlying phenomena.

- **Noise:** Multivariate models are more prone to overfitting, meaning they can capture noise in the data, and therefore the model may not be able to generalize well, and it may not be able to identify the underlying pattern or trend.
- **Data preprocessing:** Multivariate models may also require more preprocessing such as handling missing values, normalization, and feature selection, that may have an impact on the patterns and trends captured by the model.

There are a few possible explanations for why the multivariate model is not able to detect and recognize small-scaled patterns in the test set as seen in Fig. 7. One possible explanation is that the dataset is small, and the model is not able to generalize well to unseen data. In other words, the model may have overfitted to the training data and is not able to generalize to the test set. Another possible explanation is that the model is too complex for the given dataset. Having a high-dimensional dataset with multiple variables can make it challenging to identify patterns and trends, and a simpler model might be more appropriate for this dataset

(univariate). Additionally, the model may not be capturing all the important features or interactions that are present in the data. Therefore, it is essential to evaluate the feature importance of the model, and try including other variables, or transforming the data, in order to improve the performance of the model. Finally, it could be that the model is not able to correctly capture the small-scaled patterns due to the lack of enough data to learn from, thus having insufficient data points to capture small scale patterns.

Compared to the consensus of YOY CPI changes of 6.5% for December 2022, we can conclude that both of our models (univariate, multivariate) underestimate compared to the YOY CPI.

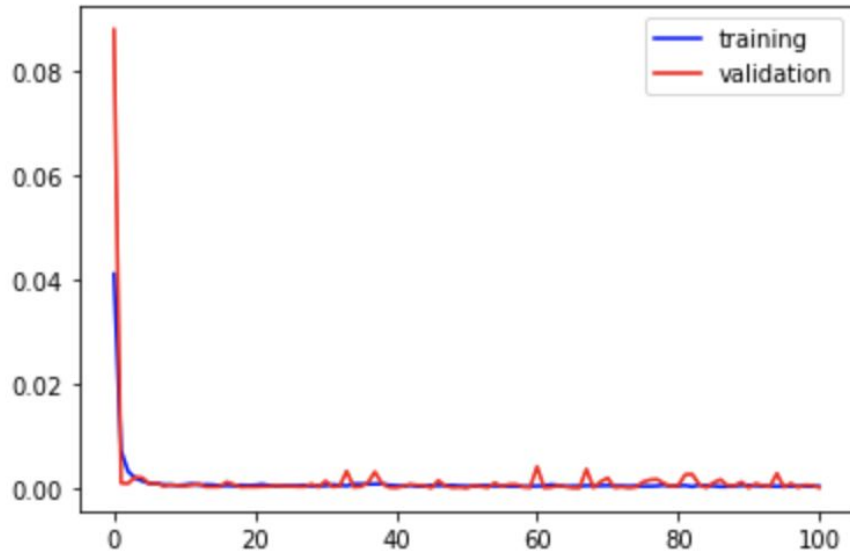


Figure 6. Multivariate Training and Validation Loss

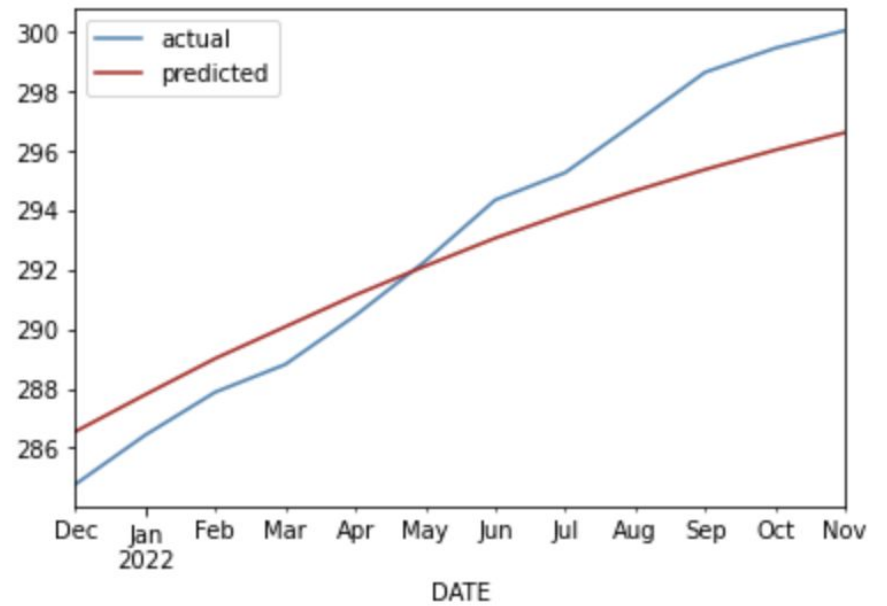


Figure 7. Multivariate Actual vs. Predicted Core CPI

VIII. DATA SCIENTIST'S REFLECTIONS AND FUTURE POSSIBILITIES

Compared to previous years, core CPI for December 2022 seems to be decreasing. A decreasing YoY change in the core CPI indicates that the prices of goods and services included in the index are not increasing as quickly as they were in the previous year. This can be attributed to a number of factors including:

There can be several reasons why the YoY change in the core CPI is decreasing:

- **Lower demand:** A decrease in consumer demand for goods and services can lead to a decrease in prices, resulting in a lower YoY change in the core CPI.
- **Increased competition:** An increase in competition among suppliers can lead to lower prices and a decrease in the YoY change in the core CPI.
- **Lower production costs:** A decrease in production costs can lead to lower prices for goods and services, resulting in a lower YoY change in the core CPI.
- **Economic downturn:** A downturn in the economy can lead to lower prices for goods and services, resulting in a lower YoY change in the core CPI.
- **Monetary Policy:** The Central Bank may lower interest rates to stimulate economic activity, this can decrease prices of goods and services, resulting in a lower YoY change in the core CPI.
- **Supply-side factors:** Factors such as increased productivity, technological advancements, and natural disasters can affect the supply of goods and services, and lead to a decrease in prices and a lower YoY change in the core CPI.

As our research has shown, LSTM networks are powerful for making predictions on time-series data. However, their complex structure and large number of variables can make training a challenge. In future research, one possibility is to explore the use of transformer-based neural networks, which are currently state-of-the-art in natural language processing. Transformer networks have been shown to perform faster than LSTM networks, and they have the added advantage of being more amenable to transfer learning.

In addition, we only used a subset of the features available in the FRED database. Expanding the feature set to include more relevant factors and contributing variables might improve model accuracy. This includes taking out features that may not be as useful and replacing them with others that affect core CPI more. We can also take a look into adding more historical time-series data for core CPI which could help our model identify smaller patterns.

Another direction to explore is reducing the temporal period used in the model. While using a longer temporal period provides more context, it also increases the complexity of the model and the amount of data required. By reducing the temporal period, we may be able to achieve similar or better performance with a simpler and more efficient model.

In conclusion, our research has highlighted the potential of LSTM networks for time-series prediction, but also its limitations. Future research might explore different architectures, like transformer-based networks, as well as other ways to improve the model performance by expanding the feature set and reducing the temporal period.

REFERENCES

- [1] Elhefnawy, N. (2022). The Consumer Price Index and the Rate of Global Economic Growth in the Twenty-First Century: A Note. Available at SSRN 4196070.
- [2] Reed, S. B., & Rippy, D. A. (2012). Consumer Price Index program, "Consumer Price Index data quality: how accurate is the U.S. CPI?" *Beyond the Numbers: Prices & Spending*, 1(12).
- [3] McCracken, M. W., & Ng, S. (2016). FREDMD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4), 574-589.
- [4] U.S. Bureau of Labor Statistics. (n.d.). CPI Home. U.S. Bureau of Labor Statistics. Retrieved January 10, 2023, from <https://www.bls.gov/cpi/>.
- [5] Lopez, J. H. (1997). The power of the ADF test. *Economics Letters*, 57(1), 5-10.
- [6] Chinn, M. D. (2006). A primer on real effective exchange rates: Determinants, overvaluation, trade flows and competitive devaluation. *Open Economies Review*, 17(1), 115–143. <https://doi.org/10.1007/s11079-006-5215-0>.

[7] Gürkaynak, R. S., Sack, B., & Wright, J. H. (2007). The U.S. treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8), 2291–2304. <https://doi.org/10.1016/j.jmoneco.2007.06.029>

[8] Garg, K. (2008). The Effect of Changes in the Federal Funds Rate on Stock Markets: A Sector-Wise Analysis. *Undergraduate Economic Review*, 4, 2.

[9] Chollet, F., et al. (2015). Keras. Retrieved from <https://github.com/fchollet/keras>.

[10] Plotly Technologies Inc. (2015). Collaborative data science. Retrieved from <https://plot.ly>.

[11] Tan, J., Yang, J., Wu, S., Chen, G., & Zhao, J. (2021). A critical look at the current train/test split in machine learning. arXiv preprint arXiv:2106.04525.

[12] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[13] Perrone, V., et al. (2020). Amazon SageMaker automatic model tuning: Scalable black-box optimization. arXiv preprint arXiv:2012.08489.

[14] Perrone, V., et al. (2021). Amazon sagemaker automatic model tuning: Scalable gradient-free optimization. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 3463-3471).