MODELLING CONSUMER PRICE INDEX DYNAMICS IN NIGERIA USING DECISION TREE AND RANDOM FOREST MACHINE LEARNING ALGORITHMS

[1]IBEKWE, U. A.; [2]MESIKE, G. C. & [3]ASHOGBON, M.B.
[1,2]Department of Actuarial Science & Insurance
Faculty of Management Sciences
University of Lagos, Nigeria
[3]Department of Banking and Finance,
Lagos State University of Science and Technology
[1]Corresponding Author: ucheibek@yahoo.com

Abstract

This study modeled Consumer Price Index (CPI) dynamics in Nigeria using historical CPI data from the Nigeria Bureau of Statistics for the period 2008-2022 inclusive. The study deployed two Machine Learning algorithms-Decision Tree and Random Forest, using Ordinary Least Squares (OLS) as benchmark. Model performances were contrasted regarding goodness of fit and prediction accuracy. The research revealed that the baseline model, OLS, turned out to give better results than both The Decision Tree and Random Forest models. The findings imply that although Random Forest performed better than Decision Tree, Ordinary Least Squares outperformed both of them.

Keywords: Consumer Price Index, Machine Learning, Decision Tree, Random Forest, OLS.

1. INTRODUCTION

The Consumer Price Index is a leading economic metric thatcomputes the mean difference in prices paid by consumers over a given time period for a basket of goods(Fernando, 2022) .

The Consumer Price Index (CPI) is the most important of the index numbers that indicate the financial health of an economy. These index numbers provide useful insights concerning business and economic situations and are used to compute the variation in prices, production, cost of living, etc. (Esan and Okafor, 2010).

The most famous measure of inflation is the Consumer Price Index (CPI), which portrays the change, in percentage terms, of the price of goods included in the basket of goods and services consumed by households. The Consumer Price Index (CPI) is frequently employed in the computation of inflation. It traces variations over time in the amounts spent by consumers to purchase a basket of goods and services.

The conceptual basis of the CPI is the theory of the cost-of-living index. The cost of living is a unique concept for each individual and is determined by the individual's preferences for different types of goods and services and the prices at which that person can purchase them.Calculating a cost-of-living index requiresthat all items in the utility function be assigned aprice.

Inflation rate is the rate of change of the general price level in a given economy from one period to another. The CPI is used to calculate inflation rate. Inflation is defined as a sustained increase in the overall price level over time. Nigeria is witnessing high inflation with economic and social implications. As real income falls, due to the eroded purchasing power of the currency, there is a reduction in the amount of goods and services each unit of the currency can buy. Confronted with an already diminished disposable income, consumers are now faced with higher prices owing to higher production costs(Folorunso and Abiola(2000) in Nse and Anietie (2018).

One of the key monetary policy tools deployed by Central Banks to address inflation issues is to raise interest rates. This has to be done with exceptional care, otherwise the outcome can be shattering and counterproductive. A case in point is the recent travails of Silicon Valley Bank (SVB), a California-based U.S. bank that provides venture-capital services mainly to startups. As of December 2022, the bank had amassed almost $200billion heavily invested in low-interest bonds. With interest-rate hikes in 2022 geared toward fighting inflation, bond values dropped dramatically and SVB incurred serious losses leading to a run on the bank (SVB IT Research, 2023).
The National Bureau of Statistics collects data for the computation of CPI by conducting a survey of family buying habits in Nigeria or in the segment of the population whose CPI is required. The NBS does not survey all the families in Nigeria; rather it uses probability sampling to generate representative samples and extracts relevant information for calculating the CPI.

Calculating CPI can involve tedious computations. This arises from the very large number of items to be included in the market basket, as well consideration of which items to include in the basket, and what to exclude. Some other challenges encountered in the computation of CPI include the choice of base year. The base year has to be a 'normal' i.e. a year when economic conditions are neither too good nor too bad, changes in buying habits, patterns and quality of goods over time(Esan and Okafor, 2010).

Formulas for Calculating Consumer Price Index (CPI).
Two formulas are employed. First, to find the current cost of the weighted-average basket of products, and second, to determine year-on-year change.
CPI (annual) = (value of basket in current year) / (value of basket in previous year).
Current year's CPI and previous year's CPI are then used to calculate the inflation rate.
Inflation Rate = [(New CPI - Previous CPI) / (Previous CPI)] x 100.

2. REVIEW OF LITERATURE

A number of previous studies have modeled the consumer price index in Nigeria and other countries.
Shu'ara and Olaolu (2022) examined the implications of the second quarter 2022 Consumer Price Index in Nigeria on consumers in the country. The paper showed that during this period, food, cooking gas and fuel prices increased beyond the purchasing power of the average Nigerian family. The study also revealed that high poverty and unemployment levels affected the poorer consumers disproportionately.

The authors suggested that the Nigerian government should reduce the level of corruption and put in place social safety net initiatives for the poor.
Diewert (2021) considered several approaches to The Theory of Index Numbers, including the test approach using price levels; tests for Bilateral Price Indexes; the Fisher Ideal Index and Test approach as well as the test performance of other indexes using the Circularity Test and concludes that under fairly weak regularity conditions, the only price index satisfying the circularity test is a weighted geometric average of all the individual price ratios, the weights being constant through time.

Diewert and Fox (2020) observed the lack of consensus on two key issues regarding the use of multilateral indexes in Consumer Price Index theory-the best multilateral method to use and the best way of extending the resulting series in the face of new observations. The authors presented theoretical and simulation evidence on the extent of substitution biases in alternative methods. Their results suggest the use of 'mean splicing' for updating.

Cross and Fare (2015) show how to use value data (price times quantity) to construct Fisher's price and quantity indexes, especially, revenue and expenditure data. The model extends previous work by the authors which showed how to recover relative prices from value data without explicit price or quantity information. Model accuracy was assessed over a range of price changes, firm sample sizes, and response variation in a Monte Carlo experiment. The model outperformed component indexes with accuracy levels that increase with response variation.

Rather than regarding the CPI as measuring changes in the cost of purchasing a fixed basket of goods, Schultze (2003) argues that the CPI should measure a cost-of-living index. This more ambitious view aims to base the CPI on measuring the changes in the cost of maintaining a household's standard of living at some specified level.
Adams, Awujola and Alumgudu (2014) fit a time series model to the consumer price index (CPI) in Nigeria and provided a five-year forecast for the expected CPI. The Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model was used to obtain the post-sample forecasts. The authors found that the best fitted model was ARIMA(1,2,1).

Eurostat (2018) provides a comprehensive overview of the methods that are deployed in the compilation process for the harmonized index of consumer prices (HICP) and provides a practical guide to all steps necessary to produce an HICP Fernando (2022) explained what the Consumer Price Index is, and how it is used. According to him, the CPI is an important metric that measures the average change in prices paid by consumers over a specified period of time for a basket of goods and services. In Nigeria, the index is calculated and published monthly by the National Bureau of Statistics. It indicates the health and direction of the economy, and it is a common measure of inflation. The index, being an indicator of inflation is closely watched by policy makers and the financial markets.

Gautier et al.(2022) documented new evidence on the adjustment of consumer prices at the micro level in the euro area. The authors also investigated how patterns of price adjustment have evolved over the period 2005 to 2019 and how the have contributed to inflation dynamics. They found that the overall size of non-zero price changes contributed a lot to changes in inflation over time, while the contribution the overall frequency of price changes is much smaller. The study also assessed how several aggregated shocks i.e. monetary, oil supply, demand, VAT and unemployment shocks are transmitted to prices, and the results of the study confirmed that price setters respond to shocks by adjusting the overall size rather than the overall frequency of price changes.

3. MATERIALS AND METHODS

3.1 DATA

This article analyzes monthly data on the composite Consumer Price Index and some other pertinent macroeconomic variables obtained from the Central Bank ofNigeria database from 2008 to 2021 inclusive. The variables selected for the study include Balance of payments, Government expenditure, interest rates, inflation rates, import cover, exports, external reserves, deposit and lending rates. The variables are extracted from the literature (Ibekwe and Shiro, 2022;. Rahayu et al., 2017; Oke et al., 2015).

3.2 METHODS

Machine Learning Python codes are used to carry out exploratory data analysis. Two contrasting algorithms-Decision Tree and Random Forest-were developed and deployed to analyze historical time series data. The performance of each model is assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean as well as Absolute Percentage Error (MAPE) to evaluate model fit and prediction accuracy. Other evaluation metrics utilized include R-squared and Adjusted R-squared.

R-squared: This measure quantifies how much of the variation in the target variable is explained by the model. Frequentlyy, this is the first metric calculated when evaluating linear regression model performance Bigger R-squared values are preferred.

Adjusted R-squared: In concept somewhat similar to R-squared, but is not significantly adversely affected with increasing number of variables, unlike the case of R-squared which generally increases as the number of variables increases. Higher values of adjusted R-squared are preferred.
Mean Absolute Error (MAE): Most straightforward metric for evaluating prediction accuracy. It is not responsive to outliers. Lower values of MAE are preferred.
Root Mean Square Error (RMSE): Also used in evaluating prediction accuracy. Responsive to outliers. Smaller values preferred.
Bias:The difference between predicted and actual value.
Variance: Shows the spread of the data.
Variance and bias are used to portray goodness of fit.

3.3 CROSS-VALIDATION

Used toaddress over-fitting.Cross-validationevaluates the prediction accuracy. Cross-validation is a technique for determining the extent to which a model generalizes to new data. In Cross-validation, the training data is repeatedly partitioned into a fixed number of folds. The data in each fold is analyzed, and the error estimates from the different folds are averaged.Repeat these steps for several values of the tuning parameters and choose the best tuning parameter that minimizes Average Mean Squared Prediction Error (Shayan et al., 2018).

3.4 DECISION TREES (DTS)

Decision Trees (DTs) are a predictive modeling technique used in classification and regression. They are used to obtain the value of a target variable by making use of decision rules deduced from features of the data. Decision tree technique is a data mining method employed in classification and prediction systems .To avoid over-fitting, the decision tree has to be pruned. An algorithm called minimal cost-complexity pruning is used to prune the tree. (Song and Lu, 2015)

3.5 RANDOM FOREST REGRESSOR

Random Forest is an estimator that deploys more than one decision tree on different sub-samples of the data set, and is used for performing regression and classification tasks. The output of the Random Forest is the mean of the outputs of the component decision trees. Using the mean values of the sub-samples improves the predictive accuracy of the model and helps reduce over-fitting. Each decision tree normally has high variance, but when combined in parallel the overall variance is low since each decision tree is perfectly trained on its very own sub-sample. (Fawagreh et al., 2014)

4. RESULTS AND DISCUSSION

The dataset contains 1,848 data points arranged in 168 rows and 11 columns.
All the columns are numeric and have non-null values.

Table 1. Summary statistics

	Count	Mean	Std	Min	25%	50%	75%	Max
Money Suppply (M2)	168	7.18E+06	1.98E+06	4304355	5.63E+06	6974993	8.14E+06	13369301
Import Cover(months)	168	1.01E+01	3.70E+00	4.43	7.50E+00	9.035	1.15E+01	20.83
Fuel Pump Price/litre	168	1.07E+02	4.12E+01	40	6.50E+01	97	1.45E+02	190.87
Foreign Reserves Position	168	3.87E+04	8.19E+03	23689.87	3.30E+04	37130.23	4.34E+04	62081.86
Crude Oil Price (US$/Barrel)	168	7.76E+01	2.75E+01	14.28	5.54E+01	73.555	1.06E+02	138.74
Commercial Banks Interest Rate on Time Deposits Maturing in 12 months	168	8.28E+00	2.84E+00	3.53	5.64E+00	8.015	1.06E+01	16.47
Prime Lending Rate of Commercial Banks	168	1.61E+01	2.03E+00	11.13	1.55E+01	16.595	1.71E+01	19.66
Maximum Lending Rate of Commercial Banks	168	2.60E+01	3.65E+00	17.58	2.32E+01	26.07	2.87E+01	31.56
Imports (CIF)	168	4.47E+03	1.23E+03	2502.6	3.52E+03	4441.59	5.18E+03	8574.64
Balance of Trade	168	1.35E+03	1.66E+03	-3480.64	0.00E+00	1050.955	2.36E+03	5996.98
Food>Year-on-Year change (%)	168	1.40E+01	3.77E+00	7.88	1.02E+01	13.46	1.66E+01	22.95
Average Exchange Rate	168	2.26E+02	9.09E+01	117.72	1.54E+02	169.68	3.06E+02	414.34
All Items>Year-on-Year change (%)	168	1.23E+01	3.02E+00	7.71	9.59E+00	12.065	1.44E+01	18.72

Observations:
The distribution of the variables and all the columns are plotted.

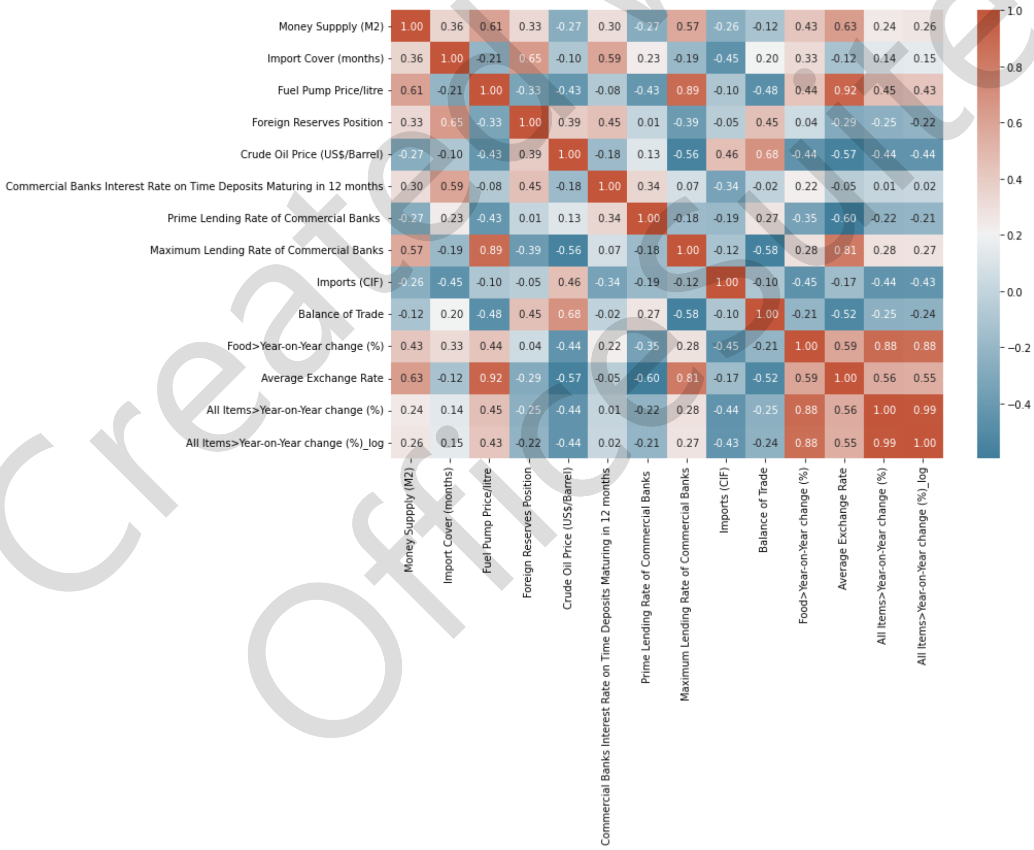
UNIVARIATE ANALYSIS

Distribution plots using Python code sns.distplot showed that Money Supply (M2), Import Cover, and Average Exchange Rate are highly right skewed. The target variable isAll Items>Year-on-Year change (%).

CORRELATIONS

There are not many correlations among the variables as is clearly seen in the heat map. The heatmap is obtained by the code for correlation matrix. As expected, some positive correlation is indicated between fuel pump price and average exchange rate. This means that careful choice of independent variables has largely averted the problem of multicollinearity.

Fig. 1sns.heatmap output



4.1 MODEL BUILDING

Data is prepared and split into training and testing sets in the ratio 70:30. The model is built and trained on the train data. Cross validation technique is deployed, and the model is tested on the test data. This is standard Machine Learning procedure.

SPLITTING THE DATASET

The data is also divided into explanatory and target variables.The initial 5 rows of the train set are displayed in the table using Python code.
X_train.head().

Table 2. X_train.head()

69	1	6795688	10	97	44155.1	112.3	4.7	17.1	24.9	5547.4	1919.8	9.2	157.4
28	1	5493793.3	9.9	65	38815.8	77.5	6.3	18.8	22.6	4400.7	2164.2	13	150.3
58	1	6639086.6	9.4	97	42568.3	111.1	6.2	16.5	24.7	4639.2	3814.3	11.6	157.3
153	1	8718197.9	8.1	161.2	35580.5	39.74	5	11.3	28.4	6434.2	0	17.4	381
66	1	7011005.8	9.5	97	45834.1	109.8	5.8	16.5	24.6	4968.7	3498.2	9.99	157.3

Calculate VIF for each feature to check for multicollinearity.

Table 3. VIF for each feature

	Feature
0	582.090410
1	4.237776
2	3.650925
3	18.290992
4	6.104886
5	6.332106
6	2.508052
7	5.043398
8	13.308795
9	2.659070
10	3.853895
11	19.594565

Table 4. Model 1 Summary

OLS Regression Results			
Dep. Variable:	All Items>Year-on-Year change (%)	**R-squared:**	0.653
Model:	OLS	**Adj. R-squared:**	0.616
Method:	Least Squares	**F-statistic:**	17.95
Date:	Tue, 14 Feb 2023	**Prob (F-statistic):**	1.55e-19
Time:	14:45:50	**Log-Likelihood:**	-234.53
No. Observations:	117	**AIC:**	493.1
Df Residuals:	105	**BIC:**	526.2
Df Model:	11		
Covariance Type:	nonrobust		

Omnibus:	2.161	**Durbin-Watson:**	2.120
Prob(Omnibus):	0.339	**Jarque-Bera (JB):**	2.058
Skew:	0.244	**Prob(JB):**	0.357
Kurtosis:	2.571	**Cond. No.**	1.81e+08

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.81e+08.

Table 5. Model 2 Summary

OLS Regression Results			
Dep. Variable:	All Items>Year-on-Year change (%)	**R-squared:**	0.928
Model:	OLS	**Adj. R-squared:**	0.920
Method:	Least Squares	**F-statistic:**	111.5
Date:	Tue, 14 Feb 2023	**Prob (F-statistic):**	1.20e-53
Time:	15:17:55	**Log-Likelihood:**	-142.58
No. Observations:	117	**AIC:**	311.2
Df Residuals:	104	**BIC:**	347.1
Df Model:	12		
Covariance Type:	nonrobust		

Omnibus:	14.112	**Durbin-Watson:**	2.129
Prob(Omnibus):	0.001	**Jarque-Bera (JB):**	15.297
Skew:	0.812	**Prob(JB):**	0.000477
Kurtosis:	3.709	**Cond. No.**	1.90e+08

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.9e+08.
Checking model performance.

Table 6. targets.head()

0	8.57
1	8.03
2	7.78
3	8.17
4	9.7

Name: All Items>Year-on-Year change (%), dtype: float64

Table 7. Performance of the Model

	Data	RMSE	MAE	MAPE
0	Train	0.818495	0.635353	5.579796
1	Test	1.068327	0.791205	6.435310

Observations:
RMSE,MAE and MAPE of train and test data are a bit close (not too different), indicating that the model is over-fitting somewhat, but has generalized well.
Applying cross validation and building the regression model using sklearn linear regression, dataset is divided into 10 folds to obtain 10 r-squared values.
RSquared:0.885(+/-0.098)
MSE:0.825(+/-0.636)

Observation: MSE seems tolerably good. Higher values preferred.

Table 8. Model 2 Parameters

const	4.194278E+00
MoneySuppply(M2)	2.161725E-08
Import Cover (months)	-4.890006E-02
Fuel Pump Price/litre	2.078728E-02
Foreign ReservesPosition	-5.715679E-05
Crude Oil Price (US$/Barrel)	4.826105E-03
ComBanks Int. rates on 12 months Deposits	-1.562304E-01
Prime Lending Rate of Commercial Banks	5.817595E-01
Maximum Lending Rate of Commercial Banks	-4.362293E-01
Imports (CIF)	-1.307006E-04
Balance of Trade	-2.667020E-04
Food>Year-on-Year change (%)	7.026831E-01
Average Exchange Rate	1.064379E-02
dtype: float64	

Equation of fit
Equation="log(All Items>Year-on-Year change (%)=)"
log(All Items>Year-on-Year change (%)=) (4.19427815405283)* const +(2.161724860242283e-08)* Money Suppply (M2) +(-0.04890006133695108)* Import Cover (months) +(0.020787283153176395)* Fuel Pump Price/litre +(-5.715679390233588e-05)* Fgn Rsv Position +(0.004826104753872469)* Crude Oil Price (US$/Barrel) +(-0.15623039867161723)* Commercial Banks Interest Rate on Time Deposits Maturing in 12 months +(0.5817595057888069)* Prime Lending Rate of Commercial Banks +(-0.4362292513088659)* Maximum Lending

Rate of Commercial Banks +(-0.00013070062976862641)* Imports (CIF) +(-0.00026670203581393776)* Balance of Trade +(0.7026830966839727)* Food>Year-on-Year change (%) +(0.010643792927498275)* Average Exchange Rate +
To interpret Regression Coefficients, all other dependent variables are held constant.

4.4 Build the Non-Linear models (Decision Tree and Random Forest) and checktheir performance. We therefore

Import Decision Tree Regressor using sklearn
Split the data in 70:30 ratio of train to test data.
Separate the dependent and independent variables
Redefine the Decision tree regressor
Fit Decision Tree Regressor to train dataset
Check model performance on the train and test set.

Table 9. Model Performance

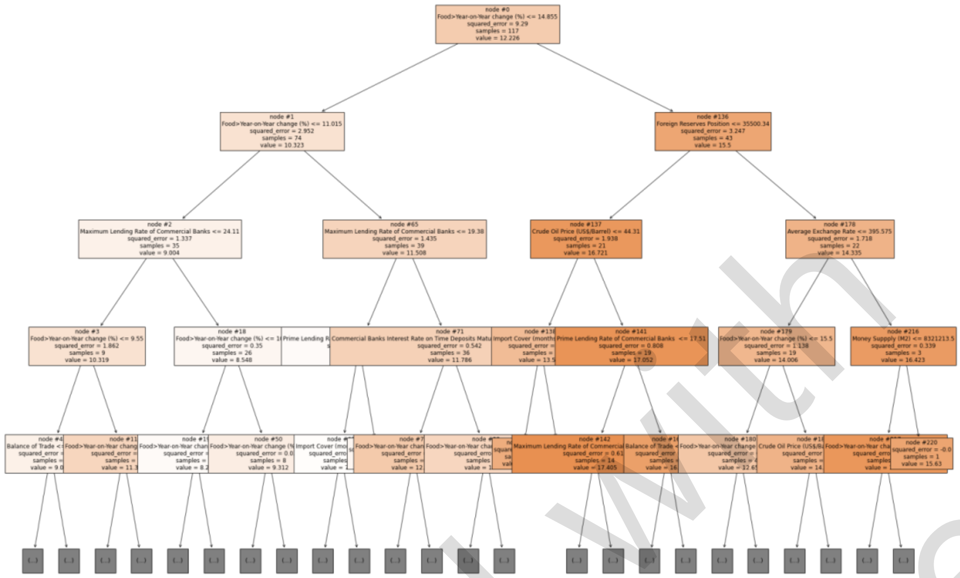
	Data	RMSE	MAE	MAPE
0	Train	0	0	0
1	Test	1.013648	0.746667	6.024069

Observations: Test set values are clearly of much higher order than the train set. The model is evidently over-fitting and does not generalize well.
Decision trees tend to have high variance, but they are not difficult to interpret.
Splitting is based on maximum information gain. It is based on the variable that gives least variance.

The following code yields the Decision Tree:

```
features=list(X1.columns)
plt.figure(figsize=(35,25))
plot_tree(dt,max_depth=4,feature_names=features,filled=True,fontsize=12,node_ids
=True,class_names=True)
plt.show()
```

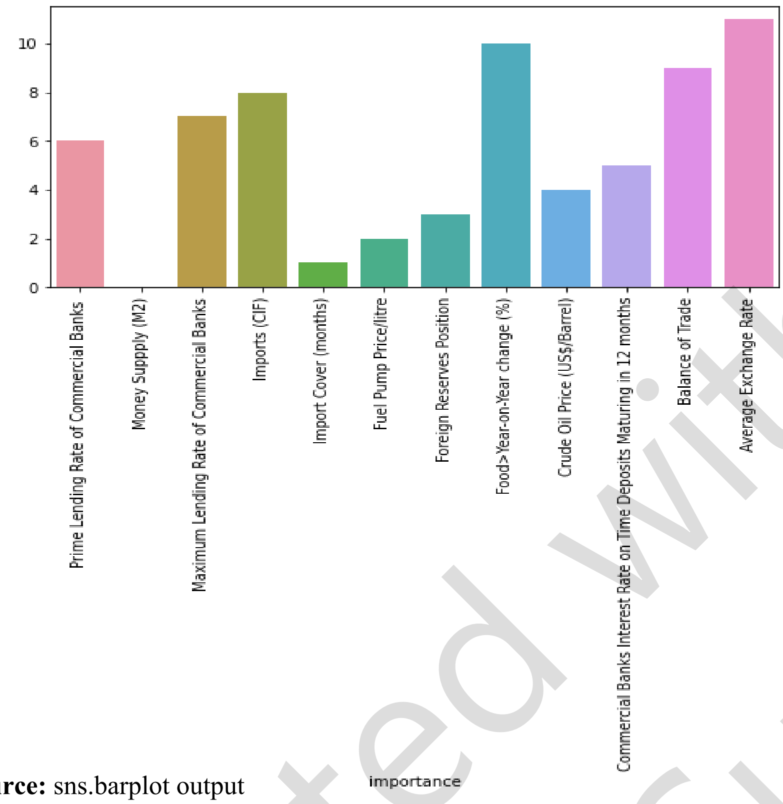
Fig.2



The tree has depth 5. The more complex the relationships, the greater the required depth.

Analyze the variables and plot the feature importance for each variable.

Fig.3 Feature Importance



Source: sns.barplot output

Table 10. Importance of each variable.

6	Prime Lending Rate of Commercial Banks
0	Money Suppply (M2)
7	Maximum Lending Rate of Commercial Banks
8	Imports (CIF)
1	Import Cover (months)
2	Fuel Pump Price/litre
3	Foreign Reserves Position
10	Food>Year-on-Year change (%)
4	Crude Oil Price (US$/Barrel)
5	Commercial Banks Interest Rate on Time Deposit...
9	Balance of Trade
11	Average Exchange Rate
Name: importance, dtype: object	

importance_inf.importance

4.5 BUILDING RANDOM FOREST

Initial choice of hyperparameters is random, but alterations are introduced subsequently for improved results. Model building involves:

Splitting the data.

Checking model performance on the train and test dataset.

Score gives the accuracy of the test data and labels. Here, the score was **0.909929328006268**

Observations: Good score. Higher scores preferred.

Table 11. Model Performance on Train and Test set.

	Data	RMSE	MAE	MAPE
0	Train	0.497316	0.397564	3.330865
1	Test	0.876095	0.659758	5.348924

Observations: RMSE,MAE and MAPE for the random forest are not small and are not close for both train and test dataset. Hence, the model may not be performing very well and likely not giving generalized results.

5.0 CONCLUSION

This study compared two techniques for modelling Consumer Price Index (CPI) dynamics in Nigeria. The models were Decision Tree and Random Forest machine learning algorithms, using Ordinary Least Squares as baseline model.The three models were compared in terms of model fit and prediction accuracy. Model performance was measured using RMSE, MAE and MAPE.

The study found that OLS results for train and test data were not very different, indicating that the model generalized well and was not over-fitting.
The Decision Tree model, however, was high in variance and did not generalize well. The Decision Tree model was highly overfitting as the order of differences between train and test data using RMSE, MAE and MAPE was rather large.

Random Forest algorithm returned very high accuracy level (90.99 percent) on test data and labels. RMSE, MAE and MAPE for the random forest are not small and not close for both train and test data set. Hence, the model, although giving high accuracy may not be generalizing well.

The baseline model, OLS, turned out to give better results than both The Decision Tree and Random Forest models. The study therefore suggests that although Random Forest performed better than Decision Tree, Ordinary Least Squares can outperform both of them.

References

Adams, S.O., Awujola, A. and Alamgudu, A.I. (2014). Modeling Nigeria's Consumer Price Index using ARIMA model. *International Journal of Development and Economic Sustainability,* 2(2), 37-47, Published by European Centre for Research Training and Development UK (www.ea-journals.org).

Cross, R.M. and Fare, R. (2015). Value Data and the Fisher Index. *Theoretical Economics Letters,* 5(2) 262-267. doi: 10.4236/tel.2015.52031.

Diewert, W.E. (2021), "Quality Adjustment Methods", Draft Chapter 8 in Consumer Price Index Theory, Washington D.C.: International Monetary Fund, published online at: https://www.imf.org/en/Data/Statistics/cpi-manual

Diewert, W.E. and K.J. Fox (2020) "Substitution Bias in Multilateral Methods for CPI Construction Using Scanner Data," *Journal of Business and Economic Statistics*, published online at: https://doi.org/10.1080/07350015. 2020.1816176

Esan, E.O. and Okafor, R.O. (2010). Basic Statistical Methods, Revised Edition.

Eurostat (2018), *Harmonized Index of Consumer Prices (HICP)Methodological Manual*, Luxembourg: Publications Office of the European Union.

Fernando, J. (2022). Consumer Price Index (CPI) Explained: What it is and How it is used. Investopedia:https://www.investopedia.com/contributors/53746/

Gautier, E., Conflitti, C., Faber, R.P., Fabio, B., Fadejeva, L., Jouvanceau, V. (2022). New Facts on Consumer Price Rigidity in the Euro Area. *ECB Working Paper Series No. 2669/ June 2022. JEL codes: D40, E31. 1-116.*

Ibekwe, U. A., and Shiro, A.A. (2022). Modelling the Determinants of Naira/U.S. Dollar Currency Exchange Rates Using Principal Component Analysis (PCA) and Singular Value Decomposition (SVD). *Nigeria Journal Of Risk and Insurance,12(*1), 77-79.

Nse, S.U. and Anietie, S.I. (2018). A Predictive Model for Inflation in Nigeria. *CBN Journal of Applied Statistics, 9* (2) 103-129.

Oke, M.O., Adetan. T.T. (2015). An Empirical Analysis of the Determinants of Exchange Rate in Nigeria. *International Journal of Scientific Research and Management (IJSRM), 6* (5) 47-57.

Rahayu, S., Sugiarto, T., Madu, L., Holiawati, and Subagyo, A. (2017). Application of Principal Component Analysis (PCA) to Reduce Multicollinearity. *International Journal of Educational Methodology, 3*(2), 75-83. https://doi.org/10.12973/ijem.3.2.75

Schultz, C.L. (2003). The Consumer Price Index: Conceptual Issues. *Journal of Economic Perspectives, 17* (1)3-22.

Shu'ara, J. And Olaolu, O.E. (2022). Implications of the 2022 Q2 Consumer Price Index (CPI) on Nigerian Consumers-Way out. *American Journal of Multidisciplinary Research in Africa,* 3(1) 1-8www.mprijournals.com

Silicon Valley Bank (2023). SVB IT Research.https://www.svb.com >signature-research.