Mesopotamian journal of Big Data
Vol. **(2024)**, 2024, **pp**. 102–117
DOI: https://doi.org/10.58496/MJBD/2024/008 ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BigData

Research Article

Advanced Machine Learning Approaches for Enhanced GDP Nowcasting in Syria Through Comprehensive Analysis of Regularization Techniques

Khder Alakkari[1,*], Bushra Ali[2], Mostafa Abotaleb [3], Rana Ali Abttan[4], Pushan Kumar Dutta [5]

[1] *Faculty of Economics, Tartous, Statistics and Programming, Tishreen University, Syria.*

[2] *Department of Banking & Financial Sciences, Faculty of Economics, Tartous University, Tartous, Syria.*

[3] *School of Electronic Engineering and Computer Science, Department of System Programming, South Ural State University, 454080 Chelyabinsk, Russia.*

[4] *Department of Electrical Engineering, College of Engineering, University of Baghdad, Baghdad, Iraq.*

[5] *School of Engineering and Technology, Amity University, Kolkata, India.*

ARTICLE INFO

Article History
Received 07 May 2024
Accepted 08 Jul 2024
Published 02 Aug 2024

Keywords
GDP Nowcasting
Machine Learning
Lasso Regression
Ridge Regression
Elastic Net Regularization
Data-Scarce Environments
Economic Forecasting
Google Trends
Macroeconomic Indicators
High-Frequency Data

ABSTRACT

This study addresses the challenge of nowcasting Gross Domestic Product (GDP) in data-scarce environments, with a focus on Syria, a country facing significant economic and political instability. Utilizing a dataset from 2010 to 2022, three machine learning algorithms Elastic Net, Ridge, and Lasso were applied to model GDP dynamics based on macroeconomic indicators, commodity prices, and high-frequency internet search data from Google Trends. Among these, the Lasso regression model, noted for its variable selection and sparsity promotion, proved most effective in capturing Syria's complex economic realities, achieving the lowest Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). This accuracy highlights the Lasso model's capability to identify robust economic relationships despite limited data, thereby reducing overfitting and improving forecast generalizability. The study underscores the significant impact of non-traditional indicators, such as Google Trends Agriculture (GTA) and Google Trends Consumption (GTC), on GDP growth, offering valuable insights for policymakers and analysts in data-scarce environments. The findings support the use of machine learning techniques, particularly Lasso regression, as powerful tools for economic forecasting, enhancing informed decision-making in challenging settings.

1. INTRODUCTION

Nowcasting, the process of estimating the current state of an economic variable, such as Gross Domestic Product (GDP), is crucial for policy-makers, analysts, and investors to make informed decisions in a timely manner [1]. Traditional methods of GDP nowcasting rely on historical data and econometric models, which often face limitations due to data availability, model assumptions, and the rapidly changing nature of the global economy.

Predicting the short-term dynamics of the economy is a crucial input into various economic agents' decision-making processes. However, accurately nowcasting key macroeconomic indicators can be challenging for various reasons. For instance, official estimates are released with a substantial delay, and the uncertainty in the data and estimates can lead to multiple revisions, sometimes years after their first release. Additionally, various data series are required for accurately nowcasting of macroeconomic indicators — further complicating the process

*Corresponding author. Email: khderalakkari1990@gmail.com

In recent years, machine learning techniques have gained traction in various fields, including economics, due to their ability to capture complex patterns and handle large-scale datasets.
Machine learning (ML) has revolutionized decision-making in a variety of fields by providing new tools for forecasting. At its core, ML involves formulating a loss or cost function for forecasting rules. A forecasting rule, denoted as $\boldsymbol{f(x_t)}$, predicts the value of a target variable, $\boldsymbol{y_t + h}$ at a future horizon, $\boldsymbol{h}$ based on information available at time $\boldsymbol{t}$. The loss function, $\boldsymbol{\ell(y_t + h, f(x_t))}$ quantifies the error incurred by the forecasted value compared to the actual outcome. The central goal is to approximate the optimal decision rule, $\boldsymbol{f *}$ which minimizes the expected loss, $\boldsymbol{E[\ell(y_t + h, f(x_t))]}$. This approach has its roots in decision theory [2], and is adopted in statistical learning and economic forecasting [3]. For example, when using a cubic loss function, $\boldsymbol{\ell(y_t + h, f(x_t)) = (y_t + h - f(x_t))^3}$ the optimal decision rule corresponds to the (non-linear) regression, $\boldsymbol{f * (x_t) = E[y_t + h|x_t]}$ with respect to $\boldsymbol{f(x_t)}$. Specifically, this means that the optimal forecasting rule is the one that minimizes the squared difference between the forecasted value and the actual outcome.
Decision rules used for forecasting are based on the available data. However, there is a trade-off between bias and variance in forecasting performance. Overfitting may result from flexible nonparametric approaches that decrease bias at the expense of increased variation. In order to lessen variance, regularization and dimensionality reduction also introduce bias. A variety of high-dimensional, nonparametric tools are made possible by machine learning, which can be used to improve forecasting performance, adjust to the bias-variance trade-off, and provide flexible and accurate approximations of optimal decision rules. This means that machine learning can help us develop better decision rules for forecasting by appropriately balancing bias and variance.
Many widely used machine learning tools are based on statistical methods that are well-known and well-established. For example, deep learning can be understood as a regression model with non-linearity that is created by a multilayer neural network [4]. As a new generation of regression and classification trees, random forests and gradient boosting make sense [5]. Penalized regression has its roots in the concept of shrinkage [6], which is the process of adding bias to a model in order to reduce its variance.
Nowcasting GDP using machine learning methods has shown promising results in various studies. Researchers have explored different machine learning algorithms, such as ridge regression, boosting, and elastic net, and random forest, to estimate GDP growth in different countries [7] [8] [9] These studies have compared the performance of machine learning models with traditional time series models like Vector Autoregression (VAR) and found that machine learning models outperform VAR in terms of predictive accuracy [10]. Additionally, incorporating high-frequency macroeconomic indicators, financial market data, and economic uncertainty indices in the nowcasting models has further improved their performance. The use of machine learning methods has also enabled timely predictions of economic growth, addressing the issue of lag in official GDP figures. Overall, machine learning techniques have proven to be effective in nowcasting GDP and providing accurate and timely insights into economic trends.
Examine the issue of Nowcasting the US GDP growth for each quarter through the use of more frequent macroeconomic and financial data, it is discovered that the machine learning forecasts are either better than or comparable to those released by the Federal Reserve Bank of New York [11]. Further benefits can be obtained by utilizing the information derived from the textual analysis of business news [12],
This research aims to explore the application of machine learning algorithms for nowcasting GDP in Syria, a country facing significant economic challenges and political instability. For Syria, there have been some studies that have attempted to real GDP nowcasting using available data. among which we mention [13] research that uses the MIDAS Almon Polynomial Weighting model to Nowcasting Syria's annual GDP based on the monthly inflation rate data and [14] In which the Bayesian shrinkage function of the mixed VAR model was used for real-time forecasting of real GDP in Syria based on five high-frequency macroeconomic variables.
Syria's economic landscape has been greatly impacted by the ongoing war, which has resulted in severe disruptions to economic activities, infrastructure, and data collection systems. As a result, traditional GDP estimation methods may not accurately reflect the current economic situation. By leveraging machine learning techniques, this study seeks to provide a more accurate and timely estimation of Syria's GDP. The research will utilize a diverse range of data sources, including macroeconomic indicators, And financial market indicators textual data from Google Trend. These non-traditional data sources offer valuable insights into economic activities, even in the absence of official statistics, and can help overcome data limitations in war-affected regions. The machine learning models will be trained on historical data from Syria, the models will be designed to capture both linear and non-linear relationships between the predictors and GDP, allowing for a more comprehensive understanding of the economic dynamics.

2. RELATED WORK

[15] he researches analyses the use of sparse methods to forecast the actual (in the chain-linked volume meaning) expenditure components of the US and EU GDP in the short-run sooner than national statistics authorities officially release

the data. Use LASSO in conjunction with its most recent improvements for variable selection and forecast estimate. In order to enhance the forecasting performance, suggest a modification that combines main components analysis with LASSO scenarios. Using benchmark ARMA and factor models as a comparison, conducted pseudo-real-time experiments for gross fixed capital formation, private consumption, imports, and exports across a sample period from 2005 to 2019 in order to assess the predicting performance. The major findings imply that sparse approaches are capable of identifying suitable subsets of explanatory variables and outperforming the benchmarks.

[11] Structured machine learning regressions for high-dimensional time series data that may be collected at various frequencies are introduced in this article. provide oracle inequalities for the sparse-group LASSO estimator in a framework that takes mixing processes into account and acknowledges the possibility of heavier than exponential tails in the macroeconomic and financial data. The estimator outperforms other options in an empirical application to nowcasting US GDP growth, and text data can be a helpful supplement to more conventional numerical data.

[16] test the nowcast performance of popular algorithms in a full "real-time" setting, that is, with real-time vintages of New Zealand GDP growth (our target variable) and real-time vintages of approximately 600 predictors, in order to contribute to the growing body of literature on forecasting macroeconomic variables using machine-learning algorithms. The outcomes demonstrate that machine-learning algorithms can outperform a dynamic factor model and a basic autoregressive benchmark by a large margin. additionally demonstrate the potential for machine-learning algorithms to enhance, and in one instance even surpass, the official Reserve Bank of New Zealand estimates.

[17] this paper reviews it the required steps for Nowcasting, taking Luxembourg as a considered both standard and alternative indicators, used as inputs in several nowcasting methods, including various factor and machine learning models. Overall, mixed frequency dynamic factor models and neural networks perform well, both in absolute terms and in relative terms with respect to a benchmark autoregressive model. The gains are larger during problematic times, such as the financial crisis and the recent Covid period.

[18] this paper proposed bridge models to nowcast French gross domestic product (GDP) quarterly growth rate. The bridge models, allowing economic interpretations, are specified by using a machine learning approach via Lasso-based regressions and by an econometric approach based on an automatic general-to-specific procedure. These approaches allow to select explanatory variables among a large data set of soft data. A recursive forecast study is carried out to assess the forecasting performance. It turns out that the bridge models constructed using the both variable-selection approaches outperform benchmark models and give similar performance in the out-of-sample forecasting exercise.

[19] The text data is derived from fifteen well-known European newspapers. Daily sentiment metrics are created from these news articles and assessed their value for nowcasting; by comparing to competitive and rigorous benchmarks, it is found that newspaper text is helpful in nowcasting GDP growth, especially in the first half of the quarter when other lower-frequency soft indicators are not available. A non-linear machine learning model can help capture extreme movements in growth. The choice of sentiment measure matters when tracking economic shocks such as the Great Recession and the Great Lockdown.

[20] This research aims to use alternate, faster-to-available data to improve the accuracy of GDP nowcasting models. Specifically, created nowcasting models that integrate sparse estimation using Elastic Net employing weekly retail sales data and hundreds of daily Internet search traffic data, in addition to typical monthly economic data. Using the forecast combination approach, which combines various forecasting models, a large number of candidate models were generated for the model formulation and data selection. The "Best models" were chosen based on their ability to minimize forecast error, incorporating data collected after the COVID-19 pandemic. The investigation demonstrates that using alternate data considerably raises the model's forecasting accuracy.

[21] In this paper, projection models for Peru's monthly GDP rate growth are presented. These models combine high-frequency unstructured sentiment variables with structured macroeconomic data. The window sample, which spans 91 variables in all, runs from January 2007 to May 2023. The most accurate predictions for each model were found by evaluating six ML algorithms. In comparison to traditional time series models, the results show that each machine learning model with unstructured data has a high capacity to produce more accurate and anticipated predictions. The models that performed best were Gradient Boosting Machine, LASSO, and Elastic Net, which were able to reduce prediction errors by 20% to 25% when compared to the AR and Dynamic Factor Models (DFM) models.

3. METHODOLOGY

The research methodology consists of using three machine learning algorithms (Lasso, Ridge, Elastic Net) to predict the real-time Gross Domestic Product (GDP) in Syria and obtain information about economic growth in the last two years in the absence of the ability to collect new information about this most important variable in the economy. We use these algorithms because of their ability to deal with the large number of variables with limited temporal availability. These techniques are a popular remedy for the overfitting issue, which occurs when a model fits training data well but struggles

to generalize to fresh test data. Some or all of the coefficients are kept, and their magnitude is shrunk, contingent on the particular parameters selected for our model. to calculate the special cases of the Lasso and Ridge models, as well as the elastic net regression net. We explore the network over several lambda penalty parameters and estimate one lambda penalty parameter at a time. In order to select the parameter with the lowest error rate, we additionally employ cross validation methods. Following estimation, we display further validation statistics and diagnostics, graphs of coefficient evolution with respect to the penalty parameter, and customized views of coefficient tables and other summary data [22] [23].
Even though the ordinary least squares estimator has many good qualities, like unbiasedness, it can have significant volatility in some situations. The least squares estimates are highly susceptible to random errors and may have a high variance, for instance, if the data have many correlated regressors or more regressors than the length of the dataset (often referred to as the large, small problem). While the Ordinary Least Squares (OLS) estimator, which minimizes the residual sum of squares (RSS), is widely used, it encounters significant challenges when dealing with multicollinearity and limited data—characteristics inherent to the Syrian context. The OLS estimator seeks to find the coefficients, β, that minimize:

$$\mathrm{RSS}(\beta) = \Sigma(y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta p x_{ip})^2 \qquad (1)$$

where: y_i represents the observed GDP for the i-th observation. x_{i1}, ..., x_{ip} are the corresponding values of the predictor variables. $\beta_0, \beta_1, \ldots, \beta p$ denote the coefficients to be estimated. Multicollinearity, a phenomenon where predictor variables are highly correlated, inflates the variance of OLS estimates, leading to unreliable coefficients and difficulty in isolating the individual effects of predictors. The Variance Inflation Factor (VIF) quantifies the severity of multicollinearity:

$$VIF_i = \frac{1}{1-R_i^2} \qquad (2)$$

Here, R_i^2 represents the R-squared value obtained from regressing the i predictor on all remaining predictors. High VIF values (generally exceeding 5 or 10) signal strong multicollinearity, necessitating the use of regularization techniques to mitigate its impact. Regularization techniques introduce a penalty term to the OLS cost function, effectively shrinking the magnitude of coefficients and reducing model complexity. This approach combats overfitting, allowing the model to generalize better to new data. This study focuses on three prominent regularization techniques: Ridge Regression, Lasso Regression, and Elastic Net Regression. Using an elastic net, lasso, ridge regression model, regularization can be used to lessen this variance by creating bias and lowering the total error. Penalized regression techniques such as ridge regression, Lasso, and elastic net all function by reducing the model's regressors' magnitudes. Adding a penalty component to the standard cost function for linear regression is the normal procedure:

$$J = \frac{1}{2m}\sum_{i=1}^{m}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\left[\frac{(1-\alpha)}{2}\sum_{j=1}^{p}\beta_j^2 + \alpha\sum_{j=1}^{p}\left|\beta_j\right|\right] \qquad (3)$$

Equation (1) transforms into a ridge regression model, a Lasso model, or an elastic net model based on the value of in the penalty term. The penalty's impact is determined by the size of the punishment parameter, $\lambda \geq 0$. In the event that a "large" value for λ is selected, the minimizing of this cost function:

$$min_\beta J \qquad (4)$$

will cause the β levels to decrease or possibly go to zero. Less complicated and less prone to overfitting is a model with fewer (or zero) coefficients.

Ridge:

The conventional least squares estimator with an L2 penalty term added is known as the ridge estimator:

$$J = \frac{1}{2m}\sum_{i=1}^{m}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2 \qquad (5)$$

The coefficients' sizes are decreased by the penalty, but they are not made zero. More shrinkage occurs when the cost function J is minimized with a larger λ. We employ cross-validation in the process of selecting the penalty parameter. approach entails dividing the data into test and training sets, iterating over the parameter list, and using the training data to estimate a set of coefficients. The dependant is then predicted using these coefficients on the test data set, and an error measure is computed. λ_{min} is the penalty parameter with the least amount of mistake [24]. We additionally compute the standard error of the error measures across the training and test sets for cross-validation processes involving multiple test sets (providing the maximum penalty parameters between one and two standard errors of λ_{min} as $\lambda_{1\ SE}$ and $\lambda_{2\ SE}$).

Equation (2)'s right-hand side can be solved for the coefficients using the conventional method, which results in:

$$\beta = (X'X + \lambda I)^{-1}(X'Y) \qquad (6)$$

It should be noted that when the OLS coefficients are represented by Equation (3), the coefficients are more severely penalized ($\beta \to 0$) as the penalization parameter increases ($\lambda \to \infty$). Conveniently, the matrix $X'X + \lambda I$ becomes nonsingular and invertible when the positive constant λ is added to the diagonal [25].

Variance is the degree of uncertainty in those estimations, whereas bias is the discrepancy between the estimated and actual values. Reducing model complexity to the point where the model both fits training data and generalizes well to test data is the aim of regularization. More complicated, low-bias models typically fit training data better, whereas less complex, low-variance models generally generalize better to test data. For instance, in OLS, the number of regressors and model complexity are correlated. The estimator's variance is decreased by reducing the number of regressors, but bias is introduced in the process. We refer to this as the bias-variance tradeoff. For models such as elastic net, ridge, and Lasso, the comparable decrease in complexity is accomplished not just via the deletion of individual coefficients but also through a reduction in the magnitudes of the coefficients. By using Equation (3) to determine the expectation of the ridge coefficient, we can examine this in greater detail:

$$\begin{aligned} E\left(\hat{\beta}^{\text{ridge}}\right) &= E\left((X'X + \lambda I)^{-1}(X'Y)\right) \\ &= (X'X + \lambda I)^{-1}(X'X)\beta \end{aligned} \qquad (7)$$

If $E\left(\hat{\beta}^{\text{ridge}}\right) \neq \beta$ the ridge estimator is biased. And see also $\lambda \to \infty$, $E\left(\hat{\beta}^{\text{ridge}}\right) \to 0$ as predictably. The bias and variance of the ridge estimator are provided by, albeit we won't derive them here:

$$\text{bias}\left(\hat{\beta}^{\text{ridge}}\right) = -\lambda(X'X + \lambda I)^{-1}\beta \qquad (8)$$

$$\text{var}\left(\hat{\beta}^{\text{ridge}}\right) = \sigma^2(X'X + \lambda I)^{-1}X'X\{(X'X + \lambda I)^{-1}\}' \qquad (9)$$

Equation (5)'s bias rises as lambda rises, whereas Equation (6)'s variance (σ^2 is the error variance from the residuals) falls.

Lasso:

An L1 penalty term added to the OLS estimator is called the Lasso (Least Absolute Shrinkage and Selection Operator) estimator [26]:

$$J = \frac{1}{2m}\sum_{i=1}^{m}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j| \qquad (10)$$

There is no analytical solution since the penalty term is a sum of absolute values, making the answer nonlinear. It is necessary to solve the Lasso equation quantitatively. The coefficients may approach 0, in contrast to ridge regression. It is important to note that Lasso will typically decrease the other regressors in the group and favor the associated one [27]. Ridge regression will cause the group's total coefficients to decrease proportionately.

Elastic Net:

The ridge and Lasso models are combined to create the elastic net model. Iterating Equation (1):

$$J = \frac{1}{2m}\sum_{i=1}^{m}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\left[\frac{(1-\alpha)}{2}\sum_{j=1}^{p}\beta_j^2 + \alpha\sum_{j=1}^{p}|\beta_j|\right] \qquad (11)$$

The regularization term is comprised of the L1 and L2 penalties, with the parameter α regulating the degree of mixing. This transforms into a ridge regression model when α = 0. It becomes a Lasso model at α = 1. Since the Lasso term drives the correlated regressors towards zero while the ridge term shrinks them proportionally, the compromise between these two [28].

Time Series Cross Validation:

One or more tuning parameters must be specified for the practical application of ML algorithms. It is customary to use K-fold cross validation for data, which may or may not be suitable for time series data. According to [29], the conventional K-fold cross-validation technique is still appropriate for autoregressive models with errors. However, traditional cross-validation fails with correlated errors because of the correlation between the training and test samples, such as when the regression has only projection interpretation due to misspecification.

In order to prevent this, we first look at the autocorrelation of errors and select a range of cross-validation techniques, such as:

- **K-fold:** The Random generator and Seed fields dictate the specifics of the shuffling. The dataset is split into K equally spaced "folds," and the ordering is shuffled. The remaining K-1 folds are merged to create the training set, and one fold is kept out as the test set. Next, repeat this process with a test set consisting of each fold being held out sequentially. The data are averaged over all K folds where:

$$cv_{(k)} = \frac{1}{k}\sum_{i=1}^{k} MSE_i \qquad (12)$$

where $\mathrm{MSE_i}$ is loss function. And:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \qquad (13)$$

- **Leave One Out**: With the exception of holding out a test set of size P and using the remaining data as the training set, this is comparable to K-Fold. Repeat this procedure for each of the remaining combinations [30]. Utilize the omitted observations to evaluate the model:

$$CV(\lambda) = \frac{1}{T}\sum_{t=1}^{T} \ell\left(y_t - \hat{f}_{\lambda,-t,l}(x_t)\right) \qquad (14)$$

Where ℓ is loss function (MSE) and $\hat{f}_{\lambda,-t,l}(x_t)$ serve as a machine learning model's prediction rule.

- **Rolling Windows:** Following the selection of a window size for the dataset, the window is partitioned into training and test sets, with the test set in each window always following the training set (the default test set size is 1, but we can alternatively define the test set size as a fraction). Until it reaches the end of the dataset, the window "rolls" through it. We can also select a starting period in the dataset to exclude from cross-validation (the initial period) and how far ahead of the training set you want the test set to be (the horizon).

Following cross-validation, we assess the models' correctness and contrast them using the performance metrics (Root Mean Square Error – Mean Absolute Percentage error – Normalized Root Mean Square Error) listed below:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_i - y_i')^2}{n}} \qquad (15)$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - y_i'}{y_i}\right| \times 100\% \qquad (16)$$

$$NRMSE = \frac{RMSE}{Mean} \qquad (17)$$

where $\boldsymbol{y_i'}$ is the predicted value; $\boldsymbol{y_t}$ is the actual value; and $\boldsymbol{n}$ is number of fitted observed.

Data and Results:

Table 1, "Information about Variables," systematically catalogs the datasets utilized from multiple sources spanning the period from 2010 to 2023, with the exception of the Gross Domestic Product (GDP) data which is recorded until 2022. The table enumerates various economic and financial indicators that have been gathered for the purpose of employing machine learning techniques to nowcast Syria's GDP. These indicators include market indices, trading volumes, consumer price indices, and macroeconomic data derived from high-frequency internet searches, detailed as follows:

- **Market Data**: Collected from the Damascus Stock Exchange, data includes market stock price indices, the number of traded companies, market and trading values, the average daily trading value, the number of shares traded, the number of trades executed, and the number of trading days.

- **Macroeconomic Indicators**: Compiled from the Central Bureau of Statistics and the Central Bank of Syria, these include consumer prices, the exchange rate, and the interest rate.
- **Commodity Prices**: Gold price data obtained from the Craftsmanship Association.
- **Google Trends Data**: Internet search frequencies for terms related to gold, stocks, education, gas, oil, employment, consumption, agriculture, industry, prices, investment, exchange rates, decrees, and law.

Each entry in the table specifies the source, measurement unit, time span, and the specific economic variable tracked. This extensive collection of data provides a robust framework for analyzing the intricate dynamics of Syria's economy through advanced computational models, enhancing the predictive accuracy of GDP nowcasting in a context marked by significant data limitations due to ongoing regional instability. This methodological approach aligns with contemporary research practices in economic forecasting, leveraging non-traditional data sources to compensate for gaps in official statistical reporting According to the following table:

TABLE I. INFORMATION ABOUT VARIABLES

Variable	Time	Measuring unit	Code	Source
Market stock price index	2010-2023	Point	**DWX**	**Invalid source specified.**
Number of traded companies	2010-2023	Company	**NTC**	**Invalid source specified.**
Market value	2010-2023	Billion Syrian pounds	**MV**	**Invalid source specified.**
Trading value	2010-2023	Billion Syrian pounds	**TV**	**Invalid source specified.**
Average daily trading value	2010-2023	Billion Syrian pounds	**ADTV**	**Invalid source specified.**
Number of shares traded	2010-2023	Billion Syrian pounds	**NST**	**Invalid source specified.**
Number of trades executed	2010-2023	Thousand deals	**NET**	**Invalid source specified.**
Number of trading days	2010-2023	Day	**NTD**	**Invalid source specified.**
Consumer prices	2010-2023	Point	**CPI**	**Invalid source specified.**
Exchange rate	2010-2023	Bound	**EXR**	**Invalid source specified.**
Interest rate	2010-2023	Point	**INTM**	**Invalid source specified.**
Gold Price	2010-2023	Bound	**Gold**	**Invalid source specified.**
Google Trend - Gold	2010-2023	Point	**GTG**	Google Trend
Google Trend - Stock	2010-2023	Point	**GTS**	Google Trend
Google Trend - Education	2010-2023	Point	**GTE**	Google Trend
Google Trend - Gaz	2010-2023	Point	**GTGg**	Google Trend
Google Trend - Oil	2010-2023	Point	**GTO**	Google Trend
Google Trend - Employment	2010-2023	Point	**GTEM**	Google Trend
Google Trend - Consumption	2010-2023	Point	**GTC**	Google Trend
Google Trend - Agriculture	2010-2023	Point	**GTA**	Google Trend
Google Trend - Industry	2010-2023	Point	**GTI**	Google Trend
Google Trend - Prices	2010-2023	Point	**GTP**	Google Trend
Google Trend - Investment	2010-2023	Point	**GTINV**	Google Trend
Google Trend - Exchange Rate	2010-2023	Point	**GTEX**	Google Trend
Google Trend - Decree	2010-2023	Point	**GTD**	Google Trend
Google Trend - Law	2010-2023	Point	**GTL**	Google Trend
Gross Domestic Product	2010-2022	Million Syrian pounds	**GDP**	**Invalid source specified.**

Table 2, "Elastic Net Regularization Model for GDP Nowcasting," delineates the modeling framework and resultant coefficients obtained from applying Elastic Net regularization to nowcast Syria's GDP using an array of macroeconomic and digital indicators spanning the period from 2010 to 2022. This table is central to demonstrating the methodological rigor and analytical precision of the nowcasting model employed in this study.

TABLE II. MODEL FOR GDP SYRIA USING ELASTIC NET REGULARIZATION ALGORITHM

Dependent Variable: GDP
Method: Elastic Net Regularization
Sample (adjusted): 2010 2022

Included observations: 13 after adjustments			
Penalty type: Elastic Net (alpha = 1)			
Lambda at minimum error: 0.04974			
Regressor transformation: L1			
Cross-validation method: Leave P Out (number left out = 2)			
Selection measure: Mean Squared Error			
	(minimum)	**(+ 1 SE)**	**(+ 2 SE)**
Lambda	0.04974	0.1047	0.1519
Variable		**Coefficients**	
CPI	-0.034685	-0.018533	-0.007315
ADTV	0.000342	0.000270	0.000109
DWX	-0.003563	-1.11E-08	0.000000
EXR	-0.002458	-0.000640	-0.002608
GOLD	1.15E-10	0.000000	-5.80E-06
GTA	0.947115	0.763708	0.708154
GTC	0.863529	0.481011	0.237641
GTD	0.451748	0.455984	0.438940
GTE	-0.225158	-0.169157	-0.125911
GTEM	0.000283	0.000000	0.000000
GTEX	1.08E-10	0.007929	2.66E-06
GTG	0.088256	0.077746	0.064897
GTGG	0.156817	0.143617	0.093736
GTI	-1.25E-09	0.000000	0.000000
GTINV	0.000477	0.000000	0.000000
GTL	0.088745	0.085268	0.049383
GTO	0.076771	0.106583	0.000000
GTP	-0.000803	-3.27E-08	0.000000
GTS	-2.76E-05	4.51E-10	0.000000
INTM	-1.10E-08	-6.88E-08	-2.35E-05
MV	1.90E-09	7.06E-08	1.22E-05
NET	2.13E-10	0.001641	0.000551
NC	0.125438	0.097283	0.095302
NST	0.000724	0.000173	5.20E-05
NTC	0.362562	0.247276	0.222495
NTD	0.000452	0.000728	7.88E-07
TV	-0.001902	0.000000	0.000000
C	-32.86687	-30.73022	-25.53476
d.f.	27	22	19
L1 Norm	36.29873	33.38776	27.58189
R-squared	0.935875	0.857619	0.771405

The table provides a detailed overview of an Elastic Net regularization model designed to nowcast Syria's GDP. The model employs a Lasso approach (alpha = 1), prioritizing variable selection and sparsity. It analyzes data from 2010 to 2022, using 13 observations. The optimal lambda value for minimizing error is determined to be 0.04974, with additional analyses conducted at one and two standard errors of lambda to assess model sensitivity. Notably, coefficients for variables like Google Trends Agriculture (GTA) and Google Trends Consumption (GTC) are high, signifying their strong predictive power for GDP. Conversely, coefficients for CPI and EXR decrease significantly as lambda increases, highlighting their susceptibility to regularization. The Leave P Out cross-validation method, with 2 observations left out, ensures robust model validation. The model aims to minimize forecast errors, as indicated by the Mean Squared Error selection measure. The decreasing L1 norm across higher lambda values suggests a reduction in variable influence as regularization intensifies. This decreasing influence is consistent with the economic intuition that as regularization increases, the model becomes less complex and relies less heavily on individual predictors, potentially sacrificing some explanatory power for greater generalizability and robustness. The high initial R-squared value of 0.935875, which declines as lambda increases, illustrates the inherent trade-off between model complexity and fit, underscoring the importance of finding an optimal balance between capturing specific economic relationships and avoiding overfitting.

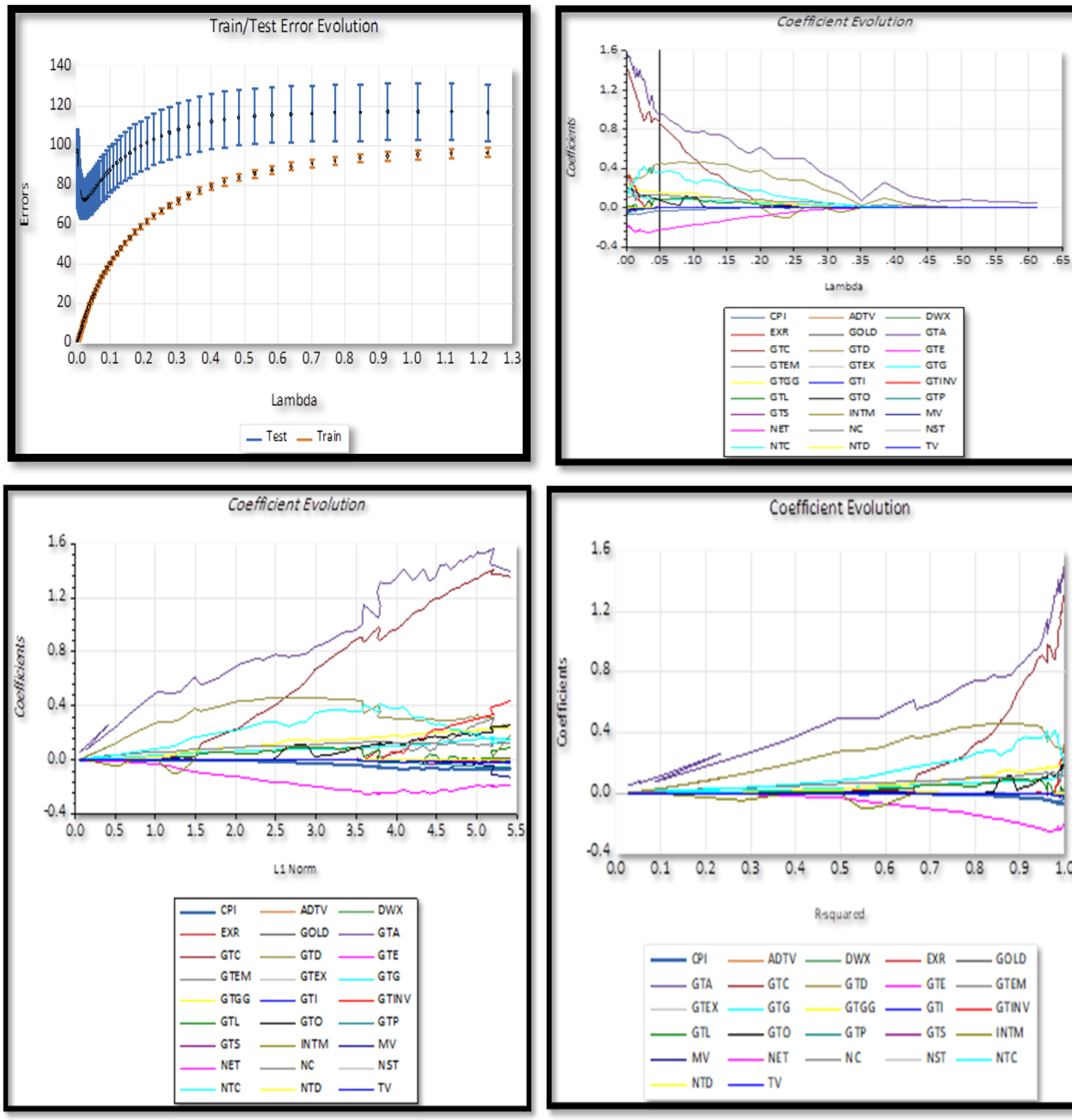

Fig .1. Error and Coefficient Evolution for (Lambda-L1 Norm-R Squared) For Elastic Net Regularization

The "Coefficient Evolution" graph, depicting the trajectory of predictor coefficients under increasing L1 Norm (a measure of regularization strength), provides valuable insights into the model's behavior and variable selection process. The graph illustrates how the Elastic Net regularization method selectively shrinks coefficients, balancing bias and variance to optimize model performance and mitigate overfitting. The distinctive trajectories of coefficients, notably the persistent influence of Google Trends Agriculture (GTA) and Consumer Price Index (CPI) even at high regularization levels, highlight the robustness and significance of these predictors in the GDP nowcasting model. The visualization allows for a nuanced understanding of the model's sensitivity to different variables under varying regularization pressures, ultimately aiding in validating the model's configuration and ensuring its effectiveness in real-world economic forecasting applications.

TABLE III. MODEL FOR GDP SYRIA USING RIDGE ALGORITHM

Dependent Variable: GDP			
Method: Elastic Net Regularization			
Sample (adjusted): 2010 2022			
Included observations: 13 after adjustments			
Penalty type: Ridge (alpha = 0) *analytic			
Lambda at minimum error: 0.1299			
Regressor transformation: L1			
Cross-validation method: Leave P Out (number left out = 2)			
Selection measure: Mean Squared Error			
	(minimum)	(+ 1 SE)	(+ 2 SE)
Lambda	0.1299	0.5757	1.212
Variable		Coefficients	
ADTV	0.001268	0.000860	0.000616
CPI	-0.022832	-0.008645	-0.004707
DWX	-0.002397	0.000485	0.000803
EXR	-0.010395	-0.004599	-0.002695
GOLD	0.008478	0.004820	0.002844
GTA	0.494727	0.212716	0.123901
GTC	0.603615	0.238982	0.126546
GTD	0.368382	0.188546	0.112902
GTE	-0.136522	-0.062843	-0.036122
GTEM	-0.017275	-0.006740	-0.001250
GTEX	0.008665	0.010397	0.007656
GTG	0.062302	0.030204	0.017909
GTGG	0.128114	0.058053	0.032806
GTI	-0.015098	-0.007129	-0.003335
GTINV	0.127896	0.027037	0.011412
GTL	0.097359	0.059817	0.037831
GTO	0.112133	0.040578	0.021938
GTP	0.009988	0.015301	0.010953
GTS	0.004238	0.021508	0.015850
INTM	-0.261341	-0.213655	-0.148273
MV	0.027737	0.032237	0.022945
NC	0.069920	0.033739	0.020612
NET	0.001042	0.001689	0.001290
NTC	0.209351	0.109011	0.067209
NST	0.001477	0.000981	0.000709
NTD	0.004934	0.003621	0.002475
TV	-0.002253	0.001584	0.001644
C	-503.4156	-238.6134	-141.1699
d.f.	28	28	28
L1 Norm	506.4626	240.1203	142.0719
R-squared	0.841347	0.551935	0.376930

The table (3) presents a Ridge regression model for nowcasting Syria's GDP. The shrinking of coefficient magnitudes with increasing lambda highlights the bias-variance tradeoff inherent in Ridge regression, where model complexity is reduced to improve generalizability and prevent overfitting. The optimal lambda value of 0.1299 minimizes forecast error, as indicated by the Mean Squared Error selection measure. The L1 Norm, reflecting the sum of coefficient magnitudes, decreases as lambda increases, suggesting diminishing individual variable influence with greater regularization. Notably, while coefficients for variables like GTA (Google Trends Agriculture) and GTC (Google Trends Consumption) remain relatively strong predictors, coefficients for CPI and EXR are more susceptible to shrinkage, indicating their potentially less robust relationship with GDP. The R-squared value, representing the model's goodness of fit, declines as lambda increases, demonstrating the tradeoff between model complexity and explanatory power. From an economic perspective, the strong predictive power of GTA and GTC suggests the importance of consumer behavior and agricultural activity in driving GDP growth in Syria. The susceptibility of CPI and EXR to regularization could point to market distortions or policy interventions that impact their direct influence on GDP. This nuanced analysis, derived from the Ridge regression model, provides valuable insights for understanding the drivers of economic growth in Syria and highlights the potential of leveraging non-traditional data sources in economic forecasting.

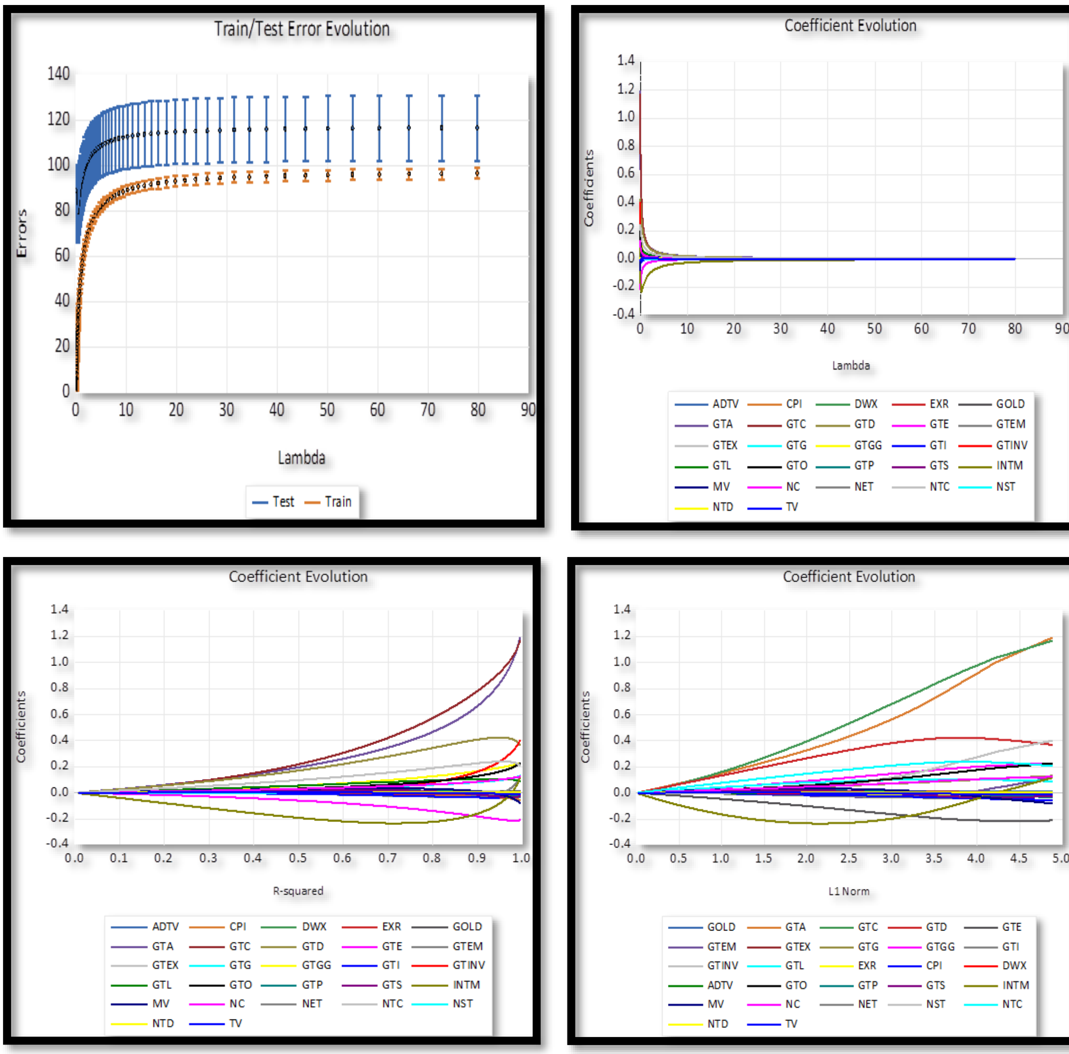

Fig .2. Error and Coefficient Evolution for (Lambda – L1 Norm-R Squared) for Ridge Algorithm

The figure (3) depicting coefficient evolution as a function of R-squared in the Ridge model of Syria's GDP provides a nuanced understanding of variable importance under varying model fit scenarios. The trajectories of coefficients, notably the sharp positive trends for GOLD, GTA, and GTC, indicate their increasing influence as the model explains more variance in GDP. Conversely, the relative stability of other coefficients suggests their consistent yet minor role in predicting GDP. This differential sensitivity to R-squared underscores the model's ability to discern key economic drivers, highlighting variables like GOLD and GTA as potentially significant contributors to economic growth in Syria. This visualization aids not only in interpreting model dynamics but also in optimizing model specifications for enhanced predictive accuracy by focusing on the most impactful variables.

TABLE IV. MODEL FOR GDP SYRIA USING LASSO ALGORITHM

Dependent Variable: GDP			
Method: Elastic Net Regularization			
Sample (adjusted): 2010 2022			
Included observations: 13 after adjustments			
Penalty type: Lasso (alpha = 1)			
Lambda at minimum error: 0.04974			
Regressor transformation: L1			
Cross-validation method: Leave P Out (number left out = 2)			
Selection measure: Mean Squared Error			
	(minimum)	**(+ 1 SE)**	**(+ 2 SE)**
Lambda	0.04974	0.1047	0.1519
Variable		**Coefficients**	
ADTV	0.000342	0.000270	0.000109
CPI	-0.034685	-0.018533	-0.007315
DWX	-0.003563	-1.11E-08	0.000000
EXR	-0.002458	-0.000640	-0.002608
GOLD	1.15E-10	0.000000	-5.80E-06
GTA	0.947115	0.763708	0.708154
GTC	0.863529	0.481011	0.237641
GTD	0.451748	0.455984	0.438940
GTE	-0.225158	-0.169157	-0.125911
GTEM	0.000283	0.000000	0.000000
GTEX	1.08E-10	0.007929	2.66E-06
GTG	0.088256	0.077746	0.064897
GTGG	0.156817	0.143617	0.093736
GTI	-1.25E-09	0.000000	0.000000
GTINV	0.000477	0.000000	0.000000
GTL	0.088745	0.085268	0.049383
GTO	0.076771	0.106583	0.000000
GTP	-0.000803	-3.27E-08	0.000000
GTS	-2.76E-05	4.51E-10	0.000000
INTM	-1.10E-08	-6.88E-08	-2.35E-05
MV	1.90E-09	7.06E-08	1.22E-05
NET	2.13E-10	0.001641	0.000551
NC	0.125438	0.097283	0.095302
NST	0.000724	0.000173	5.20E-05
NTC	0.362562	0.247276	0.222495
NTD	0.000452	0.000728	7.88E-07
TV	-0.001902	0.000000	0.000000
C	-32.86687	-30.73022	-25.53476
d.f.	27	22	19
L1 Norm	36.29873	33.38776	27.58189
R-squared	0.935875	0.857619	0.771405

This table (4) presents a Lasso regression model for nowcasting Syria's GDP, highlighting the impact of increasing regularization (lambda) on coefficient magnitudes and model fit. The optimal lambda value of 0.04974, determined through Leave-One-Out cross-validation, minimizes the Mean Squared Error, indicating the best balance between model complexity and predictive accuracy. As lambda increases, the L1 Norm decreases, reflecting the Lasso model's tendency to shrink coefficients towards zero, promoting sparsity. Variables like GTA (Google Trends Agriculture) and GTC (Google Trends Consumption) maintain high coefficient magnitudes even at higher lambda values, suggesting their strong and consistent influence on GDP. Conversely, coefficients for CPI and EXR shrink significantly with increasing lambda, indicating their potentially less robust relationship with GDP. The declining R-squared values reflect the tradeoff between model complexity and explanatory power, where higher regularization prioritizes generalizability over fitting the training data perfectly. From an economic standpoint, the persistent influence of GTA and GTC suggests the importance of agricultural activity and consumer behavior in driving GDP growth. The sensitivity of CPI and EXR to regularization might point to market distortions or policy interventions that affect their direct influence on GDP.

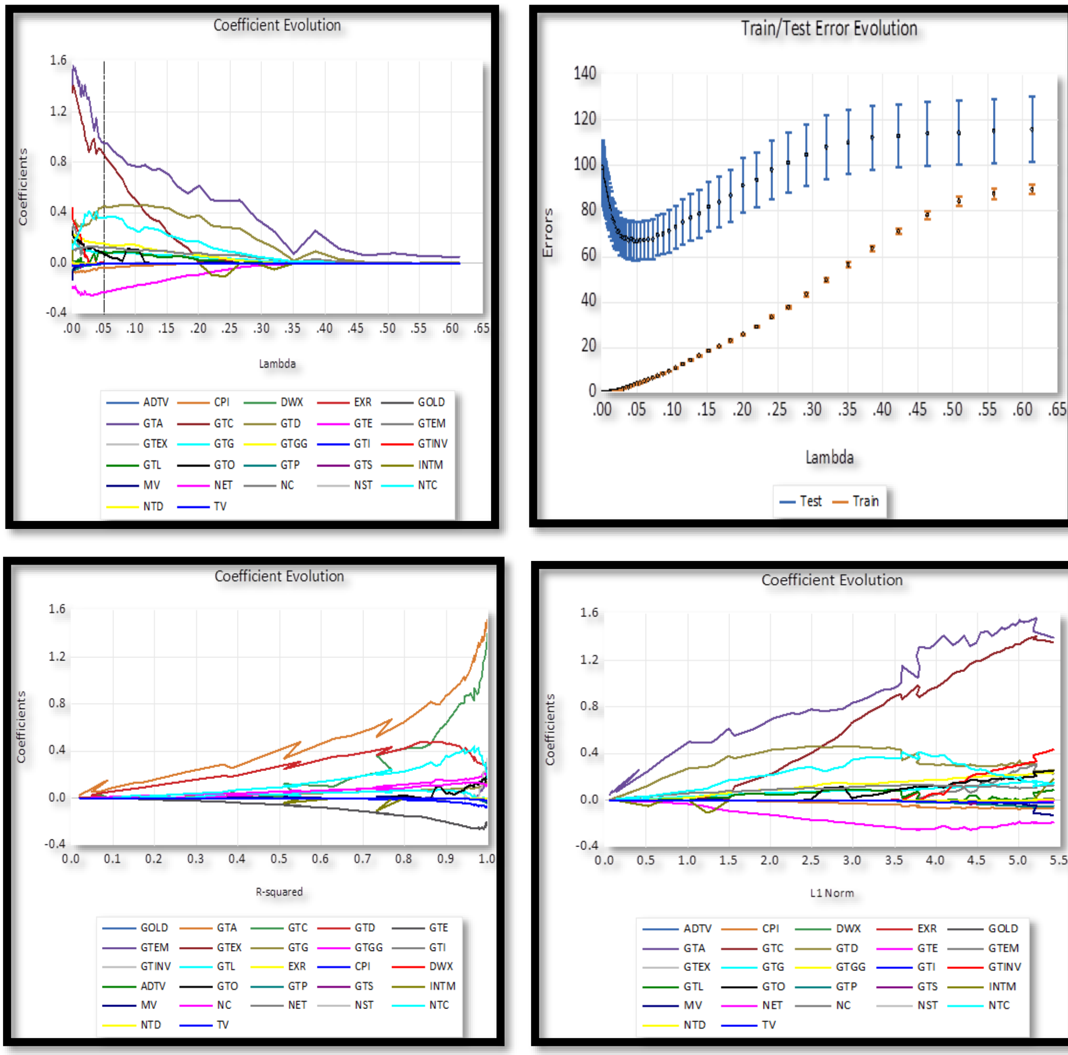

Fig .3. Error and Coefficient Evolution for (Lambda – L1 Norm-R Squared) for Lasso Algorithm

The figure (3) depicting the evolution of training and testing errors, alongside coefficient paths, across varying lambda values in the Lasso model provides a comprehensive view of the regularization process and its impact on model performance. The convergence of low training and testing errors around a lambda of 0.15 highlights the optimal balance between model complexity and generalizability. As lambda increases, the steady decline in training error and initial decrease in testing error, followed by a slight increase with growing variability, illustrate the bias-variance tradeoff inherent in regularization. Coefficient paths reveal that variables like GTA (Google Trends Agriculture) and CPI retain high magnitudes even at higher lambda values, signifying their robust influence on GDP, while other coefficients shrink, indicating their lesser importance. This visualization effectively communicates the dynamics of regularization in the Lasso model, guiding the selection of an appropriate lambda value that minimizes overfitting while maintaining predictive power for robust economic forecasting.

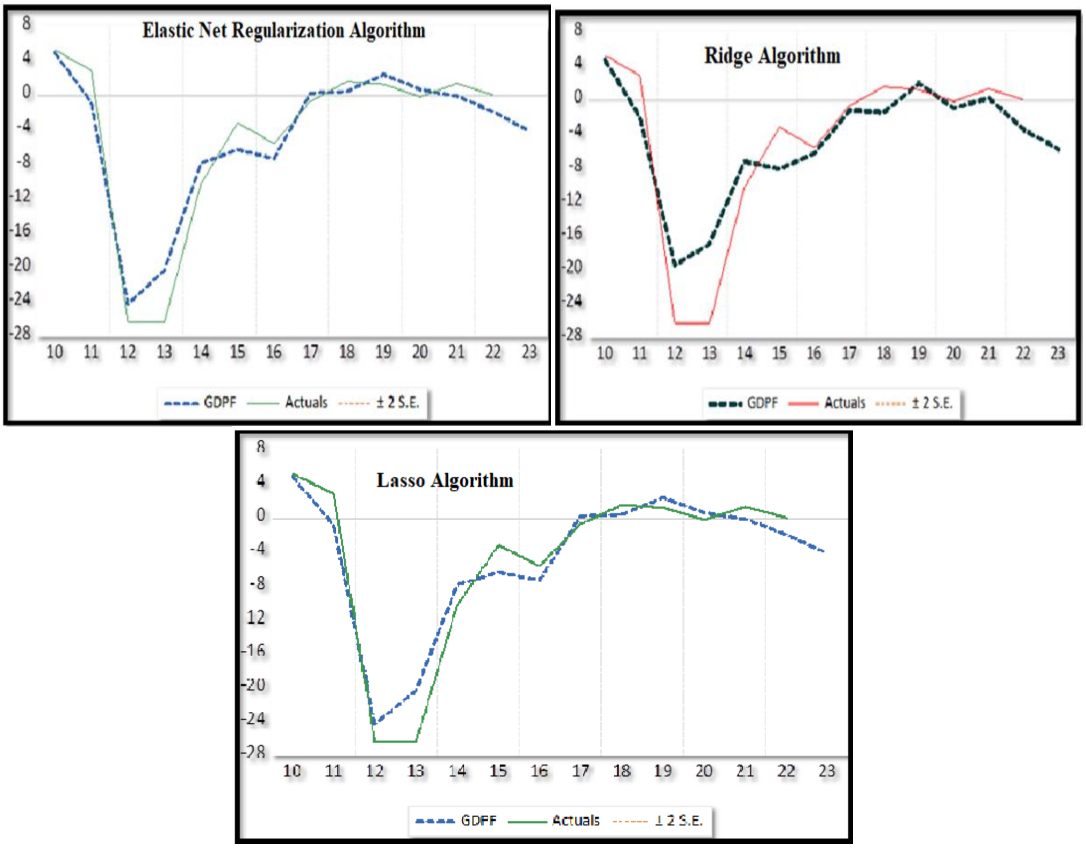

Fig .4. GDP Forecast vs. Actuals" graph

TABLE V. MODELS ACCURACY INDICATORS

Indicators	Model1 (Elastic Net)	Model2 (Ridge)	Model3 (Lasso)
Root Mean Squared Error	0.008887	4.083316	0.004487
Normalized Root Mean Squared Error	1.519924	1.626829	1.018924
Mean Absolute Percentage Error	3.993454	10.44233	2.113421

The accuracy indicators for the three machine learning models used for GDP nowcasting in Syria, namely Elastic Net, Ridge, and Lasso, reveal crucial insights into their performance. While all three models demonstrate reasonable accuracy, the Lasso model emerges as the most effective based on the presented metrics and figures. The Lasso model exhibits the lowest Root Mean Squared Error (RMSE) at 0.004487, signifying the smallest average deviation of its predictions from actual GDP values. It also boasts the lowest Mean Absolute Percentage Error (MAPE) at 2.113421%, suggesting its forecasts are, on average, within 2% of the actual GDP, a commendable level of accuracy in economic forecasting. Although its Normalized Root Mean Squared Error (NRMSE) of 1.018924 is marginally higher than ideal, it's still the lowest among the three models, indicating better consistency in handling data fluctuations. The figures depicting the predicted GDP against the actual values further confirm the Lasso model's superior performance. The line representing Lasso's predictions closely aligns with the actual GDP line, reinforcing its accuracy. Conversely, while the Elastic Net model also demonstrates good accuracy, its slightly higher RMSE and MAPE suggest marginally less precise predictions compared to Lasso. The Ridge model exhibits the lowest accuracy among the three, reflected in its significantly higher RMSE and MAPE values. This disparity in performance stems from the models' inherent characteristics. Lasso's ability to completely eliminate insignificant variables through coefficient shrinkage allows it to create a more parsimonious and efficient model, especially in a setting with limited data and potential multicollinearity, such as the Syrian economic context. From a statistical standpoint, the Lasso model's lower error metrics indicate its superior ability to capture the underlying relationships within the data and generate reliable predictions. Economically, the Lasso model's accuracy holds significant implications for policymakers and analysts, providing a more trustworthy tool for understanding and forecasting Syria's economic trajectory amidst challenging conditions. These insights are invaluable for guiding economic policy decisions, resource allocation, and strategic planning in a context where reliable economic data is scarce.

4. CONCLUSION AND RECOMMENDATIONS

This research aimed to nowcast Syria's Gross Domestic Product (GDP) using a set of macroeconomic and digital indicators, addressing the scarcity of reliable economic data in a context marked by conflict and instability. Utilizing a sample spanning from 2010 to 2022, the study employed three machine learning algorithms – Elastic Net, Ridge, and Lasso – to model GDP dynamics based on a range of variables including market indices, commodity prices, consumer price indices, and internet search trends. While all three models exhibited reasonable accuracy, the Lasso regression model emerged as the most effective in capturing the intricacies of Syria's economic landscape. The Lasso model, characterized by its variable selection and sparsity promotion, achieved the lowest Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE), demonstrating its superior ability to generate accurate and reliable GDP predictions. Statistically, the Lasso model's lower error metrics highlight its capacity to effectively discern the underlying relationships within the data, mitigating overfitting and enhancing the generalizability of its forecasts. From an economic perspective, the Lasso model's accuracy underscores the significant influence of variables like Google Trends Agriculture (GTA) and Google Trends Consumption (GTC) on GDP growth, providing valuable insights for policymakers and analysts seeking to understand and anticipate economic trends in Syria. The findings suggest that consumer behavior and agricultural activity play crucial roles in driving economic growth, particularly in a context where traditional economic indicators may be less reliable due to data limitations. This study advocates for the adoption of machine learning techniques, particularly Lasso regression, as robust and adaptable tools for economic forecasting in data-scarce environments, providing crucial information for evidence-based policymaking, resource allocation, and strategic planning.

Conflicts Of Interest

The authors should pledge that they don't have any conflict of interest in regards of their research. If there are no conflict of interest then authors can declare the following "The authors declare no conflicts of interest".

Funding

The funding section of your journal paper template should provide a concise and transparent declaration of the financial support received to carry out the research presented in your paper.

Acknowledgment

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R735), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

[1] Guma , M. M. Mijwil, B. A. Buruga and M. Abotaleb, "A Comprehensive Review on Cybersecurity Issues and Their Mitigation Measures in FinTech," Iraqi Journal For Computer Science and Mathematics, vol. 5, no. 3, pp. 45-91, 2024.
[2] A. Wald, "Statistical decision functions," The Annals of Mathematical, pp. 165-205, 1949.
[3] C. W. Granger and M. H. Pesaran, "Economic and statistical measures of forecast accuracy," Journal of Forecasting, vol. 19, no. 7, pp. 537-560, 2000.
[4] X. Chen, "Large sample sieve estimation of semi-nonparametric models," Handbook of Econometrics, vol. 6, p. 5549–5632, 2007.
[5] L. Breiman, C. Friedman, J. Stone and A. Olshen, "Classifcation and regression trees," CRC press, 1984.
[6] W. James and C. Stein, "Estimation with quadratic loss," Proceedings Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1961.

[7] S. Ghosh and A. Ranjan, “A machine learning approach to gdp nowcasting: an emerging market experience,” Buletin Ekonomi Moneter dan Perbankan, pp. 33-54, 2023.
[8] Q. Zhang, H. Ni and X. Hao, “Nowcasting Chinese GDP in a data-rich environment: Lessons from machine learning algorithms,” Economic Modelling, vol. 122, pp. 106204-106204, 2023.
[9] N. Malik and B. Agarwal, “Time Series Nowcasting of India’s GDP with Machine Learning,” International Conference on Artificial Intelligence of Things, pp. 1-6, 2022.
[10] M. Gareev and A. Polbin, “Nowcasting Russia’s key macroeconomic variables using machine learning,” Voprosy Economiki, pp. 133-157, 2022.
[11] A. Babii, E. Ghysels and J. Striaukas, “Machine learning time series regressions with an application to nowcasting,” Journal of Business and Economic Statistics, vol. 40, no. 3, pp. 1049-1106, 2022.
[12] L. Bybee, B. Kelly, A. Manela and D. Xiu, “The structure of economic news,” National Bureau of Economic Research, p. Discussion paper, 2020.
[13] Y. Mansour, R. Ismail and K. Alakkari, "Nowcasting As An Approach To Mitigate Economic Uncertainty In Syria," Tishreen University Journal- Economic and Legal Sciences Series, vol. 42, no. 6, pp. 1-20, 2021.
[14] K. Alakkari, “Nowcasting GDP in Parsimony Data Environment Using Bayesian Mixed Frequency VAR (Empirical Evidence from Syria),” Syrian Journal for Science and Innovation, vol. 1, no. 2, pp. 1-21, 2023.
[15] S. Jokubaitis, D. Celov and R. Leipus, “Sparse structures with LASSO through principal components: Forecasting GDP components in the short-run,” International Journal of Forecasting, vol. 37, no. 2, pp. 759-776, 2021.
[16] A. Richardson, T. van Florenstein Mulder and T. Vehbi, “Nowcasting GDP using machine-learning algorithms: A real-time assessment,” International Journal of Forecasting, vol. 37, no. 2, pp. 941-948, 2021.
[17] M. Marcellino and V. Sivec, “Nowcasting GDP growth in a small open economy,” National Institute Economic Review, vol. 256, pp. 127-161, 2021.
[18] O. Darné and A. Charles, “Nowcasting GDP growth using data reduction methods: Evidence for the French economy,” Economics Bulletin, 2020.
[19] J. Ashwin, E. Kalamara and L. Saiz, “Nowcasting euro area GDP with news sentiment: a tale of two crises,” . ECB Working Paper No, pp. 1-52, 2021.
[20] A. Nakazawa, "Constructing GDP Nowcasting Models Using Alternative Data," Bank of Japan, pp. No. 22-E-9, 2022.
[21] J. Tenorio and W. Pérez, "GDP nowcasting with machine learning and unstructured data to Peru," Working Paper No, pp. 1-31, 2023.
[22] T. Hastie, R. Tibshirani and H. Friedman, The elements of statistical learning: data mining, inference, and prediction, vol. 2, New York: springer, 2009, pp. 1-758.
[23] J. Friedman, T. Hastie and R. Tibshirani, "Applications of the lasso and grouped lasso to the estimation of sparse graphical models," Technical report, pp. 1-22, 2010.
[24] C. T. Carneiro, A. P. Rocha and M. L. Ramírez, “Ridge regression ensemble of machine learning models applied to solar and wind forecasting in Brazil and Spain,” Applied Energy, vol. 314, p. 118936, 2022.
[25] P. Patil, A. Rinaldo and R. Tibshirani, “Estimating functionals of the out-of-sample error distribution in high-dimensional ridge regression,” in International Conference on Artificial Intelligence and Statistics , 2022.
[26] S. Reid, R. Tibshirani and J. Friedman, “A study of error variance estimation in lasso regression,” Statistica Sinica, pp. 35-67, 2016.
[27] J. Ranstam and A. Cook, “LASSO regression,” Journal of British Surgery, vol. 105, no. 10, pp. 1348-1348, 2018.
[28] F. Amini and G. Hu, “A two-layer feature selection method using genetic algorithm and elastic net,” Expert Systems with Applications, vol. 166, p. 114072, 2021.
[29] C. Bergmeir, R. Hyndman and B. Koo, “A note on the validity of cross-validation for evaluating autoregressive time series prediction,” Computational Statistics and Data Analysis, vol. 120, pp. 70-83, 2018.
[30] K. Chu and J. Marron, “Comparison of two bandwidth selectors with dependent errors,” Annals of Statistics, vol. 19, no. 4, p. 1906–1918, 1991.