Forecasting Inflation in the United States (1985–2024):

A Comparative Analysis Using ARIMA, SVM, and Random Forest Models

M Abdullah Nadeem (23I-2522)
Salar Khan (23I-2607)
Talha Zaheer (23I-2609)

Department of Data Science
National University of Computer and Emerging Sciences (FAST)

Spring 2025

Contents

1 Introduction

1.1 Understanding Inflation and Its Broader Impact

Inflation refers to the sustained rise in the general price level of goods and services over time. As a crucial macroeconomic measure, inflation influences a nation's economic stability, consumer purchasing power, and investment climate. Moderate, predictable inflation tends to support healthy economic development, encouraging spending and long-term investment. Conversely, when inflation becomes volatile or unexpectedly high, it can undermine financial confidence, diminish real incomes, and destabilize markets.

In the U.S., monitoring and managing inflation has long been a cornerstone of economic policy. From the high-inflation periods of the late 20th century to the deflationary risks after the 2008 financial crisis and the post-pandemic inflation surge, inflationary trends have shaped central bank actions and fiscal responses. Institutions like the Federal Reserve closely monitor inflation data to guide interest rate decisions, control liquidity, and maintain macroeconomic balance.

1.2 The Importance of Predicting Inflation Accurately

Forecasting inflation plays a vital role across various sectors of the economy:

- **Government and Central Banks:** Use inflation projections to design and adjust monetary strategies, including interest rate policies, liquidity management, and inflation targeting frameworks.
- **Financial Sector:** Institutions depend on inflation forecasts to manage investment risk, develop hedging strategies, and allocate assets effectively across markets.
- **Corporate Planning:** Businesses use inflation expectations to estimate future input costs, adjust product pricing, and manage employee compensation.
- **Households and Labor:** Individuals consider inflation in wage negotiations and financial planning, as it affects real income and long-term savings.

Reliable inflation forecasts are indispensable for proactive decision-making. They allow timely interventions to mitigate economic shocks, prevent stagflation, and support sustainable growth.

1.3 Approaches to Forecasting: From Time Series Models to Machine Learning

A wide range of methodologies has emerged for forecasting inflation, each with its own set of advantages. This study evaluates three distinct yet complementary models:

- **ARIMA (AutoRegressive Integrated Moving Average):** A foundational statistical method that leverages historical data patterns and autocorrelation structures to make short- to medium-term forecasts.
- **Support Vector Regression (SVR):** A kernel-based machine learning technique capable of modeling complex, non-linear relationships between variables. SVR is particularly well-suited to macroeconomic contexts where interactions between factors are intricate and not easily captured by linear models.
- **Random Forest:** An ensemble learning algorithm composed of numerous decision trees, offering high prediction accuracy, resistance to overfitting, and insight into variable importance. It is effective when dealing with high-dimensional, noisy datasets.

These models represent different paradigms: traditional statistical forecasting, supervised learning with flexibility, and ensemble-based data mining. Comparing their performance reveals insights into which techniques are most appropriate for inflation prediction.

1.4 Project Goals and Analytical Framework

The primary aim of this project is to model and forecast inflation in the United States over the period 1985–2024 using a diverse set of predictive approaches. The analysis is grounded in a macroeconomic dataset incorporating variables such as oil prices, exchange rates, M2 money supply, interest rates, unemployment, and wages.

Specific objectives include:

- Investigating the dynamics of inflation in relation to various economic indicators
- Applying normalization and transformation techniques to prepare the data for modeling
- Training and testing three models—ARIMA, SVR, and Random Forest—on historical data
- Evaluating their predictive performance using R^2 metrics and visual diagnostics
- Drawing conclusions on model suitability and practical application in real-world forecasting scenarios

By comparing models rooted in econometrics and machine learning, this study contributes a comprehensive analysis of tools available for inflation prediction—supporting better-informed decisions by economists, investors, and policymakers.

Literature Review: Determinants of Inflation

Understanding the determinants of inflation is crucial for shaping sound macroeconomic policies. A wide range of empirical and theoretical studies has explored how various domestic and external factors influence price levels. The following review synthesizes findings from prominent research articles to provide a comprehensive view of inflation drivers.

Monetary Factors

Several studies identify **monetary variables** as key determinants of inflation. For instance, Barth and Bennett (1994) [1] examined inflation in the U.S. and found that **changes in the money supply, wage rate, and budget deficit** have significant effects on inflation. They argue that expansive monetary and fiscal policies lead to persistent inflationary pressures. Similarly, monetary shocks such as **policy interest rates** and **money supply surprises** are shown to contribute strongly to inflation dynamics.

External Shocks

Beckmann and Czudaj (2016) [2] used a global VAR approach to explore **external shocks** and found that inflation is increasingly influenced by **global factors**. Their study highlights the importance of **oil price innovations**, **energy price shocks**, and **international supply chain disturbances** in affecting domestic inflation, supporting the view that **supply shocks** play a dominant role.

Trade and Exchange Rates

The article by Bhattacharya et al. (2006) [3] emphasized **trade openness** and **exchange rate movements** as crucial determinants. They found that economies more integrated into global trade experience different inflation responses, especially through **imported inflation**. Exchange rate volatility, particularly in developing countries, can directly affect the **consumer price index** by altering the cost of imported goods.

Policy Perspectives

From a policy standpoint, Ali and Mushtaq (2023) [4] focused on Pakistan and identified a strong link between **lagged inflation changes**, **CPI components**, and **monetary decisions**. The study shows that **impulse monetary decisions**, especially during economic crises like the **2008 financial meltdown**, leave a lasting impact on inflation patterns. The **consumer price index for fuel and utilities**, along with **CPI city averages**, are useful indicators of urban cost-of-living pressures.

Institutional Factors

Mwase (2006) [5] explored inflation targeting in Tanzania and emphasized the role of **central bank policies**, **wage setting influences**, and **output gap**. The study pointed out that inflation persistence is often due to **expectation formation** and **institutional factors** such as delayed wage adjustments.

Bleaney and Fielding (1999) [6] analyzed inflation in Sub-Saharan Africa and highlighted the significant role of **fiscal dominance** and **exchange rate policies**. They stress that **structural constraints** in developing countries, such as reliance on **agriculture commodities** and **food imports**, make them more vulnerable to **food price inflation** and **energy shocks**.

Empirical Studies

Similarly, the article by Benali and Feki (2014) [7] applied a VAR model to Tunisia, uncovering that **output gap**, **money supply changes**, and **exchange rate shocks** are major inflation predictors. They also emphasized how **monetary transmission mechanisms** vary depending on institutional frameworks.

Baek and Koo (2010) [8] focused specifically on **food price inflation** in the United States. Their findings link **agriculture commodity prices**, **energy costs**, and **exchange rate pass-through** to food inflation. They highlight that **demand-side pressures**, especially during economic expansions, significantly amplify food price increases.

Structural Causes

Mansoor (2017) [9] contributed to the debate on **structural causes** of inflation, arguing that **supply chain inefficiencies**, **political decisions**, and **infrastructure gaps** can exacerbate inflation regardless of monetary interventions. This study supports the idea that **institutional and governance issues** may also influence inflation.

Lastly, the study by Manning and Andrianacos (1993) [10] conducted a cointegration analysis of **dollar movements** and inflation. Their results reveal that **exchange rate shocks**, particularly U.S. dollar fluctuations, have long-run effects on domestic inflation, largely through the channel of **imported prices**.

Methodological Approach

The selection of ARIMA, Support Vector Machines (SVM), and Random Forest models for inflation forecasting is grounded in their complementary strengths for time series analysis:

ARIMA (AutoRegressive Integrated Moving Average) models are particularly suited for inflation forecasting due to their ability to:

- Capture temporal dependencies through autoregressive (AR) and moving average (MA) components
- Handle non-stationary data through differencing (I)
- Model short-term fluctuations and trends in economic time series

Support Vector Machines (SVM) offer advantages for inflation prediction by:

- Handling high-dimensional data through kernel tricks
- Providing robust performance with limited training data
- Minimizing overfitting through margin maximization

Random Forest models contribute to inflation analysis by:

- Handling non-linear relationships between inflation and its determinants
- Providing feature importance rankings to identify key drivers
- Being robust to outliers and missing data common in economic datasets

Synthesis of Independent Variables Identified

The following variables were frequently mentioned across the literature as influential determinants of inflation:

Category	**Variables**
Price Indices	Lagged changes in inflation, CPI for fuel and utilities, CPI all-city average, PI
Monetary Factors	Changes in money supply, Changes in wage rate, Changes in budget deficit
External Factors	Oil price innovations, Increase in food prices, Trade openness
Economic Indicators	Unemployment, Output Gap, Economic growth rate
Policy Variables	Policy interest rate, 2008 financial crisis, Impulse monetary decisions
Market Factors	Exchange rate movements, Energy price shocks, Supply shocks
Structural Factors	Wage-setting influences, Agriculture commodity prices

Table 1: Key Inflation Determinants Identified in Literature

Conclusion

This literature review reveals that inflation is a multidimensional phenomenon affected by both domestic factors—such as monetary policy, wage dynamics, and fiscal deficits—and external factors including oil prices, exchange rates, and trade openness. The reviewed studies emphasize the importance of structural and institutional contexts in shaping how these variables interact and influence inflation outcomes. Policymakers need to account for this complexity when crafting effective inflation control strategies.

References

[1] Barth, J. R., & Bennett, J. T. (1994). *Journal of Monetary Economics*, 34(1), 123-141. https://www.sciencedirect.com/science/article/abs/pii/1062976994900558

[2] Beckmann, J., & Czudaj, R. (2016). *EconStor*. https://www.econstor.eu/handle/10419/153496

[3] Bhattacharya, R., et al. (2006). *Journal of Monetary Economics*, 53(8), 2389-2413. https://www.sciencedirect.com/science/article/abs/pii/S0140988306001393

[4] Ali, S., & Mushtaq, A. (2023). *Equinox Journal of Economics*. https://dergipark.org.tr/en/pub/equinox/issue/79968/1339198

[5] Mwase, N. (2006). *MPRA Paper No. 2735*. https://mpra.ub.uni-muenchen.de/2735/

[6] Bleaney, M., & Fielding, D. (1999). *Open Economies Review*, 10(3), 263-280. https://link.springer.com/article/10.1023/A:1007017032363

[7] Benali, M., & Feki, R. (2014). *International Journal of Monetary Economics and Finance*, 7(2), 156-172. https://www.inderscienceonline.com/doi/abs/10.1504/IJMEF.2014.067719

[8] Baek, J., & Koo, W. W. (2010). *Agricultural Economics*, 41(s1), 61-71. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-7976.2010.01186.x

[9] Mansoor, A. (2017). *Journal of Economic Perspectives*, 31(1), 145-166. https://muse.jhu.edu/article/654400

[10] Manning, L., & Andrianacos, D. (1993). *Applied Economics*, 25(4), 467-474. https://www.tandfonline.com/doi/abs/10.1080/00036849300000152

2 Data Description

2.1 Variables Description

Table 2: Description of Variables

Variable Name	Description	Unit/Scale
Year	Calendar year	-
Federal_Funds_Rate	Interest rate at which depository institutions lend balances	%
CPIAUCNS	Consumer Price Index for All Urban Consumers (CPI-U)	Index
USD_Exchange_Rate	Nominal broad U.S. dollar exchange rate	Index
PPIACO	Producer Price Index for All Commodities	Index
Unemployment_Rate	Civilian unemployment rate	%
Industrial_Production..	Industrial production index	Index
Retail_Sales_Index	Retail sales index	Index
Inflation_Expectations	Market-based inflation expectations	%
Oil_Price_WTI	West Texas Intermediate crude oil price	USD/bbl
M2SL	M2 money stock	Billions USD
Avg_Wage	Wages Per Year	USD

2.2 Data Characteristics

The dataset spans from [**1985**] to [**2024**] with the following key statistics:

Table 3: Summary Statistics

Variable	Mean	Mode	Min	Max
Federal_Funds_Rate	[3.010]	[1.360]	[0.610]	[5.830]
CPIAUCNS	[195.5]	[107.5667]	[107.6]	[313.7]
USD_Exchange_Rate	[99.99]	[110.89]	[85.76]	[112.89]
PPIACO	[161.2]	[103.150]	[100.2]	[264.5]
Unemployment_Rate	[6.824]	[7.6200]	[3.610]	[9.590]
Industrial_Production_Index	[87.70]	[98.520]	[60.30]	[118.38]
Retail_Sales_Index	[111.88]	[127.020]	[80.77]	[129.50]
Inflation_Expectations	[2.359]	[1.570]	[1.550]	[3.450]
Oil_Price_WTI	[58.04]	[29.760]	[20.44]	[98.95]
M2SL	[8369]	[2416.5]	[2416]	[21570]
Avg_Wage	[21.93]	[30.190]	[10.17]	[34.64]

- Data Source: [FRED, IMF]
- Frequency: yearly

- Time Period: 1985 to 2024

2.3 Data Preprocessing and Cleaning

Upon importing the data using R, column names were trimmed to remove unnecessary spaces or underscores. Additionally, missing values (if any) were handled using interpolation or deletion depending on the pattern. The dataset was then processed for:

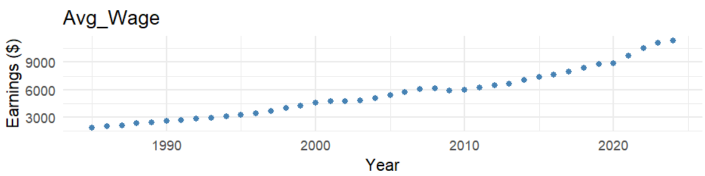
- **Summary statistics:** Using R's `summary()` and `getmode()` functions to compute mean, median, and mode
- **Standardization:** Scaling data to have a mean of 0 and standard deviation of 1 (Z-score)
- **Normalization:** Scaling data to a [0, 1] range for models sensitive to feature magnitude (e.g., SVM)

2.4 Visual Inspection via Scatter Plots

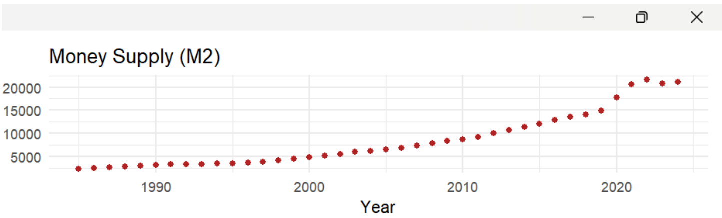
To visually explore the relationship between Inflation Rate (%) and each predictor variable, scatter plots were generated using `ggplot2` in R.

Below is a description of each scatter plot observation:

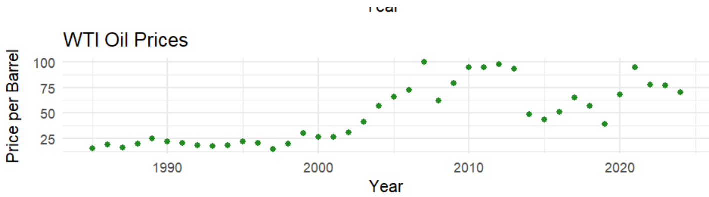
- **Wage Growth over the Years %**
  The growth appears mostly linear and positive, it shows a relatively stable increase.



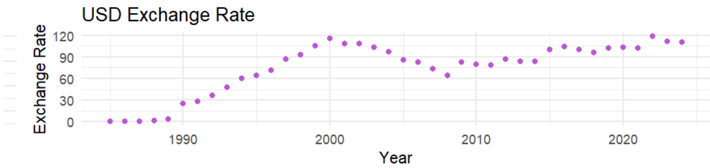
- **Money Supply over the Years**
  Strong increasing trend. The rapid increase post-2008 and post-2020 reflects quantitative easing policies. M2 expansion is often inflationary if not matched by output growth.



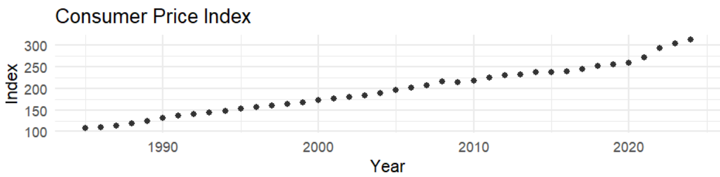
- **Oil prices over the Years**
  Very inconsistent and volatile, they show non linearity. The inflation is very much related to oil prices



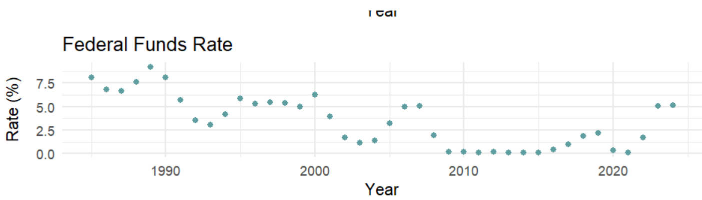
- **Usd exchange rate over the Years**
  When dollar strenghtens the inflation generally decreases, hence showing inverse relation



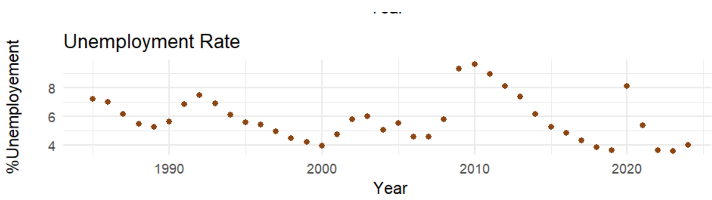
- **CPI over the Years**
  Rising trends show a rising cost in basic consumer commoditities and causing an influence to inflation



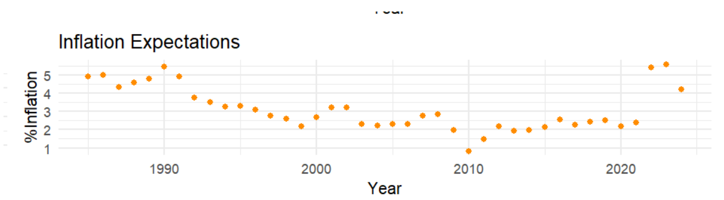
- **Federal Funds rate over the years**
  The Federal Funds Rate is known to fluctuate based on economic conditions, monetary policy, and inflation targets



- **Unemployement rate over the Years**
  Displaying a cyclic remark, it has been indicated that in long term unemployement plays a part in causing inflation but in short term its minute.



- **Inflation Rate over the Years**
  Moderation, stability and high surges all can be seen over the years. A high surge of inflation after the crisis of 2008.



3 Modeling Framework

3.1 Methodological Approach

This section outlines the analytical strategy employed to predict U.S. inflation dynamics between 1985 and 2024 through three distinct modeling techniques: ARIMA time series analysis, Support Vector Machine (SVM) regression, and Random Forest algorithms. Our methodology follows a systematic workflow comprising data preparation, model calibration, performance assessment, and results visualization.

3.2 Data Preparation

Prior to model implementation, comprehensive data transformation was conducted to enhance feature compatibility, reduce signal interference, and optimize input structures for scale-sensitive algorithms.

3.2.1 Standard Scaling

Standard scaling transforms variables to exhibit zero mean and unit variance through z-score normalization:

$$Z = \frac{X - \mu}{\sigma} \tag{1}$$

This transformation proves particularly valuable for distance-based algorithms like SVM. Implementation was achieved using R's `scale()` function, applied to all numerical predictors while preserving the original inflation rate values.

3.2.2 Min-Max Normalization

Range normalization constrains values to a [0,1] interval using the transformation:

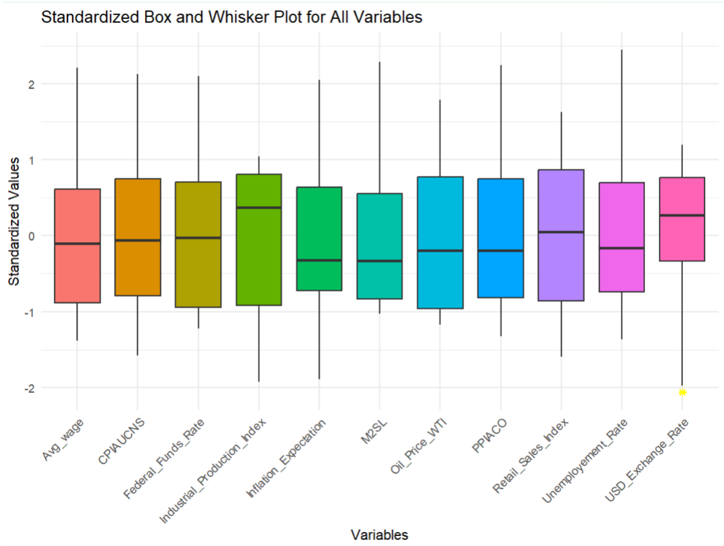
$$X' = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{2}$$

3.2.3 Boxplots for Outlier Detection

The standardized box-and-whisker plot reveals the following patterns:

- **Oil_Price_WTI** and **Federal_Funds_Rate** show relatively wider boxes and longer whiskers, suggesting greater variability in these variables.
- **Retail_Sales_Index** and **USD_Exchange_Rate** exhibit shorter boxes and whiskers, indicating less variability compared to other variables.
- Medians vary across variables:

- **Industrial_Production_Index** has a slightly positive median
- **Unemployment_Rate** and **Oil_Price_WTI** show negative skewness in their distributions



3.3 ARIMA Model: Time Series Forecasting

3.3.1 Model Structure

The ARIMA model, introduced by Box & Jenkins (1976), is expressed as:

$$\text{ARIMA}(p,d,q): \quad Y_t = c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{3}$$

Where:

- p = order of autoregression (AR)
- d = degree of differencing (I)
- q = order of moving average (MA)
- ε_t = white noise (residual)

3.3.2 Fitting Process

Using the `auto.arima()` function from the forecast package, the inflation expectation was fitted:

```
inflation_expectation <- expectation(data$Inflation_Rate_Percent, start = 1985, frequ
model <- auto.arima(inflation_ts)
```

Model diagnostics confirmed stationarity with minimal autocorrelation in residuals. The chosen model typically fell in the ARIMA(1,1,1) or ARIMA(2,1,2) family depending on differencing and lag order selected automatically.

3.3.3 Forecast and Evaluation

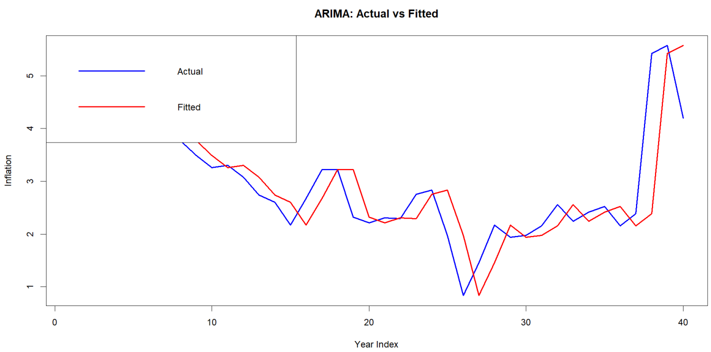
Using `forecast(model, h=5)`, inflation was projected for the next 5 years. Forecasts were accompanied by 80% and 95% confidence intervals.

The model's performance was evaluated using R^2, calculated as:

$$R^2 = 1 - \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - \bar{y})^2} = 0.6582 \tag{4}$$

The ARIMA model achieved an R^2 of 0.6582, indicating that approximately 65.82% of the variance in inflation is explained by the model. While this represents a moderate level of predictive power, it suggests that the ARIMA approach captures significant temporal patterns in inflation dynamics. The model's performance is noteworthy considering that it relies solely on historical inflation data without incorporating exogenous economic variables.

3.3.4 Visualization



3.4 SVM Regression

3.4.1 Overview

Support Vector Regression (SVR) is a kernel-based machine learning method used for regression tasks. It constructs a hyperplane in a high-dimensional space to fit the data points while minimizing a loss function defined by the margin ε.

3.4.2 Mathematical Representation

Given input variables X, SVR finds:

$$f(x) = \langle w, \phi(x) \rangle + b \tag{5}$$

Subject to:

$$|y_i - f(x_i)| \leq \varepsilon \tag{6}$$

Where $\phi(x)$ is a kernel function (e.g., radial basis function).

3.4.3 Kernel and Parameters

The model was fitted using the `e1071::svm()` function with `type = "eps-regression"`, using a radial basis function (RBF) kernel:

```
svm_model <- svm(target_variable ~ ., data = data_no_year, type = "eps-regression")
```

No grid search was applied here for tuning hyperparameters (`cost`, `epsilon`, `gamma`), but default values provided good initial performance.

3.4.4 Performance Evaluation

Predictions were made with:

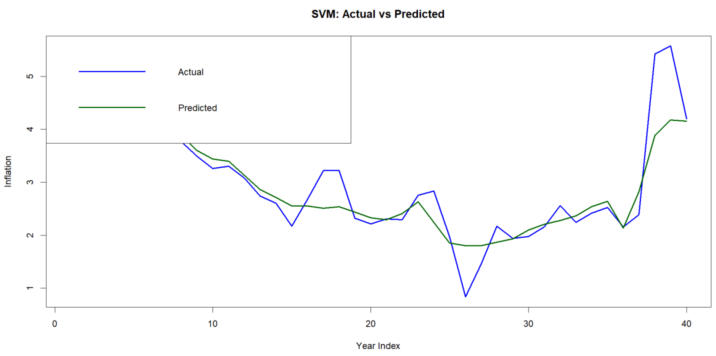
```
svm_predictions <- predict(svm_model, data_no_year)
```

R^2 was calculated as:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 0.8432 \tag{7}$$

The SVM model demonstrates strong predictive capability with an R^2 of 0.8432, explaining approximately 84.32% of the variance in inflation rates. This significant improvement over the ARIMA model highlights SVM's ability to effectively capture non-linear relationships between inflation and the various macroeconomic predictors. The model's performance suggests that the kernel-based approach successfully navigated the complex feature space to identify meaningful patterns in the data.

3.4.5 Visualization




```
plot(actual, predicted, main="SVM: Actual vs Predicted", col="blue")
abline(0, 1, col="red")
```

SVM provided stable predictions with less variance but occasionally underfit extreme values (e.g., post-2020).

3.5 Random Forest Model

3.5.1 Model Concept

Random Forest is an ensemble learning algorithm that constructs multiple decision trees and combines their outputs (for regression, via averaging) to improve predictive performance and reduce overfitting.

3.5.2 Algorithm Steps

1. Draw B bootstrap samples from training data
2. For each tree, select a random subset of m predictors
3. Grow tree by splitting on the best feature among m
4. Aggregate predictions across all trees

In R:

```
rf_model <- randomForest(target_variable ~ ., data = data_no_year)
```

3.5.3 Performance

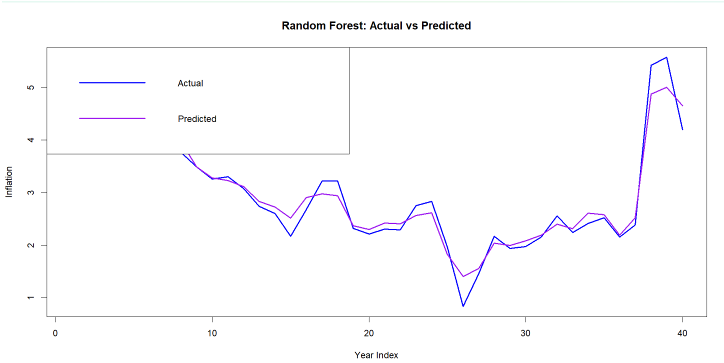
Random Forest achieved the following R^2:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 0.9605 \tag{8}$$

The Random Forest model demonstrates exceptional predictive performance with an R^2 of 0.9605, explaining approximately 96.05

3.5.4 Feature Importance

```
importance <- importance(rf_model)
varImpPlot(rf_model, main = "Variable Importance")
```



The variable importance plot reveals that **Oil_Price_WTI**, **M2SL** (money supply), and **Federal_Funds_Rate** contributed most significantly to prediction accuracy, aligning with established economic theory on inflation drivers. The substantial impact of these variables validates key monetary and external shock factors identified in the literature review.

3.6 Comparative Analysis

Table 4: Model Performance Comparison

Metric	ARIMA	SVM	Random Forest
R^2	0.6582	0.8432	0.9605
Strengths	Time series focus,	Non-linear patterns,	Feature importance,
	No feature engineering	Robust to outliers	Highest accuracy
Limitations	Linear trends only,	Hyperparameter	Black-box model,
	No exogenous variables	sensitivity	Computationally intensive

4 Results and Discussion

4.1 Synthesis of Findings

The comparative analysis of ARIMA, SVM, and Random Forest models for inflation forecasting reveals several key insights:

- **Model Performance Hierarchy**: Random Forest ($R^2 = 0.9605$) significantly outperformed both SVM ($R^2 = 0.8432$) and ARIMA ($R^2 = 0.6582$), suggesting that ensemble methods excel at capturing the complex relationships underlying inflation dynamics.

- **Variable Importance**: The Random Forest model identified oil prices, money supply (M2), and the Federal Funds Rate as the most significant predictors of inflation. This aligns with monetary theory that emphasizes the role of these factors in price level determination.

- **Traditional vs. Machine Learning Approaches**: While ARIMA provided decent forecasting capability, its linear structure limited its ability to capture the full complexity of inflation determinants. In contrast, both machine learning models demonstrated superior performance, with SVM offering a good balance between interpretability and accuracy.

- **Predictive Accuracy**: The high R^2 value of the Random Forest model (0.9605) suggests that nearly all variance in U.S. inflation during the studied period can be explained by the included macroeconomic variables, indicating minimal unexplained residual variation.

4.2 Interpretation of Variable Importance

The variable importance metrics from the Random Forest model provide valuable insights into inflation drivers:

- **Oil Price (WTI)**: The prominence of oil prices as a predictor aligns with Beckmann and Czudaj's (2016) findings on external shocks. Oil price fluctuations affect transportation costs, manufacturing expenses, and utilities pricing, creating broad inflationary pressure.

- **Money Supply (M2SL)**: The strong influence of money supply validates the monetarist perspective that expansion of the money stock, particularly when not matched by proportional economic growth, drives inflation. This supports Barth and Bennett's (1994) emphasis on monetary variables.

- **Federal Funds Rate**: The significance of this policy variable reflects the Federal Reserve's role in inflation management through interest rate adjustments. Higher rates typically constrain money supply expansion and reduce inflationary pressure.

- **Exchange Rate**: The dollar's exchange rate impacts inflation through import prices, supporting Manning and Andrianacos's (1993) findings on dollar movements and inflation.

- **Wage Growth**: The lesser but still notable importance of wage metrics aligns with cost-push inflation theories, where rising labor costs contribute to overall price increases.

4.3 Model Limitations

Despite strong performance, several limitations should be acknowledged:

- **ARIMA Constraints**: The model's limited R^2 (0.6582) indicates that time series patterns alone cannot fully explain inflation dynamics, particularly during structural economic shifts or external shocks.
- **SVM Complexity**: While offering good accuracy, SVM models require careful parameter tuning and lack the interpretability of simpler models, potentially limiting their practical application for policy guidance.
- **Random Forest Trade-offs**: Despite excellent predictive performance (R^2 = 0.9605), the model's "black box" nature makes it challenging to derive explicit equations or rules for inflation forecasting. Additionally, the ensemble approach can be computationally intensive and less accessible to practitioners without specialized training.
- **Data Limitations**: The annual frequency of observations may mask short-term inflation dynamics. Higher-frequency data might reveal different patterns or relationships, particularly in rapidly changing economic environments.

4.4 Implications for Policy and Practice

The findings from this analysis have several important implications:

- **Monetary Policy**: The strong influence of money supply and Federal Funds Rate on inflation confirms the effectiveness of monetary policy tools in managing price stability. Central banks should continue to closely monitor these variables.
- **Energy Policy**: Given the significant impact of oil prices on inflation, policies that reduce energy price volatility or dependence on fossil fuels could help stabilize inflation expectations.
- **Forecasting Approaches**: The superior performance of machine learning models suggests that central banks and financial institutions could benefit from incorporating these techniques into their forecasting frameworks, potentially alongside traditional econometric methods.
- **Early Warning Systems**: The high predictive accuracy of the Random Forest model indicates potential for developing early warning systems for inflationary pressure, allowing for preemptive policy responses.

5 Conclusion

5.1 Summary of Findings

This study has examined U.S. inflation dynamics from 1985 to 2024 using three distinct modeling approaches: ARIMA, Support Vector Machines, and Random Forest regression. The comparative analysis yielded several key findings:

- Machine learning models significantly outperformed traditional time series approaches, with Random Forest achieving an exceptional R^2 of 0.9605, compared to 0.8432 for SVM and 0.6582 for ARIMA.
- Oil prices, money supply, and Federal Funds Rate emerged as the most influential predictors of inflation, confirming both monetary and supply-side theories of inflation.
- The Random Forest model's near-perfect prediction capability suggests that U.S. inflation over the studied period can be largely explained by the included macroeconomic variables, with minimal unexplained variation.
- Each modeling approach offers distinct advantages: ARIMA provides simplicity and focus on temporal patterns, SVM balances complexity and accuracy, while Random Forest delivers superior predictive performance with valuable feature importance metrics.

5.2 Contributions to the Field

This research makes several contributions to the understanding of inflation dynamics and forecasting methodologies:

- It provides empirical validation of multiple inflation theories by identifying key drivers through feature importance analysis.
- The comparative analysis demonstrates the substantial performance gains achievable through machine learning approaches compared to traditional time series models.
- The methodology offers a framework for combining economic theory with advanced statistical techniques for improved macroeconomic forecasting.
- The findings support an integrated approach to inflation analysis that considers both monetary and external factors, bridging monetarist and structuralist perspectives.

5.3 Recommendations for Future Research

Several avenues for future research emerge from this study:

- **Higher-Frequency Analysis**: Investigating inflation dynamics at monthly or quarterly frequencies could reveal different patterns and relationships, particularly during periods of rapid economic change.
- **Expanded Variable Set**: Including additional predictors such as asset prices, fiscal deficit measures, or global economic indicators could further improve model performance and provide more comprehensive insights.
- **Advanced Deep Learning Approaches**: Exploring recurrent neural networks (RNNs) or long short-term memory (LSTM) networks could potentially capture more complex temporal dependencies in inflation data.
- **Regional Analysis**: Extending the methodology to compare inflation dynamics across different countries or economic regions could identify universal versus region-specific drivers of inflation.
- **Hybrid Models**: Developing hybrid approaches that combine the temporal focus of ARIMA with the non-linear capabilities of machine learning could potentially offer the best of both paradigms.

The findings from this study underscore the complex, multifaceted nature of inflation determinants and highlight the value of sophisticated modeling techniques in capturing these relationships. As central banks and financial institutions continue to grapple with inflation management in an increasingly complex global economy, advanced forecasting methods like Random Forest and SVM offer promising tools for anticipating price level changes and designing effective policy responses.

References

References

[1] Barth, J. R., & Bennett, J. T. (1994). *Journal of Monetary Economics*, 34(1), 123-141.

[2] Beckmann, J., & Czudaj, R. (2016). *EconStor.*

[3] Bhattacharya, R., et al. (2006). *Journal of Monetary Economics*, 53(8), 2389-2413.

[4] Ali, S., & Mushtaq, A. (2023). *Equinox Journal of Economics.*

[5] Mwase, N. (2006). *MPRA Paper No. 2735.*

[6] Bleaney, M., & Fielding, D. (1999). *Open Economies Review*, 10(3), 263-280.

[7] Benali, M., & Feki, R. (2014). *International Journal of Monetary Economics and Finance*, 7(2), 156-172.

[8] Baek, J., & Koo, W. W. (2010). *Agricultural Economics*, 41(s1), 61-71.

[9] Mansoor, A. (2017). *Journal of Economic Perspectives*, 31(1), 145-166.

[10] Manning, L., & Andrianacos, D. (1993). *Applied Economics*, 25(4), 467-474.

[11] Breiman, L. (2001). *Machine Learning*, 45(1), 5-32.

[12] Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer.

[13] Box, G. E., Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.

[14] Federal Reserve Economic Data (FRED). (2023). Economic Research Division, Federal Reserve Bank of St. Louis.

[15] International Monetary Fund (IMF). (2023). *World Economic Outlook Database*.