



**KIIT Deemed to be University**

## **Assignment 2**

30 October 2025

**Submitted By:**

**Satyam Kumar**

Roll No.: 22051615 (CSE)

B.Tech in Computer Science Engineering

**Submitted To:**

**Prof. N Sangita Achary**

KIIT Deemed to be University

# Time Series Forecasting: All India Consumer Price Index Analysis

## Dataset

Source: All India CPI Index Data (Government of India - Rural+Urban Combined Sector)

Dataset Link: [Dataset](#)

Notebook File: [CPI\\_TimeSeries\\_Analysis\\_1.ipynb](#)

## 1. Dataset Overview

The dataset contains monthly Consumer Price Index observations from January 2013 to June 2020, covering a period of 7.5 years with 87 sequential data points. The CPI is a crucial economic indicator measuring average price changes for consumer goods and services.

Attribute	Description
Time Period	January 2013 to June 2020
Total Observations	87 monthly data points
Sector	Rural + Urban Combined
Data Quality	Clean, no missing values
Initial CPI	104.6 (Jan 2013)
Final CPI	151.8 (Jun 2020)

The dataset exhibits a clear upward trend with consistent inflation patterns throughout the observed period.

## 2. Model Selection Rationale

The ARIMA (AutoRegressive Integrated Moving Average) model was selected based on systematic analysis:

### (a) Stationarity Testing

The Augmented Dickey-Fuller (ADF) test revealed non-stationarity in the original series, necessitating first-order differencing to achieve stationarity.

### (b) Parameter Selection via Grid Search

A comprehensive grid search was conducted over ARIMA parameters (p, d, q) ranging from 0 to 5. The Akaike Information Criterion (AIC) was used to identify the optimal model configuration.

### (c) ARIMA(1,1,0) Model

The selected model combines one autoregressive term with first-order differencing. This configuration effectively captures the inflation trend while maintaining model parsimony.

**Model Selection Criteria:**

- Lowest AIC value among all tested configurations
- Balance between model complexity and predictive accuracy
- Strong performance on both training and testing sets
- Interpretability and ease of implementation

3. Forecasting Results

Metric	Value
MAE (Mean Absolute Error)	7.23
RMSE (Root Mean Square Error)	8.18
MAPE (%)	4.61%
Model Accuracy	95.39%

Interpretation:

- MAE measures average forecast error in absolute terms (7.23 CPI points).
- RMSE penalizes large deviations, indicating minimal outlier errors.
- MAPE of 4.61% demonstrates strong predictive accuracy.
- Model accuracy of 95.39% confirms excellent performance.

4. Visualization of Results

The following visualizations were generated to validate model performance:

- Original Time Series Plot: Displays CPI trend from 2013-2020 with clear upward trajectory.
- Rolling Statistics Plot: Shows increasing rolling mean and stable rolling standard deviation.
- Seasonal Decomposition: Separates trend, seasonal, and residual components.
- ACF and PACF Plots: Guided parameter selection for ARIMA model.
- Training vs Testing Forecast: Demonstrates model accuracy on hold-out test set.
- Residual Analysis Plot: Confirms random distribution with no systematic patterns.
- Future Forecast Visualization: Shows 12-month predictions maintaining upward trend.

These visualizations confirm that the ARIMA model effectively captures the underlying inflation patterns with minimal systematic bias.

5. Error Analysis

a. Historical Inflation Analysis

Metric	Value
Total Inflation (2013-2020)	45.12%
Annual Average Inflation	6.09%

Metric	Value
Training Observations	70 months (80%)
Testing Observations	17 months (20%)

## b. Key Observations

- The model successfully captures the upward inflation trend without systematic bias.
- Residual analysis shows random distribution centered at zero, validating model assumptions.
- Test set predictions closely match actual values with minimal deviation.
- Future forecasts maintain consistent inflation patterns observed historically.

## c. Future Forecast (12 Months Ahead)

The model generated 12-month forecasts (July 2020 - June 2021) with CPI values ranging from 153.73 to 156.66, representing a projected inflation of approximately 1.9% over the forecast period. These predictions can inform economic planning and monetary policy decisions.

## 6. Conclusion

This analysis successfully demonstrates the application of time series forecasting techniques to India's Consumer Price Index data. The key findings are:

### Model Performance:

- ARIMA(1,1,0) achieved 95.39% accuracy with MAPE of only 4.61%.
- Systematic grid search and AIC optimization ensured optimal parameter selection.
- Residual analysis validated model assumptions and confirmed adequate fit.

### Economic Insights:

- Total inflation of 45.12% over 7.5 years (annual average: 6.09%).
- Consistent upward trend in CPI indicating sustained inflationary pressure.
- Future forecasts suggest continued inflation, valuable for policy planning.

### Practical Applications:

The forecasting framework developed can be applied to:

- Monetary policy formulation and interest rate decisions.
- Economic planning and resource allocation strategies.
- Business forecasting for pricing and cost management.

### Future Enhancements:

While the current model performs excellently, potential improvements include:

- Exploring SARIMA models if seasonal patterns are detected in longer time series.
- Incorporating exogenous variables (GDP growth, oil prices) using SARIMAX.
- Implementing machine learning approaches (LSTM, Prophet) for comparison.
- Developing ensemble methods combining multiple forecasting techniques.