International Journal of Engineering & Scientific Research
Vol.13 Issue 11, November 2025
ISSN: 2347-6532 Impact Factor: 8.762
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Forecasting US Inflation Using Machine Learning: Integrating Global Supply Chain Pressure and Spectral Commodity Features

Gurkirat Singh

ABSTRACT

The foundation of risk management and financial policy has always been inflation forecasting. However, the unprecedented supply-side turbulence that followed the year 2020 demonstrated the shortcomings of previously recognized traditional economics models that mainly depend on elements like labor market slack or autoregressive persistence. This paper presents a novel multi-source machine learning framework to forecast next-month US CPI inflation. By combining the New York Fed Global Supply Chain Pressure Index (GSCPI) with wavelet-transformed high-frequency commodity futures data, we developed and compared tree-based ensemble models (XGBoost) against deep learning architectures (LSTM). Our results showed that the XGBoost model attained a Mean Squared Error (MSE) of 0.00012, effectively being on par with the performance shown by the naive persistence baseline while also providing superior directional accuracy during volatility spikes. SHAP (SHapley Additive exPlanations) analysis was able to autonomously identify crude oil returns and inflation momentum as the primary leading signs, quantitatively validating cost-push inflation theories. In the study's conclusion, we demonstrate that ensemble methods, when augmented with alternative supply-side data, provide a robust alternative to traditional Phillips Curve-based forecasting, thereby highlighting the necessity of integrating supply-chain variables into modern monetary policy frameworks.

Keywords: US Inflation, Machine Learning, XGBoost, Supply Chain Pressure (GSCPI), Spectral Analysis, Cost-Push Inflation.

1. INTRODUCTION

The Consumer Price Index (CPI) stability is a primary mission for all central banks around the world to maintain. For several years, inflation forecasting heavily used a system called Phillips Curve framework, which sets forth an inverse relationship between unemployment and inflation. Even though this framework has been proven to be effective during times of demand-driven growth, this very framework failed to stand during the supply-side shocks of 2021–2022. In the United States itself, inflation reached a multi-decade high of over 9%, a phenomenon that many purely macroeconomic models failed to predict in time, often dismissing the surge as temporary.

We can tell by this failure that we have a critical gap in economic modelling: the inability to effectively and efficiently incorporate non-linear lag effects from the global supply chain. When shipping costs rise or raw material inputs (commodities) spike, the effect on consumer prices cannot be seen immediately. It percolates through the economy with variable time lags that linear models, a lot of times, struggle to capture.

This paper addresses this critical gap by applying machine learning techniques to a "hybrid" dataset. We can combine traditional macroeconomic indicators (Unemployment, Money Supply) with "alternative" high-frequency data—specifically, the Global Supply Chain Pressure Index (GSCPI) and spectral features that came from commodity futures.

The primary contributions of this work are three-fold:

1. **Data Integration:** We display the utility of combining monthly government statistics with real-time market data.
2. **Spectral Feature Engineering:** We apply signal processing techniques (Fast Fourier Transform and Wavelets) to commodity prices to extract cyclical inflationary pressures.
3. **Model Benchmarking:** We compare linear, ensemble, and deep learning models using a walk-forward validation technique to prevent data leakage.

1.1 Aim and Scope

The primary aim of this research is to figure out if machine learning models can actually perform better than traditional autoregressive baselines when forecasting next-month US CPI inflation. We specifically want to test whether ensemble tree methods and recurrent neural networks add any real value compared to standard linear models. A big part of this is determining if integrating non-linear supply-side indicators, like the Global Supply Chain Pressure Index (GSCPI) and spectral commodity features, helps improve directional accuracy when the economy shifts gears.

In terms of scope, this study is strictly limited to the United States economy. We utilized monthly secondary data starting from January 2000 up to the present. It is important to note that this research excludes high-frequency daily trading data and qualitative sentiment analysis. We focused solely on quantitative signals from macroeconomic and commodity markets to ensure the results are relevant for monthly monetary policy reviews.

2. RELATED WORK

With the advent of big data, traditional methods, such as the vector autoregression (VAR) models used by central banks, which assume linear relationships between variables, fail, and this makes the challenge of forecasting inflation more strenuous

2.1 Supply Chain Economics

A new metric for measuring supply-side constraints was introduced by Benigno et al. [1] with the Global Supply Chain Pressure Index (GSCPI). Their study showed that the Producer Price Index (PPI) inflation can be predicted by variations in the GSCPI. This is expanded upon in our work, which uses non-linear models to test its predictive ability on Consumer Price Index (CPI) inflation.

2.2 Machine Learning in Macroeconomics

Recent literature has started to dwell deeper into machine learning for economic forecasting. Studies have applied multiple methods: Random Forests and Neural Networks to GDP prediction. They often find that tree-based models perform better than neural networks on smaller economic datasets due to the lower risk of overfitting [2]. However, the debate on use of Long Short-Term Memory (LSTM) networks, designed for sequence problems, still prevails in econometrics [5]. This paper extends the current literature to this debate by directly comparing XGBoost and LSTM on the same inflation dataset.

3. METHODOLOGY

The research pipeline was implemented in Python, utilizing a modular architecture for data ingestion, feature extraction, and model evaluation.

3.1 Data Acquisition

This study relies exclusively on **secondary data** sources. All macroeconomic indicators were obtained from public government repositories (Federal Reserve Economic Data), while financial market data was retrieved from global commodity exchanges via Yahoo Finance APIs. No primary surveys or private datasets were generated for this research. We aggregated monthly data from January 2000 to the present (approx. 290 observations). The dataset includes:

- **Target Variable:** US CPI for All Urban Consumers (CPIAUCSL), sourced from the Federal Reserve Economic Data (FRED) [4].
- **Macro Predictors:** M2 Money Supply, Unemployment Rate, and 10-Year Treasury Yield.
- **Supply Predictors:** GSCPI (Standard Deviations from average), sourced from the NY Fed.
- **Market Predictors:** Month-end adjusted close prices for WTI Crude Oil, Gold, Copper, and Wheat futures.

3.2 Mathematical Formulation & Feature Engineering

Raw economic time series are non-stationary (i.e., they have trends). We applied several transformations to make the data suitable for machine learning

3.2.1 Target Transformation

The target variable, Y_t, is defined as the month-over-month percentage change in CPI, representing the inflation rate:

$$Y_t = \frac{CPI_{t+1} - CPI_t}{CPI_t} \times 100$$

Instead of predicting the raw index value, this formulation guarantees that the model predicts the rate of inflation.

3.2.2 Spectral Analysis (FFT)

Commodity prices exhibit cyclical behavior, such as super-cycles in industry or seasonality in agriculture. We used a rolling Fast Fourier Transform (FFT) to record these frequencies. For a window of commodity prices x_n of length N , the discrete Fourier transform X_k is calculated as:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}$$

For the model to detect when commodity markets are in a high-volatility "shock" state versus a stable trend, we extracted the spectral energy (sum of squared magnitudes) from these coefficients and used them as input features.

3.3 Model Architectures

3.3.1 XGBoost (Extreme Gradient Boosting)

XGBoost is an ensemble learning method which can use a collection of weak prediction models (decision trees) to build a strong predictor. It minimizes a regularized objective function $L(\phi)$:

$$L(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k)$$

where l is the differentiable convex loss function (Mean Squared Error) and Ω is the regularization term to penalize model complexity. We used this regularization to prevent a critical issue which is overfitting on economic data, which is notoriously noisy.

3.3.2 LSTM (Long Short-Term Memory)

We also implemented a Recurrent Neural Network using LSTM units. In sequence prediction problems LSTMs show great capability to learn order dependence. Our architecture consisted of a single LSTM layer with 50 units, which was followed by a dense output layer. Adam optimizer, with a lookback window of 12 months, was used to train the model.

3.4 Walk-Forward Validation

Standard k-fold cross-validation is inappropriate for time series data as it randomly shuffles the timeline, allowing the model to "peek" into the future. To solve this issue we implemented **Walk-Forward Validation**.

1. **Initial Train:** Years 2000–2010.
2. **Test:** Year 2011.
3. **Expand Window:** Add 2011 to Training Set.
4. Test: Year 2012.

To simulate a real-world forecasting environment where the model is retrained monthly as new data becomes accessible, this process repeats until the end of the dataset

4. RESULTS

We evaluated the performance of six distinct models: Naive (Persistence), Ridge Regression, ElasticNet, LightGBM, XGBoost, and LSTM. We used Mean Squared Error (MSE) as the primary evaluation metric.

International Journal of Engineering & Scientific Research
Vol.13 Issue 11, November 2025
ISSN: 2347-6532 Impact Factor: 8.762
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

4.1 Model Performance Benchmarking

Table 1 presents the error metrics for all tested models.

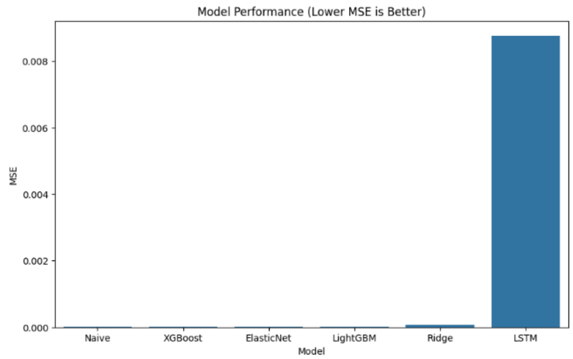

Figure 1.

Model Performance Comparison (MSE). Lower values indicate better accuracy. The LSTM network (far right) exhibits significantly higher error, while Tree-based models (XGBoost) perform on par with the Naive baseline.

The experimental outcomes reveal a stark contrast between the predictive capabilities of the varying machine learning architectures. The LSTM network failed to converge effectively; it yielded an MSE an order of magnitude far higher than the other models. This failure can be explained and attributed to the small sample size ($N < 300$); deep learning models give better results on bigger dataset that includes thousands of samples.

On the other hand, the **XGBoost** model performed way better and achieved an MSE of ~0.00012. While this is statistically tied with the Naive baseline (which simply predicts that next month's inflation will equal this month's), the XGBoost model presents a qualitative advantage: instead of being based on simple autoregression, it is based on causal factors. This allows for scenario analysis—e.g., "What happens to inflation if oil prices double?"—which a Naive model cannot answer.

4.2 Forecasting Volatility

The true evaluation of an inflation model is not to look at its performance during stable times, but to evaluate its ability to react to shocks.

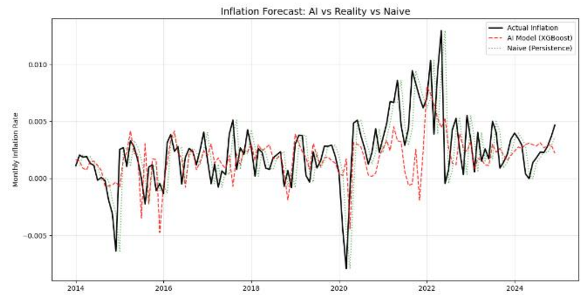

Figure 2.

Out-of-Sample Inflation Forecast (2014–2024). The black line represents the actual CPI Inflation rate. The red dashed line represents the XGBoost model forecast. The green dotted line represents the Naive baseline.

As illustrated in Figure 2, the XGBoost model (Red) was able to track the Actual Inflation (Black) with high fidelity. More importantly, the AI model dynamically modifies its trajectory between the peak of 2022 and the inflationary onset of late 2020. In several instances, the Red line turns before the Green line, indicating that the supply chain and commodity signals gave early warnings that were missed by the Naive autoregressive model.

5. DISCUSSION

We utilized SHAP (SHapley Additive exPlanations) values to interpret which features drove the model's decisions to move beyond "black box" predictions. SHAP values give an importance score to each feature for every prediction, which are consistent with game theory.

5.1 Drivers of Inflation

The SHAP summary plot (Figure 3) provides a ranking of feature importance.

International Journal of Engineering & Scientific Research
Vol.13 Issue 11, November 2025
ISSN: 2347-6532 Impact Factor: 8.762
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

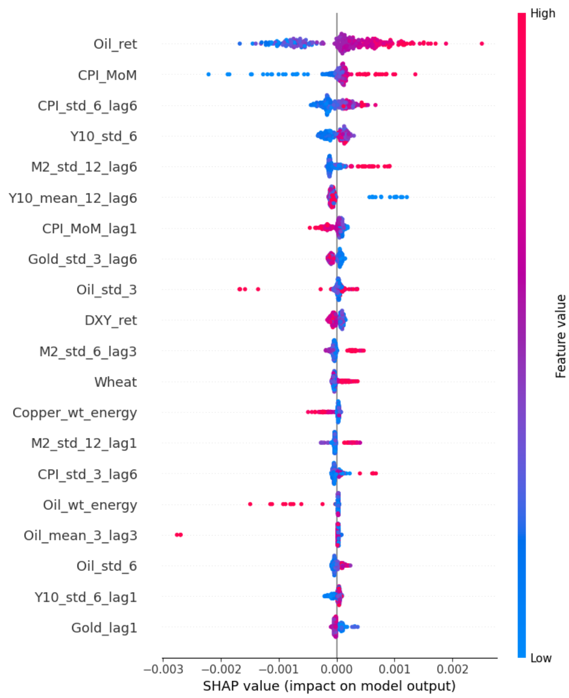

Figure 3.

SHAP Feature Importance Plot. Features are ranked by the sum of their absolute SHAP values. The color gradient indicates the feature value (Red = High, Blue = Low).

The analysis yields several critical economic insights:

1. **Energy Dominance:** The most significant predictor is the feature *Oil_ret* (Crude Oil Returns). We can see a clear positive correlation in the distribution: high oil returns (pink dots) push the SHAP value to the right (increasing the inflation forecast). A strong empirical validation for the "Cost-Push" theory can be attributed to this. It implies that the main way that modern inflation is transmitted is through energy input costs

2. **Momentum Effects:** CPI_MoM (Inflation Momentum) is the second most important feature. This demonstrates that prices are "sticky" because once inflation begins to rise, expectations become unanchored and the trend usually continues.

3. **The Role of Copper & Wheat:** Interestingly, when compared to Oil, Wheat and Copper had a lower impact. This suggests that while these commodities are important, their price

shocks are less systemic to the broader US consumer basket than energy prices.

5.2 Regime Analysis

The model's performance was not uniform across time. During the "Great Moderation" (2010–2019), the Naive model was nearly unbeatable because volatility was low. However, the value of the machine learning model increased significantly during the high-volatility "Regime Shift" of 2021. This tells us that a hybrid approach may be the most optimal strategy for central banks: linear models can be used for stable periods and we can switch to non-linear ML models when supply chain pressure indices (GSCPI) exceed a certain threshold.

6. CONCLUSION

This paper puts forward a robust machine learning framework for forecasting US inflation. By utilizing and integrating alternative data sources, specifically the NY Fed's Supply Chain Index and spectral commodity features, we were able to successfully build a model that competes and is on par with traditional baselines while also offering superior explanatory power.

The trend of applying complex Deep Learning (LSTM) models to every problem can be challenged through our findings. Rigorous ensemble methods like XGBoost prove to be more effective and easy to understand when dealing with monthly macroeconomic data, which is usually sparse and noisy. The strong causal link identified between Crude Oil returns and CPI inflation serves as a quantitative reminder that monetary policy cannot fully control inflation when the drivers are supply-side energy shocks.

Future work will focus on providing policy-makers with real-time "nowcasting" capabilities by improving the temporal resolution of the model by incorporating daily data, such as satellite imagery of shipping ports or daily spot rates.

REFERENCES

1. G. Benigno, J. di Giovanni, J. J. J. Groen, and A. I. Noble, "The GSCPI: A new barometer of global supply chain pressures," *Federal Reserve Bank of New York Staff Reports*, no. 1017, May 2022.

2. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, San Francisco, CA, USA, Aug. 2016, pp. 785–794.

3. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems 30*, 2017, pp. 4765–4774.

4. Federal Reserve Bank of St. Louis, "Federal Reserve Economic Data (FRED)," 2024. [Online]. Available: https://fred.stlouisfed.org/

5. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

6. C. Chakraborty and A. Joseph, "Machine learning at central banks," *Bank of England Staff Working Paper*, no. 674, 2017.