International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5012 - 5016
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0659

AI Powered Business Forecasting System Using ML and Dashboard

Swathi P[1], Shwetha KR[2], Divya G S[3]
[1]Student, Department of Computer Science and Engineering, AMCEC, India.
[2,3]Assistant Prof, Department of Computer Science and Engineering, AMCEC, India.
Emails: *Swathiprabhakar7113@gmail.com[1], kr.shwetha12@gmail.com[2], divyags.siddaraj@gmail.com[3]*

Abstract
This research work presents an advanced forecasting framework that aims to address inefficiencies in the supply chain and prevent any inventory imbalance using machine learning algorithms. Built upon an optimally designed XGBoost algorithm, this framework analyzes multi-year transactional information along with high-frequency time-series information for capturing volatile market conditions. The backend algorithmic framework is effectively integrated using a web-based Streamlit application which consolidates all past data along with future sales trends into live metrics. Through experimental analysis, it is demonstrated that the proposed framework empowers retail store managers, independent of programming skills, to set their optimal inventory levels.
Keywords: *Machine Learning, XGBoost, Predictive Analytics, Business Intelligence, Demand Forecasting, Streamlit Interface.*

1. Introduction
Any company requires predicting its further purchases. However, if one makes an inaccurate forecast, he can lose stock or buy extra goods that will never be sold. Manual analysis of the sales history is ineffective since consumer demands usually change. To solve the above problems, we have designed a robust prediction model capable of processing high speed data flows using machine learning algorithms. We have applied an advanced XGBoost algorithm that can find patterns quite easily. It also takes into account the past sales statistics, seasonal variations, and unusual events such as payday or holidays to predict further trends. Our results are presented in an easily readable graphical and tabular format on a convenient Streamlit web page. It is not necessary to know programming languages and coding. It is enough just to perform appropriate actions using forecasts.

1.1.Background and Motivation
Today's business environment generates an enormous amount of data on a daily basis that requires advanced analytics to facilitate long-term planning and allocation of organizational resources. The ability to forecast sales accurately is very helpful because it will enable companies to maintain track of their inventories, allocate funds properly, and predict their financial outcomes based on actual figures. As traditional time series analysis is often unable to handle random market volatility and seasonal dependencies,machine-learning models use multidimensional feature sets to build accurate and stable predictions.

1.2.Objective of Study
The main objective of the research is to develop a decoupled method of forecasting that will allow converting raw data into actionable business intelligence. The process can be accomplished through the following five objectives:

1.2.1. Decoupled Architecture
Creating a separation between intense computation at the backend to prevent any performance issues on the frontend side.

1.2.2. Data Filtering Pipeline
Building a pipeline of data cleansing for removing random noises and extracting seasonal patterns in the data.

1.2.3. Modelling
Using the optimized XGBoost algorithm that can capture behavioural change while maintaining the consistency of data structure.

1.2.4. Simulation Scenarios
Modelling the market behaviour using machine learning baseline in cases when the data is insufficient.

1.2.5. Dashboard

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5012 - 5016
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0659

Combining all the above models in one web dashboard.

2. Literature Survey

Pervasive computing's wide application has led to a significant restructuring of the processes underlying retail logistics and stock management. The use of the old-fashioned manual system of updating accounts through spreadsheets often results in important execution delays and tracking mistakes. On the other hand, the shift towards automated computing models helps retailers design objective systems of predicting future needs for products. According to industry studies, the integrated model of automated data pipelines, cloud services, and business intelligence reporting systems effectively resolves all logistical challenges resulting from the dynamic changes in consumer shopping behaviours. Decision-tree ensembles represent an ideal methodology in this particular field of analysis.

2.1.Summary of Literature Survey

This shift from manual spreadsheet solutions towards automation in forecasting and analysis has transformed how retail planning works. XGBoost [1] has mathematically proven that tree-based modelling is efficient at handling big datasets, although it requires human intervention in terms of practical decision-making. Implementing forecasting pipelines through Python allows one to tackle changes in the market much more efficiently [2], [3]. Incorporating advanced analytics raises issues related to the accessibility of systems to managers without technical knowledge. For this reason, the modern literature focuses on designing user-friendly dashboards based on transactional data [4],[6] without the need for data scientists' expertise. In addition, the predictive power of any algorithmic framework will be essentially limited by the quality and structure of input datasets used. Although decision tree machine learning systems substantially increase the precision of predictions, data cleansing still is a major obstacle in practice [5]. Resolving this problem in real-world implementations becomes possible due to automated data pipeline creation [7]. Regarding temporal pattern analysis, sequential networks implemented in Python are often integrated with boosting algorithms which fix errors in local optimizations. On the backend side, working with data flow in large volumes without computational overload calls for using Pandas and NumPy frameworks. Breaking down the metrics related to datetime into fine-grained features including diurnal, weekly, and seasonal components prevents mathematical errors and helps to minimize computational time [11]. Finally, strict validation practices should be mandatory to test the model's generalizability to avoid overfitting [12].

3. Related work

3.1.Business Forecasting and Sales Prediction

Business forecasting enables organizations to anticipate future sales in order to plan their operations accordingly. Machine learning algorithms, as pointed out by Brownlee [1], enhance the accuracy of such forecasts by considering past sales data and determining significant patterns therein. Singh [2] developed an artificial intelligence-based system for prediction of sales which outperforms any conventional method. As highlighted by Liu [3], machine learning models help to determine sales trends and seasonality through analysing data.

3.2.XGBoost-Based Forecasting Models

Business forecasting enables organizations to anticipate future sales in order to plan their operations accordingly. Machine learning algorithms, as pointed out by Brownlee [1], enhance the accuracy of such forecasts by considering past sales data and determining significant pattern therein Singh [2] developed an artificial intelligence based system for prediction of sales which outperforms any conventional method. As highlighted by Liu [3], machine learning models help to determine sales trends and seasonality through analyzing data.

3.3.Dashboard-Based Business Analytics

Business dashboards help users interpret forecasting results. They use charts, graphs and tables. Interactive dashboards make it easy to evaluate sales data and predictions. Streamlit is a tool. It is simple and easy to use. By combining machine learning predictions, with dashboard visualization businesses can track performance. They can make decisions. The proposed system uses a Streamlit dashboard. It presents forecast results and business insights in a way. This makes it easy to comprehend the data.Shown as Figure 1 A Smart Sales Forecasting System Built on Data, Patterns, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5012 - 5016
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0659

Predictions.

4. System Architecture and Design

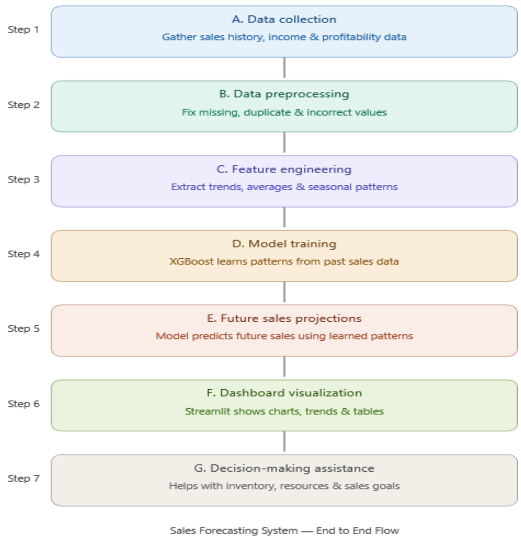

Figure 1 A Smart Sales Forecasting System Built on Data, Patterns, and Predictions

4.1.1. Collection of Sales History Data and Other Data

Initially, it is important to gather sales history data as well as other data related to income in order to assess how well the business is doing in regard to profitability.

4.1.2. Data Preprocessing

The data is refined. Missing, repeated and wrong values are eliminated.

4.1.3. Feature Engineering

Sales data is examined to make predictions more accurate. Past sales, averages and day, month, year patterns are considered.

4.1.4. Model Training

XGBoost looks at all the data and learns from it.

4.1.5. Future Sales Projections

After learning from old data, the model guesses what sales might look like next.

4.1.6. Dashboard Visualization

Results are shown on a Streamlit dashboard with graphs and tables.

4.1.7. Decision Making Assistance

Everything is displayed on Streamlit in a simple way. Users can see sales movement through graphs and tables that are easy to read.

4.2.Integration of the AI Model

We extract details like year, month, day and weekday from the data. We also verify if it is the last day of the month. After that we analyse old sales numbers from a few days, a week and a month back. Then we derive moving averages to measure how sales varied over time[8].

4.3.Interactivity and Visualization

Anyone can use this app. No technical knowledge is required. Just select a time period, provide some basic details and the results appear on their own. Everything is presented in a simple and clear way.

4.4.Efficiency Estimation of the Proposed Model

The data was divided into two sections. One section was used to train the model and the other section was used to evaluate how good its guesses were against real numbers. Two things were measured to assess the model. First is RMSE which shows by how much the predicted numbers were different from the real ones. The smaller the number the better it is. Second is R^2 which shows how well the model captured what was going on in the data. The higher this number the better the model has performed[9].

5. Testing and Performance Analysis

The evaluation of the efficiency of the implemented XGBoost algorithm implies using real sales data rather than fake one in the testing phase. This test is how we confirm if the model is working correctly.

5.1.Technology Stack

The project was developed using Python. Before forecasting, a few things were organized to help the model like sales patterns, averages, and the month and day details.Old data was used for training. New data was used for testing. This is just like real life you gain knowledge from the past to predict the future.Three main tools were utilized are Pandas and NumPy, XGBoost[10].

5.2.Performance Metrics

Table 1 Performance Evaluation of XGBoost Model

Metric	Value
RMSE	66.5822
R^2 Score	0.7703
Algorithm	XGBoost

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5012 - 5016
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0659

The empirical performance results of the optimized XGBoost regression model from the validation data set revealed that the RMSE was 66.5822 while the R2 was 0.7703. R2 is the indicator of how much variance is accounted for by the fitted model. Thus, according to the above R2 value, it means that the fitted model covers 77% of the variance in the data stream. In addition, the lower RMSE value indicates that there is very little difference between the actual data and the predictions from the model.

5.3.Forecasting Results

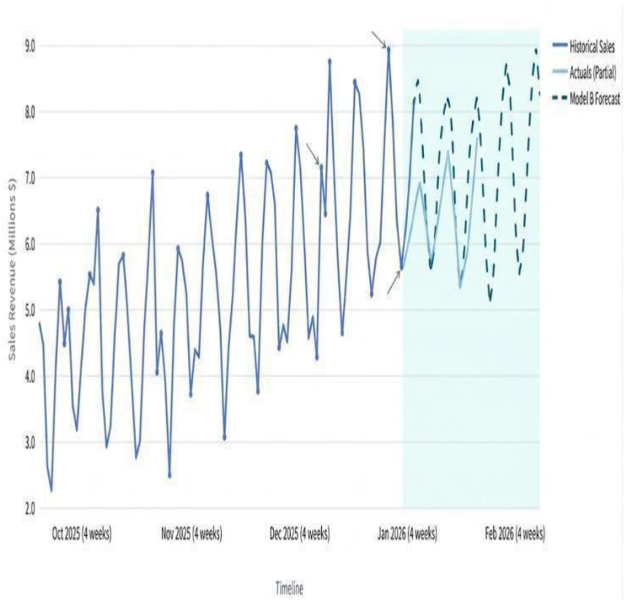

Figure 2 Sales forecasts and time trends from XGBoost Optimisation

The figure below illustrates the visualisation of the forecasting results on two different lines of data. The blue line represents past sales. The red line is what sales might appear like in the next 30 days. You can see sales rise at the start of every week and the forecast reflects that same pattern. The chart makes it easy to identify what is approaching next[13].

5.4.Dashboard Analysis

A simple dashboard was developed using Streamlit to display the forecast results. It is easy to use and needs no technical setup. Sliders were incorporated so the user can just drag and choose the time period they want to view. That is all they need to do.As shown in Figure 3 Interface of Streamlit for Business Forecasting and Predictive Analytics

5.5.Performance Discussion

Figure 3 Interface of Streamlit for Business Forecasting and Predictive Analytics

The interface works as a control plane for all of the processes in the architecture. At the top of the interface, three metric cards display indicators of the system such as total number of records in the active dataset, sales speed[14] in the last 30 days, and time passed since the last database synchronization. Under the metric widgets, the main visualization interface presents the comparison of the predictions from the model and real values of transactions. For the logistics and supply chain analysis, there is a table with detailed information under the visual plots.

Conclusion and Future Scope

Findings from the empirical analysis in this study reveal that using an enhanced gradient boosting approach together with appropriate temporal feature engineering techniques produces a highly reliable methodology for demand forecasting. The developed approach does not depend on the analysis of time series data in its pure form; rather, it constructs structural features in an automated way using a lag matrix, rolling windows, and calendar indicators. As such, the automated transformation of data by the system helps in identifying different seasonalities at multiple levels contained in dense transactions data. Through validation against actual operational history, minimal error rates were recorded, demonstrating the reliability of the XGBoost engine.

Scope for the Future

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5012 - 5016
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0659

Further improvements of this study will involve incorporating multiple external datasets in a multimodal way, such as competitor market activity, macroeconomic indicators, and consumer purchases described in our architecture, directly into the process of semantic feature engineering. As a means of further improvement of the tree-based optimization algorithm of XGBoost regressor, in future iterations, sequential networks will be introduced in order to refine features locally and to address complex temporal dependencies. Lastly, our architecture will be enhanced further by including the data processing orchestration layer in order to support language model calculations for the vector database.

Reference

[1]. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
[2]. J. Brownlee, Machine Learning for Time Series Forecasting with Python, Machine Learning Mastery, 2020.
[3]. H. Liu, "Time Series Forecasting for Business Intelligence Using Machine Learning Models," Elsevier Journal of Information Systems, vol. 28, no. 4, pp. 210–220, 2021.
[4]. R. Gupta, "Intelligent Business Forecasting Using XGBoost and Predictive Analytics," International Journal of Computer Science and Engineering Research, vol. 12, no. 2, pp. 101–109,2023.
[5]. I. H. Witten, E. Frank, and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2017.
[6]. A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow, O'Reilly Media, 2019.
[7]. F. Chollet, Deep Learning with Python, Manning Publications, 2021.
[8]. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, 2009.
[9]. W. McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy and Jupyter, O'Reilly Media, 2022.
[10]. M. Kuhn and K. Johnson, Applied Predictive Modeling, Springer, 2018.
[11]. G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning, Springer, 2021.
[12]. C. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[13]. L. Breiman, "Random Forests,"
[14]. Machine Learning Journal, vol. 45, no.1, pp. 5–32, 2001.J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2011.