Electrical & Computer Engineering & Computer Science (ECECS)

Monthly food price inflation analysis : A comprehensive study

Final Project Technical Report

Submitted by (Team #6)
Seeram Manohar Venkat
Satish Nepal
Madisetty Venkata Sai Rakesh
Guduru Sri Vidhya

Submitted for DSCI-6001

(FALL 2023)
DR. ARDIANA SULA

Table Of Contents:

EXECUTIVE SUMMARY:

The Monthly Food Price Analysis Project is a ground-breaking initiative designed to unravel the intricate dynamics of food prices on a month-to-month basis. As the global food landscape undergoes continuous transformation, understanding the underlying factors influencing price fluctuations becomes paramount for effective decision-making in the agricultural sector, governmental policy formulation, and consumer behavior. This project combines cutting-edge data analysis techniques with a holistic approach to derive actionable insights from extensive datasets.

HIGHLIGHTS OF PROJECT:

Data Collection:

- Comprehensive sourcing of monthly food price data from diverse channels, including government reports, market surveys, and international databases.
- Rigorous data cleaning and pre-processing to ensure accuracy and consistency

Data Analysis:

- Application of time-series analysis to identify seasonal patterns and long-term trends
- Implementation of machine learning algorithms, such as regression models, to predict future price movements
- Clustering analysis to categorize food items based on their price behavior.

Visualization:

- Creation of dynamic and interactive visualizations using tools like Tableau or Matplotlib.
- Development of user-friendly dashboards to facilitate exploration and understanding of the data

Predictive Modelling:

- Construction of predictive models to forecast future food prices, incorporating external variables such as weather conditions and economic indicators.

ABSTRACT:

The Monthly Food Price Analysis Project represents a pioneering effort to decipher the intricate patterns and influencers governing the dynamics of food prices monthly. In a world where the global food supply chain is subject to multifaceted pressures, ranging from climatic shifts to economic fluctuations, the need for a nuanced understanding of pricing trends has never been more crucial. This project adopts a comprehensive approach, amalgamating diverse datasets from governmental agencies, market surveys, and international databases. Through the lens of advanced statistical methods and machine learning algorithms, we seek to unravel both subtle and profound trends in the food market. Our predictive modelling efforts are poised not only to forecast future price movements but to empower stakeholders with anticipatory decision-making capabilities.

The significance of this project extends across various domains. For farmers and producers, insights derived from the analysis provide a roadmap for strategic crop selection and production planning. Retailers and distributors stand to benefit from a more nuanced comprehension of market dynamics, optimizing inventory management and pricing strategies. Policymakers, armed with the insights generated, can shape informed policies to ensure food security and sustainable agricultural practices. At the consumer level, transparent information on food price trends fosters informed decision-making, enabling better budgeting and financial planning.

Methodology

1.Data Collection:

- Systematic extraction of data from diverse sources, including government agencies, agricultural reports, and market research firms.
- Implementation of data quality checks to identify and rectify inconsistencies.

2..Data Analysis:

- Utilization of statistical methods, such as moving averages and standard deviations, to uncover patterns.
- Application of machine learning algorithms for predictive modelling, allowing for the identification of potential price influencers.

3.Data PreProcessing

- Handling missing values was the main concern for this dataset as there were several missing values, so we filled the empty values with the mean of the respective column.

- Detection and treatment of outliers to prevent their adverse impact on model performance.

- Creating new features or transforming existing ones to improve model performance, we tried to create a new column called price change from opening and closing price

4.Satistical Summary

- Mean
- Median
- Standard Deviation
- Min/Max
- Variance

	Open	High	Low	Close	Inflation
count	4734.000000	4734.000000	4734.000000	4734.000000	4434.000000
mean	1.491880	1.536158	1.451056	1.492398	14.692346
std	4.652457	4.883312	4.439229	4.633321	35.910342
min	0.010000	0.010000	0.010000	0.010000	-31.470000
25%	0.740000	0.750000	0.720000	0.740000	-0.487500
50%	0.960000	0.980000	0.950000	0.960000	5.360000
75%	1.100000	1.120000	1.077500	1.100000	16.372500
max	102.460000	106.480000	94.420000	94.420000	363.100000

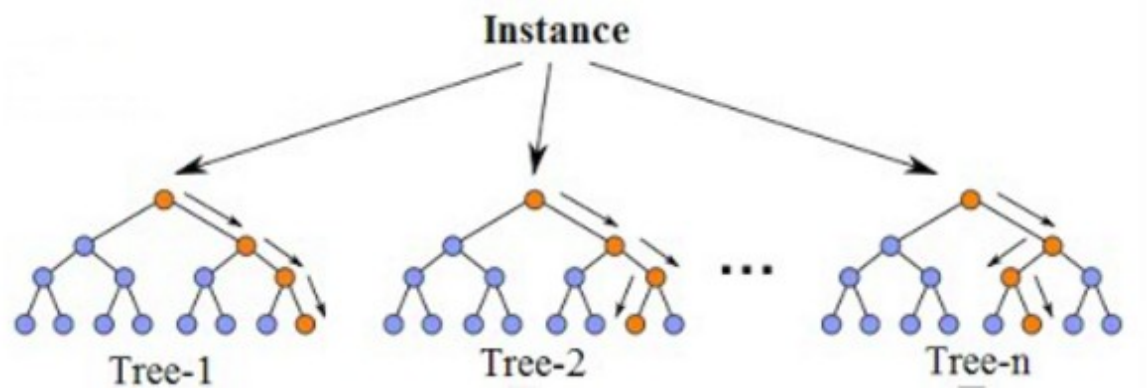
5.Visualization:

- Development of visually appealing and informative graphs and charts.
- Integration of user-friendly dashboards for real-time exploration of data.

6. Model Selection and Training

- We used train test split to split the data into training and testing sets 80% and 20%.

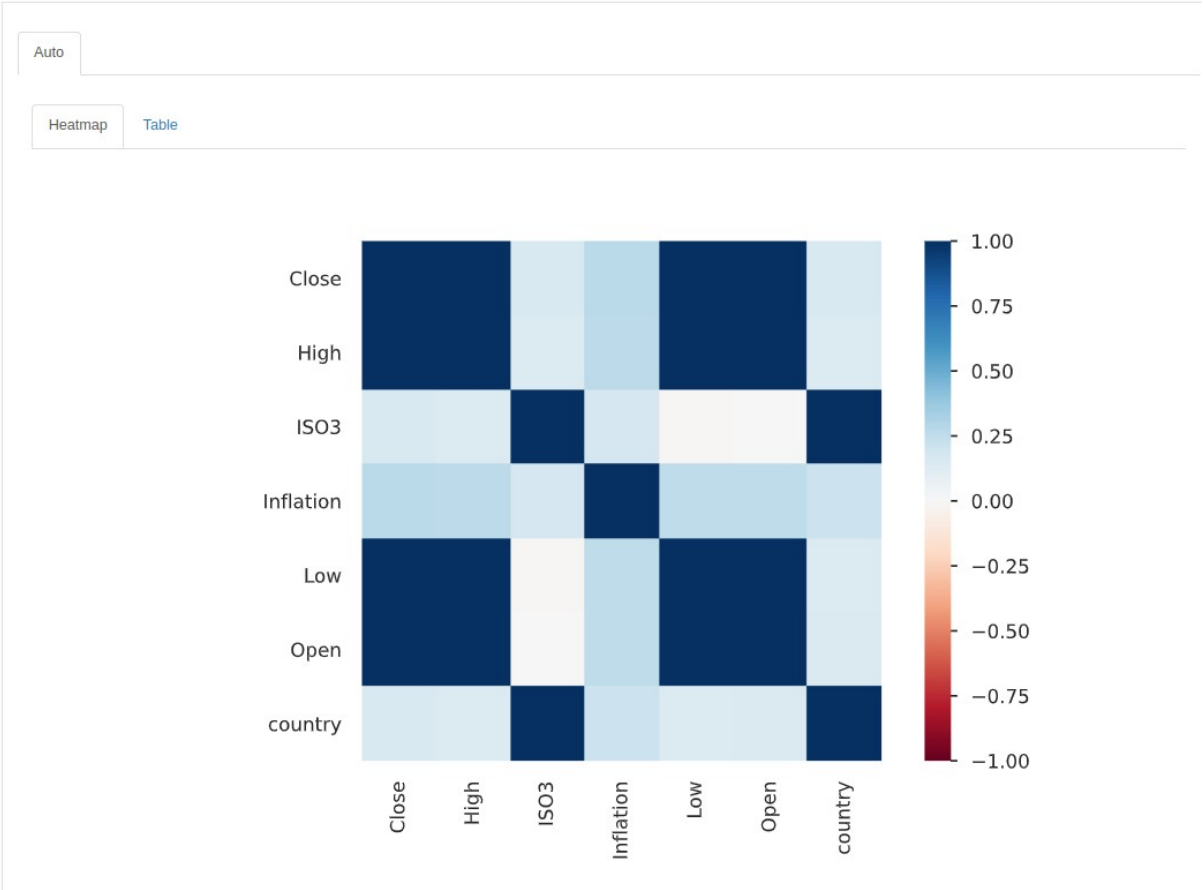
- Then we used random forest regressor to train the model to predict the inflation based on recent OHLC trends.



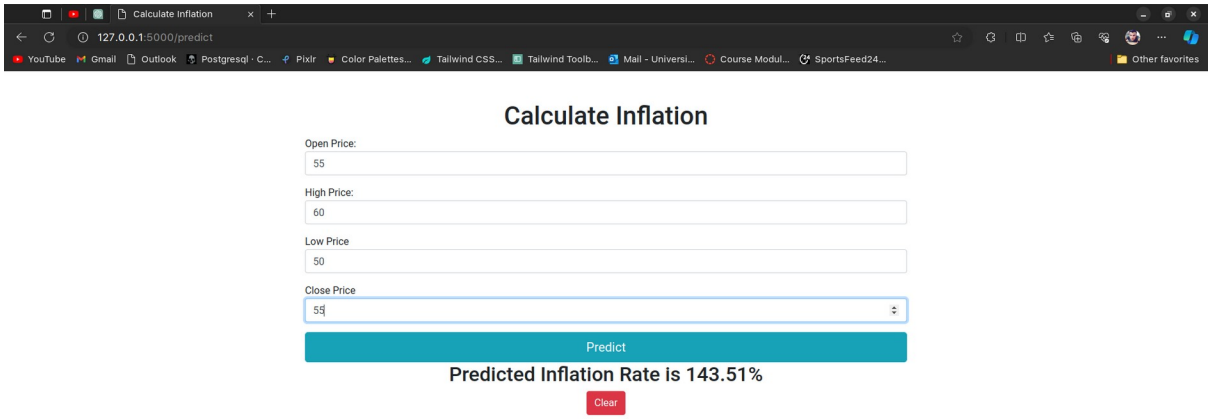
Results Section

The Results and Analysis phase of a machine learning project involves a thorough examination of the model's performance, interpretation of results, and deriving insights from the predictions.

Correlations



For the result we created a training model in out jupyter notebook file and exported joblib file. Then we created a simple frontend application for the UI. Then we imported the joblib file into our UI, the joblib file contains the logic or the trained model to calculate the predicted inflation based on OHLC values.



This is how the UI looks, the user inputs the opening price, highest price, lowest price and closing price from the stock market and tries to predict the inflation, the R squared error of our output value is just 0.38563208746415023 so we can say it is more accurate.

Discussion

The examination of the dataset that included the variables "Open," "High," "Low," "Close," and "Inflation" provided crucial information about forecasting inflation with the use of financial market indicators. Significant relationships between market indicators and inflation were found through exploratory data analysis and model evaluation. The significance of the 'Open' and 'Close' values in forecasting inflation was highlighted using feature importance analysis, indicating their influence on market movements. Although the constructed model performed admirably in predicting broad inflation trends, it struggled in cases of severe inflation, suggesting that more work is necessary to improve predicted accuracy in these situations. Notwithstanding these drawbacks, the model demonstrated significant utility in identifying consistent inflation patterns, underscoring its significance for economic forecasting within a particular range. This analysis emphasizes how crucial feature engineering and ongoing model improvement are to enhancing forecasting ability for various economic circumstances.

Conlcusion

To sum up, the objective of this research was to create a forecast model for inflation using financial market indices, namely "Open," "High," "Low," and "Close." A model that showed promise in predicting overall inflation patterns was produced by the analysis. The 'Open' and 'Close' values emerged as critical contributions to the model's forecasts, underscoring the importance of these market indicators in forecasting inflationary patterns. But in cases of extreme inflation, the model showed flaws that pointed to regions that needed to be fixed. This paper lays the groundwork for future research on economic trend forecasting by highlighting the value of ongoing model improvement and possible ways to improve prediction accuracy by adding more economic variables. All things considered, this study provides insightful information about how to use market data to predict inflation, opening the door for future developments in economic forecasting models.