International Journal of Surgery and Anesthesia Research

PRIME OPEN ACCESS

Volume 2, Issue 1

Research Article

Date of Submission: 12 Mar, 2026
Date of Acceptance: 10 Apr, 2026
Date of Publication: 20 Apr, 2026

Prediction of Drug Procurement in Healthcare Supply Chains Using Machine Learning Algorithms

Simachew Demisse* and Mekonnen Wagaw

Bahir Dar University Institute of Technology, Ethiopia

***Corresponding Author:** Simachew Demisse, Bahir Dar University Institute of Technology, Ethiopia.

Citation: Demisse, S., Wagaw, M. (2026). Prediction of Drug Procurement in Healthcare Supply Chains Using Machine Learning Algorithms. *Int J Surg Anesth Res, 2*(1), 01-08.

Abstract
The effective forecasting of the demand for drugs and their lead time is important in order to ensure that their critical supply is always available without incurring additional healthcare costs and wastage within the healthcare supply chain. In this paper, we investigate the implementation of machine learning techniques utilizing XGBoost, Random Forest, Extra Trees, and Linear Regression models for increased accuracy and flexibility of demand and lead time prediction. Sales figures, seasonal trends, and other external events serve as real-time data inputs for the proposed model which tries to respond to the changes in the demand for the drugs in an automated fashion.

The models are evaluated with performance indicators which include Mean Squared Error (MSE), Mean Absolute Error (MAE), R squared, and Percentage Error, where XGBoost showed the best results. Moreover, SHAP explains the impacts of certain features so that relevant inventory control decisions can be made to further enhance management efficiency. With a prototype developed in Streamlit, the proposed solution is found to be usable and scored 'Good' on the System Usability Scale. This work highlights the opportunity machine learning presents to healthcare logistics and serves to inspire further work on automated adaptive forecasting systems.

Keywords: Drug Demand Prediction, Explainable AI (XAI), Machine Learning, Healthcare Supply Chains, XGBoost

Introduction
Maintaining appropriate levels of essential medicines when needed, while avoiding waste and reducing costs, requires efficient management of healthcare supply chains. Therefore, accurate forecasting of drug demand and lead time helps to avoid stockouts and overstocking, which cause treatment delays, additional costs, and compromised patient care [1,2]. Conventional forecast methods like time series analyses or regression models often rest on historical data and assume demand patterns are predictable. Quite the opposite, as the healthcare sector is experiencing seasonal fluctuations, unexpected disease outbreaks, and evolving demographics, making realizing these goals using traditional methods highly infeasible [3,4].

Integrating real-time elements like sales, seasonal context, and external observations into more extensive datasets enables machine learning models to detect patterns and non-linear correlations hidden within drug demand. Utilizing advanced strategies in XGBoost, Random Forests, Extra Trees, and Linear Regression, this study shifts its focus on explainable AI to enhance the flexibility and precision of micro-supply chain forecasting accuracy in healthcare. The study also places focus on the application of SHAP explanations alongside inventory prediction models to empower decision-makers with transparent insights for more effective inventory management.

Streamlit was the chosen platform to present the prototype created for the project, which serves as proof-of-concept demonstrating these advanced machine learning approaches. The assessment of usability for the prototype based on System Usability Scale (SUS) testing returned a score suggesting a user-friendly experience, demonstrating the prototype's readiness for implementation into practice [5]. The research seeks to improve healthcare logistics through better predictive models of drug procurement, which would minimize waste and enhance patient care and health outcomes.

Related Works

There has been an increasing focus in the research concerning the use of machine learning techniques especially when integrating forecasting and inventory management aspects across the logistics supply chain. A number of studies have proven the ability of machine learning techniques in the enhancement of supply chain management processes; however, there are very few that have looked at supply chains that are in the healthcare centres where the prediction of the Procurement (demand and lead time) for drugs is fundamental. Various researchers carry out their studies looking into the application of machine learning models in demand and lead time forecasting within normal, general supply chains. As an example, examine how effective forecasting of demand using the Random Forest (RF) and Support Vector Machines (SVM). These models were found to be superior to their ARIMA and exponential smoothing counterparts in terms of predictive accuracy and ease of implementation. Even so, Xun and Li 2024 used LSTM (Long Short-Term Memory) networks to model time series data and also found that models based on LSTM are best suited to forecast models for general supply chain demand as they have the capacity to characterize demand data with non-linear characteristics. Joseph et al. 2022 proposes a hybrid CNN-LSTM model to forecast supply chain demand, in which they use CNN for extracting features and LSTM for time-series predictions.

Their research shows considerable improvements in prediction accuracy in regards to forecasting over the use of single models. Jahin, Shahriar, and Amin 2024 proposed another approach in which multi-channel deep fusion networks (MCDFN) take multiple channels and time series as input to facilitate dynamic-state prediction in patient-specific environments. While most of the studies reviewed tend to focus on the wider context of the supply chain, a few studies seem to target healthcare supply chain related issues.

According to Motamedi et al. 2021, machine learning models were employed in predicting the demand for medical consumables but still mainly rely on sales data retrospectively without utilization of real time data or variables associated with patients. Further, Zheng et al. 2022 proposed an ensemble approach that consolidates decision trees and GBM in order to improve drug demand prediction capabilities. They worked on the complexity of service demand prediction but fell short in tackling the issues brought about by the external environment for instance patients or health political systems.

Pulling in live data is essential to enhancing the accuracy and adaptability of forecasting models. According to, research on the impact of availability of real-time data streams in forecasting demand in supply chains and argues that potential benefits exist where current information is always incorporated into forecasting models as they significantly increase the accuracy especially during periods of turbulence. Still, the real-time integration in healthcare supply chains remains a rather poorly researched area of inquiry.

In addition, highlight the challenge of model explainability which in most cases is the case in most healthcare applications where the models have to be interpretable to the stakeholders that sit behind the decision-making processes. Although there appears to be an increasing interest in explainable AI in the literature, the use of these methods in healthcare supply chains is still scarce.

Despite these technological advances, there are still considerable deficiencies in the use of machine learning in healthcare supply management. Many studies apply the traditional forecasting techniques or employ generic supply chain data without addressing the special features of the healthcare industry such as drug stockouts and their implications, the fluctuations in demand for the drugs, and how crises such as wars and pandemics affect the level of demand. Moreover, even though these challenges are key in demand planning, very few studies report on their resolution in relation to healthcare supply chains. There is clearly a lack of emphasis on real-time data context, dynamic lead time and data streams associated with the unique aspects of healthcare and even the provision of an interpretable model to potential data users such as clinicians and supply chain managers presents more questions than answers.

Methods

We adopt a data-driven experimental research design to create and validate machine learning models tailored to drug procurement forecasting in healthcare supply chains. The pipeline involves key stages including data collection, data preprocessing and feature engineering, model development and evaluation, and implementation of prototype and usability testing. The outcome is a predictive system that is accurate, interpretable, and deployable, helping stakeholders make better procurement decisions. Figure 1 illustrates the proposed model architecture for the overall study.

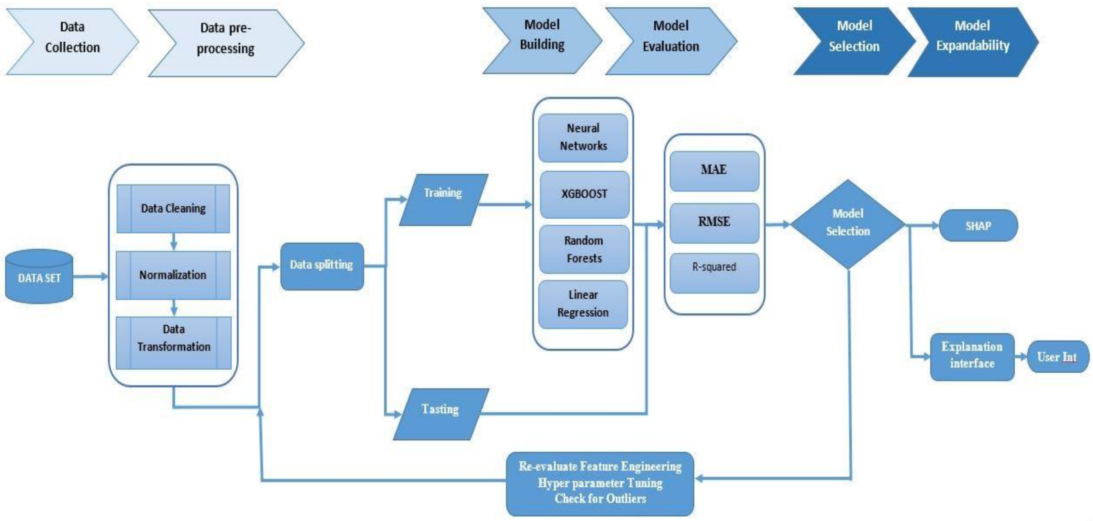


Figure 1. Model Architecture

Data Collection And Sources

This was done using historical data from the Ethiopian Pharmaceutical Supply Agency (EPSA) and the procurement records. Drug consumption behavior, inventory stock/level, lead time, supplier behavior, and context information like demographics and epidemiological information are packed in the dataset. Given the extent of the current pandemic, there is a burgeoning need to conduct research to combat the extent of this challenge. The two and a half years of data were selected to make sure that all seasonal and trend based variations related to drug demand were captured [6].

Data Pre-processing and Feature Engineering

To enhance data quality and modelling effectiveness, extensive preprocessing steps were applied

- **Handling Missing Values:** Missing data points were addressed through imputation techniques, such as mean imputation for numerical attributes and mode imputation for categorical variables.

- **Outlier Detection and Treatment:** Outliers were identified using statistical methods and domain-specific knowledge, ensuring that extreme values did not disproportionately affect model performance.

- **Feature Normalization:** Continuous variables were scaled using Min-Max normalization to ensure consistency across features and prevent scale dominance in certain models.

- **Feature Selection:** Correlation analysis, recursive feature elimination (RFE), and domain expertise were used to select the most relevant predictors influencing drug demand. Key features such as order day, unit cost, and total amount were identified as significant contributors [7].

- **Feature Extraction:** Temporal features such as month, quarter, and seasonality indicators were derived to capture cyclic demand patterns [8].

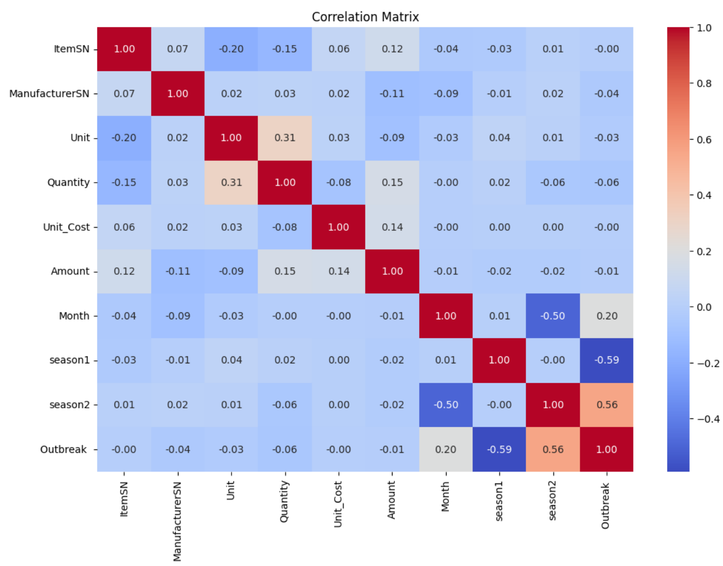


Figure 2. The Correlation Matrix of the Features

Figure 2. shows the correlation matrix elucidating the interrelationship co-relation between features. There exists a moderate negative correlation between the variables "Month" and "season2" to -0.50 possibly due to seasonality effects Yes. From the data, one observes that "season1" and "Outbreak" exhibited a high positive correlation (0.59) and "season2" and "Outbreak" showed moderate positive correlation (0.56). It appears, from the seasonal timings perspective, that there is a correlation of seasonality and disease outbreaks in Ethiopia. It is an important factor for health care management and resource management. It can be assumed that "season1" is during the period of the rains which is associated with high incidence of malaria and Cholera diseases. The moderate correlations of other factors indicate that these variables are rather dislocated and tend to affect the end variable "Quantity" via different paths. This data is useful in explaining

Model Development and Evaluation

A comparative analysis of multiple machine learning algorithms was conducted to determine the most effective approach for drug demand forecasting. The models considered include

• **Linear Regression (LR):** A baseline model for benchmarking performance [9].

• **Random Forest (RF):** An ensemble learning method capable of capturing complex interactions between features [10].

• **Extra Trees (ET):** A variation of RF designed to enhance variance reduction and predictive accuracy [11].

• **XGBoost (XGB):** A gradient boosting model known for its efficiency in handling structured data [12].

• **Hyperparameter Tuning**

Model tuning hyperparameters through grid search and cross-validation for optimal results We determined the best configuration for each model by iterating key parameters like number of estimators, tree depth, learning rate, and regularization terms.

• **Model Evaluation Metrics**

The models were assessed using multiple evaluation metrics

• **Mean Absolute Error (MAE):** Measures average absolute differences between actual and predicted values [13].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

• **Mean Squared Error (MSE):** is one of the most commonly used metrics for evaluating the performance of a prediction model. It calculates the average of the squared difference between each predicted value and the actual value. This measure is commonly used in regression model development where many phenomena are the result of predicted continuous numbers.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

• **R-Squared (R^2):** Indicates the proportion of variance in the dependent variable explained by the model [14].

$$R^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})}$$

• **Percentage Error (PE):** A relative error metric based on MAE, providing insights into the prediction error magnitude.

$$Percentage\ Error = \left(\frac{MAE}{Average\ Actual\ Value}\right) \times 100$$

The best-performing model was selected based on its ability to minimize prediction errors and maximize explanatory power.

Prototype Development and Usability Testing

To demonstrate the applicability of the model to real-life scenarios, Streamlit was used to create a prototype system which serves as a front-end interface for users to interface with. This interface allows uploading of datasets for model inference, visualizing demand forecasts and prediction confidence intervals, and exploring contributions of the most important features via explainability methods, like SHAP [15]. A System Usability Scale (SUS) evaluation with domain experts scored 79.58, indicating high usability and user satisfaction.

In general, this methodological framework involves data preprocessing, machine learning modelling, and an interactive decision-support prototype. This approach enables flexibility in a constantly changing healthcare landscape and delivers actionable insights to enhance procurement strategies. For healthcare supply chain management, this allows greater trust as well as informed decisions due to model interpretability and real-time usability.

Results

Dataset Splitting

We also explored using different machine learning algorithms to see how they perform for drug demand prediction through three train-test split ratios — 90-10, 80-20 and 70-30 [16]. The demand data include a total of 46702 examples while lead time data have a total of 9657. Table 1 gives an overview of the dataset partitions used for each experiment:

Experiment	Train (%)	Test (%)	Demand Data (Train/Test)	Lead Time Data (Train/Test)
Exp. 1	90%	10%	42,031 / 4,670	8,691 / 966
Exp. 2	80%	20%	37,362 / 9,340	7,725 / 1,932
Exp. 3	70%	30%	32,692 / 14,011	6,759 / 2,898

Table 1: Train and Test Sizes by Experiment

Model Performance Evaluation

We assessed four machine learning models Linear Regression, Random Forest, Extra Trees, and XGBoost using Mean Squared Error (MSE), Mean Absolute Error (MAE), R-squared (R^2), and Percentage Error.

Model	MSE	MAE	R^2	% Error
XGBoost	3,291.62	20.87	0.99	3.59%
Random Forest	12,780.53	48.16	0.98	8.29%
Extra Trees	26,459.72	91.27	0.96	15.71%
Linear Regression	243,406.62	269.72	0.79	46.42%

Table 2: Model Performance Metrics for Demand Prediction

XGBoost demonstrated superior performance in demand prediction, achieving the lowest MSE and MAE, as well as the highest R^2 score.

Lead Time Prediction

For lead time forecasting, Random Forest outperformed the other models with the lowest MAE and percentage error, indicating its robustness in handling temporal dependencies.

Model	MSE	MAE	R^2	% Error
XGBoost	73.20	0.16	0.99	0.70%
Random Forest	0.25	0.02	0.99	0.10%
Extra Trees	0.11	0.98	0.97	4.32%
Linear Regression	2.36	7.37	0.79	32.27%

Table 3: Model Performance Metrics for Lead Time Prediction

Hyperparameter Tuning

To enhance model accuracy, hyperparameter tuning was performed using Grid Search and Randomized Search techniques. The optimal configurations for XGBoost, Random Forest, and Extra Trees were determined, significantly improving model efficiency [17].

```
# Create and train XGBoost models for each target
xgb_params = {
    'max_depth': 7,
    'learning_rate': 0.1,
    'n_estimators': 100,
    'objective': 'reg:squarederror',
    'random_state': 72
}
```

```
# Lead time Model Training and Tuning with GridSearchCV
Lead_time_params = {
    'max_depth': [3, 4, 5],
    'learning_rate': [0.01, 0.05, 0.1],
    'n_estimators': [100, 200, 300],
    'subsample': [0.7, 0.8, 0.9],
    'colsample_bytree': [0.7, 0.8, 0.9],
    'reg_alpha': [0, 1],
    'reg_lambda': [0, 1]
}
```

Demand Parameter

Lead Time Parameter

Figure 1: Hyperparameter Optimization Results Xgboost Parameter

Explainability Analysis Using Shap

SHAP (SHapley Additive exPlanations) was employed to understand the contributions of different features to model predictions. The analysis revealed that key variables such as order_day, total amount, and unit cost had the highest impact on demand predictions.

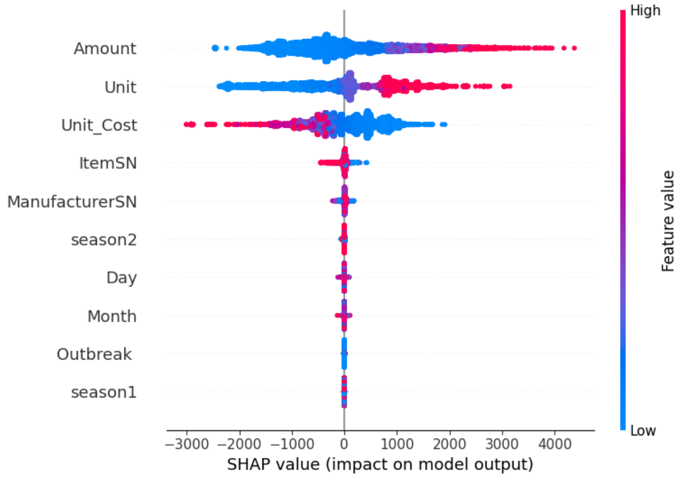
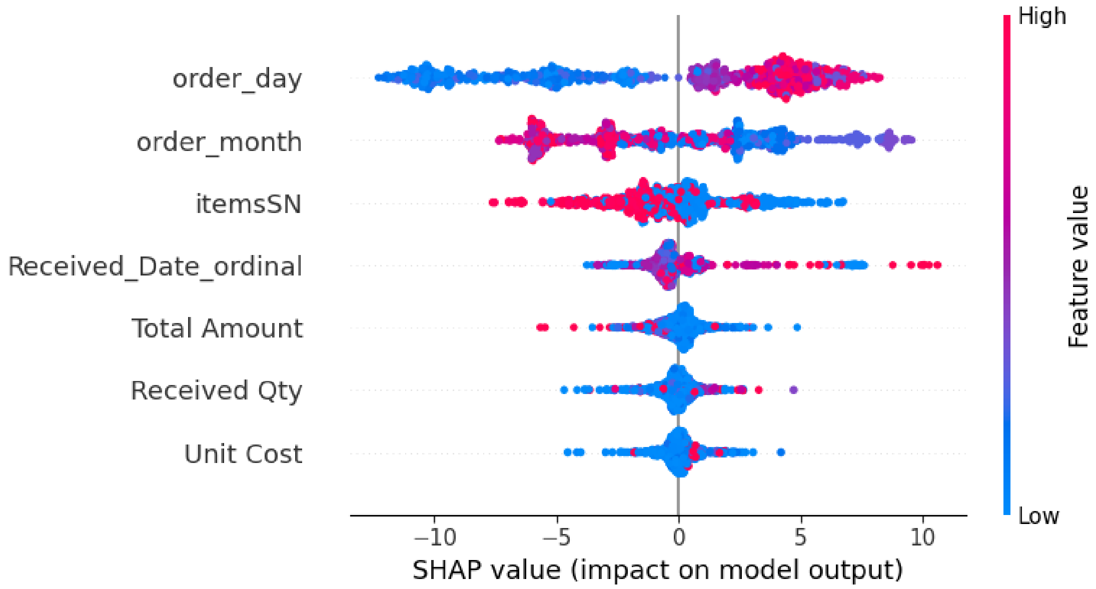


Figure 2: SHAP Summary Plot for Demand Prediction



XGBoost SHAP summary plot for Lead Time

Prototype Development and Usability Evaluation

A Streamlit-based web application was developed for drug demand forecasting. The usability of the prototype was evaluated using the System Usability Scale (SUS), with an average score of 79.58, indicating good usability.

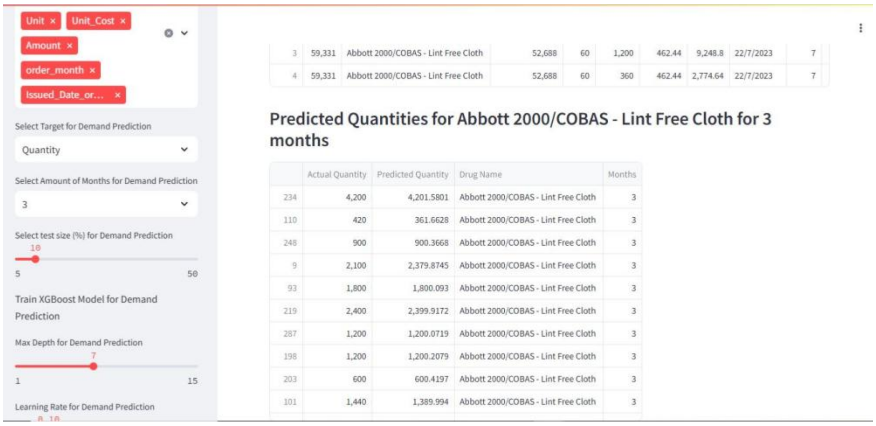


Figure 3: Prototype Interface for Model Training and Visualization/ Sample Prediction of Quantity for Selected Drug from the Developed Prototype.

Experimental results show that the finest model for demand prediction is XGBoost while randomly Forest works the best for lead time forecasting. The research introduces explainability techniques for the black-box nature of the predicted models and a prototype that addresses health supply chain stakeholders interactively.

Discussion

Model Performance Analysis

Machine learning models comparison showed that XGBoost was consistently able to predict drug demand better than other algorithms. Not only did the model have the lowest Mean Absolute Error (MAE), but it also had the highest R-squared (R^2) score, indicating superior predictive capability. This is consistent with results in previous studies, as tree-based ensemble models perform well on structured time-series data. One of the conclusions from the analysis was that Random Forests were the best performers in terms of predictive power and lead time prediction which suggests that the model is capturing the variability in supply chain delays. Such results justify the need to utilize suitable models for each predictive task in healthcare supply chains.

Implications for Healthcare Supply Chain Optimization

An accurate prediction of the demand for drugs is crucial to avoid stockouts and wastage. With the help of machine learning, healthcare institutions can timely adapt procurement systems by predicting the future demand. This also has an additional aspect of lead time prediction that contributes towards increased efficiency in terms of time-sensitive operations. This study's insights lay the groundwork for data-informed planning in the pharmaceutical supply chain, which can drive cost savings and facilitate superior patient service.

Explainability and Interpretability of Predictions

An important part of this research is the use of SHAP (SHapley Additive exPlanations) which improves model explainability. The analysis of feature importance revealed some major drivers like order_day, total amount, and unit cost, which significantly affected the demand estimation accuracy. This level of detail is vital for healthcare practitioners and supply chain experts who need to validate and have faith in the AI suggestions. Explainable AI is crucial with regard to ethical considerations on support systems integration.

Practical Application of the Web-Based Prototype

The web-based prototype designed in Streamlit allows users to inspect demand forecasts, analyze predicted value confidence, and modify model inputs. The usability test returned a highly rated System Usability Scale (SUS) score, showing great acceptance from users. This prototype with the ability to conduct real-time interactive forecasting is an effective tool that enables healthcare supply chain stakeholders to make responsible purchasing decisions.

Conclusion

The objective of this study was to estimate the demand of drugs in a healthcare supply chain utilizing machine learning techniques. The study demonstrated that, among all the models assessed, XGBoost outperformed the rest in demand forecasting due to having the highest R^2 score and the lowest MAE. Moreover, Random Forest was proven to be the most accurate in lead time prediction, which improves procurement planning and operational activities.

The incorporation of SHAP (SHapley Additive exPlanations) as one of the explainability methods provided adequate reasoning for feature importance, thus adding transparency into the business choices made. The created web-based prototype further provided the stakeholders in the health sector with the ability to predict the demand with the provided real-time data using the predictive models.

Even though the study presented promising findings, there remain some drawbacks. The models are based on past procurement data, meaning external components such as an outbreak of a disease or a new regulation being implemented are not taken into account.

Future directions could include incorporating real-time data streams as well as implementing deep learning approaches for more accurate predictions. Usability testing with healthcare professionals on a broader scale is needed to better understand implementation issues in real-world settings.

The findings of this study thus aid in the evolution of AI-enabled forecasting for pharmaceutical supply chains, by integrating predictive modeling, explainability, and real-world deployment. Machine learning methods have the potential to have a major impact on drug order fulfillment, shortages, and patient care, said the findings, which believe that healthcare supply chains, including those for drugs, can become more resilient and driven by data.

References

1. Shen, H., Deng, Q., Lao, R., & Wu, S. (2016). A case study of inventory management in a manufacturing company in China. Nang Yan Business Journal, 5(1), 20-40.
2. Bilal, A. I., Bititci, U. S., & Fenta, T. G. (2024). Challenges and the Way Forward in Demand-Forecasting Practices within the Ethiopian Public Pharmaceutical Supply Chain. Pharmacy, 12(3), 86.
3. Jobira, T., Abuye, H., Jemal, A., & Gudeta, T. (2021). Evaluation of pharmaceuticals inventory management in selected health facilities of West Arsi Zone, Oromia, Ethiopia. Integrated Pharmacy Research and Practice, 1-11.
4. Abbas, A., Cheng, L., & Raza, F. (2023). AI in Healthcare: Applications, Challenges, and Future Prospects. Cosmic Bulletin of Business Management, 2(1), 145-162.
5. Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of management information systems, 24(3), 45-77.
6. F. Pedregosa et al., 'Scikit-learn: Machine learning in python', *J. Mach. Learn. Res.,* vol. 12, pp. 2825–2830, 2011.
7. Dhal, P., & Azad, C. (2022). A comprehensive survey on feature selection in the various fields of machine learning. Applied intelligence, 52(4), 4543-4581.
8. M. L. Waskom, 'Seaborn: statistical data visualization', J. Open Source Softw., vol. 6, no. 60, p. 3021, 2021.
9. Maulud, D., & Abdulazeez, A. M. (2020). A review on linear regression comprehensive in machine learning. Journal of applied science and technology trends, 1(2), 140-147..
10. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
11. Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. Machine learning, 63(1), 3-42.
12. . Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794)
13. Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International journal of forecasting, 22(4), 679-688.
14. Kvålseth, T. O. (1985). Cautionary note about R 2. The American Statistician, 39(4), 279-285.
15. Aas, K., Jullum, M., & Løland, A. (2021). Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. Artificial Intelligence, 298, 103502.
16. Musgrave, K. (2023). Correctly evaluating machine learning algorithms. Cornell University.
17. Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of machine learning research, 13(2).