

## Article

# Applying Machine Learning and Statistical Forecasting Methods for Enhancing Pharmaceutical Sales Predictions

Konstantinos P. Fourkiotis and Athanasios Tsadiras \* 

School of Economics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; fourkiok@econ.auth.gr

\* Correspondence: tsadiras@econ.auth.gr

**Abstract:** In today's evolving global world, the pharmaceutical sector faces an emerging challenge, which is the rapid surge of the global population and the consequent growth in drug production demands. Recognizing this, our study explores the urgent need to strengthen pharmaceutical production capacities, ensuring drugs are allocated and stored strategically to meet diverse regional and demographic needs. Summarizing our key findings, our research focuses on the promising area of drug demand forecasting using artificial intelligence (AI) and machine learning (ML) techniques to enhance predictions in the pharmaceutical field. Supplied with a rich dataset from Kaggle spanning 600,000 sales records from a singular pharmacy, our study embarks on a thorough exploration of univariate time series analysis. Here, we pair conventional analytical tools such as ARIMA with advanced methodologies like LSTM neural networks, all with a singular vision: refining the precision of our sales. Venturing deeper, our data underwent categorisation and were segmented into eight clusters premised on the ATC Anatomical Therapeutic Chemical (ATC) Classification System framework. This segmentation unravels the evident influence of seasonality on drug sales. The analysis not only highlights the effectiveness of machine learning models but also illuminates the remarkable success of XGBoost. This algorithm outperformed traditional models, achieving the lowest MAPE values: 17.89% for M01AB (anti-inflammatory and antirheumatic products, non-steroids, acetic acid derivatives, and related substances), 16.92% for M01AE (anti-inflammatory and antirheumatic products, non-steroids, and propionic acid derivatives), 17.98% for N02BA (analgesics, antipyretics, and anilides), and 16.05% for N02BE (analgesics, antipyretics, pyrazolones, and anilides). XGBoost further demonstrated exceptional precision with the lowest MSE scores: 28.8 for M01AB, 1518.56 for N02BE, and 350.84 for N05C (hypnotics and sedatives). Additionally, the Seasonal Naïve model recorded an MSE of 49.19 for M01AE, while the Single Exponential Smoothing model showed an MSE of 7.19 for N05B. These findings underscore the strengths derived from employing a diverse range of approaches within the forecasting series. In summary, our research accentuates the significance of leveraging machine learning techniques to derive valuable insights for pharmaceutical companies. By applying the power of these methods, companies can optimize their production, storage, distribution, and marketing practices.

**Keywords:** sales forecasting; machine learning; time series analysis; pharmaceutical industry; seasonality effects; Anatomical Therapeutic Chemical (ATC) Classification System



**Citation:** Fourkiotis, K.P.; Tsadiras, A. Applying Machine Learning and Statistical Forecasting Methods for Enhancing Pharmaceutical Sales Predictions. *Forecasting* **2024**, *6*, 170–186. <https://doi.org/10.3390/forecast6010010>

Academic Editors: Konstantinos Nikolopoulos and Luigi Grossi

Received: 31 December 2023

Revised: 11 February 2024

Accepted: 14 February 2024

Published: 16 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In this transformative era that we are going by, the pharmaceutical industry emerges as the edge of global healthcare. As we start, our analysis indicates that the global pharmaceutical sector's revenues surged to an estimated USD 1.4 trillion in 2021, with projections suggesting a potential doubling by 2030 [1]. This abrupt growth underscores the necessity for accurate sales forecasting, especially considering challenges posed by global events, notably the COVID-19 pandemic during the 2019–2021 period [2].

Historically, the pharmaceutical industry has depended on traditional forecasting models [3]. Yet, these models, focused on historical data and basic statistical methods, often

fall short of capturing the intricate dynamics of drug sales. Factors such as seasonality, influenced by factors from weather patterns to global health crises, highlight the need for a more agile and adaptive forecasting approach [4].

Our study aims to leverage artificial intelligence, specifically machine learning, to analyse a dataset of 600,000 transactions from 2014 to 2019. We use traditional methods and modern techniques like Facebook Prophet, LSTM Neural Networks, and XGBoost to create accurate sales forecasts.

Our dataset is categorized into eight groups comprising 57 different products based on the Anatomical Therapeutic Chemical (ATC) Classification System [5]. Our study provides insights into pharmaceutical sales across various ATC categories, including M01AB (acetic acid derivatives related to anti-inflammatory products) and N05B (anxiolytic drugs), among others.

Clarifying the specific objectives of our research, our central aim is to precisely forecast sales for subsequent years, drawing on data from 2014 to 2019. By analysing historical sales data, we aim to anticipate the cyclical illnesses that manifest throughout the year and ensure we are adequately stocked with the appropriate pharmaceutical products to address these conditions [6]. Our goals extend beyond accurate forecasting to adeptly regulate inventory within our outlets. This involves curtailing expenses linked to excessive stock or potential stock shortages and judiciously directing our marketing endeavours, discerning products poised for a surge in demand, enabling the astute allocation of marketing assets and crafting of nuanced promotional campaigns [7].

The aim of the article is to present a robust methodology, detailing its strengths, limitations, and pivotal role in advancing the field. We will compare traditional forecasting methods with advanced machine learning techniques to achieve more reliable predictions. This improvement in forecasting will aid the industry in optimizing the supply chain process, reducing waste, and fostering greater consumer trust and loyalty.

Our article follows a structured approach: Section 2, the Literature Review, precisely outlines the existing body of knowledge, detailing the selection process of relevant works and exemplifying the research questions driving our analysis. Moving forward, Section 3, the Methodology section, provides a comprehensive overview of the research approach, including the selection criteria for studies and the identification of research objectives. In Section 4, consisting of the Research Results and Discussion, we delve into the findings derived from our analysis, addressing challenges, limitations, and emerging trends while effectively responding to the research questions posed. Finally, Section 5, the Conclusions, Proposals, and Recommendations, synthesizes key insights, proposes applicable recommendations, and outlines avenues for future research, thus offering a comprehensive conclusion to our study.

## 2. Literature Review

The pharmaceutical industry, while it is standing at the heart of global healthcare, relies heavily on forecasting and shows that its process is pivotal to shaping managerial decisions in areas like operations, finance, marketing, and intricate models with respect to anticipating future trends [8]. To address the challenges of traditional forecasting, a new generation of advanced algorithms has been developed in recent years.

Berrar's paper describes the naive Bayes classifier, emphasizing its foundation on Bayes' Theorem. This approach is highlighted due to its ability to classify data based on the conditional probability of an event, assuming independence between predictors, and this classifier is praised for its simplicity and effectiveness as it can provide robust and insightful predictive analyses in various fields [9]. Given the specific pharmacological focus of our study, we find that naive Bayes, while effective in various fields, might not fully capture the complex correlations present in pharmaceutical sales dynamics. Our methodology builds upon this understanding and explores alternative models that align more closely with the complexities of drug sales forecasting in our domain. Aburto and Weber present the Seasonal Naive theory, which is a refined forecasting approach centred

on specific time intervals, and it enhances the predictive model by comparing sales data from equivalent days in previous weeks, allowing for a more nuanced analysis [10].

In the study by Mancuso et al., we find important insights about how ARIMA, exponential smoothing models, and the ANN method compare, including the use of combination models. The research points to an interesting conclusion that combined forecasting methods, although not widely used, lead to better predictions [11].

Pamungkas researched exponential smoothing methods and explained that if a drug's sales have been steady, Single Exponential Smoothing would be used, but if there is a noticeable trend, double exponential smoothing comes into play. For drugs with sales that both rise and fall seasonally, the Triple Exponential Smoothing, which is also known as the Holt–Winters equation, is employed [12]. In a similar direction, IMECE and BEYCA explored the Holt–Winters model, analysing the trend, level, and seasonality in forecasting [13].

In the research paper by Sushama Rani Dutta, ARIMA was employed as a time series model to analyse past data for predicting future trends, leveraging its capability to use lagged moving averages to smooth the time series data, making it particularly suitable for sales predictions and technical analysis [14].

While traditional methods have their benefits, newer techniques have been created to address the complexities of modern pharmaceutical forecasting. Zunic and his team presented Facebook's Prophet model, a tool adept at capturing complex sales patterns ranging from daily to yearly rhythms [15]. Emphasizing the potential of neural networks, Bandara highlighted the capabilities of long short-term memory (LSTM) networks. These networks, a type of recurrent neural network, are designed to handle long-term dependencies, effectively remembering and retrieving information and data over extended periods [16].

In the research conducted by Yuxuan Han, the LSTM model's effectiveness in pharmaceutical sales forecasting was notably demonstrated. This advanced approach outperformed traditional models like ARIMA in capturing complex data patterns over time, showcasing its potential to significantly improve sales forecasting in the pharmaceutical industry [17].

XGBoost is recognized for its efficiency and superior performance, utilizing both exact and approximate algorithms to find optimal tree splits, along with features such as handling sparse data and out-of-core computation, making it a powerful and scalable tree boosting system [18]. Given the superior performance of XGBoost as demonstrated in our study, we have seamlessly integrated this algorithm into our methodology, showcasing its effectiveness in pharmaceutical sales forecasting.

Seasonality, integral to pharmaceutical sales, dictates that drug demands oscillate with the seasons. This underscores the need for forecasting models to adeptly incorporate these seasonal nuances [19].

The *Best Practice Guide* by the BioPhorum Operations Group reveals the necessity of accurate forecasting, transparent communication, and strategic alignment in improving supply chain efficiency, which are vital for ensuring consistent patient supply and effectively responding to the dynamic demands of the biopharmaceutical market [20].

In the paper by Moosivand, Rajabzadeh Ghatari, and Rasekh, the challenges of forecasting and supply chain planning in pharmaceutical manufacturing are being explored. They identify specific challenges such as demand variability, regulatory compliance, and the need for precise coordination between different stages of the supply chain. The research underscores the necessity of advanced forecasting techniques and strategic planning in mitigating these challenges, thereby enhancing overall supply chain effectiveness [21].

Pharmaceutical companies confront significant challenges in managing their supply chains. Moosivand et al. [21] examined these issues, proposing strategies for improvement, including collaborative supplier relationships and technology investment. Similarly, Yani and Aamer [22] focused on demand forecasting accuracy in the pharmaceutical supply chain, offering insights into machine learning techniques for enhanced precision.

Using more in-depth analysis, Zhu et al. [23] address this challenge by proposing a novel demand forecasting framework that leverages advanced machine learning models.

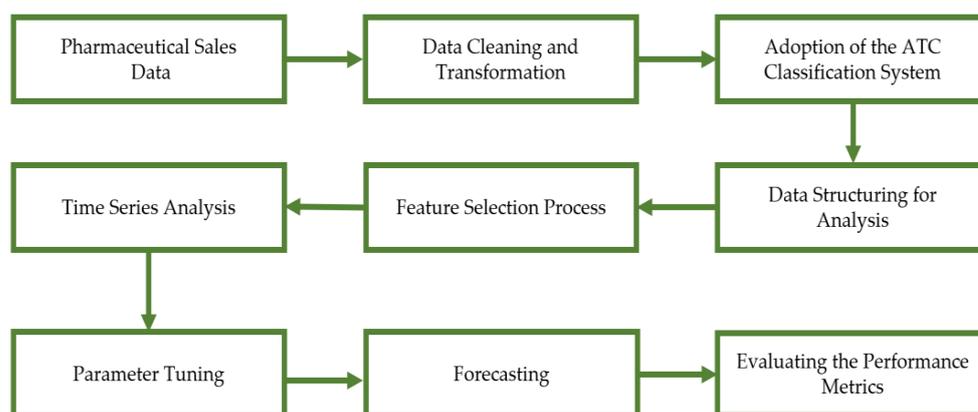
Their approach involves cross-series training using time series data from multiple products and incorporating downstream inventory information and supply chain structure data.

In the study conducted by Zdravković et al. [24], the effectiveness of univariate time series analysis in forecasting pharmaceutical products' sales is highlighted, emphasizing its value in strategic planning for pharmacies.

KPMG's "Pharma 2030: From evolution to revolution" report delves into the innovative impact of AI and big data analytics on pharmaceutical industry forecasting, since it emphasizes that these technologies will augment demand forecasting accuracy and resource allocation efficiency, significantly improving supply chain management. In conclusion, the report showcases the potential for these advanced technologies to reform traditional practices, pointing to a future dominated by data-driven decision making in the industry [25].

### 3. Methodology

Our approach, as shown in Figure 1, starts with a dataset of 600,000 pharmaceutical sales records from 2014 to 2019, and these are organized into eight ATC groups. We have simplified the data for analysis across eight drug categories. Our method involves three main steps: cleaning and preparing the data, analyzing the sales over time, and forecasting future sales.



**Figure 1.** Machine learning workflow.

#### 3.1. Feature Engineering and Data Preparation

In this section, we will thoroughly explore the data analysis and database utilized for programming and developing both traditional and advanced machine learning algorithms. This step is pivotal for ensuring the suitability and validity of our analysis, where we have meticulously selected options from a diverse array of ten representative forecasting methods (statistical, boosting, or based on neural networks) that are specifically tailored to address the intricate dynamics of pharmaceutical sales. The processes detailed here are vital components of our methodology, laying robust groundwork for accurate and dependable forecasting in our study.

##### 3.1.1. Data Cleaning and Transformation

Our objective was to ensure the integrity and quality of our dataset. To achieve this, we implemented a range of techniques to address and rectify issues such as missing values, outliers, and inconsistencies. In collaboration with pharmacists, we tackled these challenges using a systematic approach.

For missing values, we employed a variety of imputation methods, including mean, median, mode, and K-nearest neighbors (KNN) imputation, selecting the most appropriate technique based on the nature of the data [26]. To identify and correct outliers, we consulted with pharmacy experts to discern whether an anomaly was a true outlier or a data entry

error. Statistical methods like Z-scores were then applied to either adjust or remove these outliers [27].

Inconsistencies were addressed through meticulous data standardization, ensuring uniform metrics throughout the dataset. This process also involved textual data cleaning to harmonize categorical data and logical checks to eliminate any illogical or contradictory entries.

Finally, the refined dataset underwent a rigorous validation process, with pharmacy experts reviewing and confirming the accuracy and consistency of our data transformations. This thorough approach to data cleaning and transformation forms the cornerstone of our analysis, guaranteeing a reliable and robust dataset for our forecasting endeavors.

### 3.1.2. Adoption of the ATC Classification System

We employed the Anatomical Therapeutic Chemical (ATC) Classification System developed by the World Health Organization to systematically categorize the pharmaceutical products sold in the pharmacy over a six-year period. This strategic classification was essential not only for organizing the data but also for ensuring the interpretability and validity of our results. We condensed a diverse range of 57 different pharmaceutical products into eight distinct categories based on the ATC drug classification system. This approach was pivotal in facilitating a more focused and meaningful analysis. The ATC system's relevance is particularly evident in drug utilization research, as it allows for a comprehensive understanding of drug sales patterns and therapeutic uses. For example, in our study, we analyzed drug groups like M01AB, which includes acetic acid derivatives used for treating pain and inflammation, and R03, comprising medications for obstructive airway diseases such as asthma and chronic obstructive pulmonary disease (COPD) [28].

This systematic categorization under the ATC framework not only streamlined our data but also enhanced the clarity and accuracy of our conclusions, thereby significantly contributing to the depth and relevance of our research in pharmaceutical forecasting.

### 3.1.3. Data Structuring for Analysis

In the next data structuring phase, we transitioned from an hourly to a weekly data framework based on the decision, which was driven by several key considerations, each playing a vital role in enhancing the accuracy and applicability of our analysis [29]:

1. Smoothing out daily variability;
2. Capturing significant trends;
3. Aligning with operational cycles;
4. Enhancing forecasting models;
5. Enabling comparative analysis.

### 3.1.4. Feature Selection Process

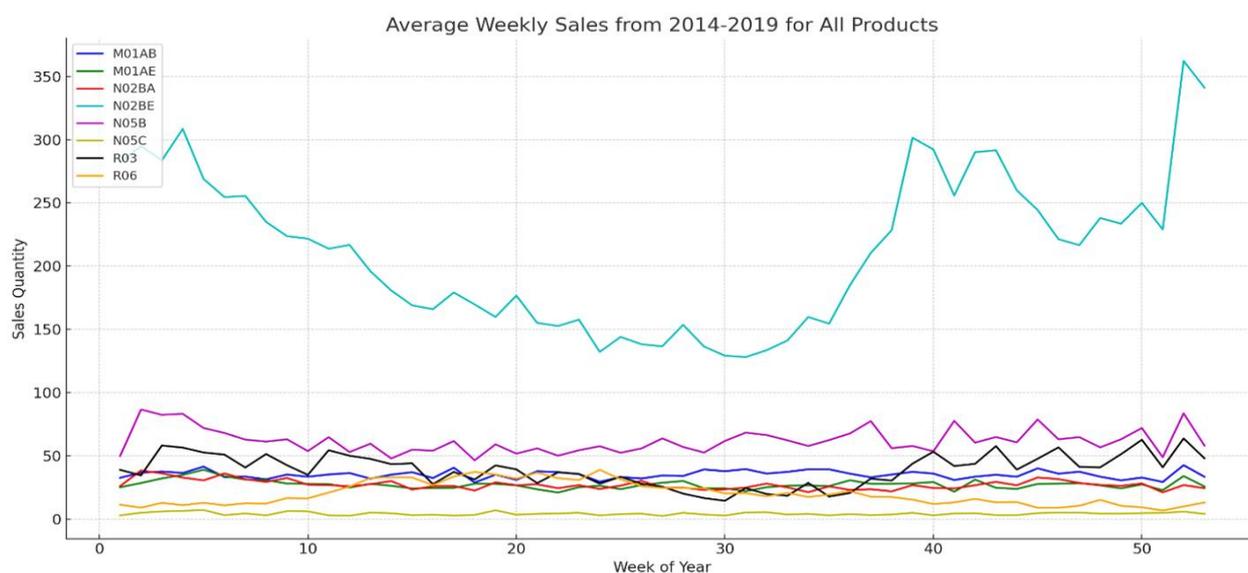
A pivotal component of our methodology involved the careful selection of features for our machine learning models. This process integrated both advanced analytical techniques and domain-specific expertise. Key factors considered included historical sales data, seasonal trends, and category-specific characteristics, ensuring a comprehensive approach to our analysis [30]. A noteworthy aspect of our dataset selection, focusing on the years 2014 to 2019, was the deliberate exclusion of data from 2020 and onwards. By limiting our dataset to this pre-pandemic period, we aimed to maintain the integrity and consistency of our analysis, enabling more accurate and reliable forecasting in a more stable and predictable market environment.

In summary, the feature engineering and data preparation phase was a meticulously executed process that is essential for setting a strong foundation for our machine learning workflow. This phase ensured that the data were clean and relevant and structured in a way that supports effective analysis and forecasting.

### 3.2. Time Series Analysis

Having prepared the data, we proceeded to visually analyse the sales patterns through various visualizations. A series of vivid visualizations was created using the capabilities of Python libraries, including Pandas for data orchestration and Matplotlib and Seaborn for graphics. Time series charts have evolved into timelines that represent a continuum of sales over a specific period of an era. These timelines explained the cyclicity and variation of the sales records by reflecting their temporal nuances. Improved trend analysis was performed by combining daily sales with their 30- and 365-day equivalents. This holistic view highlights both fleeting and persistent sales trends and provides a panoramic view of sales dynamics. Before displaying, the dates feature was converted to date/time format to ensure chronological accuracy. This layered visualization approach not only represented the distribution of data but also recorded its origin in time, providing a comprehensive overview of the sales environment [31].

Figure 2 depicts weekly drug sales from 2014 to 2019, revealing a notable surge in year-end sales, particularly in the last quarter, contrasting with quieter sales at the year's start. Detailed analysis shows increased sales of drugs like M01AB and M01AE during winter for conditions like arthritis and flu, with demand shifting as winter transitions to spring [32]. The sales of N02BA and N02BE, commonly prescribed for headaches and migraines, are peaking, and the reason is probably because of the change in seasons [33]. The sales of psycholeptics, particularly N05B and N05C, remain consistent throughout the year, indicating consistent demand, especially among people struggling with anxiety and sleepless nights [34]. Drugs such as R03 and R06, which are often in demand for asthma and allergies, also show a consistent sales pattern, although with less seasonal fluctuations. Prescription rates for R03 (drugs for obstructive airway diseases) drugs may be higher during the spring and fall months when pollen counts are typically higher [35], while prescription rates for R06 (antihistamines for systemic use) drugs may also be higher during the spring and fall months, as well as during the summer months, when insect bites are more common [36].



**Figure 2.** Average weekly sales from 2014 to 2019 for all products.

The statistics plot in Figure 3 illustrates the average monthly prescriptions for each product category from 2014 to 2019. Notably, prescription rates peak in the fourth quarter and decline in the first, and this is possibly due to seasonal illness patterns. Anti-inflammatory and antirheumatic drugs consistently emerge as the most prescribed.

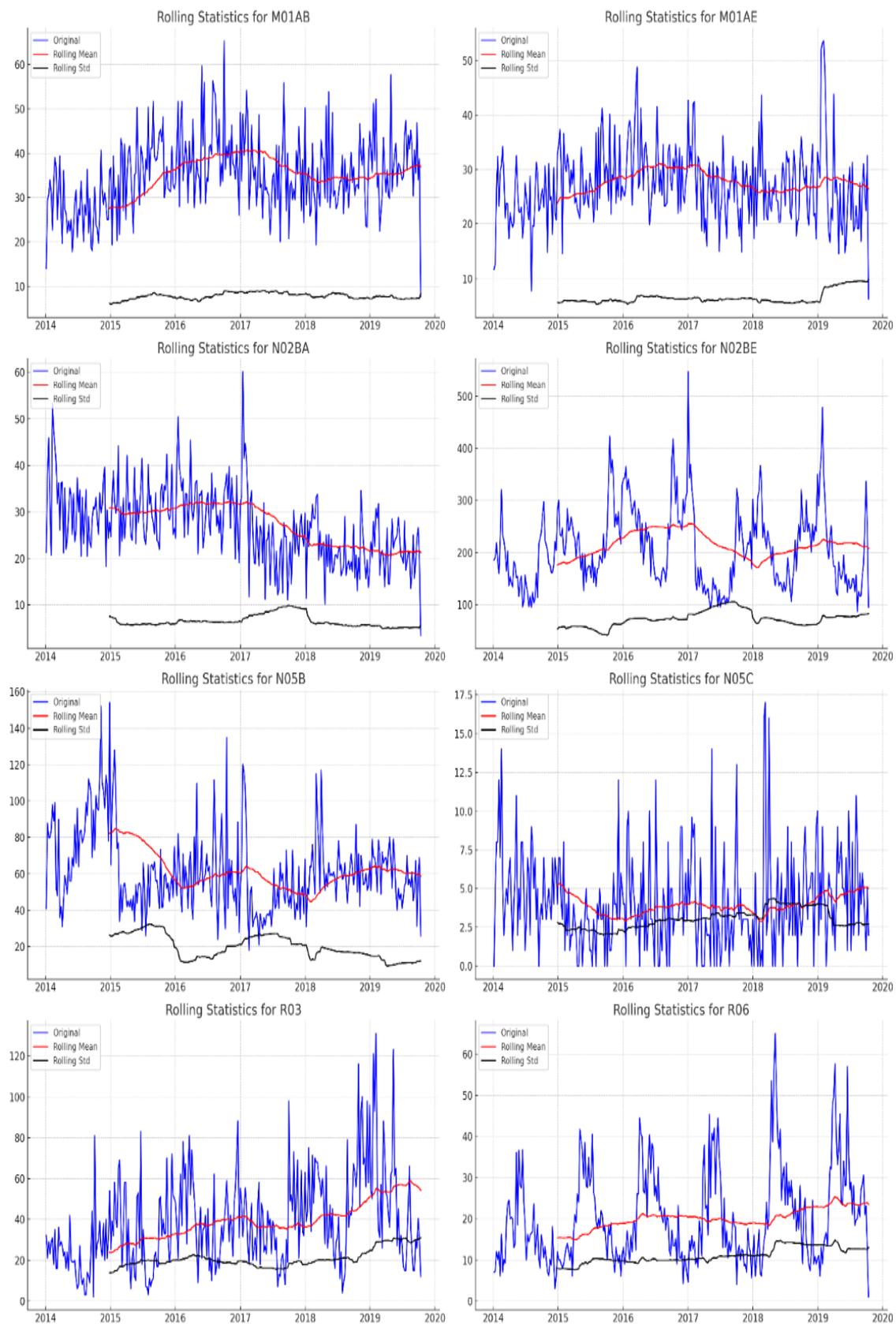
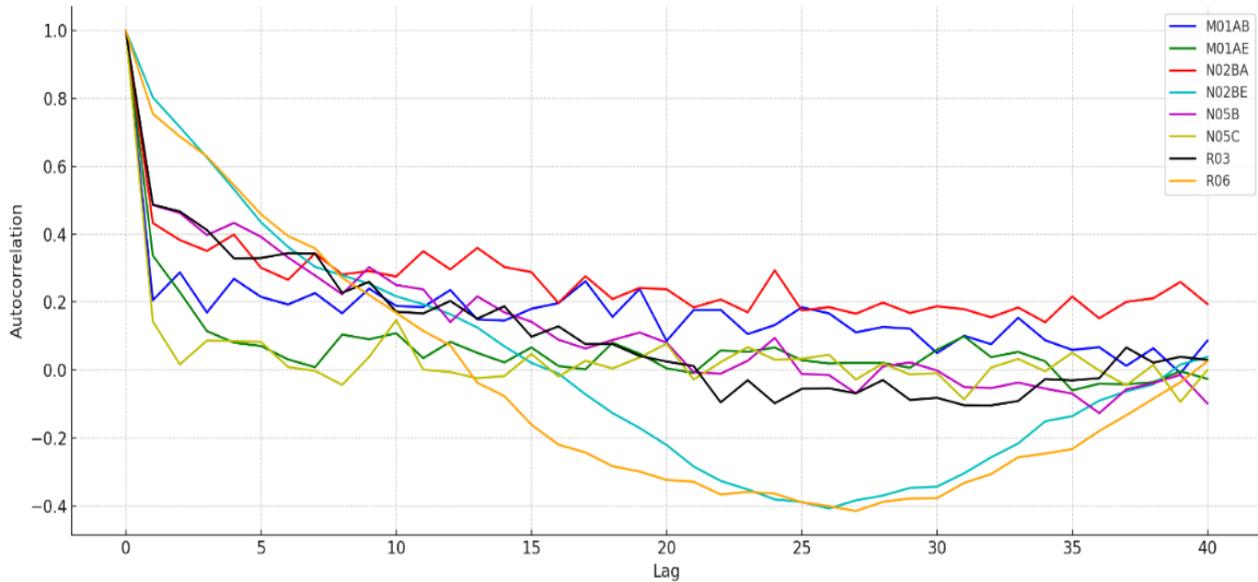


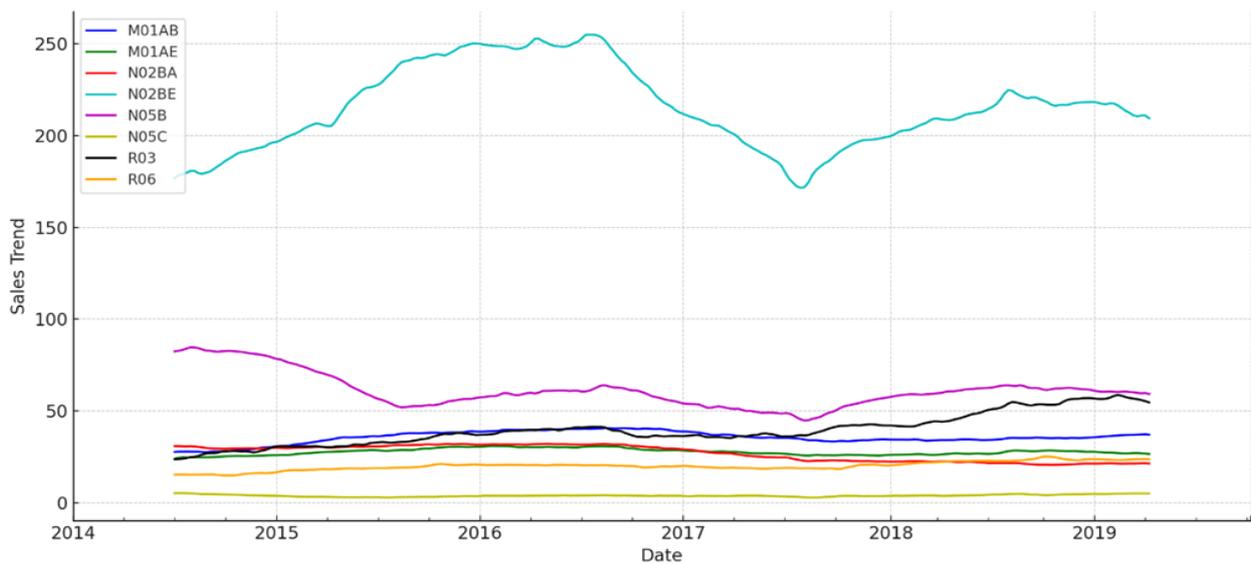
Figure 3. Statistics for all products.

To strengthen the basis for accurate predictions, a thorough assessment of the stationarity of the data was carried out. In Figure 4, the autocorrelation function (ACF) was implemented [37].



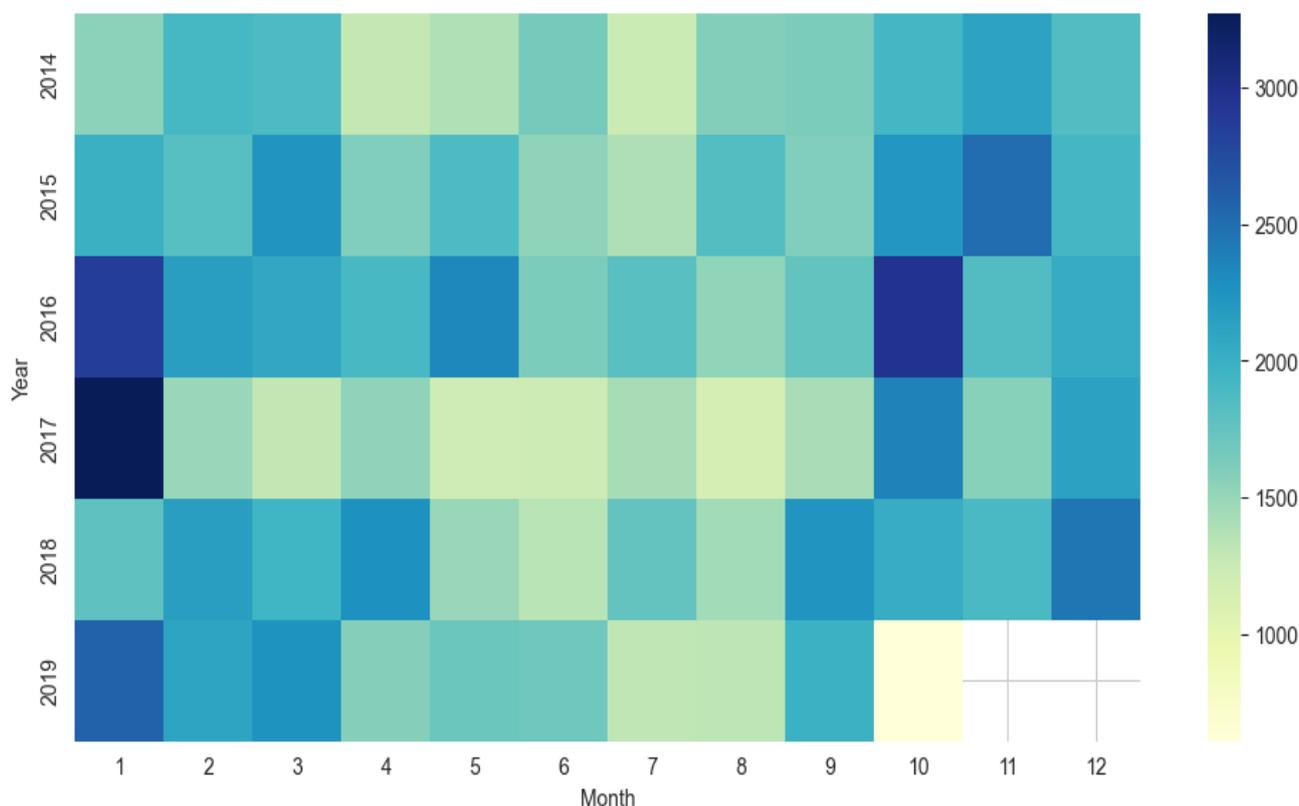
**Figure 4.** Autocorrelation function (ACF) of listed products.

To decode the intricate dynamics of pharmaceutical sales across various ATC drug categories, in Figure 5, a robust time series analysis was embarked upon. Initially, the data underwent seasonal decomposition, adopting both additive examples. This revealed not only the raw data but also unmasked underlying trends, seasonal fluctuations, and anomalies or residuals that might skew interpretations.



**Figure 5.** Additive decomposition trends for listed products.

A deeper dive into the data was facilitated through diverse visualizations, such as heatmaps, in Figure 6 for instance, depicting a vivid picture of sales patterns across months and years [38].



**Figure 6.** Month-year heatmap for total sales across all products.

The heatmap in Figure 6, capturing data from 2014 to 2019, intuitively displays sales fluctuations, with darker tones indicating higher sales. Seasonal trends are evident, such as increased sales in winter months and a decrease in warmer months.

### 3.3. Forecasting

In our economic research's time series analysis, we thoroughly examined the contrast between conventional forecasting approaches such as Naïve, Seasonal Naïve, and ARIMA, juxtaposed with contemporary methodologies like LSTM Neural Network and XGBoost. This comparison aimed to evaluate the effectiveness and predictive capabilities of these different techniques across various forecasting scenarios and timeframes, ensuring a comprehensive understanding of their respective strengths and limitations. Classical models like ARIMA and exponential smoothing rooted in econometric principles offer a foundational basis, while contemporary tools such as Facebook Prophet and LSTM neural networks provide advanced computational depth, enabling us to capture intricate seasonal nuances and interpret extended data sequences with precision and insight.

In our study, refining forecasting models involved optimizing hyperparameters through a grid search. This step notably enhanced drug demand prediction accuracy, decreasing the chance of shortages. Ultimately, this method guarantees a steady supply of pharmaceuticals, elevating customer satisfaction and loyalty.

In Table 1, a detailed table of optimized hyperparameters for a range of forecasting algorithms is showcased, including Naïve, Seasonal Naïve, exponential smoothing, ARIMA, Facebook Prophet, and advanced models like LSTM and XGBoost, each with specified parameter values for precision in pharmaceutical sales forecasting. These parameters encompass aspects like test sizes, weights, and alpha/beta/gamma ranges, ensuring a tailored approach to model tuning and evaluation across various product categories.

**Table 1.** Optimized hyperparameter values for algorithms.

Algorithm	Hyper Parameters	Values
Naïve	None	Not Applicable
Seasonal Naïve	Test Size Weights	52 weeks [0.4, 0.3, 0.2, 0.1] (more recent years have more weight)
Single Exponential Smoothing	Test Size Alpha Values Range Optimal Alpha	52 weeks 0.01 to 1 (100 values) Varies per product (determined through cross-validation)
Double Exponential Smoothing	Alpha Values Range Beta Values Range Optimal Alpha Optimal Beta	0.01 to 1 (10 values) 0.01 to 1 (10 values) Varies per product (determined through optimization) Varies per product (determined through optimization)
Triple Exponential Smoothing	Alpha Values Range Beta Values Range Gamma Values Range Optimal Alpha Optimal Beta Optimal Gamma	0.01 to 1 (10 values) 0.01 to 1 (10 values) 0.01 to 1 (10 values) Varies per product (determined through optimization) Varies per product (determined through optimization) Varies per product (determined through optimization)
Arima Rolling Forecast	$p$ Values Range $d$ Values Range $q$ Values Range Split Ratio	0 to 5 (integers) 0 to 1 (integers) 0 to 5 (integers) 80% training, 20% validation
Arima Long-Term Forecast	$p$ Range $d$ Range $q$ Range Test Size Forecast Steps Best Parameters Split Ratio	0 to 5 0 to 2 0 to 5 52 weeks 52 Varies per product (determined through grid search) Last 50 observations for testing
Facebook Prophet—Long Term	Growth Changepoint Prior Scale Seasonality Prior Scale Interval Width	'linear' (for all products) Varies by product (e.g., (10, 30, 50) for M01AB) Varies by product (e.g., (150, 170, 200) for N02BE) 0.0005 (for all products)
XGBoost Model	learning_rate n_estimators max_depth Subsample colsample_bytree colsample_bylevel Gamma Split Ratio	(0.001, 0.01, 0.05, 0.1, 0.5) (30, 50, 100, 150, 200, 300) (2, 3, 5, 7, 9, 11) (0.7, 0.8, 0.9, 1) (0.7, 0.8, 0.9, 1) (0.7, 0.8, 0.9, 1) (0, 0.1, 0.2, 0.3, 0.4, 0.5) (0.2) (corresponding to 20% validation, 80% training split)
LSTM Neural Network—Long Term	Number of Steps Number of Features Model Type Layers Activation Dropout Optimizer Loss Split Ratio	3 1 Sequential LSTM layers with varying units (50, 100, 150) 'relu' 0.2, 0.3 (varies by layer) Adam (learning rate 0.001 for 'N02BA', otherwise default 'adam') 'mse' (0.3) (corresponding to 30% validation, 70% training split)

In Table 1, the tuning of hyperparameters for all eight groups of pharmaceutical products is further depicted:

- For the **Seasonal Naïve model**, the hyperparameters focus on capturing seasonal patterns with weights that give more importance to recent years. This approach, tailored for a 52-week cycle, allows the model to emphasize recent trends that can be more indicative of future patterns.
- The **Single Exponential Smoothing model** adapts to data by calculating the optimal alpha value, which determines the level of weight given to the most recent observation in forecasting. The range of alpha is broad, providing flexibility to model various rates of change in data trends.
- **Double exponential smoothing** extends upon single smoothing by not only considering the level but also the trend of the time series data. It determines the optimal alpha and beta values through optimization, allowing for a nuanced understanding of both recent changes and underlying trends.
- The **triple exponential smoothing model** incorporates seasonality on top of level and trend, making it suitable for data with seasonal fluctuations. The optimization process finds the best alpha, beta, and gamma values to balance the level, trend, and seasonal components of the time series.
- **ARIMA rolling forecast** is designed for short-term forecasting, with hyperparameters  $p$ ,  $d$ , and  $q$  defining the model's structure. It uses a split of 80% training data to capture the underlying process and 20% validation data to ensure the model's predictive accuracy.
- For long-term predictions, the **ARIMA long-term forecast** model extends the rolling forecast approach with additional forecast steps, allowing for an extended horizon in the predictions and employing a grid search to identify the best combination of  $p$ ,  $d$ , and  $q$  parameters.
- The **Facebook Prophet—Long-Term** model is employed for its robust handling of time series with irregular trends, seasonality, and holidays. It uses a linear growth model with adjustable changepoint and seasonality parameters to adapt to each product's unique characteristics, ensuring comprehensive forecasting.
- Among 10 various forecasting algorithms, the **XGBoost model**, a machine learning model, includes hyperparameters such as learning rate,  $n\_estimators$ ,  $max\_depth$ ,  $subsample$ ,  $colsample\_bytree$ ,  $colsample\_bylevel$ , and  $gamma$ . The specified split ratio of 20% for validation and 80% for training indicates a conscious approach to balancing model training with validation, optimizing the model's forecasting performance.
- The **LSTM Neural Network—Long Term** model incorporates hyperparameters such as the number of steps (3), features (1), and LSTM layers with varying units (50, 100, and 150), utilizing 'relu' activation and 'mse' for loss calculation. Dropout rates are set at 0.2 or 0.3 to aid in model generalization, and an Adam optimizer with a learning rate of 0.001 for 'N02BA' enhances optimization. The training and validation split of 70/30 is meticulously chosen to validate the model's efficacy.

In our research analysis, we use two common evaluation metrics, mean squared error (MSE) and mean absolute percentage error (MAPE).

#### Mean Squared Error (MSE)

MSE is a valuable metric for assessing the general precision of a forecasting model. Its consideration of all errors, regardless of direction, makes it suitable for identifying areas of improvement in predictions, while MSE also quantifies the average of the squared differences between predicted and actual values [39].

$$\text{Typology: MSE} = \frac{\sum (y_i - p_i)^2}{n}$$

where:

$n$  is the number of observations;

$y_i$  is the actual (observed) value for observation  $i$ ;  
 $p_i$  is the predicted value for observation  $i$ .

### Mean Absolute Percentage Error (MAPE)

MAPE calculates the average percentage difference between predicted and actual values. It offers a straightforward representation of forecasting accuracy in percentage terms, making it easily interpretable, while MAPE also evaluates forecast performance, particularly in scenarios where percentage accuracy is crucial [39].

$$\text{Typology: MAPE} = \frac{\sum \frac{|y_i - p_i|}{y_i}}{N} 100$$

where:

$N$  is the number of observations;  
 $y_i$  is the actual (observed) value for observation  $i$ ;  
 $p_i$  is the predicted value for observation  $i$ .

The entire coding framework, as outlined in our workflow—which includes stages like data cleaning, ATC classification adoption, time series analysis, feature selection, data structuring, parameter tuning, forecasting, and performance evaluation—was implemented on a personal computer equipped with 16 GB RAM, an Intel Core i5 8th generation processor, an SSD, and four cores. The total runtime for executing the complete code amounted to 54 h. This information provides transparency for our computational setup, facilitating the reproducibility and understanding of the computational resources required for similar analyses.

As we continue our analysis, Tables 2 and 3 are presented next, which consist of comprehensive evaluation tables of various forecasting models, showcasing the mean square error (MSE) and mean absolute percentage error (MAPE) of different groups of pharmaceutical products, which serve as critical measures of predictive accuracy for each model.

**Table 2.** Forecasting mean square error (MSE) of groups of products.

MSE	M01AB	M01AE	N02BA	N02BE	N05B	N05C	R03	R06
Naïve Models								
Naïve	72.49	86.98	75.57	2404.25	239.42	10.95	1087.25	86.5
Seasonal Naïve	49.19	58.75	55.27	2037.18	176.38	7.85	963.29	65.17
Statistical Models								
Smoothing Models								
Single Exponential Smoothing	71.11	94.91	31.59	6753.29	146.81	7.19	941.86	165.06
Double Exponential Smoothing	69.57	88.63	31.18	4805.13	140.76	7.24	775.75	152.47
Triple Exponential Smoothing	71.48	76.39	39.81	2190.53	241.21	8.58	738.68	59.77
Autoregressive Integrated Moving Average (ARIMA) Models								
Arima Rolling Forecast	60.7	68.32	28.34	2480.99	147.14	7.49	663.06	69.18
Arima Long-Term Forecast	75.4	90.86	33.15	5778.34	149.92	7.26	784.99	168.76
Prophet Model								
Facebook Prophet—Long Term	73.92	88.6	28.81	3149.32	184.17	9.03	1205.37	113.04
Machine Learning Models								
LSTM Neural Network—Long-Term	65.04	58.04	38.27	2214.6	250.98	13.29	577.66	84.95
XGBoost	54.81	28.8	40.69	1518.56	260.52	7.9	350.84	63.85

**Table 3.** Predictive model MSE benchmarking.

Categories	M01AB	M01AE	N02BA	N02BE	N05B	N05C	R03	R06
Our Study's Best Algorithm/MSE	Seasonal Naive 49.19	XGBoost 28.80	Arima Rolling Forecast 28.34	XGBoost 1518.56	Double Exponential Smoothing 140.76	Single Exponential Smoothing 7.19	XGBoost 350.84	Triple Exponential Smoothing 59.77
Zdravković et al. [25] Best Algorithm/MSE	Facebook Prophet 69.62	ARIMA 76.57	ARIMA 31.94	Auto-ARIMA 2147.07	Auto-ARIMA 147.13	ARIMA 7.98	Auto-ARIMA 666.68	Stacked LSTM 66.84

After the examination of MSE, in Table 2, we can summarize and notably showcase the dominance of the Extreme Gradient Boosting (XGBoost) model:

- **Machine learning models:**
  - **The Extreme Gradient Boosting (XGBoost) Model** outperforms the models for M01AE anti-inflammatory and N02BE/B analgesic drugs, with MSE values of 28.8 and 1518.56, showcasing adeptness in unraveling complex sales trends, while it also stands out for R03 drugs for airway diseases.
- **Statistical models:**
  - **The Autoregressive Integrated Moving Average (ARIMA) Rolling Forecast Model** is the most accurate for N02BA Analgesic Drugs, with an MSE of 28.34.
  - **The Double Exponential Smoothing (DES) and Single Exponential Smoothing (SES) models** are preferred for psycholeptic drugs, specifically N05B anxiolytics and N05C sedatives, reflecting their capacity to smooth erratic sales data.
  - **The Triple Exponential Smoothing (TES) model** is proven to be effective for R06 antihistamines, emphasizing the importance of selecting the right model for effective inventory management.
- **Naïve models:**
  - **The Seasonal Naïve (Naïve) model** is identified as notably effective for M01AB anti-inflammatory drugs, with a minimal MSE of 49.19, indicating strong seasonal sales patterns.

The above MSE results from our study can be compared to those reported by the study of Zdravković et al. [24], which used the same dataset. As observed in Table 3, our results outperform those of Zdravković et al.'s [24] study in all eight drug categories. The best algorithms based on MSE for each drug category are shown in Table 3 for both studies.

After the examination of the mean absolute percentage error (MAPE) outlined in Table 4, focusing on different product groups, several noteworthy observations emerge:

- **Machine learning models: Regarding the XGBoost model:**
  - *M01AB and M01AE anti-inflammatory drugs:* Demonstrates remarkable performance with the lowest MAPE of 17.89% and 16.92%, respectively, highlighting its robust ability to model complex, non-linear relationships in pharmaceutical sales data;
  - *N02BA analgesic drugs:* Maintains dominance with the lowest MAPE value of 17.98%, showcasing consistent and accurate forecasting for these categories;
  - *N02BE analgesic drugs:* Sustains exceptional performance, achieving the lowest MAPE of 16.05%;
  - *R06 antihistamines:* Continues excellence with the lowest MAPE at 36.26%.
- **Statistical models: Regarding the Facebook Prophet—Long-Term model:**
  - *N05B anxiolytics:* Demonstrates specialty in forecasting sales, achieving a MAPE of 18.39%;

Regarding the **Triple Exponential Smoothing (TES) model:**

- *R03 drugs for obstructive airway diseases:* Stands out with a MAPE of 39.91%, indicating its capability to capture trends and seasonality.

- **Naïve Models: Regarding the Naïve model:**
  - *N05C sedatives*: Surprisingly proven to be effective with a low MAPE of 12.12%.

**Table 4.** Forecasting mean absolute percentage error (MAPE) of groups of products.

MAPE	M01AB	M01AE	N02BA	N02BE	N05B	N05C	R03	R06
Naïve Models								
Naïve	22.96%	36.79%	28.41%	44.57%	56.17%	12.12%	72.60%	93.51%
Seasonal Naïve	27.48%	34.77%	37.24%	24.42%	31.78%	17.73%	61.90%	72.02%
Statistical Models								
Smoothing Models								
Single Exponential Smoothing	22.08%	33.16%	29.08%	36.22%	19.87%	86.81%	69.94%	100.14%
Double Exponential Smoothing	23.27%	30.97%	28.44%	21.70%	19.17%	86.66%	64.32%	98.78%
Triple Exponential Smoothing	24.50%	29.04%	30.86%	18.89%	21.90%	>100%	39.91%	63.68%
Autoregressive Integrated Moving Average (ARIMA) Models								
Arima Rolling Forecast	19.85%	30.22%	27.08%	19.99%	18.49%	inf%	45.91%	53.91%
Arima Long-Term Forecast	23.04%	32.77%	29.51%	20.15%	19.70%	inf%	47.83%	57.33%
Prophet Model								
Facebook Prophet—Long Term	23.11%	30.45%	28.02%	20.38%	18.39%	Inf%	41.11%	66.08%
Machine Learning Models								
LSTM Neural Network—Long Term	20.87%	26.97%	31.36%	19.59%	19.19%	Inf%	46.68%	48.98%
XGBoost	17.89%	16.92%	17.98%	16.05%	24.76%	Inf%	54.78%	36.26%

#### 4. Research Results and Discussion

Based on the above analysis, our examination of various models in forecasting pharmaceutical sales yields insightful conclusions and noteworthy results:

In our examination of **Naïve models**, the lowest mean absolute percentage error (MAPE) was observed for N05C, which was categorized as psycholeptics drugs used as hypnotics and sedatives, registering at 12.12%. However, in the Naïve model for obstructive airway diseases (R06), the highest error rate was recorded at 93.51%. These results emphasize the varied performance of Naïve models across distinct drug categories, suggesting a potential correlation between the pharmacological properties of the drugs and their predictability in sales. This diversity underscores the importance of tailoring modelling approaches to specific drug categories.

Transitioning to **statistical time series analysis models**, both the Triple Exponential Smoothing model and the ARIMA forecast demonstrated promising results, with the former, especially for M01AE (propionic acid derivatives), boasting a MAPE of 29.04% and the latter excelling for M01AB (acetic acid derivatives) with an error rate of 19.85%. Additionally, our long-term forecasting with the Facebook Prophet model showcased superior results, especially when applied to N05B—the psycholeptic drugs category, exhibiting a mean absolute percentage error (MAPE) of 18.39%. The discussion surrounding these statistical models delves into their effectiveness in complex patterns within pharmaceutical sales, with the Facebook Prophet model standing out in long-term predictions.

In the field of **machine learning models**, our incorporation of the XGBoost algorithm yielded impressive results, proving valuable across multiple drug categories, notably achieving the lowest mean squared error (MSE) of 1518.56 for N02BE/B, which includes pyrazolones and anilides. The algorithm consistently outperformed traditional methods by capturing complex, non-linear relationships within the data. For instance, in the M01AB category, XGBoost displayed a low MAPE of 16.92%. The discourse here focuses on

the adaptability and effectiveness of machine learning techniques in handling diverse drug categories.

As we have already mentioned in Section 3, our research showed that our forecasting methods produced better results, as reflected by the mean squared error (MSE), when compared to the outcomes from Zdravković et al. [25].

Also, in the evaluation of the success of our forecasting models, we considered the mean absolute percentage error (MAPE) to be as critical as the mean squared error (MSE) in determining their accuracy and reliability. MAPE provides a clear picture of predictive precision by expressing forecast errors as a percentage, which is particularly useful in the diverse landscape of pharmaceutical sales. Within our study, the XGBoost algorithm proved to be exceptionally effective, achieving the lowest MAPE values in five distinct drug categories. This highlights its robustness and solidifies its position as a top-performing model in our predictive toolkit.

Through the optimization of hyperparameters using grid search, we significantly enhance the accuracy of our forecasts for pharmaceutical demand in retail pharmacies. This advancement has the potential to positively impact other aspects of pharmacy operations as well, because it optimizes the supply chain effectively, ensuring that drug shortages are pre-emptively addressed. By precisely calibrating the quantity of each medicine, we can minimize waste and avoid excess stock in warehouses, while simultaneously meeting the needs of the community. Furthermore, ensuring the consistent availability of pharmaceutical products enhances customer satisfaction and builds a foundation of trust and loyalty among consumers.

In conclusion, our findings in this section emphasize the strengths of Naïve, statistical (including Facebook Prophet), and machine learning models in pharmaceutical sales forecasting. XGBoost demonstrated proficiency in modelling complex, non-linear relationships within pharmaceutical sales data. The complexity of sales patterns is underscored by distribution histograms, which reveal irregular sales trends, seasonal effects, and outliers, demonstrating the intricate dynamics that influence pharmaceutical sales. These insights not only advance our methodologies but also lay the foundation for future research to build upon these improved modelling techniques. The discussion encourages further exploration of model enhancements and applications in the evolving landscape of pharmaceutical sales forecasting.

## 5. Conclusions, Proposals, and Recommendations

The growing importance of the pharmaceutical industry requires changes in sales forecasting techniques, and while traditional models serve as a foundation, they often fail to capture the detailed complexities of drug sales to the same degree as the machine learning models. By comparing traditional methods with machine learning (ML) models, such as the Extreme Gradient Boost (XGBoost) and LSTM neural network models, our research has produced some interesting results: For example, forecasting in the category of product M01AB, the XGBoost model reduced our mean absolute percentage error (MAPE) from 27.48% to 17.89% and the Seasonal Naïve method has carried out. Similarly, for a different category of products, XGBoost for N02BA achieved a metric MAPE score of 17.98%, a significant improvement over Seasonal Naïve by 37.24%. These outcomes underscore the superiority of machine learning techniques over traditional methods in capturing the intricacies of sales dynamics.

While LSTM neural networks exhibited promising results, their full potential was hindered by the constraints of a limited dataset, leading to suboptimal and less efficient outcomes that are evident in Tables 2 and 3. The results highlight the importance of considering dataset size in optimizing the performance of LSTM neural networks, suggesting that with larger datasets, these models could potentially outperform the majority of forecasting models.

Our research has uncovered seasonal variations, revealing higher sales during winter months for certain medications and transitional seasons impacting the demand for others.

This information could prove invaluable for inventory planning and targeted marketing campaigns. These seasonal insights provide practical implications for pharmaceutical companies aiming to align their strategies with the temporal dynamics of medication demand.

In essence, while comprehending sales trends remains crucial, integrating advanced forecasting models is imperative for the future, and focusing on that, our study indicates that machine learning techniques like XGBoost and LSTM neural networks offer enhanced prediction accuracy, facilitating timely access to medication globally. Future research should explore additional machine learning algorithms for pharmaceutical forecasting, with a focus on LSTM neural networks, which may yield superior results, particularly with larger datasets. Combining these algorithms with external datasets, such as demographic or climate data, could further enhance prediction precision. These recommendations provide a roadmap for future research to build upon our findings and explore new avenues for refining pharmaceutical sales forecasting methodologies.

Understanding the best-performing models helps pharmaceutical companies ensure that essential medications are consistently available, and this allows them to predict future sales that can ensure that people attain the medicines they need whenever they need them.

**Author Contributions:** For this research article, the contributions are as follows: Conceptualization, K.P.F. and A.T.; methodology, K.P.F. with guidance from A.T.; software, K.P.F.; validation, K.P.F., with oversight from A.T.; formal analysis, K.P.F.; investigation, K.P.F.; resources, A.T.; data curation, K.P.F.; writing—original draft preparation, K.P.F.; writing—review and editing, A.T.; visualization, K.P.F.; supervision, project administration, and funding acquisition, A.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data is contained within the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Voumvaki, J.; Koutouzou, A. *Greek Pharma Industry: In Position to Capitalize on EU Shift towards More Self-Reliance*; Sectoral Report April 2022; National Bank of Greece, Economic Analysis Department Eolou: Athens, Greece, 2022; Volume 86.
2. Ghaffar, A.; Rashidian, A.; Khan, W.; Tariq, M. Verbalising importance of supply chain management in access to health services. *J. Pharm. Policy Pract.* **2021**, *14* (Suppl. S1), 91. [[CrossRef](#)] [[PubMed](#)]
3. Lee, K.; Joo, S.; Baik, H.; Han, S.; In, J. Unbalanced data, type II error, and nonlinearity in predicting M&A failure. *J. Bus. Res.* **2020**, *109*, 271–287.
4. Ray, S.; Nikam, R.; Vanjare, C.; Khedkar, A.M. Comparative Analysis of Conventional and Machine Learning Based Forecasting Of Sales In Selected Industries. *IJFANS Int. J. Food Nutr. Sci.* **2022**, *11*, 3780–3803.
5. Lim, C.M.; Yusof, F.A.M.; Selvarajah, S.; Lim, T.O. Use of ATC to Describe Duplicate Medications in Primary Care Prescriptions. *Eur. J. Clin. Pharmacol.* **2011**, *67*, 1035–1044. [[CrossRef](#)]
6. Martinez, M.E. The Calendar of Epidemics: Seasonal Cycles of Infectious Diseases. *PLoS Pathog.* **2018**, *14*, e1007327. [[CrossRef](#)]
7. Govindan, K.; Kannan, D.; Jørgensen, T.B.; Nielsen, T.S. Supply Chain 4.0 Performance Measurement: A Systematic Literature Review Framework Development and Empirical Evidence. *Transp. Res. Part E* **2022**, *164*, 102725. [[CrossRef](#)]
8. Rathipriya, R.; Abdul Rahman, A.A.; Dhamodharavadhani, S.; Meero, A.; Yoganandan, G. Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model. *Neural Comput. Applic.* **2023**, *35*, 1945–1957. [[CrossRef](#)]
9. Berrar, D. Bayes' Theorem and Naive Bayes Classifier. *PLoS Pathog.* **2018**, *14*. [[CrossRef](#)]
10. Aburto, L.; Weber, R. A Sequential Hybrid Forecasting System for Demand Prediction. *Transp. Res. Part E* **2022**, *164*. [[CrossRef](#)]
11. Mancuso, A.C.B.; Werner, L. A Comparative Study on Combinations of Forecasts and Their Individual Forecasts by Means of Simulated Series. *Acta Sci. Technol.* **2019**, *41*, e41452. [[CrossRef](#)]
12. Pamungkas, A.; Puspasari, R.; Nurfiarini, A.; Zulkarnain, R.; Waryanto, W. Comparative Analysis of Exponential Smoothing Methods for Forecasting Marine Fish Production in Pekalongan Waters, Central Java. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *934*, 012016. [[CrossRef](#)]
13. İmece, S.; Beyca, Ö.F. Demand Forecasting with Integration of Time Series and Regression Models in Pharmaceutical Industry. *Int. J. Adv. Eng. Pure Sci.* **2022**, *34*, 415–425. [[CrossRef](#)]
14. Dutta, S.R.; Das, S.; Chatterjee, P. Smart Sales Prediction of Pharmaceutical Products. In Proceedings of the 2022 8th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 21–22 April 2022; pp. 1–6.

15. Zunic, E.; Korjenic, K.; Hodzic, K.; Donko, D. Application of Facebook's prophet algorithm for successful sales forecasting based on real-world data. *arXiv* **2020**, arXiv:2005.07575.
16. Bandara, K.; Shi, P.; Bergmeir, C.; Hewamalage, H.; Tran, Q.; Seaman, B. Sales demand forecast in e-commerce using a long short-term memory neural network methodology. In *Neural Information Processing, Proceedings of the 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, 12–15 December 2019*; Proceedings, Part III 26; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 462–474.
17. Han, Y. A Forecasting Method of Pharmaceutical Sales Based on ARIMA-LSTM Model. In Proceedings of the 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT), Shenyang, China, 13–15 November 2020.
18. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
19. Goh, C.; Law, R. Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tour. Manag.* **2002**, *23*, 499–510. [[CrossRef](#)]
20. BioPhorum Operations Group. Forecasting and Supply Planning: A Best Practice Guide for the Biopharmaceutical Industry. *BioPhorum*, 5 April 2018.
21. Moosivand, A.; Rajabzadeh Ghatari, A.; Rasekh, H.R. Supply Chain Challenges in Pharmaceutical Manufacturing Companies: Using Qualitative System Dynamics Methodology. *Iran. J. Pharm. Res.* **2019**, *18*, 1103–1116. [[PubMed](#)]
22. Yani, L.P.E.; Aamer, A. Demand forecasting accuracy in the pharmaceutical supply chain: A machine learning approach. *Int. J. Pharm. Healthc. Mark.* **2023**, *17*, 1–23. [[CrossRef](#)]
23. Zhu, X.; Ninh, A.; Zhao, H.; Liu, Z. Demand Forecasting with Supply-Chain Information and Machine Learning: Evidence in the Pharmaceutical Industry. *Prod. Oper. Manag.* **2021**, *30*, 3231–3252. [[CrossRef](#)]
24. Zdravković, M.; Đorđević, J.; Catić-Đorđević, A.; Pavlović, S.; Ivković, M. *Univariate Time Series Analysis and Forecasting of Pharmaceutical Products' Sales Data at Small Scale*; Information Society of Serbia—ISOS Serbia: Belgrade, Serbia, 2020.
25. KPMG Global Strategy Group. *Pharma 2030: From Evolution to Revolution*; KPMG International Cooperative: Amstelveen, The Netherlands, 2017.
26. Adam, M.B.; Baba, I.; Ali, N.; Mohammed, M.B.; Zulkafli, H.S. Comparison of Five Imputation Methods in Handling Missing Data in a Continuous Frequency Table. *AIP Conf. Proc.* **2021**, *2355*, 040006.
27. Singh, K.; Upadhyaya, S. Outlier Detection: Applications and Techniques. *Int. J. Comput. Sci. Issues* **2012**, *9*, 3.
28. Hollingworth, S.; Kairuz, T. Measuring Medicine Use: Applying ATC/DDD Methodology to Real-World Data. *Pharmacy* **2021**, *9*, 60. [[CrossRef](#)]
29. Sarker, I.H. Data Science and Analytics: An Overview from Data-Driven Smart Computing Decision-Making and Applications Perspective. *SN Comput. Sci.* **2021**, *2*, 377. [[CrossRef](#)]
30. Ensafi, Y.; Hassanzadeh Amin, S.; Zhang, G.; Shah, B. Time-Series Forecasting of Seasonal Items Sales Using Machine Learning: A Comparative Analysis. *Int. J. Inf. Manag. Data Insights* **2022**, *2*, 100058. [[CrossRef](#)]
31. Shmueli, G.; Bruce, P.C.; Gedeck, P.; Patel, N.R. *Data Mining for Business Analytics: Concepts, Techniques, and Applications in Python*; John Wiley & Sons: Hoboken, NJ, USA, 2019; 608p.
32. Lewis, E.J.; Bishop, J.; Aspinall, S.J. A Simple Inflammation Model That Distinguishes Between the Actions of Anti-Inflammatory and Anti-Rheumatic Drugs. *Inflamm. Res.* **1998**, *47*, 26–35. [[CrossRef](#)]
33. Twycross, R.G. Analgesics. *Postgrad. Med. J.* **1984**, *60*, 876–880. [[CrossRef](#)]
34. John, U.; Baumeister, S.E.; Völzke, H.; Grabe, H.J.; Freyberger, H.J.; Alte, D. Estimation of Psycholeptic and Psychoanaleptic Medicine Use in an Adult General Population. *Int. J. Methods Psychiatr. Res.* **2008**, *17*, 220–231. [[CrossRef](#)] [[PubMed](#)]
35. Lareau, S.C.; Fahy, B.; Meek, P.; Wang, A. Chronic Obstructive Pulmonary Disease (COPD): A Comprehensive Overview. *Am. J. Respir. Crit. Care Med.* **2019**, *199*, P1–P2. [[CrossRef](#)] [[PubMed](#)]
36. Church, D.S.; Church, M.K. Pharmacology of Antihistamines. *World Allergy Organ. J.* **2011**, *4*, S22–S27. [[CrossRef](#)] [[PubMed](#)]
37. Dürre, A.; Fried, R.; Liboschik, T. Robust Estimation of (Partial) Autocorrelation. *Wiley Interdiscip. Rev. Comput. Stat.* **2015**, *7*, 205–222. [[CrossRef](#)]
38. Zhao, S.; Guo, Y.; Sheng, Q.; Shyr, Y. Advanced Heat Map and Clustering Analysis Using Heatmap3. *BioMed Res. Int.* **2014**, *2014*, 986048. [[CrossRef](#)] [[PubMed](#)]
39. Kim, S.; Kim, H. A new metric of absolute percentage error for intermittent demand forecasts. *Int. J. Forecast.* **2016**, *32*, 669–679. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.