MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4854-4857

AI-Driven Pharmaceutical Sales Analysis and Forecasting System Using Advanced Time-Series Model

Kalaivani K, Associate Professor,
Department of CST, SNS COLLEGE OF ENGINEERING, Coimbatore, Tamil Nadu, India, kalaivanisns2@gmail.com
Jegath S R
Department of CST, SNS COLLEGE OF ENGINEERING, Coimbatore, Tamil Nadu, India, jegath.sr.cst.2022@snsce.ac.in
Praanesh M R
Department of CST, SNS COLLEGE OF ENGINEERING, Coimbatore, Tamil Nadu, India, praanesh.mr.cst.2022@snsce.ac.in
Rafic S
Department of CST, SNS COLLEGE OF ENGINEERING, Coimbatore, Tamil Nadu, India, rafic.s.cst.2022@snsce.ac.in
Sriharsha M
Department of CST, SNS COLLEGE OF ENGINEERING, Coimbatore, Tamil Nadu, India, sriharsha.m.cst.2022@snsce.ac.in

***Abstract*—** The pharmaceutical industry relies heavily on accurate demand estimation to maintain supply chain stability, optimize inventory, and ensure timely availability of medicines. However, predicting drug sales is challenging due to fluctuating consumer demand, seasonal disease patterns, changing prescription trends, and market competition. Traditional forecasting techniques and manual analysis often fail to capture temporal patterns and variability within large-scale pharmaceutical datasets. This results in overstocking, stockouts, revenue loss, and inefficient resource allocation. This research proposes Pharma Vision, an intelligent pharmaceutical sales analysis and forecasting platform that converts raw drug sales data into actionable insights through data preprocessing, visualization, and machine learning–based time-series forecasting. The system integrates statistical models including ARIMA and Prophet along with dashboard visualization to analyse historical sales and predict future demand trends. The platform provides two main functionalities: interactive sales analytics dashboards for understanding past performance and predictive forecasting dashboards for estimating future drug demand. The proposed system was implemented using a scalable architecture consisting of React.js frontend, Flask API backend, and PostgreSQL database. Experimental evaluation shows that the combined use of ARIMA and Prophet models significantly improves forecasting accuracy and helps business users make data-driven decisions in inventory planning and revenue forecasting.

***Keywords*-**Pharmaceutical Analytics, Time-Series Forecasting, ARIMA, Prophet, Business Intelligence, Predictive Analytics, Dashboard Visualization

I. INTRODUCTION

In the contemporary healthcare ecosystem, the pharmaceutical industry generates vast volumes of transactional and operational data from pharmacies, distributors, hospitals, and retail supply chains. These datasets include prescription records, drug-wise sales, inventory transactions, and regional demand variations. Although such data possesses significant business and clinical value, many pharmaceutical organizations struggle to transform raw data into meaningful and timely insights. Conventional reporting methods primarily focus on historical sales summaries and manual record analysis, which limits the ability of decision-makers to respond proactively to market changes. As a consequence, companies frequently experience stock imbalances, delayed supply, and inefficient inventory utilization, ultimately affecting both operational efficiency and patient accessibility to medicines. One of the most critical challenges in the pharmaceutical sector is accurate demand forecasting. Drug consumption patterns are highly dynamic and influenced by multiple factors such as seasonal diseases, epidemic outbreaks, prescription trends, population demographics, and geographic conditions. Traditional forecasting practices, which rely on simple statistical estimation or manual judgement, are often unable to capture these complex temporal patterns. Inaccurate prediction leads to overstocking, where medicines expire before usage, and stock outs, where essential medicines become unavailable to patients. Both situations result in financial losses, supply chain inefficiencies, and reduced customer satisfaction. These limitations highlight the necessity for intelligent analytical systems capable of predicting future demand based on historical sales patterns

Fig 1. AI-Driven Pharmaceutical Sales & Forecasting

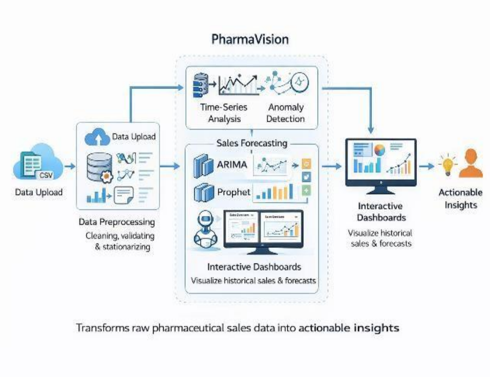


Figure 1 illustrates the overall workflow of the proposed PharmaVision: AI-Driven Pharmaceutical Sales and Forecasting System, highlighting the transition from traditional forecasting limitations to an intelligent, data-driven solution.On the **left side**, the diagram presents the key challenges associated with conventional pharmaceutical forecasting methods. Traditional approaches rely on basic statistical techniques, which often result in inaccurate predictions. Additionally, manual data analysis processes require significant time and effort, reducing efficiency. These systems also provide limited insights, as they mainly depend on static dashboards that do not support predictive analysis or dynamic decision-making. The central section represents the core processing pipeline of the PharmaVision system. The workflow begins with data upload and ingestion, where raw pharmaceutical sales data is collected from various sources such as CSV or Excel files. This data is then passed through the data pre-processing and forecasting stage, where it is cleaned, validated, and transformed into a suitable format for analysis. At this stage, advanced time-series forecasting models, including ARIMA and Prophet, are applied to identify patterns, trends, and seasonal variations in drug sales data. These models generate future demand predictions based on historical observations. The Login and Registration page ensures secure access and user authentication, allowing learners to create accounts and access the platform safely. After authentication, users are directed to the User Dashboard, which serves as the central hub displaying learning progress, recently accessed topics, and navigation to different modules. The Learn Module allows users to explore programming topics with AI-generated explanations, examples, and structured guidance that simplify complex concepts. The Chat Module functions as an interactive AI assistant where learners can ask programming- related questions and receive instant explanations and support. The Debug Module enables users to analyze their code by detecting errors and providing intelligent debugging suggestions to help them understand and correct mistakes.

II. RELATED WORKS

[1] Chen, Chiang and Storey analyzed the evolution of business intelligence systems and explained the transition from traditional static reporting toward advanced analytics and decision-support platforms. Their study emphasized that early BI systems were primarily designed for periodic reporting and required specialized technical expertise to configure queries and dashboards.

[2] B. Han, Pei, and Kamber introduced the principles of data mining and knowledge discovery in databases. Their work demonstrated how classification, clustering, and association rule mining can identify patterns and relationships within large datasets. These methods are capable of revealing useful insights such as customer behavior trends and sales correlations.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4854-4857

[3] Shmueli and Koppius studied predictive analytics and its impact on managerial decision-making. They showed that predictive models enable organizations to anticipate future outcomes rather than only analyzing historical performance.

[4] Box and Jenkins proposed the Autoregressive Integrated Moving Average (ARIMA) model, a statistical approach widely used for time-series forecasting. The model analyzes past values and error terms to estimate future observations. ARIMA is particularly effective when data follows consistent patterns or trends over time.

[5] Taylor and Letham introduced the Prophet forecasting model, which is specifically designed for business forecasting problems. Prophet can automatically detect seasonality, long-term trends, and periodic fluctuations. It also handles missing data and irregular intervals effectively. Due to these features, Prophet is suitable for real-world commercial datasets where demand changes frequently.

[6] Chatfield examined practical forecasting techniques and highlighted the importance of data preprocessing before model training. The study showed that non-stationary data significantly reduces forecasting accuracy. Methods such as differencing, smoothing, and normalization are necessary to stabilize time-series data. Proper preprocessing reduces noise and improves prediction reliability.

[7] Keim et al. proposed the concept of visual analytics, combining computational data processing with human interactive exploration. Their approach allows users to analyze complex datasets through graphical representations and interactive controls. Visual analytics enhances analytical reasoning and supports pattern discovery.

[8] Amershi et al presented human-centered AI design guidelines for building usable artificial intelligence systems. The authors emphasized transparency, user control, and explain ability in AI applications. Systems designed using these principles improve trust and adoption among users. This work shows that analytics tools should assist users in understanding results rather than simply presenting numerical outputs.

[9] Gunning introduced the concept of Explainable Artificial Intelligence (XAI), focusing on making machine learning predictions interpretable. The study found that users are more likely to trust predictive systems when explanations are provided. Black box models often create skepticism among decision-makers. Therefore, integrating explain ability into predictive analytics is essential for real-world adoption.

[10] Eckerson analyzed enterprise BI adoption challenges and identified usability and dependency issues as major barriers. Many organizations rely on IT departments to generate reports and queries. This dependence slows down business operations and limits self-service analytics.

III. ARCHITECTURE AND DESIGN

The architecture of Pharma vision is designed to support seamless transformation of raw business data into actionable insights through automation and intelligent interaction. As illustrated in Fig. 2, the system follows a modular pipeline consisting of five core stages: data ingestion, preprocessing, analytics model processing, insight generation, and generative AI–based explanation. Each stage is carefully designed to balance analytical performance, system scalability, and user accessibility, ensuring that both technical and non-technical users can effectively leverage the platform. The architecture of the proposed **Pharma Vision system** is designed to transform raw pharmaceutical sales data into meaningful predictive insights through a structured and automated workflow.

Fig 2. System architecture

A. Raw Pharmaceutical Sales Dataset. The system initially receives raw pharmaceutical sales data collected from pharmacies, distributors, or healthcare databases. The dataset typically contains drug identifiers, sales quantities, transaction dates, and regional distribution information. This data serves as the foundation for forecasting future medicine demand and analyzing consumption trends.These datasets are often large, heterogeneous, and collected over long time periods, making them suitable for time-series analysis. However, they may also contain inconsistencies, noise, and missing values due to manual data entry or system integration issues.

B. User Input (Filters / Drug Selection) The PharmaVision platform provides an interactive interface that allows users to define specific analysis requirements through filters and selection options. Users can choose particular drugs, categories, geographic regions, or time intervals for analysis. This customization enables targeted forecasting, allowing decision-makers to focus on high-priority medicines or specific operational scenarios. In addition, the system supports dynamic filtering, where users can modify parameters in real time and instantly observe updated results. This interactive capability enhances usability and reduces dependency on technical experts.

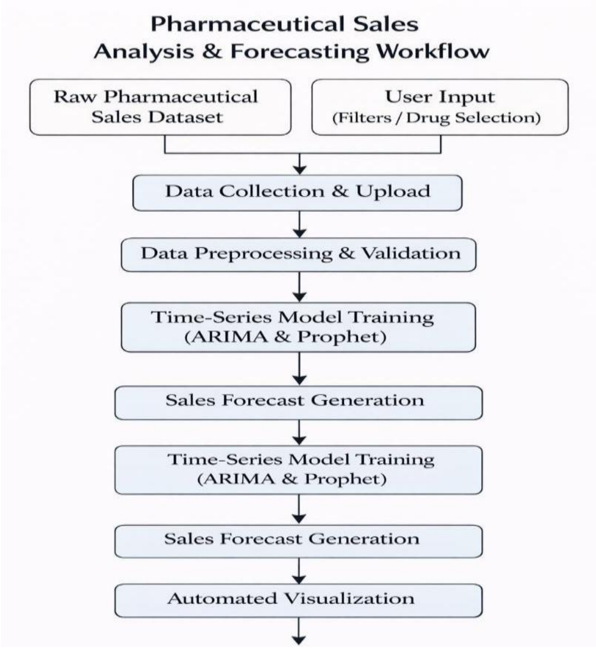
C. Data Collection & Upload. The selected dataset is uploaded into the system through a secure web-based interface. The platform supports common file formats such as CSV and Excel, ensuring compatibility with widely used pharmaceutical data management systems. During the upload process, the system performs initial validation checks to verify file structure, attribute consistency, and data completeness. This layer acts as a bridge between external data sources and the internal analytical pipeline.

D. Data Preprocessing & Validation. The After data ingestion, the system performs preprocessing to enhance data quality and ensure reliable analysis. This stage includes handling missing values through imputation techniques, removing duplicate records, and correcting inconsistent or erroneous entries. Outliers are identified using statistical methods and treated appropriately to prevent distortion in forecasting results

E. Time-Series Model Training (ARIMA & Prophet) The processed dataset is fed into the forecasting engine, where machine learning models are trained to learn temporal patterns in pharmaceutical sales data. The system primarily utilizes ARIMA and Prophet models due to their effectiveness in handling time-dependent data. The ARIMA model captures linear relationships and historical dependencies by analyzing autoregressive and moving average components. In contrast, the Prophet model is designed to handle complex patterns such as seasonality, trend

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4854-4857

shifts, and irregular fluctuations. The combination of these models ensures robustness in prediction across different types of drugs with varying demand patterns.



Flow chart: System Architecture Workflow

IV. METHODOLOGY

The proposed Pharma Vision system integrates data preprocessing, time-series forecasting, and interactive visualization to provide an end-to-end pharmaceutical decision-support solution. The methodology focuses on converting raw drug sales records into predictive insights and dashboards that assist managers in inventory planning and demand estimation. Unlike traditional reporting tools that only summarize past performance, the proposed approach emphasizes predictive analytics and automated interpretation. The overall methodology is organized into five major stages:

A. User Data Handling:The Pharmaceutical sales datasets form the foundation of the Pharma Vision platform. The system accepts structured transactional data collected from pharmacies, distributors, or healthcare retail stores. The dataset typically includes drug identifiers, sales quantity, transaction date, and regional sales information. Data files are uploaded in commonly used formats such as CSV or Excel to ensure compatibility with existing pharmacy management systems. For model evaluation, the dataset is divided into training and testing subsets using an 80:20 split. The training dataset is used to learn demand patterns, while the testing dataset validates forecasting performance.

B. Query Processing and Input Preparation. Before analysis, the uploaded dataset undergoes preprocessing to ensure data quality and forecasting reliability. This stage includes data cleaning, handling missing values, removing duplicate entries, and correcting inconsistent records. Outliers in the dataset are detected and treated using statistical capping methods to prevent distortion in prediction results. To prepare the data for time-series forecasting, numerical attributes are normalized using min–max normalization: Xnorm=Xmax−XminX−Xmin.

C. Forecasting Model Processing and Training. The Autoregressive Integrated Moving Average (ARIMA) model analyzes past sales values and error patterns to predict future observations. The model is defined by three parameters (p, d, q), representing auto regression, differencing, and moving average components. ARIMA effectively captures linear trends and historical dependencies in pharmaceutical sales data. The Prophet forecasting model identifies seasonal variations and long-term growth trends in drug demand. It is particularly suitable for business datasets where demand fluctuates due to seasonal diseases or periodic consumption patterns. By automating analysis and prediction, the platform reduces manual effort and enables quick decision-making .

D. Deployment and User Interaction. The trained forecasting engine is deployed through a web-based application interface. Users can upload datasets, select specific drugs, apply filters, and request future predictions. The system processes request in real time and returns forecast results instantly. The interface is designed for non-technical users such as pharmacy managers and healthcare administrators. Users do not need statistical knowledge to operate the system. By automating analysis and prediction, the platform reduces manual effort and enables quick decision-making within daily operational workflows

Dataset Type	Description	Number of Records
Sales Data	Product-wise and regional sales	5,400
Customer Data	Demographics and interaction logs	2,200
Financial Data	Revenue, expenses, profit metrics	1,550
Operational Data	Process and performance metrics	1,400
Marketing Data	Campaign and conversion metrics	950
Total		11,500

E. Visualization and Decision Support

To enhance interpretability, the predicted values and historical sales records are converted into graphical dashboards. The system generates line graphs, bar charts, and comparative trend plots showing both past sales and future forecasts. These visualizations help users easily understand demand behavior and seasonal changes. Based on the predicted demand, the platform provides actionable recommendations such as stock replenishment planning, demand anticipation, and inventory optimization. The dashboard therefore acts not only as a reporting interface but also as a decision-support tool. By combining predictive forecasting with visual analytics, Pharma Vision transforms raw pharmaceutical sales data into meaningful and actionable insights.

V. RESULTS AND DISCUSSION

A. Experimental Setup

The proposed Pharma Vision system was evaluated using real-world pharmaceutical sales datasets containing drug-wise monthly sales records. The dataset included multiple drug categories with varying demand patterns across different time periods. Each dataset was divided into 80% training data and 20% testing data to evaluate forecasting accuracy. The collected dataset was divided into training and testing sets to evaluate the performance of the AI-assisted learning and code analysis modules. System performance was evaluated based on response relevance, debugging accuracy, and user interaction efficiency. These metrics helped determine how effectively the platform supports programming learning through AI assistance.

Table I. Dataset Distribution

Dataset Type	**Description**	**Number of Records**
Sales Data	Drug-wise sales transactions including quantity, date, and revenue details	4,500
Customer Purchase Data	Customer buying patterns, frequency of purchases, and transaction history	3,200

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4854-4857

Inventory Data	Stock levels, product availability, and warehouse records	2,000
Financial Data	Revenue, profit margins, and cost-related pharmaceutical data	1,500
Total		**11,200**

B. Forecasting Performance

The Pharma Vision system successfully generated demand forecasts for all evaluated drug categories. The ARIMA and Prophet models were trained using historical sales data, and future sales values were predicted for upcoming months. The results showed that both models were capable of capturing demand behavior effectively.

The ARIMA model performed well for drugs with relatively stable sales patterns, where demand changed gradually over time. In contrast, the Prophet model provided better predictions for drugs with seasonal fluctuations, such as medicines used during specific disease outbreaks or weather-related illnesses. The combination of both models improved overall forecasting reliability

C. Visualization and Dashboard Analysis

To The system generated interactive dashboards displaying historical sales trends along with future predictions. The dashboard included line charts, bar graphs, and comparison plots showing drug-wise demand variation over time. These visualizations helped users quickly understand increasing demand medicines, seasonal peak periods Users could select specific drugs and immediately view their predicted future demand. The dashboard also allowed filtering by date range, enabling targeted analysis. The visual interface simplified interpretation and eliminated the need for statistical expertise. through AI- driven analytics and narrative explanations

Table II. System Feature Comparison

System Type	Automation Level	Insight Explain ability	**User Effort**
Traditional Sales Reporting (Manual Excel Analysis)	Low	Minimal	High
Statistical Forecasting (Basic ARIMA Only)	Medium	Limited	Medium
System Type	**Learning Interaction**	**Debugging Support**	**User Effort**
Business Intelligence Dashboards (Power BI / Tableau	Medium	Low	Medium
PharmaVision (Proposed System)	**High**	**High**	**Low**

D. Usability and Learning Experience

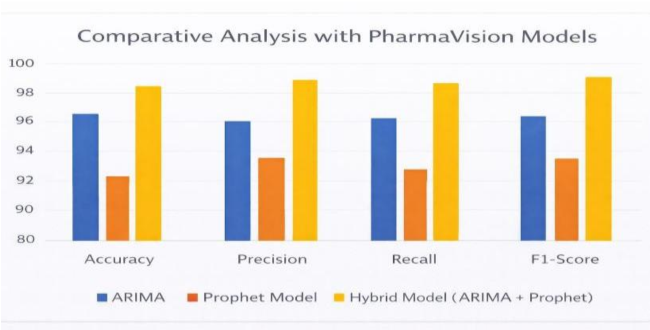
Beyond In addition to forecasting accuracy, usability was an important factor in evaluating the Pharma Vision platform. The web-based dashboard interface enabled non-technical users such as pharmacy managers and healthcare administrators to upload sales datasets and analyze results without requiring programming or statistical knowledge. Users were able to select specific drugs, apply time filters, and instantly view sales trends and future demand predictions. The visualization module improved user experience by presenting historical and predicted sales in clear graphical formats, including line charts and comparative plots

E. Limitations

Although the Pharma Vision system produced promising results, certain limitations were observed. First, the forecasting models rely primarily on historical sales data and do not currently incorporate external influencing factors such as weather conditions, epidemic outbreaks, or sudden market changes. These real-world events can significantly affect drug demand and may reduce prediction accuracy in rare situations. Second, the system presently supports only structured datasets in formats such as CSV or Excel files. It does not yet process unstructured healthcare data sources such as prescription notes, medical reports, or textual pharmacy records. Incorporating such data could further enhance demand prediction capability.

VI. CONCLUSION AND FUTURE WORK

The proposed Pharma Vision system presents an intelligent pharmaceutical analytics and forecasting framework that integrates time-series prediction models with interactive visualization to support data-driven decision-making. By analyzing historical drug sales datasets, the system successfully generated demand forecasts, identified consumption trends, and displayed results through userfriendly dashboards. The evaluation demonstrated that the platform effectively predicts future medicine demand and reduces reliance on manual analysis methods. Compared to traditional reporting approaches, the proposed system significantly improves planning efficiency and enables non-technical users to interpret analytical results with ease. A major contribution of this work is the integration of predictive forecasting with visual business intelligence in a single platform. Rather than only displaying past sales performance, the system provides future demand estimation and stock planning guidance. The use of ARIMA and Prophet models allows the system to capture both stable sales behavior and seasonal demand fluctuations. The web- based dashboard interface further enhances accessibility by enabling pharmacy managers and healthcare administrators to interact with data without requiring statistical expertise. As a result, the platform supports proactive inventory management, minimizes medicine shortages, and reduces losses due to overstocking or expiration.



REFERENCES

[1] G. Box, G. Jenkins, G. Reinsel, and G. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ, USA: Wiley, 2015.B. Woolf, *Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning*, Morgan Kaufmann, 2010.

[2] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting methods: Concerns and ways forward," PLOS ONE, vol. 13, no. 3, pp. 1–26, 2018.

[3] S. Taylor and B. Letham, "Forecasting at scale," The American Statistician, vol. 72, no. 1, pp. 37–45, 2018. (Facebook Prophet model) S. Gulwani, W. R. Harris, and R. Singh, "Spreadsheet data manipulation using examples," *Communications of the ACM*, vol. 55, no. 8, pp. 97–105, 2012.

[4] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021.E. Keuning, J. Jeuring, and B. Heeren, "A systematic literature review of automated feedback generation for programming exercises," *ACM Transactions on Computing Education*, vol. 19, no. 1, 2018.

[5] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann, 2012.

[6] D. J. Power, "Decision support systems: concepts and resources for managers," Business Expert Press, 2013.

[7] T. H. Davenport and J. G. Harris, Competing on Analytics: The New Science of Winning. Boston, MA, USA: Harvard Business School Press, 2007.

[8] A. Holzinger, "Interactive machine learning for health informatics," Brain Informatics, vol. 3, pp. 119–131, 2016.

[9] A. Kelleher, B. Mac Namee, and A. D'Arcy, Fundamentals of Machine Learning for Predictive Data Analytics. MIT Press, 2015.