Forecasting The Consumer Price Index: A Comparative Study of Machine Learning Methods

Article *in* International Journal of Computing and Digital Systems · February 2024

DOI: 10.12785/ijcds/150137

Fadi Sibai
Gulf University for Science & Technology

Ali El-Moursy
University of Sharjah

International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (Feb-24)

http://dx.doi.org/10.12785/ijcds/150137

Forecasting The Consumer Price Index: A Comparative Study of Machine Learning Methods

Fadi N. Sibai[1], Ali El-Moursy[2] and Ahmad Sibai[3]

[1]*Electrical and Computer Engineering Department, Gulf University for Science and Technology, Mishref, Kuwait*
[2]*Electrical and Computer Engineering Department, University of Sharjah, Sharjah, U.A.E.*
[3]*Computer Science and Engineering Department, University of South Florida, Tampa, Florida, U.S.A.*

Received 11 Apr. 2023, Revised 7 Jan. 2024, Accepted 21 Jan. 2024, Published 1 Feb. 2024

Abstract: The Consumer Price Index (CPI) is an indicator of inflation and is tracked by many government and economic agencies to make decisions of major importance. Its prediction is a valuable input into government policies such as taxation, and it greatly impacts the cost of borrowing money. The CPI has been traditionally predicted with statistical methods such as the Autoregressive Integrated Moving Average (ARIMA) model. In this paper, we forecast the Saudi Arabian Consumer Price Index with six machine learning (ML) methods, using the Orange 3 data mining and analytics tool, and based on the published historical January 2013 to November 2020 CPI data. We compare the performances of Decision Tree (Tree), k-Nearest Neighbors (kNN), Linear Regression (LR), Neural Networks (NN), Random Forest (RF), and Support Vector Machine (SVM), all applied to the 2013-2020 Saudi Arabian CPI dataset. Multiple experiments were conducted to vary the training and testing sets, optimize the machine learning parameters, and improve the MSE and R^2 metrics. The predicted CPI values of these ML methods were also compared to the 2021-2024 International Monetary Fund (IMF) CPI forecast and the actual 2021-2024 CPIs (post mortem). The results indicate that the multilayer perceptron neural network model outperforms the other ML models, is nearest to the actual CPI, and may be used to forecast the CPI for up to 3 years from the latest CPI data in the training dataset. The kNN model follows the neural network model in second place. The best fitting Excel trend line underperformed all ML methods in forecasting the Saudi Arabian CPI.

Keywords: Consumer Price Index, Forecasting, Machine Learning, Data Science, Orange 3

1. Introduction

The Consumer Price Index (CPI) reflects the weighted average of consumer goods and services, is a good indicator of inflation, and helps in weighing the purchasing power of a certain country in terms of its currency, and its cost of living. Countries like the USA and Saudi Arabia monthly publish the CPI. The movement of the CPI affects how the government adjusts spending and interest rates and is a pillar of a government's macroeconomic and fiscal policies. This in turn affects matters such as taxation and the cost of borrowing. Indeed, taxes are added in the CPI, affecting future taxation levels. The CPI thus affects the cost of living and how the economy may be macro-controlled. For instance, when the CPI decreases, the government may issue lower social security checks. When the CPI increases, the government may raise interest rates, thereby lowering the volume of borrowing, and reducing spending in favor of saving. From consumers and investors perspectives, accurately forecasting the CPI helps in making decisions related to spending and investment. A job seeker receiving two similar offers from firms located in two different regions, may choose the region associated with the lower CPI.

In this work, we investigate the accuracy of six machine learning (ML) methods in forecasting the consumer price index (CPI) of Saudi Arabia based on published historical CPI data. We employ Orange 3 [1], an open source tool for data mining and visualization with a set of machine learning and data mining functions to train and test these six ML methods on the 2013-2020 Saudi Arabian CPI dataset. Forecasting is a difficult task involving extensive trend analysis, and guessing the predicted values of several variables. Economic data has been traditionally forecasted with statistical and time series techniques [2]. ML models operate differently from these statistical techniques. When trained with historical data, ML methods analyze the data for identifying patterns and relations that are not obvious to detect. For instance, neural network models with each data vector in the training set strengthen or weaken synaptic weights interconnecting neurons in the various layers. Moreover, ML models can handle larger dataset from multiple sources than traditional statistical methods used in forecasting business and economic data. Importantly, with their nonlinear models, when data changes abruptly, ML models can capture these abrupt changes better than traditional statistical methods. For all of the above reasons, ML methods tend to perform well in time series predictions and forecasting.

Predicting the CPI has a great impact on economic planning, investment distributions, and the people wellbeing. Complementing the success of ML methods in classification [3-5], and given their success in predicting time series, the main contribution of this work is its application and investigation of the accuracy of six ML techniques namely, Random Forest (RF), Support Vector Machine (SVM), k-Nearest Neighbors (kNN), Neural Networks (NN), Decision Tree, and Linear Regression (LR), in forecasting the Saudi Arabian CPI based on the historical 7-year CPI dataset in the time period between November 2013 and November 2020. To the best of our knowledge, this is the first research work to develop and analyze ML models in forecasting the Saudi Arabian CPI.

This paper is organized as follows: Section 2 reviews related work. The adopted methodology is explained in Section 3. The experimental results are presented in Section 3, then analyzed in Section 4. The paper concludes in Section 5.

2. Prior Related Work

In [2], several forecasting methods were described including qualitative techniques such as market research and consensus, quantitative techniques such as moving average, Box-Jenkins, and trend projections based on time series, and causal methods such as regression models and economic input-output models. Forecasting based on time series helps in predicting the future based on historical values. Several versions of time series forecasting exist including linear vs. non-linear, parametric vs. non-parametric, and univariate vs. multivariate methods. Good forecasting methods consider the overall trend of the data, as well as variations from seasonal cycles, and other variations due to noise or one-time events.

Meyer and Pasaogullari [6] analyzed various forecasting models for inflation, and concluded that none does particularly better than the others, although some may outshine others during fixed periods of time. Qin et al [7] combined a genetic algorithm with SVM to predict the CPI. The genetic algorithm was employed to optimize the SVM parameters such as the initial width of the RBF function, the penalty factor, and the kernel function, and then SVM was used to forecast the CPI, resulting in yearly CPI errors between 1.4-26.4

Autoregressive Integrated Moving Average (ARIMA) [8], and Autoregressive Moving Average (ARMA) models are two popular methods for forecasting time series. While sharing some similarities, ARMA is a stationary model, while ARIMA is an integrated model which requires taking the differences between observations before ARIMA reaches stationarity. Stationarity refers to the statistical properties of a process generating a time series not changing over time. In [9] and [10], the Albanian CPI was forecasted based on multiple, regression models, Seasonal Autoregressive Integrated Moving Average (SARIMA), and ETS, with the SARIMA model shown to be more than accurate than the ETS model. The SARIMA time series model was also shown to be a better CPI predictor than multiple regression models. Zhang et al [11] analyzed a Chinese CPI forecast model for the years 1995–2008, and concluded that the ARMA model had a good forecast accuracy. Other authors also proposed time series model for CPI. Ngailo [12] fitted the predicted data to the data provided by the Tanzania Ministry of Health and Social Welfare.

A few researchers concluded that ARIMA was the best method for forecasting the CPI. Junior [13] compared the ARIMA model to predict the time series of the Bovespa Index with other models, with the former model scoring a mean absolute error percentage of 0.052%, lower and better than the other models. In [14], the ARIMA(1, 1, 0) model was used to predict the CPI of the Bandar Lampung City with only 6-month forecasted data and was shown to be very near to the actual data (under 1% error). No data was provided for the next 5 years. Mohamed [15] concluded that ARIMA (0, 1, 3) was the most suitable model for predicting CPI in Somalia based on time series data for the years 2013-2020. Subhani and Panjwani [16] used ARIMA to conclude that CPI had a significant association with government bonds. Adam et al [17] studied forecasted the CPI in Nigeria and found the ARIMA(1, 2, 1) model to work best (R^2=0.767) when considering the data for the years 1980-2010. In [18], the gray prediction model was used to predict the CPI in 3 months in 2020. In [19], the CPI of Mauritius was predicted with an ARIMA (0, 2, 3) model yielding an RMSE of 1.43. In [20], the Ecuadorian CPI was forecasted with Support Vector Regression, particle filter, SARIMA, Fast Fourier Transform, Theta Method, with the last method scoring a mean absolute error of 0.3453.

More recently, researchers who used machine learning methods include Huong et [21] who forecasted the CPIs of AU, Spain and OECD countries (550 data values) with an ensemble learning model and the NSGA-II multi-objective evolutionary algorithm, and only reported MSE numbers to evaluate the goodness of their forecast. Zahara et al [22] used the long short-term memory (LTSM) deep learning technique to predict Indonesian CPI. For optimization, they used stochastic gradient descent (SGD), and a few other methods and concluded that the Adaptive moment (Adam) optimization algorithm resulted in the best RMSE. Harris [23] forecasted the Canadian CPI to predict food prices and compared multi-layer perceptron (MLP), M5P tree, sequential minimal optimization (SMO), and linear regression, with the former resulting in the best mean absolute percentage error (MAPE). Yang and Guo [24] used deep learning based on a recurrent neural network to predict the CPI, and consequently inflation, and achieved a mean square error (MSE) of 0.359. In [25], SARIMA was compared to deep neural network with 20 hidden layers, with the latter one achieving a lower MSE of 1.75.

In summary, the above references have employed various methods to forecast the CPI of several countries. However, their obtained prediction errors can be further reduced by exploring a variety of machine learning methods to forecast the CPI, as targeted by our work. Unlike the above references, we investigate the application of 6 machine learning methods, including a deep neural network with 4 hidden layers, to forecast the Saudi Arabian CPI, optimize the ML method parameters to minimize the errors, and identify the ML methods which perform best. Several regression measures can be used to evaluate how good the prediction is, such as the following metrics which we calculate in our work.

$$MAE = Mean\ Absolute\ Error = \sum_{i=1}^{n} \frac{|Y_i - \hat{Y}_i|}{n} \qquad (1)$$

$$MSE = Mean\ Square\ Error = \sum_{i=1}^{n} \frac{(Y_i - \hat{Y}_i)^2}{n} \qquad (2)$$

$$RMSE = Root\ Mean\ Square\ Error = MSE^{1/2} \qquad (3)$$

$$R^2 = Coeff.\ of\ Determination = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2} \qquad (4)$$

where $\hat{Y}_i$ is the predicted value of Y_i, and $\bar{Y}$ is the mean of the Y_i's. R^2, coefficient of determination, is between 0 and 1, and indicates an excellent fit when 1, and a bad fit when 0. A negative R^2 value means that the prediction is less accurate that the mean value of the dataset over time.

3. Experimental Results

A. Research Methodology

The methodology adopted in this work follows.

a. The dataset consisting of Saudi Arabian CPI numbers between January 2013 and November 2020 is prepared. Initially, the month and year are separate features.

b. The Microsoft Excel application is used to graph the CPI data trend between 2013-2020 and through curve fitting, generate the trendline which predicts the CPI between 2021-2024.

c. The Orange 3 application and CPI dataset are used to train and test the 6 ML methods.

d. The performances of the 6 ML methods are evaluated by analyzing the MSE and R^2 numbers.

e. Step d is repeated multiple times by varying the ML method parameters and stopping when the MSE and R^2 numbers cannot be further improved. The ML methods with best MSE and R^2 numbers are identified.

f. Steps a-e are repeated after combining the month and year into one feature.

g. For the single datum case of step f, the training and testing of the 6 ML methods are repeated with various distributions of data rows for training and testing, including random sampling.

h. The Orange 3 tool is used to predict the CPI between years 2021-2024 for each of the 6 ML methods. Note that this is data not present in the dataset used for training and testing.

i. The 2021-2024 CPI predictions by the 6 ML methods are compared to the 2021-2024 IMF predictions, and (post mortem) to the actual published 2021-2023 Saudi Arabian CPIs. The best performing ML methods are identified.

B. Dataset

The Saudi CPI numbers from Jan. 2013 to Nov. 2020 are published in [26] and serve as the dataset for training and testing our ML methods. As the Saudi CPI numbers, are around 100, the first step is to normalize them to be around 10, by dividing the raw CPI numbers by 10. The normalized CPI numbers are plotted in Fig. 1. The trendline generated by Microsoft Excel is

$$CPI_{Norm.} = 2\ 10^{-7}\ x^4 - 3\ 10^{-5}\ x^3 + 0.0013\ x^2 - 0.0072\ x + 9.318 \qquad (5)$$

with R^2=0.8361, and x =1 representing the first month date 1/2013, x =2 representing the second month date 2/2013, etc. The Microsoft Excel application was also used to curve-fit the data and forecast the CPI for the next 24 months as shown in Fig. 2, based on Equation (5). The 24-month forecast appears to be grossly exaggerated based on the 4th degree polynomial of Equation (5).

The International Monetary Fund (IMF) [27] forecasted the Saudi CPI for 4 years following 2020 as shown in Table I [28]. These numbers appear to be more believable and reasonable than those forecasted by Equation (5). This striking difference justifies the investigation of data mining and machine learning methods to better forecast the CPI. Another objective of this work is therefore to identify that the Machine Learning methods whose forecasts of the CPI are more accurate and reliable than the Excel trendline of Equation (5).

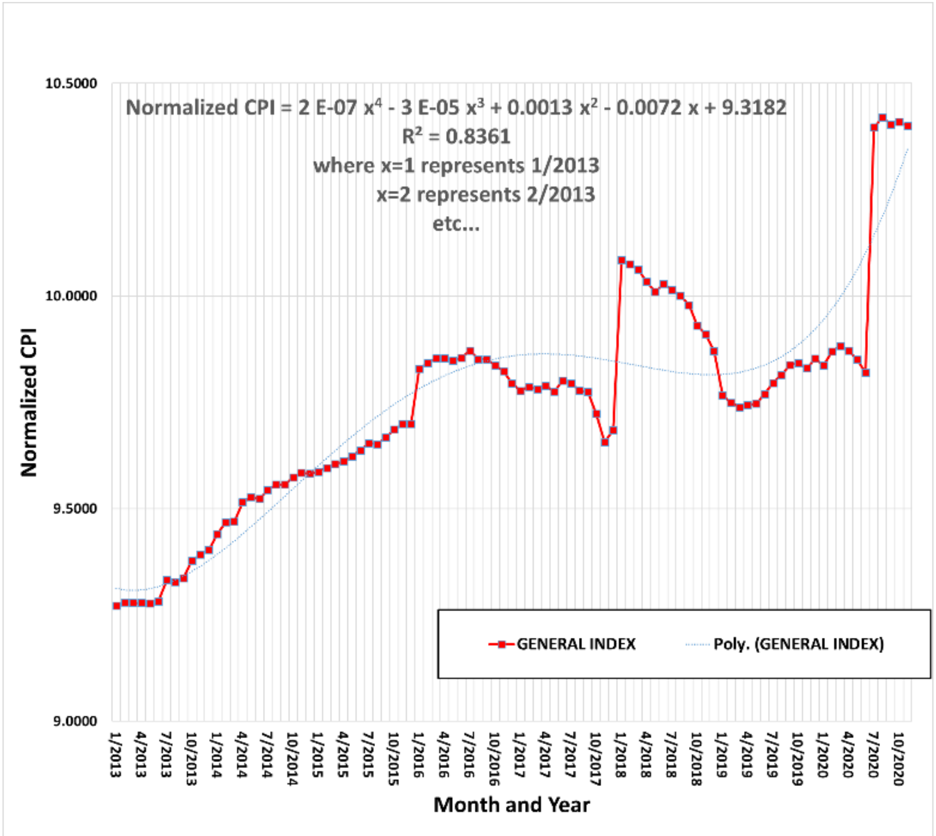


Figure 1. Historical Saudi CPI [17] (Normalized)

TABLE I. International Monetary Fund Forecast

Year	Saudi CPI Growth (IMF Forecast)	Normalized CPI (IMF Forecast; 1/2020 CPI=0.9837)
2020	2.240%	10.058
2021	2.148%	10.274
2022	2.107%	10.491
2023	2.054%	10.706
2024	2.055%	10.926

C. Two-Input Data

The 84 rows of the normalized Saudi CPI from 1/2013 to 11/2020 corresponding each to the normalized CPI in the corresponding month were uploaded onto Orange [1] data mining and machine learning tool. At first, the inputs to the ML models consisted of 2 input data, month and year, and the target was the normalized CPI Index. Thus, the dataset was organized into 3 columns: month, year, and the corresponding CPI, as follows.

Month	Year	CPI
1	2013	...
...	...	...
11	2020	...

Initially, the training set was composed of all 84 rows for the years 2013-2018. The testing set was composed of 11 rows for the data between Jan. 2019 and Nov. 2019. With this dataset, Orange was configured to train and test the following methods: Random Forest (RF), Support Vector Machine (SVM), k-Nearest Neighbors (kNN), Neural Networks (NN), Decision Tree, and Linear Regression (LR). These methods were set up with the default Orange parameter configurations of Orange, and the training was performed with these parameters. Orange allows a number of machine learning tools in its rich library to be graphically selected to be trained and tested on the chosen dataset. Upon execution run completion, Orange displays the results similar to Table II below. Table II shows the MSE, RMSE, MAE and R^2 (refer to Equations 1-4) obtained by testing the ML methods with the training data set (84 rows).

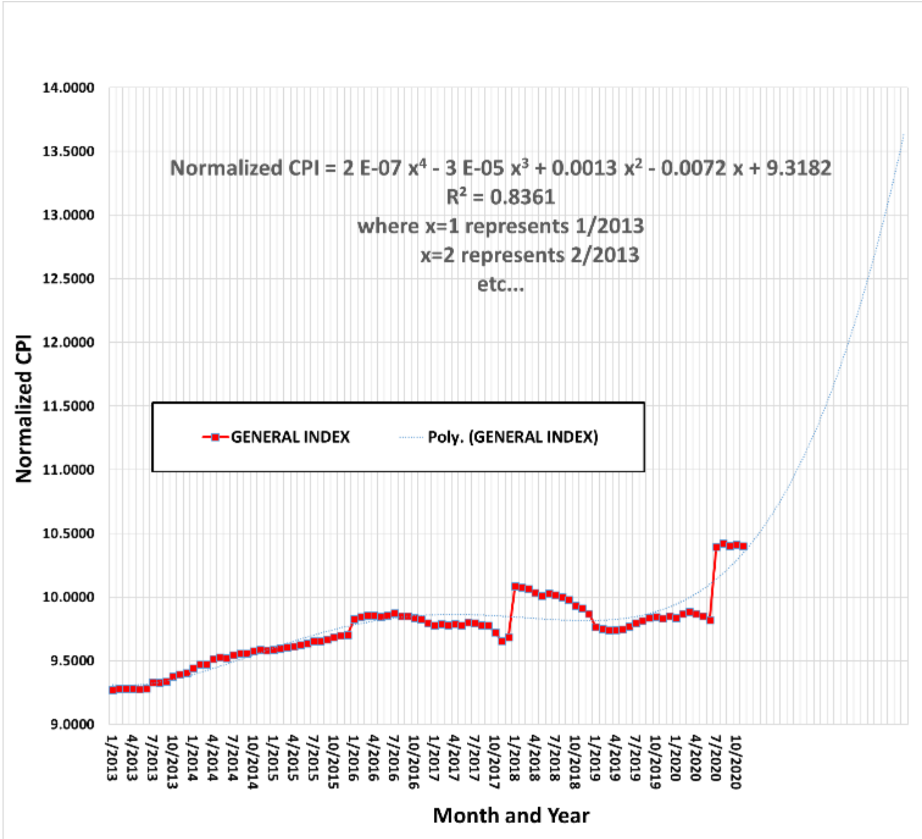


Figure 2. 2013-2020 Saudi CPI (Normalized) with 24 months in 2021-2022 forecasted by Microsoft Excel

TABLE II. Metrics For Testing with 84-row Training Dataset - Two-Input Data

ML Method	MSE	RMSE	MAE	R^2
Random Forest	0.000	0.021	0.015	0.990
SVM	0.025	0.159	0.126	0.438
kNN	0.010	0.101	0.089	0.771
Neural Networks	0.022	0.148	0.112	0.511
Decision Tree	0.000	0.011	0.008	0.998
Linear Regression	0.013	0.116	0.094	0.699

Table III shows the MSE, RMSE, MAE and R^2 obtained by testing the ML methods with the testing data set (11 rows). As expected, compared to testing with the training dataset, the performance metrics decline when testing the ML methods with the testing dataset, which contains data which was not used during training. The small or even negative values of R^2 indicate a poor fit. SVM and LR outperform the rest with the smallest MSE and the largest R^2.

Fine tuning the ML method parameters to improve the results involved varying the ML parameters to obtain the best possible performance metrics. After this lengthy exercise was completed, the obtained optimal parameters are displayed in Table IV. Table V shows the optimal MSE, RMSE, MAE and R^2 metrics obtained by testing with the 11-row testing dataset. Compared to Table III, Table V shows smaller MSEs for RF, kNN, NN, and Decision Tree, and slightly higher R^2 numbers indicating a minor improved fit over the initial experiment.

D. One-Input Data - 11-Row Testing Set

In the next experiment, the month and year values were combined into one number. Precisely, the first point January 2013 became 1, February 2013 became datum 2, etc. The single output target remained the normalized Saudi CPI. Thus, the dataset was organized into 2 columns: the number representing the combined month and year, and the corresponding CPI. Thus, the dataset was organized into 2 columns: a number (representing the row number,

TABLE III. Metrics For Testing with 11-row Testing Dataset - Two-Input Data

ML Method	MSE	RMSE	MAE	R^2
Random Forest	0.163	0.404	0.322	-1.160
SVM	0.071	0.267	0.235	0.057
kNN	0.123	0.350	0.249	-0.623
Neural Networks	0.121	0.348	0.296	-0.608
Decision Tree	0.159	0.398	0.319	-1.102
Linear Regression	0.075	0.274	0.261	0.009

TABLE IV. Optimal ML Parameters - Two-Input Data

ML Method	Parameter Settings
Random Forest	# trees=4; Limit depth of individual trees=2; Do not split subsets smaller than 3.
SVM SVM SVM	SVM Cost =1, Regression epsilon=0.1; Kernel: Polynomial, kernel: $(gxy + c)^d$, g =auto, c=0, d=3; Numerical tolerance=0.001; Iteration limit=100.
kNN	# neighbors=8; metric =Chebyshev; Weight=Uniform
Neural Networks Neural Networks	Neurons in hidden layers= 60, 50, 30, 15; Activation=ReLU; Solver=Adam; Regularization=0.0001; Max. # iterations=444.
Decision Tree Decision Tree	Induce binary tree=checked; Min # instance in leaves=13; Do not split subsets smaller than 4; Limit the maximal tree depth=100; Stop when majority reaches 95%.
Linear Regression	Fit intercept; Elastic Net regularization L1 0.51:0.49 L2

TABLE V. Optimal Metrics For Testing with 11-row Testing Dataset - Two-Input Data

ML Method	MSE	RMSE	MAE	R^2
Random Forest	0.145	0.381	0.268	-0.918
SVM	0.071	0.267	0.235	0.057
kNN	0.119	0.346	0.254	-0.582
Neural Networks	0.071	0.266	0.209	0.060
Decision Tree	0.120	0.346	0.254	-0.584
Linear Regression	0.075	0.274	0.261	0.008

with the first row corresponding to Jan 2013, and the last row corresponding to Nov. 2020), and the corresponding CPI, as follows.

Number	CPI
1	...
...	...
84	...

The training dataset was again composed of 84 rows (rows 1-84 of the dataset) corresponding to all the data in years 2013-2018, while the testing dataset was composed of 11 rows (rows 85-95) corresponding to the months between January 2019 and November 2020. The ML parameters were again varied multiple times to achieve better performance. The best parameters are depicted in Table VI.

With the parameters of Table VI, testing the methods with the testing data set resulted in further improvements in the performance, as captured in Table VII. The R^2 and MSE values also improved in comparison to the two-input data results of Table V.

E. One-Input Data - 25-Row Testing Set

Remaining with the one input data of Subsection E, when the training dataset was reduced to 70 rows (rows 1-70), and the testing dataset increased to 25 rows (rows 71-95), and with the parameters of Table VI, the performance metrics worsened, in comparison to Table VII, as the MSE increased while R^2 dropped. Fine tuning the ML method parameters to further improve the results yielded the optimal parameter values of Table VIII. The corresponding performance metrics of Table IX, obtained by testing the methods on the 25-row testing dataset, were the best results

TABLE VI. ML Parameters - One-Input Data

ML Method	Parameter Settings
Random Forest	# trees=1; Limit depth of individual trees=2; Do not split subsets smaller than 2.
SVM SVM	SVM Cost =3.5, Regression epsilon=0.1; Kernel: x .y; Numerical tolerance=0.0035; iteration limit=205
kNN	# neighbors=24; metric =Manhattan; Weight=Uniform.
Neural Networks Neural Networks	Neurons in hidden layers= 30, 50, 50, 15; Activation=ReLU; Solver=Adam; Regularization=0.0001; Max. # iterations=430.
Decision Tree Decision tree	Induce binary tree=checked; Min # instance in leaves=3; Do not split subsets smaller than 5; Limit the maximal tree depth=100; Stop when majority reaches 100%.
Linear Regression	Fit intercept; No regularization.

TABLE VII. Metrics For Testing with 11-row Testing Dataset - One-Input Data

ML Method	MSE	RMSE	MAE	R^2
Random Forest	0.108	0.329	0.257	-0.433
SVM	0.065	0.255	0.250	0.140
kNN	0.120	0.346	0.254	-0.584
Neural Networks	0.046	0.214	0.209	0.392
Decision Tree	0.120	0.346	0.254	-0.584
Linear Regression	0.069	0.263	0.248	0.084

obtained so far, though not quite impressive.

F. One-Input Data - Random Sampling

With X% random sampling, the testing dataset is ignored by Orange, and only the training dataset is used and split into X% (=70%-80%) for training, as set below, and the rest for testing. It should be noted that random sampling improves the performance metrics, as expected, as it allows the training and testing to incorporate data from the later-dated data, thereby improving the results. With 70% random sampling, about 70% of 70 data rows were selected in the training dataset, and the remaining 30% of the 70 rows were used for testing. The corresponding results are depicted in Table X. At last, at this stage, the errors are acceptably low while R^2 is satisfactorily high. We observe that ML methods, kNN and NN, provide the best fit, i.e. the highest R^2 with the smallest MSE.

Random sampling with 80% training set and 70 rows selected in the dataset even pushed the R^2 numbers higher to indicate a better fit, and except for RL, slightly reduced the MSE numbers, as shown in Table XI.

4. Forecast Analysis

When considering the results of Table X with 70% random sampling, the NN, kNN, RF, and Decision Tree performed relatively close to each other with mean square errors in the range 0.001-0.003. NN and kNN, however, provided a better fit (R^2). In terms of MSE, SVM came next. Only LR fell in the last category given that the actual forecasts were not quite linear. Thus, in terms of fitting the data in the original dataset, NN and kNN were the best performers.

Next, we compare the forecasts projected by IMF (refer to Table I) to the ones obtained by our six ML methods trained with rows 1-70, and to the Excel trendline Equation (5) projections for the 48 months in 2021-2024. Note that these forecasts have no corresponding data in the dataset for the 2021-2024 time period, and the forecasts are pure predictions by the ML methods. Fig. 3 shows the forecasts by the 6 ML methods, accompanied by the IMF forecast and the actual CPIs (added post mortem) for the same time period. Note that the trendline projected by Excel (Equation 5) overshoots all curved in Fig. 3 and is not visible in Fig. 3. As time moves away from the date of the last data vectors contained in the training set, i.e. as we focus onto the right side of the Fig. 3 graphs, all 6 ML models performed better than the Excel fourth order polynomial trendline of Equation 5, in terms of their proximity to the IMF forecast.

For the 2021-2024 forecast, the IMF forecast (3rd curve from top), and the 6 ML method forecasts (normalized) are plotted in Fig. 3. We note that the Excel trendline forecast is largely exaggerated and is not shown in Fig. 3. The IMF-derived plot seems to be the likely target that our ML method-based forecasts should aim at approaching. Although the normalized IMF forecast data was not used in training the 6 ML methods, we observe that the SVM forecast curve followed the LR forecast curve are the nearest to the IMF curve. After the actual (real) Saudi Arabian CPIs for years 2021-2020 were available and published, these

TABLE VIII. Optimal ML Parameters - One-Input Data

ML Method	Parameter Settings
Random Forest	# trees=1; Replicable training; Do not split subsets smaller than 5.
SVM SVM	SVM Cost =4.2, Regression epsilon=0.1; Kernel: x. y; Numerical tolerance=0.0028; iteration limit=100.
kNN	# neighbors=1; metric = Euclidean; Weight=Uniform.
Neural Networks Neural Networks	Neurons in hidden layers= 30, 50, 30, 10; Activation=ReLU; Solver=L-BFGS-B; Regularization=0.0001; Max. # iterations=430.
Decision Tree Decision Tree	Induce binary tree=checked; Do not split subsets smaller than 5; Limit the maximal tree depth=91; Stop when majority reaches 100%.
Linear Regression	Fit intercept; Regularization Strength: Alpha=2; Elastic Net Regulatization: L1 0.89:0.11 L2

TABLE IX. Improved Metrics For Testing with 25-row Testing Dataset - One-Input Data

ML Method	MSE	RMSE	MAE	R^2
Random Forest	0.057	0.239	0.193	-0.002
SVM	0.049	0.221	0.210	0.139
kNN	0.057	0.239	0.184	-0.001
Neural Networks	0.028	0.169	0.143	0.501
Decision Tree	0.057	0.239	0.184	-0.001
Linear Regression	0.044	0.210	0.179	0.224

TABLE X. Metrics with 70% Random Sampling of 70 Rows

ML Method	MSE	RMSE	MAE	R^2
Random Forest	0.002	0.045	0.034	0.959
SVM	0.008	0.090	0.068	0.839
kNN	0.001	0.024	0.018	0.989
Neural Networks	0.001	0.032	0.020	0.979
Decision Tree	0.003	0.052	0.039	0.945
Linear Regression	0.016	0.125	0.097	0.691

numbers were also plotted and displayed in Fig. 3 (2nd curve from the top, post mortem). When comparing the forecasts of the 6 ML methods to the actual CPI numbers in years 2021-2023, it is clear that the NN forecast is nearest to the actual CPI curve and is the best fit. This is certainly true for years 2021 and 2022 (x values of 97-110 in Fig. 3). Starting in 2023, SVM appears to be a second competitor to NN, considering its proximity to the actual CPI curve.

5. Conclusion

The ML methods and research methodology employed herein can be applied to forecast other countries' CPIs. With two distinct input data (month and year), the predictions of all six investigated ML methods were not impressive.

With one input data (combining month and year into 1 integer), prediction metrics tremendously improved, the best predictions were obtained with 80% random sampling of 70 data rows, with NN and kNN outperforming (R^2 = 0.994) the other four ML methods, and provided good predictions after training with historical CPI data.

Comparison of the 2021-2024 forecasts reveals that, although the IMF forecast data was not used in training the ML methods, the predictions by the SVM model followed by the LR model were nearest to the IMF forecast. More importantly, when the forecasts of the 6 ML methods were compared to the actual CPI numbers (post mortem) which are absent from the dataset used for training and testing, the NN method (with 4 hidden layers, and l-BFGS-B solver) was the best performer and best fit to the actual CPI numbers in years 2021-2023, while the Excel trendline 4th order polynomial fitting curve was the furthest. The MSEs obtained in this work are better than the ones reported in [24]. However, one limitation of ML methods, such as the ones employed in thus work, is that their predictions of the CPI may be confined to a short window of a few years.

We conclude that the NN ML model may be employed to forecast the CPI for up to two or three years from the

TABLE XI. Metrics with 80% Random Sampling of 70 Rows

ML Method	MSE	RMSE	MAE	R^2
Random Forest	0.001	0.032	0.025	0.983
SVM	0.005	0.070	0.053	0.919
kNN	0.000	0.020	0.015	0.994
Neural Networks	0.000	0.019	0.015	0.994
Decision Tree	0.002	0.043	0.031	0.970
Linear Regression	0.022	0.148	0.119	0.642

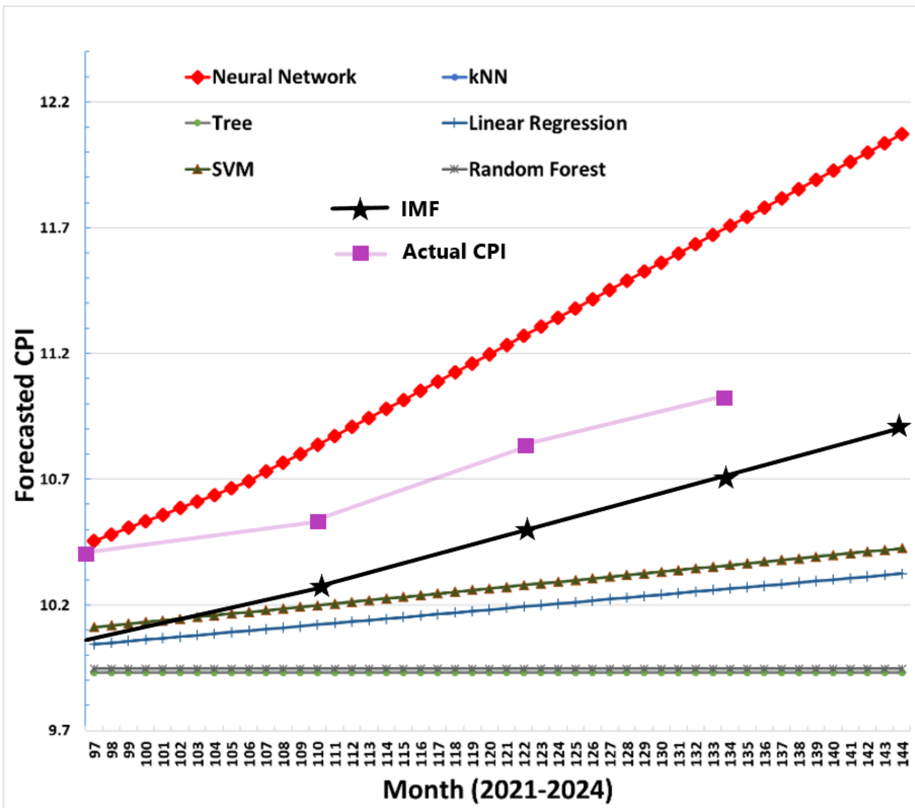


Figure 3. Comparative 2021-2024 Forecast Plots

last available economic data.

Future work should aim at further improving the accuracy and/or the time range of ML methods in forecasting economic data, by considering the addition to the dataset of feature data which has already been factored in generating the CPI numbers, such as energy, labor, food, and health costs, which will enrich the dataset and provide more insights and learning reinforcements to the ML models, thereby improving the accuracies of ML-generated CPI forecasts. Thus, the training and testing datasets will not just contain the CPI values of prior years but will also include the above suggested CPI component values. Another potential future work involves the exploration of other deep learning methods.

References

[1] J. Demsar, B. Zupan, G. Leban, and T. Curk,"Orange: from experimental machine learning to interactive data mining," PKDD, vol. 3202, pp. 537-539, 2004.

[2] J. Chambers, S. Mullick, and D. Smith, "How to choose the right forecasting technique," Harvard Business Review, Harvard Business Publishing, July 1972.

[3] G. Latif, J. Alghazo, F. N. Sibai, D. Awang Iskandar, and A. Khan, "Recent advancements in fuzzy C-means based techniques for brain MRI segmentation," Current Medical Imaging, 2021.

[4] S. Majzoub, A. Fahmy, F. N. Sibai, M. Diab,

and S. Mahmoud, ” Epilepsy detection with multichannel EEG signals utilizing AlexNet,” Circuits, Systems, and Signal Processing, 2023, Springer. https://doi.org/10.1007/s00034-023-02423-1

[5] F. N. Sibai, A. Asaduzzaman, and A. Sibai, ”A comparative study of machine learning methods for intrusion detection,” Proc. 10th Int. Conference on Electrical and Electronics Engineering, Istanbul, Turkey, 2023, IEEE.

[6] Brent H. Meyer, and Mehmet Pasaogullari, “Simple ways to forecast inflation: what works best? ,” Economic Commentary, No. 2010-17, Dec. 2010.

[7] F. Qin, T. Ma, J. Wang, H. Liang, T. Zhang, and H. Zhang, ”The CPI forecast based on GA-SVM,” Proc. 2010 Int. Conf. on Information, Networking and Automation (ICINA), Kunming, pp. V1-142-V1-147, 2010. doi: 10.1109/ICINA.2010.5636416.

[8] G. Box, and G. Jenkins, “Time series analysis: forecasting and control. San Fransisco, Holden-Day, 1976.

[9] E. Dhamo Gjika, L. Puka, and O. Zaçaj, “Forecasting consumer price index (CPI) using time series models and multiregression models (Albania case study) Proc. 10th Int Scienctific Conference Business and Management, Lithuania, 2018, DOI: 10.3846/bm.2018.51.

[10] E. Kalluci, and E. Gjika, “Difference equations and time series analysis in approximating the index prices (for Albania),” Journal of Multidisciplinary Engineering Science and Technology (JMEST), vol. 2, number 4, 2015. http://www.jmest.org/wp-content/uploads/JMESTN 42350659.pdf

[11] F. Zhang, W. Che, B. Xu, andJ. Xu, ”The research of ARMA model in CPI time series,” Proc. 2nd Int. Symposium on Computer, Communication, Control and Automation (ISCCCA-13), France, 2013. https://doi.org/10.2991/isccca.2013.8

[12] E. Ngailo, E. Luvanda, and E. Massawe, “Time series modeling with application to Tanzania inflation data,” Journal of Data Analysis and Information Processing, vol. 2(2), , Scientific Research, pp. 49-59, 2014. 10.4236/jdaip.2014.22007

[13] P. Junior, F. Salomon, and E. Pamplona, “ARIMA: An applied time series forecasting model for the Bovespa stock index,” Applied Mathematics, 5(21), Scientific Research, pp. 3383-3391, 2014. https://doi.org/10.4236/am.2014.521315

[14] F. Kharimah, M. Usman, W. Widiarti, and F. Elfaki, “Time series modeling and forecasting of the consumer price index Bandar Lampug,” Science International, PUBLICATIONS International, vol. 27(5), pp.4619-4624, 2015.

[15] Jama Mohamed, “Time series modeling and forecasting of Somaliland consumer price index: A Comparison of ARIMA and Regression with ARIMA Errors,” American Journal of Theoretical and Applied Statistics, Science Publishing Group, vol. 9, number 4, pp. 143-153, 2020. doi: 10.11648/j.ajtas.20200904.18

[16] M. Subhani, and K. Panjwani, “Relationship between consumer price index (CPI) and government bonds,” South Asian Journal of Management Sciences, Vol.3, No.1, 2009, pp.11-17, IURC. https://mpra.ub.uni-muenchen.de/36161/1/MPRA_paper_36161.pdf

[17] S. Adam, A. Awujola, and A. Alumgudu, “Modeling Nigeria’s CPI using ARIMA model. Int. Journal of Development and Economic Sustainability, EAJournals, vol.2, number 2, pp.37-47, 2014.

[18] Yifu Sheng, Jianjun Zhang, Wenwu Tan, Jiang Wu, Haijun Lin, and Guang Sun, “Application of Grey forecasting model to CPI index forecast, “ Proc. CAIS 2021: Advances in Artificial Intelligence and Security, Springer, pp 312–321, 2021.

[19] Thabani Nyoni, “Modeling and forecasting CPI in Mauritius,” Feb. 2019, https://mpra.ub.uni-muenchen.de/92423/1/MPRA_paper_92423.pdf

[20] J. Riofrio, S. Infante, and A. Hernández, “Forecasting the Consumer Price Index of Ecuador using Classical and Advanced Time Series Models,” Chapter in Information and Communication Technologies, Oct 2023. DOI: 10.1007/978-3-031-45438-7_9

[21] D. Thu Huong, V. Van Truong and B. Lam, ”Forecasting of consumer price index using the ensemble learning model with multi-objective evolutionary algorithms: Preliminary results,” Proc. International Conference on Advanced Technologies for Communications (ATC), Ho Chi Minh City, pp. 337-342, 2015. doi: 10.1109/ATC.2015.7388346.

[22] S. Zahara et al, ”Consumer price index prediction using long short term memory (LSTM) based cloud computing,” J. Phys.: Conf. Ser. 1456 012022, IOP publishing, 2020.

[23] J. Harris, “A machine learning approach to forecasting consumer food prices,” MS of Electronic Commerce Thesis, Dalhousie University, Canada, 2017.

[24] C. Yang, and S. Guo, “Inflation prediction method based on deep Learning,” Computational Intelligence

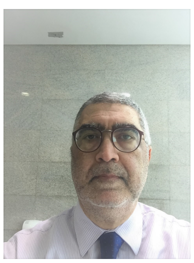
and Neuroscience vol. 4, pp. 1-6, August 2021. DOI: 10.1155/2021/1071145

[25] M. Qureshi, A. Khan, M. Daniyal, K. Tawiah, and Z. Mehmood, "A comparative analysis of traditional SARIMA and machine learning models for CPI data modelling in Pakistan," Applied Computational Intelligence and Soft Computing, vol. 5, pp. 1-10, November 2023. DOI: 10.1155/2023/3236617

[26] General Authority for Statistics, "Consumer price index, November 2020." https://www.stats.gov.sa/en/394

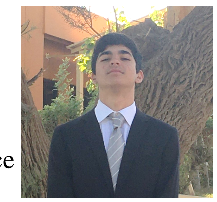
[27] International Monetary Fund, CPI home, https://data.imf.org/?sk=4FFB52B2-3653-409A-B471-D47B46D904B5

[28] "IMF, "Saudi Arabia forecast: consumer price index growth," https://www.ceicdata.com/en/indicator/saudi-arabia/forecast-consumer-price-index-growth#: :text=Saudi%20Arabia's%20Consumer%20Price %20Index,at%202.024%20%25%20in%20Dec%202025.



Fadi N. Sibai presently serves as the Associate Dean, College of Engineering and Architecture, Gulf University for Science and Technology. Kuwait. He previously served as Dean of the College of Computer Engineering and Science at Prince Mohammad Bin Fahd University, acting Dean of the School of Engineering at the American International University, and Program Director in the College of Information Technology at the UAE University. He also has extensive industrial experience with Intel Corporation and Aramco. He received the PhD and MSc degrees in Electrical Engineering from Texas A&M University, and the BSEE degree from the University of Texas at Austin. Dr. Sibai has authored or co-authored over 240 publications and reports. He holds PMP, CISSP, CQRM, and CCNA certifications.

Ali El-Moursy is a Professor and Head, Computer Engineering Department, University of Sharjah, U.A.E. He received the PhD degree from the University of Rochester, U.S.A. His research interests are in high performance computer architecture, multi-core multi-threaded micro-architecture, high performance computing, parallel computing, cloud computing, architecture modeling and simulation, and performance evaluation.



Ahmad Sibai is a senior student in the BS program in Computer Science at the University of South Florida, Florida, U.S.A. His research interests are in AI, Machine Learning, and Cybersecurity.