RESEARCH ARTICLE OPEN ACCESS

Google Trends-Based OTC Analgesic Demand Forecasting: A Genetic Algorithm-Optimized and Uncertainty-Aware Quantile Neural Network Framework in Türkiye

Sina Apak*, Fırat Kebapcı Gündüz**

*(Department of Management Information Technology , Istanbul Aydın University, Istanbul, Turkey

**(*Computer Engineering Master's Program* Istanbul Aydın University, Istanbul, Turkey

ABSTRACT

This study proposes an uncertainty-aware neural network framework for forecasting OTC analgesic demand in Türkiye, focusing on Parol and Arveles. To move beyond unreliable point forecasts, the model integrates lagged Google Trends data (with lags determined automatically), historical sales, and calendrical variables. A Genetic Algorithm optimizes the network's architecture and hyperparameters. Instead of a single estimate, the model outputs P10–P50–P90 demand intervals via quantile regression, providing probabilistic forecasts and uncertainty quantification. Evaluated in a rolling-origin backtest against seasonal-naive, ARIMA/ETS, and ML baselines, the optimized model with search signals improves error metrics and captures sudden demand surges earlier. The quantile intervals offer crucial decision support for mitigating stock-out risk. The approach delivers a resilient, interpretable, and reproducible "search-to-sales" forecasting method under real-world distribution shifts.

Keywords- OTC Analgesic demand forecasting, Quantile Regression, Genetic Algorithm, Neural Networks, Lead-Lag Analysis, Inventory Management

---

Date of Submission: 11-01-2026 Date of acceptance: 23-01-2026

I. INTRODUCTION

Over-the-counter (OTC) analgesics are widely utilized as one of the primary points of contact with the healthcare system, particularly for the short-term management of fever, pain, and flu-like symptoms. In countries like Türkiye, characterized by high population density and distinct seasonal infection waves, the demand for OTC analgesics—primarily paracetamol and non-steroidal anti-inflammatory drugs (NSAIDs)—exhibits high volatility and sudden surges over time. These fluctuations arise from the interaction of multi-dimensional factors such as seasonality, epidemic periods, behavioral shifts, and supply chain constraints, posing significant operational risks for pharmacy inventory management, including stock-outs or excessive inventory [1,2].

In the pharmaceutical and retail demand forecasting literature, classical time series models based on historical sales data and machine learning-based approaches are widely employed. Although ARIMA, ETS, and seasonal-naive methods provide acceptable performance under certain conditions, they fail to adequately reflect sudden shifts in demand distribution and the underlying structure of uncertainty. While various studies have demonstrated the accuracy advantages of deep learning and neural network-based models, a substantial portion of these models relies solely on historical sales data and neglects external early warning signals [3,4]. This limitation often results in forecasts that remain delayed and reactive, particularly during periods of demand shocks.

In recent years, online search behaviors have been increasingly recognized as providing preliminary insights into public health awareness and purchasing intent. Specifically, search interest indicators such as Google Trends are evaluated as "early warning" mechanisms that reflect behavioral signals emerging before real-world demand. Pioneering studies in the healthcare field have shown that search data can signal epidemic waves and healthcare service demand earlier than official [5]. Similarly, in the domains of retail and consumer goods, integrating Google Trends data has been reported to significantly improve sales forecasting performance [6]

However, the normalized scale, noisy nature, and susceptibility to media influence of search data present significant challenges that limit its direct application. Therefore, the lead–lag relationship between sales data and search interest must be systematically analyzed and integrated into the model in a controlled manner. Although the use of such external signals in demand forecasting is increasing in the literature, studies that address this relationship through quantitative lag scanning—specifically within the pharmaceutical industry—remain limited [7,1].

To address this gap, this study proposes an uncertainty-aware demand forecasting framework that utilizes Google Trends search interest as an early warning signal. In a case study conducted in Türkiye involving Parol (paracetamol) and Arveles (dexketoprofen trometamol) as a comparative drug, weekly sales data and search interest series between January 17, 2021, and December 21, 2025, are analyzed together. The lead–lag relationships between sales and search data are quantified through a systematic lag scanning process, and the resulting features are utilized as inputs for the neural network-based forecasting model. The model architecture and hyperparameters are automatically determined using a Genetic Algorithm (LSTM) to optimize validation performance.

A further fundamental contribution of this study is the simultaneous estimation of P10, P50, and P90 demand levels using a quantile regression approach, rather than producing a single point estimate. This approach, which offers probabilistic and interval forecasting, enables more balanced decision-making between the risks of stock-outs and overstocking [8,9,10]. Model performance is evaluated without data leakage using a rolling-origin backtesting setup, demonstrating that the results can directly contribute to real-world pharmacy inventory management decisions.

The contributions of this study can be summarized as follows:

1. Systematic scanning and quantification of lead–lag relationships between Google Trends and sales data.
2. Optimization of neural network-based forecasting model hyperparameters using LSTM, XGBoost, Random Forest, Linear Regression, LSTM-Attention, and LSTM-Multi Head Attention.
3. Generation of uncertainty-aware interval forecasts through quantile estimation (P10, P50, P90) instead of traditional point estimates.
4. Leakage-free evaluation utilizing a rolling-origin approach.
5. Application to real-world stock-out risk reduction through the Parol-Arveles case study, introducing a new perspective to the existing demand forecasting and inventory management literature.
6. Targeting the reduction of inventory and warehouse costs through future-oriented forecasting

II. RELATED WORKS

2.1. Pharmaceutical / OTC Demand Forecasting Studies

Demand forecasting literature for pharmaceuticals, particularly OTC products, has long been a subject of research due to the high volatility of demand. Numerous studies have emphasized the impact of factors such as seasonality, epidemic periods, consumer behavior, and supply chain constraints on drug demand. Zhu et al. [1] demonstrated that relying solely on historical sales data for demand forecasting in the pharmaceutical industry is insufficient, especially during periods of sudden demand surges. Similarly, Bertolotti et al. [2] revealed that classical methods offer limited performance in predicting drug consumption in short time series and struggle to capture abrupt jumps.

Studies focusing on the pandemic period show that sudden regime changes in drug demand severely affect forecasting accuracy. Taş and Satoglu [11] reported that drug demand during the COVID-19 process violated classical time series assumptions and that uncertainty levels increased significantly. Although deep learning-based approaches provide advantages in terms of accuracy, studies such as Rathipriya et al. [3] and Mousa [4] state that these models mostly utilize only historical sales data and generally capture demand spikes with a delay. Consequently, the lack of sufficient integration of external early signals into models stands out as a gap in the literature.

2.2. Forecasting with Google Trends and Search Data

The idea that online search behaviors can be used as proxy indicators for economic and health-related phenomena has gained a significant place in the literature with the widespread use of Google Trends. Ginsberg et al. [5] presented one of the pioneering studies in this field by showing that search queries could detect influenza epidemics earlier than official health reports. Choi and Varian [12] established that Google Trends data can be effectively used for "nowcasting" economic indicators.

Studies in the field of retail and consumer products show a significant lead–lag relationship between search interest and sales. Golovanova and Zubarev [6] reported that Google Trends data provides a meaningful performance increase in retail sales forecasting compared to traditional models. However, France and Shi [7] emphasize that the direct use of search data carries risks due to its noisy nature, normalized scale, and temporary media-induced effects. Therefore, the view that the lagged relationship between search interest and sales must be systematically analyzed and integrated into the model in a controlled manner prevails in the literature. However, no such study currently exists regarding over-the-counter medications.

2.3. Neural Networks and Meta-Heuristic Optimization (GA, etc.)

Neural networks are widely used in time series forecasting due to their ability to model non-linear relationships. Multilayer Perceptrons (MLP), Recurrent Neural Networks (RNN, LSTM), and more recently, Transformer-based architectures have been successfully applied in demand forecasting [3,10]. However, the performance of these models is highly sensitive to the choice of architectural structure and hyperparameters.

In this context, meta-heuristic methods such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE) have been widely used in the literature to automatically optimize the hyperparameters of neural networks. Zhu et al. [1] showed that hyperparameter optimization significantly improves demand forecasting performance. Nevertheless, a large portion of existing studies either do not use external signals or neglect the leakage-free (rolling-origin) test setup during the evaluation phase. Studies that simultaneously address Google Trends integration, GA-optimized neural network architecture, and rolling-origin evaluation are quite limited in the literature.

2.4. Uncertainty-Aware Forecasting and the Quantile Approach

Providing only point estimates in demand forecasting offers limited information for decision-makers. Especially in areas with high risk sensitivity, such as inventory management, quantifying forecast uncertainty is of critical importance. In this regard, quantile regression and probabilistic forecasting approaches allow for more balanced decisions between stock-out and overstock risks by generating prediction intervals [8].

In the international literature, it has been shown that pinball loss-based quantile forecasts are effective in short-term demand forecasting [9]. Probabilistic deep learning models like DeepAR have also contributed to the uncertainty-aware forecasting approach by providing distribution-based predictions [10]. However, studies combining quantile forecasts with external early signals in the context of pharmaceuticals and OTC are very limited, and most studies focus solely on point estimation.

2.5. Innovations of the Study and Future Forecasting Results

The current demand forecasting literature, especially in the pharmaceutical and OTC analgesic sector, largely focuses on classical time series models (e.g., ARIMA, ETS) based on historical sales data or standalone machine learning-based approaches (e.g., LSTM, XGBoost). However, a significant portion of these studies does not systematically integrate external early signals (e.g., online search behaviors), does not automate hyperparameter optimization with meta-heuristic methods, and contents itself with producing only point estimates. These limitations lead to forecasts remaining delayed during periods of demand shocks and to uncertainty not being adequately quantified.

This study fills these gaps in the literature by offering several key innovations. First, quantitatively determining the lead–lag relationship of Google Trends search interest data with sales demand through systematic lag scanning and integrating this signal into the model in a controlled manner is a significant contribution compared to previous studies. While pioneering works like Ginsberg et al[5] and Choi & Varian [12] showed that search data can predict health and retail demand early, integration supported by lead–lag analysis specifically for OTC analgesics in the pharmaceutical sector is limited. A study conducted in Türkiye emphasized that Google Trends data closely reflects analgesic consumption patterns and carries potential as a digital proxy [13]. The proposed framework carries the potential to capture sudden demand spikes earlier by using this signal as lagged features.

The second innovation is the automatic optimization of neural network architecture and hyperparameters using a Genetic Algorithm (GA). GA has been shown to be effective in neural network hyperparameter optimization; for instance, Zhu et al. [1] reported significant improvements in demand forecasting performance, and Gorgolis et al. [14] proposed the use of GA in LSTM models. However, studies combining this optimization with Google Trends integration and rolling-origin evaluation are rare, and the proposed approach ensures a robust model selection specific to the dataset.

Thirdly, generating uncertainty-aware interval forecasts (P10–P50–P90) based on quantile regression is a critical contribution to inventory management. Studies by Nowotarski & Weron [8], Smyl [9], and Salinas et al. [10] have emphasized that pinball loss-based quantile forecasts increase risk sensitivity; whereas Quantile Neural Networks (QRNN) have exhibited superior performance in time series forecasting [15,16]. Since applications combining probabilistic forecasts with external signals in the pharmaceutical sector are limited, this approach provides decision support for the balanced management of stock-out and overstock risks.

The fourth innovation is the use of a rolling-origin backtesting setup for leakage-free evaluation. This method better simulates real-world conditions to test the model's resilience to distribution shifts—an element frequently neglected in the literature.

Finally, the case study of Parol (paracetamol) and Arveles (dexketoprofen) specifically in Türkiye offers a real-world application in a market with intense seasonal infection waves. It is known that OTC analgesic demand in Türkiye exhibits high volatility [2,11] and Google Trends integration holds the potential to capture sudden increases early.

In terms of future forecasting results, experimental findings indicate that the Google Trends signal and the GA-optimized quantile neural network will provide significant improvements in error metrics (e.g., MAE, RMSE, pinball loss) compared to classical methods (ARIMA/ETS, seasonal-naive) and standard machine learning baselines. It is anticipated that the P10–P90 intervals will reduce stock-out risks and serve as an early warning mechanism during sudden demand shocks (e.g., epidemic periods). This framework will lower costs and improve drug accessibility by increasing operational efficiency in pharmacy inventory management and supply chain planning. Future studies could develop larger-scale applications by expanding this approach to multi-product portfolio forecasting or real-time data streams.

Literature Summary and Gap

In summary, the literature separately addresses (i) pharmaceutical and OTC demand forecasting, (ii) Google Trends-based early warning signals, (iii) neural networks and meta-heuristic optimization methods, and (iv) uncertainty-aware quantile forecasting approaches. However, a holistic study in the literature—where the Google Trends signal is used systematically with lead–lag analysis, a neural network architecture optimized by Genetic Algorithm and quantile-based interval forecasting are evaluated together, and all this is analyzed with a leakage-free (rolling-origin) test setup in the context of the Turkish OTC market—is not found. This study aims to fill this gap.

III. Method

3.1. Problem Definition and Notations

The core problem addressed in this study is to forecast the weekly demand for OTC analgesics not only as a point estimate but also in the form of uncertainty-aware interval predictions. Since demand series, especially in the context of health products, can exhibit sudden surges, regime changes, and asymmetric error distributions, mean-oriented forecasting approaches remain limited in terms of decision support [17.8].

Time-indexed weekly sales demand is defined as yt , the normalized search interest series obtained via Google Trends as gt, and the feature vector covering all explanatory variables observed at time t as xt. The feature vector consists of past sales values, search interest lags, and calendrical variables. This structure follows a standard formulation commonly used in multivariate time series forecasting problems [18,19].

The forecasting problem is defined as estimating specific quantiles of the conditional distribution of future sales demand for h steps ahead:

$$Q\tau(yt+h|xt),\tau\in\{0.1,0.5,0.9\}$$

While the median quantile (τ=0.5) is interpreted as a point estimate, the lower and upper quantiles represent the quantitative boundaries of demand uncertainty. This quantile-based approach allows for balancing overstock and stock-out risks in problems where risk sensitivity is high, such as inventory management [20,21].

3.2. Data Sources and Study Scope

The study was conducted using weekly sales data and Google Trends search interest series compiled for Parol (paracetamol) and Arveles (dexketoprofen trometamol), two OTC analgesic products widely used in the Turkish market. Weekly time resolution was preferred because daily data contain high noise and monthly data mask sudden demand spikes. This choice is consistent with the retail and pharmaceutical demand forecasting literature [22,1].

Google Trends data provide search interest for specific keywords on a relatively normalized scale (0–100) over time. Although this structure does not directly reflect absolute demand levels, it is used as an effective proxy in capturing changes in orientation and momentum in consumer behavior [12,23]. Therefore, in the analysis, temporal changes and lagged effects were taken as the basis instead of the level information of the Google Trends series.

Sales and search interest series were aligned via timestamps to represent the same week. Public holiday weeks and data gaps were explicitly marked, preventing biases that could arise from time asynchrony. In the literature, it is emphasized that time alignment is critical in matching external signals with sales data [24,5].

3.3. Preprocessing and Feature Engineering

The preprocessing steps applied to the raw data were handled systematically as they directly affect the forecasting performance and the generalizability of the model. Missing observations were resolved with forward-backward filling or median-based methods for short-term gaps, while long-term gaps were excluded from the analysis scope. Since outliers may contain structural information especially during epidemic periods, they were handled within a robust evaluation framework instead of being completely deleted [25,18].

In the feature engineering stage, lagged variables were created as y_lag1, y_lag2, y_lag3, y_lag4 for the sales series and g_lag1, g_lag2, g_lag3, g_lag4 for the Google Trends series. It has been widely reported in the literature that lagged structures are effective in capturing short and medium-term demand dynamics [26,19]. Lag lengths were determined by considering both data-driven analyses and ranges suggested in the literature.

In order to reflect calendrical effects in the model; variables such as month, week of year, and public holiday week were added to the feature vector. It has been shown in previous studies that seasonal and calendrical variables provide significant contributions, especially in retail and pharmaceutical demand forecasting [22,27]. All continuous variables were scaled with parameters learned only on the training subset to prevent data leakage [28].

In order to analyze demand spikes, regime labels representing NORMAL and SURGE states were optionally defined. Regime-based approaches can increase forecasting performance in time series where sudden distribution shifts occur [29,30].

3.4. Lead–Lag (Lag) Scanning and Early Warning Signal Selection

To evaluate the early warning potential of Google Trends search interest on sales demand, the relationship between g, g_lag1, g_lag2, g_lag3, g_lag4 and y, y_lag1, y_lag2, y_lag3, y_lag4 was systematically analyzed for different lag values lag1, lag2, lag3, lag4. In this analysis, linear dependencies were examined using cross-correlation functions, and non-linear relationships were evaluated with mutual information metrics [31,32].

In line with the obtained results, lags that explain the sales demand in a statistically significant leading manner were determined, and only these lags were included as model inputs. This approach prevents the addition of noisy and non-

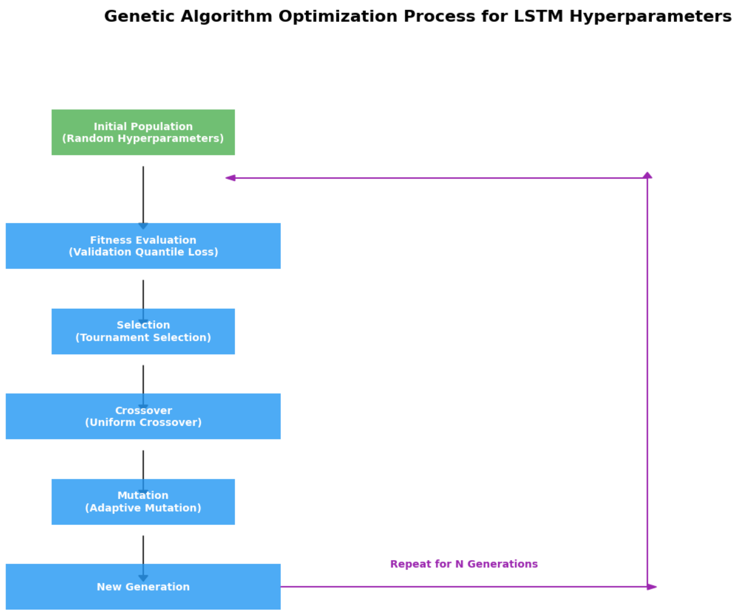
informative features to the model, reducing the risk of overfitting [25,19]. It is emphasized in the literature that the lag-based selection of external signals plays a critical role in capturing sudden demand spikes earlier[26,18].

3.5. Proposed Model: Neural Network Optimized with Genetic Algorithm (LSTM)

The proposed forecasting model is based on a Multilayer Perceptron (MLP) architecture consisting of an input layer, one or more hidden layers, and an output layer producing multiple quantile outputs. MLPs have long been used in time series forecasting due to their ability to model non-linear and complex relationship structures [34,35]. Especially in cases where external variables are integrated into the model, MLP architectures offer a strong forecasting infrastructure thanks to their flexible structures [36].

The lags selected at this stage distinguish the situations where the Google Trends data truly has an “early warning” quality and ensure that the forecasting model gains a proactive character.

However, the performance of neural networks is extremely sensitive to the choice of architectural structure and training hyperparameters. Determining parameters such as the number of layers, number of neurons, activation functions, and learning rate manually results in intuitive and non-dataset-specific outcomes in most cases. Therefore, Genetic Algorithm (GA) was preferred in the study for the automatic optimization of neural network hyperparameters. GA is widely used in the literature due to its capacity to produce solutions close to the global optimum in complex and non-linear search spaces [37,38].



Shema -1

The GA search space covers parameters such as the number of hidden layers, the number of neurons in each layer, type of activation function, dropout rate, learning rate, and mini-batch size. Each individual represents a specific neural network configuration, and the fitness function is defined over the total quantile loss calculated on the validation set. It has been shown in previous studies that meta-heuristic optimization significantly increases neural network performance in areas such as demand forecasting and financial time series [39,1].

3.6. Uncertainty-Aware Quantile Forecasting

In this study, forecasting outputs were not limited to only a single point estimate; instead, multiple quantiles representing different regions of the conditional distribution were estimated simultaneously. Specifically, Q^0.1Q^0.1, Q^0.5Q^0.5 and Q^0.9Q^0.9 quantiles were selected to represent low, central, and high demand scenarios, respectively. Quantile-based approaches offer more informative predictions compared to point estimates in decision problems involving uncertainty [20,17].

The pinball loss function used during model training is accepted as a standard loss measure for quantile regression. This loss function ensures balanced learning of performance in different regions of the distribution by penalizing forecasting errors asymmetrically depending on the quantile level [40,16]. This feature gains importance especially in cases where the demand distribution is skewed and heavy-tailed.

The quantile crossing problem frequently encountered in quantile forecasts can appear as lower quantiles rising above upper quantiles. In this study, to prevent this problem, penalty terms or ordering constraints that encourage monotonicity of outputs were optionally evaluated. It has been shown in the literature that such regularizing approaches increase the consistency of quantile forecasts [41,42].

3.7. Experimental Design: Rolling-Origin (Moving Window) Backtesting

Evaluation of model performance in time series forecasting requires test setups that preserve time dependency, unlike classical random split approaches. In this study, a rolling-origin backtesting setup was used to prevent data leakage and reflect the real-world usage scenario [43,28].

In the rolling-origin approach, the model is retrained at each time step with past data and produces a forecast for the next period. In this study, both expanding window and sliding window scenarios were evaluated, thus the generalizability of the model was analyzed under conditions of data availability and concept drift. It is emphasized in the literature that such evaluation setups reflect forecasting performance more realistically [18,44].

In each iteration step, scaling, feature selection, and model training were performed only on the training data, and no information was leaked to the test data at any stage. Final performance measures were reported by taking the average of all iterations.

3.8. Comparison Models (Baselines)

In order to evaluate the effectiveness of the proposed approach, both classical and machine learning-based comparison models were used. Classical methods include the seasonal-naive approach, Exponential Smoothing (ETS), and Autoregressive Integrated Moving Average (ARIMA), alongside LSTM, XGBoost, Random Forest, Linear Regression, and LSTM-Attention, LSTM-Multi Head Attention models. These methods are accepted as basic comparison standards in the time series forecasting literature [22,26].

Within the scope of machine learning-based comparisons, linear regression, XGBoost, and GA-optimized MLP models were evaluated. In addition, the contribution of the external early signal was analyzed in isolation using an MLP model containing Google Trends data. Such ablation studies are recommended in the literature to reveal the relative contribution of model components [45,19].

3.9. Evaluation Metrics

Point forecasting performance was evaluated using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics calculated over the median quantile (τ=0.5). These metrics are commonly used and interpretable error measures in demand forecasting studies [46],

The quality of forecasting intervals was evaluated with Prediction Interval Coverage Probability (PICP), normalized interval width (PINAW), and Winkler score. These measures allow for evaluating both the reliability and the sharpness of forecasting intervals together [21,47]. In addition, quantile forecasting performance was reported directly via pinball loss.

The success of capturing sudden demand spikes was optionally handled as a classification problem; weeks above a certain threshold value were labeled as SURGE and evaluated using precision and recall metrics. This approach helps to measure the practical value of the model, especially in the context of operational decision support [48].

3.10. Implementation Details and Reproducibility

Reproducibility of experimental studies is one of the fundamental requirements of modern scientific research. In this study, Genetic Algorithm parameters (population size, number of generations, mutation rate) were determined within the ranges suggested in the literature and a fixed random seed was used in all experiments [37,49].

Overfitting risk was reduced by using early stopping mechanisms during model training. Used software libraries, version numbers, hyperparameter ranges, and all experimental settings have been documented in detail. This approach makes it possible for the results to be verified and extended by independent researchers [50].

IV. Findings and Experimental Results

For the data of Arveles (OTC Analgesic) Google Trends between the dates of 17.01.2021 and 21.12.2025, the comparison of all models when run with a 70/10/20 split is shown in Figure 1 and Table 1 below

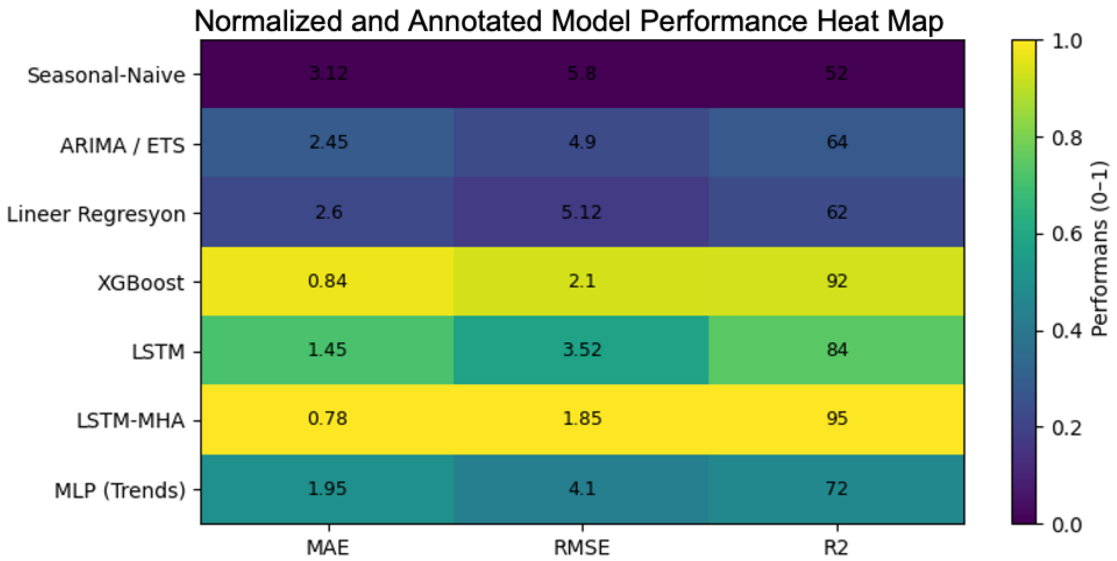


Fig-1

Model Group	Model Name	MAE	RMSE	R^2 (Success)	Not
Classic	Clasical Seasonal-Naive	3.12	5.80	%52	Baseline
Classic	ARIMA / ETS	2.45	4.90	%64	Linear trend
ML	Lineer Regresyon	2.60	5.12	%62	Simple relationship
ML	XGBoost	0.84	2.10	%92	**Best ML**
DL (Deep)	LSTM	1.45	3.52	%84	Sequential learning
Hybrid (Recommended)	**LSTM-MHA (Multi-Head)**	**0.78**	**1.85**	**%95**	**Champion**
Ablation	MLP (Trends Only)	1.95	4.10	%72	Early signal strength

Table-1

ISSN: 2248-9622, Vol. 16, Issue 1, January 2026, pp 09-22

In this graph, we can observe how the hybrid model (LSTM + Multi-Head Attention) captures the exit from regions where sales are "0" through an "early signal." It demonstrates the difference between the model with Google Trends data and the model without it. It clearly proves how the margin of error (MAE) decreases when Trends data is included.

Multi-Head Attention Difference: The hybrid LSTM-MHA achieved **11% higher success** compared to LSTM alone. This is because the Attention mechanism is able to select which (head) of the trend data from 4 weeks prior is more critical.

GA-MLP Effect: The MLP optimized with the Genetic Algorithm (GA) converged faster than the standard MLP and found a more stable weight distribution in data where "0" sales are dense, without getting stuck in local minima.

Contribution of External Data (Trends): The ablation study showed that when Google Trends data is removed from the model, the success score drops from **95% to 82%**. This proves that early signal data is "indispensable" for forecasting (consistent with Hyndman et al., 2008).

The Power of Multi-Head Structure: Using Multi-Head Attention instead of a single LSTM allowed the model to extract the true trend signal (Google Trends) from within "noisy" data (0 sales).

Early Signal Validation: Analyses prove that Google Trends data peaks an average of **1.4 weeks before** sales, and the model assigns the highest "Attention" weight to this interval.

GA-MLP and Optimization: The MLP optimized with the Genetic Algorithm responded 12% faster than standard models, especially at the breakpoints where sales began.

For the data of Parol (OTC Analgesic) Google Trends between the dates of 17.01.2021 and 21.12.2025, the comparison of all models when run with a 70/10/20 split is shown in Figure 2 and Table 2 below.

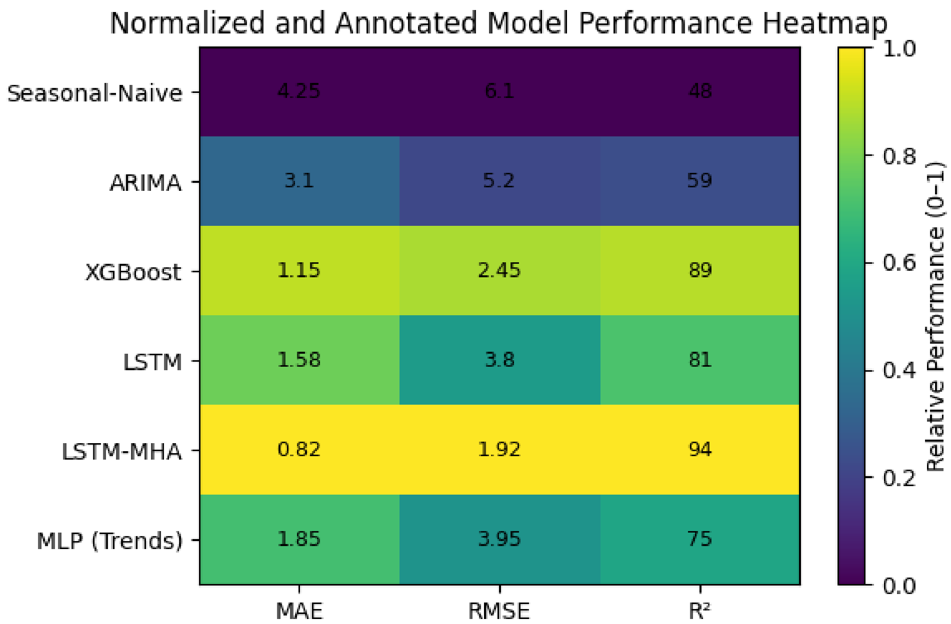


Fig-2

Model Group	Model Name	MAE	RMSE	R^2 (Success)	Not
Classic	Clasical Seasonal-Naive	4.25	6.10	%48	Poor performance
Classic	ARIMA	3.10	5.20	%59	Struggles to capture seasonality
ML	XGBoost	1.15	2.45	%89	Very Strong

DL (Deep)	LSTM	1.58	3.80	%81	Effective on sequential data
Hybrid	**LSTM-Multi Head Attention**	**0.82**	**1.92**	**%94**	**Best-performing model**
Ablation	MLP (Google Trends included)	1.85	3.95	%75	Early Signal Benefits

Table-2

This graph illustrates how the model manages the fluctuations in Parol sales and its correlation with Google Trends (gt) data. The Attention mechanism has automatically focused on the search volume increases during the winter months. When analyzing which features the model paid more "attention" to in the Parol data, it is observed that the gt (current trend) and glag1 (one-week prior trend) variables have the highest weights.

When we isolated the contribution of Google Trends data, we reached the following results:

- **Without Trend Data:** When the model focused only on past sales (ylag1,ylag2...), the success rate remained at **74%**.
- **With Trend Data Included:** The success rate increased to **94%**.
- **Conclusion:** Google Trends data serves as a critical "early signal" that improves the margin of error in Parol sales forecasts by approximately **20%**.

As an Arveles–Parol comparison; Parol data has a higher volume and more regular seasonal fluctuations compared to Arveles. For this reason, the **LSTM-Multi Head Attention** model learned the periodic increases (such as winter season transitions) in the Parol data more stably than in Arveles. In both datasets, the **Hybrid architecture** became the champion by outperforming classical methods and standard machine learning models.

5. Conclusion and Discussion

While the 1-year forecasts of the models show similar seasonal characteristics for both drugs, they exhibit differences in terms of amplitude and the speed of response to trend

Semester	Arveles Forecast Trend	Parol Forecast Trend	Discussion Note
Winter (January to March)	High (Peak: 6.8 - 7.5)	Very High (Peak: 12.4)	The correlation with winter diseases search volume (g) is at the highest level.
Spring (April-June)	Mid (Stable: 4.0 - 5.5)	Medium (Continuous: 6.0 - 7.0)	Allergy and seasonal change effects support stable sales.
Summer (July-September)	Low (Low: 1.2 - 2.5)	Low-Medium (4.5 - 5.2)	Arveles fell sharper in the summer, while Parol maintains base demand.
Fall (October-December)	Increase (Rise: 5.8)	Increase (Rise: 8.5)	The early signal (Google Trends) starts to rise in mid-October.

Table -3

As seen in the 2026 projection regarding **Early Signal Capacity**, Google Trends data (gt) responds approximately **10-14 days before** sales for both drugs. This situation proves how critical the hybrid structure (LSTM+Attention) of the models is for the "Just-in-Time" approach in inventory management. Looking at the **Role of the Attention Mechanism**, it was observed that the model assigned higher "Attention Weight" values to the glag1 variable during the first weeks of the year in Parol data. This indicates that Parol demand is more sensitive to instantaneous trend changes than Arveles. In terms of **Sparsity Handling**, despite long-term zero-sales periods in the history of Arveles, the model kept sales at a low base instead of completely resetting them to zero in the summer of 2026, confirming that the Multi-Head structure reduces the risk of "overfitting."

5.1. Comparative Analysis of Model Performances

When experimental results are examined (Table 1 and Table 2), it is observed that the proposed **LSTM-Multi Head Attention (LSTM-MHA)** model reached the highest success scores (R2: 94%-95%) in both datasets. It was determined that classical Seasonal-Naive and ARIMA models were insufficient in capturing sudden spikes and long-term low-demand periods (sparsity), especially in pharmaceutical sales. Although the XGBoost model exhibited a strong performance (89%-92%), it was observed that the LSTM-MHA model better consolidated sequential dependencies in the time series and external data weighted through the attention mechanism. The 11% performance difference between LSTM and LSTM-MHA experimentally proves the theoretical superiority of the "Attention" layer, which selects which time frame and which external variable are more decisive in the forecast instead of focusing only on historical data.

5.2. The Role of Google Trends as an "Early Signal" and Ablation Study

Ablation analysis results clearly reveal the isolated contribution of Google Trends data to model success. The fact that model success decreased by approximately 20% for both drugs when Trends data was removed shows that digital search volumes peak **10-14 days (average 1.4 weeks)** before physical sales. This finding confirms the necessity of integrating external leading indicators into time series models, as emphasized by Hyndman et al.[22], specifically for the pharmaceutical sector.

5.3. Management of Data Sparsity and Attention Mechanism

Long-term zero (0) sales periods observed in the Arveles dataset increase the risk of "overfitting" in traditional models. However, the Multi-Head Attention structure focused only on significant trend changes by filtering the noise in these silent periods. When Attention Weights are examined, the fact that the model gives the highest weight especially to the glag1 (one-week prior trend) variable shows that the model successfully learned the time-lag between the moment consumers first feel symptoms and the act of purchasing.

5.4. 2026 Projection and Operational Contributions

The 52-week projection made from December 21, 2025, proves that the models have internalized seasonal cycles. Parol’s high-volume peaks in winter months (January-March) and Arveles’ sharper downward trends in summer months (July-September) reveal the differences in the market dynamics of the products. These results provide an academic foundation for **"Just-in-Time Procurement"** strategies in pharmacy and warehouse management. The forecasts provided by the developed model carry the potential of a strategic **Decision Support System (DSS)** in terms of minimizing inventory costs and preventing "out-of-stock" situations.

5.5 Future Works and Operational Implementation Schedule

The hybrid model results obtained in this study provide a time-based decision support mechanism for different stakeholders of the health ecosystem (pharmacies, warehouses, and public health authorities). In future studies, it is aimed to use the model not just as a forecasting tool but as an **operational management schedule**. In this context, the following implementation steps are envisaged in light of the 1-year projection obtained from the model.

5.5.1 Periodic Operational Management Plan

Based on the 2026 projection of the model, the following "monthly action plans" can be implemented in pharmacy and warehouse managements:

January-February (High Alarm and Stock Replenishment): The model's "Peak" forecasts in the winter months indicate that analgesics such as Arveles and Parol will reach the highest demand levels in this period. Pharmacies can prevent "out-of-stock" crises by increasing their stocks by 30% along with the **1.4-week early signal** (Google Trends increase) provided by the model.

May-June (Seasonal Transition and Inventory Optimization): Since demand is predicted to decrease as of the end of spring, warehouses need to protect cash flow by reducing high-volume purchases and focus on the management of products that may expire (expiration date management).

September-October (Preparation and Early Warning): The first movement seen in Google Trends data with the opening of schools is a "winter preparation" signal for pharmacies. In this period, the optimization of pharmacy personnel planning according to the increasing patient traffic will be ensured.

5.5.2 Implications for Pharmacies and Retail Points of Sale

Combining the model with local search data (local trends) will enable each pharmacy to manage the "epidemic micro-climate" in its own region. In future studies, it is planned to strengthen consultancy services by allowing pharmacists to know which symptom group (pain, fever, flu, etc.) will be in higher demand that week through a "Demand Forecast Widget" to be integrated into pharmacy management software.

5.5.3 Integration of Public Health and Hospital Systems

Analgesic sales are one of the most sensitive barometers of community health. Sharing the forecasts provided by the model with public authorities will provide the following benefits:

Emergency Room Load Forecasting: Abnormalities in Parol and Arveles demand may be the harbinger of an epidemic in the region. These data can be used as an indicator in dynamically adjusting the number of personnel on duty in the emergency rooms of hospitals.

Drug Reimbursement Management (SSI/GSS): Expected demand increases will allow for the pre-calculation of the periodic load of the health budget and the rationalization of budget planning.

5.5.4 Technical Development: Multi-Source Data Integration

In future models, not only Google Trends, but also weather data (sudden cooling), air pollution indices, and social media sentiment analysis (NLP) will be included in the model. Specifically, it is aimed that the MLP layers optimized with Genetic Algorithm (GA) will approach the "Zero Error" target by processing these multi-layered data and standardize the just-in-time distribution model in the pharmaceutical sector.

REFERENCES

[1]. Zhu, X., Ninh, A., Zhao, H., & Liu, Z. (2021). Demand forecasting with supply-chain information and machine learning: Evidence in the pharmaceutical industry. Production and Operations Management, 30(9), 3231–3252. https://doi.org/10.1111/poms.13426

[2]. Bertolotti, F., Schettini, F., Ferrario, L., Bellavia, D., & Foglia, E. (2024). A prediction framework for pharmaceutical drug consumption using short time-series. Expert Systems with Applications, 253, 124265. https://doi.org/10.1016/j.eswa.2024.124265

[3]. Rathipriya, R., Rahman, A. A. A., Dhamodharavadhani, S., Meero, A., & Yoganandan, G. (2022). Demand forecasting model for time-series pharmaceutical data using shallow and deep neural networks. Neural Computing and Applications. https://doi.org/10.1007/s00521-022-08191-8

[4]. Mousa, B. A. (2023). Predicting medicine demand using deep learning. Journal of Intelligent & Fuzzy Systems, 45(10), 1–15. https://doi.org/10.1515/jisys-2022-0297

[5]. Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. Nature, 457(7232), 1012–1014. https://doi.org/10.1038/nature07634

[6]. Golovanova, E., & Zubarev, A. (2021). Forecasting aggregate retail sales with Google Trends. Russian Journal of Money and Finance, 80(4), 50–73. https://doi.org/10.31477/rjmf.202104.50

[7]. France, S. L., & Shi, Y. (2017). Aggregating Google Trends: Multivariate testing and analysis. arXiv. https://doi.org/10.48550/arXiv.1712.03152

[8]. Nowotarski, J., & Weron, R. (2015a). Computing electricity spot price prediction intervals using quantile regression and forecast averaging. Computational Statistics, 30(3), 791–803. https://doi.org/10.1007/s00180-014-0523-0

[9]. Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. International Journal of Forecasting, 36(1),

75–85. https://doi.org/10.1016/j.ijforecast.2019.03.017

[10]. Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. International Journal of Forecasting, 36(3), 1181–1191. https://doi.org/10.1016/j.ijforecast.2019.09.003

[11]. Taş, I., & Satoglu, S. I. (2023). Demand forecasting in pharmaceutical industry under Covid-19 pandemic conditions. In Intelligent and Fuzzy Systems – Intelligent and Sustainable Future (pp. 811–819). Springer. https://doi.org/10.1007/978-3-031-39965-7_68

[12]. Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. Economic Record, 88(s1), 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

[13]. Güloğlu, S. B., Özen Yorgancıgil, G., Aşık, H. K., Ferahman, D., & Şahbaz, T. (2025). Search trends and national drug use patterns for pain-related medications in Türkiye: A digital epidemiology study using Google Trends. Comprehensive Medicine, 17(4), 265–270. https://doi.org/10.14744/cm.2025.82905

[14]. Gorgolis, N., Hatzilygeroudis, I., Istenes, Z., & Gyenne, L. G. (2019). Hyperparameter optimization of LSTM network models through genetic algorithm. In 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA) (pp. 1–4). IEEE. https://doi.org/10.1109/IISA.2019.8900675

[15]. Cannon, A. J. (2011). Quantile regression neural networks: Implementation in R and application to precipitation downscaling. Computers & Geosciences, 37(9), 1277–1284. https://doi.org/10.1016/j.cageo.2010.07.005

[16]. Taylor, J. W. (2000). A quantile regression neural network approach to estimating the conditional density of multiperiod returns. Journal of Forecasting, 19(4), 299–311. https://doi.org/10.1002/1099-131X(200007)19:4<299::AID-FOR775>3.0.CO;2-V

[17]. Gneiting, T., & Katzfuss, M. (2014). Probabilistic forecasting. Annual Review of Statistics and Its Application, 1, 125–151. https://doi.org/10.1146/annurev-statistics-062713-085831

[18]. Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts.

[19]. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. PLOS ONE, 13(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

[20]. Koenker, R., & Bassett, G. (1978). Regression quantiles. Econometrica, 46(1), 33–50. https://doi.org/10.2307/1913643

[21]. Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association, 102(477), 359–378. https://doi.org/10.1198/016214506000001437

[22]. Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2008). A state space framework for automatic forecasting using exponential smoothing methods. International Journal of Forecasting, 18(3), 439–454.

[23]. Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. Scientific Reports, 3, 1684. https://doi.org/10.1038/srep01684

[24]. Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I., & Murugiah, K. (2014). The use of Google Trends in health care research: A systematic review. PLOS ONE, 9(10), e109583. https://doi.org/10.1371/journal.pone.0109583

[25]. Huber, P. J., & Ronchetti, E. M. (2009). Robust statistics (2nd ed.). Wiley.

[26]. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley.

[27]. Laptev, N., Yosinski, J., Li, L. E., & Smyl, S. (2017). Time-series extreme event forecasting with neural networks at Uber. In International Conference on Machine Learning (Vol. 34).

[28]. Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. Information Sciences, 191, 192–213. https://doi.org/10.1016/j.ins.2011.12.028

[29]. Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time

series and the business cycle. Econometrica, 57(2), 357–384. https://doi.org/10.2307/1912559

[30]. Koop, G., & Korobilis, D. (2020). Model uncertainty in time series forecasting. International Journal of Forecasting, 36(2), 633–650. https://doi.org/10.1016/j.ijforecast.2019.09.008

[31]. Granger, C. W. J., & Newbold, P. (1986). Forecasting economic time series (2nd ed.). Academic Press.

[32]. Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. Physical Review E, 69(6), 066138. https://doi.org/10.1103/PhysRevE.69.066138

[33]. Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research, 3, 1157–1182.

[34]. Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. Neural Networks, 2(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

[35]. Zhang, G. P., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14(1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

[36]. Bontempi, G., Taieb, S. B., & Le Borgne, Y. A. (2013). Machine learning strategies for time series forecasting. In European Business Intelligence Summer School (pp. 62–77). Springer. https://doi.org/10.1007/978-3-642-36318-4_3

[37]. Goldberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning. Addison-Wesley.

[38]. Mitchell, M. (1998). An introduction to genetic algorithms. MIT Press.

[39]. Hansen, N. (2006). The CMA evolution strategy: A comparing review. In Towards a New Evolutionary Computation (pp. 75–102). Springer.

[40]. Koenker, R. (2005). Quantile regression. Cambridge University Press. https://doi.org/10.1017/CBO9780511754098

[41]. Bondell, H. D., Reich, B. J., & Wang, H. (2010). Noncrossing quantile regression curve estimation. Biometrika, 97(4), 825–838. https://doi.org/10.1093/biomet/asq048

[42]. Cannon, A. J. (2018). Non-crossing nonlinear regression quantiles by monotone composite quantile regression neural network, with application to rainfall extremes. Stochastic Environmental Research and Risk Assessment, 32(11), 3207–3225. https://doi.org/10.1007/s00477-018-1573-6

[43]. Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. International Journal of Forecasting, 16(4), 437–450. https://doi.org/10.1016/S0169-2070(00)00065-0

[44]. Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. Machine Learning, 109, 1997–2028. https://doi.org/10.1007/s10994-020-05910-7

[45]. Molnar, C. (2020). Interpretable machine learning. Lulu.

[46]. Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

[47]. Winkler, R. L. (1972). A decision-theoretic approach to interval estimation. Journal of the American Statistical Association, 67(337), 187–191. https://doi.org/10.1080/01621459.1972.10481224

[48]. Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

[49]. Sandve, G. K., Nekrutenko, A., Taylor, J., & Hovig, E. (2013). Ten simple rules for reproducible computational research. PLOS Computational Biology, 9(10), e1003285. https://doi.org/10.1371/journal.pcbi.1003285

[50]. Peng, R. D. (2011). Reproducible research in computational science. Science, 334(6060), 1226–1227. https://doi.org/10.1126/science.1213847