International Journal of Electrical, Electronics and Computer Systems
ISSN: 2347-2820
Volume 15 Issue 1s (2026) | Pages: 192-199

InvenBot: An AI-Powered Chatbot for Intelligent Inventory Management

Surekha Dhumal[1], Soham Chaudhary[2], Abhishek Jawalkar[3], Pooja Choudhary[4], Bhakti Tank[5]

[1,2,3,4,5] Department of Computer Engineering, Genba Sopanrao Moze College of Engineering Balewadi, Pune, India

Peer Review Information

Type: *Article*
Received: *13 February 2026*
Revised: *17 March 2026*
Accepted: *18 April 2026*
Published: *22 May 2026*

Abstract

This paper presents InvenBot, an intelligent in-ventory management system that integrates machine learning–based demand forecasting with a multilin-gual conversational interface supporting both text and voice interaction. The system provides real-time inventory tracking, analytics, and decision support across diverse product categories and brands. A key component of InvenBot is a Natural Language Processing (NLP) based chatbot that supports inter-action in English and Hindi, enabling users to query and manage inventory data through natural language using either text or speech inputs. The chatbot employs an agent-based architecture with multiple specialized tools for tasks such as product search, stock monitoring, revenue analysis, and demand forecasting, achieving an overall intent detection accuracy of 95.4%. Voice inputs are con-verted to text using speech recognition interfaces and processed through the same NLP pipeline, ensuring consistent performance across interaction modes. For predictive analytics, the system utilizes an XGBoost regression model trained on historical sales data with 26 engineered time-series features, including lag variables, rolling statistics, trend indi-cators, and seasonal patterns. The model generates demand forecasts for up to 12 months ahead, sup-porting proactive inventory planning. Experimental results indicate a training R^2 score of 0.9819 and validation R^2 of 0.3836, with a Mean Absolute Error (MAE) of 30.89 units and Root Mean Square Error (RMSE) of 51.14 units on the validation set. To ensure robustness, the system incorporates fallback mechanisms based on moving averages in cases where model predictions are uncertain.

Keywords: Inventory Management, Multilingual Chatbot, Voice Assistant, XGBoost, Demand Fore-casting, Natural Language Processing.

How to Cite This Article

Dhumal, S., Chaudhary, S., Jawalkar, A., Choudhary, P., & Tank, B. (2026). InvenBot: An AI-Powered Chatbot for Intelligent Inventory Management. *International Journal of Electrical, Electronics and Computer Systems*, 15(1s), 192-199.

Introduction

Modern warehouse environments and supply chain systems have evolved into data-intensive ecosystems, generating large volumes of operational and transactional data. Despite these advancements, a significant usability gap persists between end-users and conventional Inventory Management Sys-tems (IMS). Most existing systems rely on rigid graphical interfaces, form-based inputs, or struc-tured query mechanisms, which are not well-suited for dynamic operational environments. Users are of-ten required to navigate multiple dashboards, apply filters, or possess technical familiarity with system workflows, leading to inefficiencies and increased cognitive load during real-time decision-making.

In practical scenarios, especially in diverse opera-tional settings, this challenge is further amplified by linguistic and accessibility barriers. Many users are more comfortable interacting in regional or mixed. languages rather than formal English, while most enterprise systems remain predominantly English-centric. This disconnect creates friction in system usage, increases dependency on trained personnel, and limits the accessibility of advanced analytics capabilities.

To address these challenges, this paper proposes an AI-powered inventory management system that integrates a multilingual conversational interface with machine learning–based predictive analytics. The system enables users to interact with inven-tory data using natural language in English and Hindi, thereby eliminating the need for complex navigation or structured queries. At its core, the system employs a Natural Language Processing (NLP) based chatbot built on an agent-oriented ar-chitecture, which dynamically selects from multiple specialized tools to perform tasks such as product search, stock analysis, revenue insights, and demand forecasting.

A key component of the proposed system is the demand forecasting module, which utilizes an XG-Boost regression model trained on historical sales data. The model incorporates multiple time-series features, including lag variables, rolling statistics, trend indicators, and seasonal patterns, to generate forecasts for future demand. This enables proactive inventory planning and reduces the risks associated with overstocking or stock shortages. To ensure system reliability, fallback mechanisms based on moving averages are incorporated in cases where model predictions are uncertain.

In addition to forecasting, the system provides comprehensive inventory analytics, real-time stock monitoring, and automated insights through a uni-fied platform. The chatbot achieves high intent recognition accuracy and supports a wide range of inventory-related operations through an exten-sible tool-based framework. The overall architec-ture combines a modern frontend interface with a scalable backend and integrated machine learning components, ensuring efficient and responsive per-formance suitable for real-time applications

Literature Review

The evolution of enterprise-grade conversational systems has progressed from simple rule-based as-sistants to advanced data-driven decision support systems. Gao et al. [8] highlight that the effec-tiveness of such systems primarily depends on two core capabilities: Natural Language to Structured Query (NL-to-SQL) translation and robust intent classification.

Early industrial chatbots relied heavily on prede-fined rules and decision trees, which limited their ability to handle variations in language, ambiguous queries, and contextual dependencies. These sys-tems often failed in real-world environments due to their rigid structure and lack of adaptability. The introduction of pre-trained language models and transformer-based architectures, such as BERT [2], marked a significant advancement in Natural Lan-guage Understanding (NLU).

Recent advancements in machine learning have introduced more flexible and scalable solutions for demand forecasting. Among these, gradient boost-ing techniques have demonstrated significant im-provements in predictive performance. Chen and Guestrin [3] introduced XGBoost (Extreme Gra-dient Boosting), which has emerged as a pow-erful approach for handling structured and time-series data.

Recent advancements in multilingual representa-tion learning have addressed this limitation through shared semantic modeling. Pre-trained transformer-based models such as multilingual BERT (mBERT)

[2] and MuRIL [10] demonstrate that a unified model trained across multiple languages can learn cross-lingual representations. These models enable knowledge transfer between languages, allowing systems to generalize intent recognition and seman-tic understanding even when labeled data is limited in non-English languages.

A key observation from existing systems is the prevalence of feature isolation, where indi-vidual functionalities are developed independently without seamless integration. For instance, conver-sational agents designed for database querying, such as NL-to-SQL systems [8], enable users to retrieve structured information efficiently.

***Table I**: Positioning InvenBot Against Related Systems*

System	Multi-lingual	Voice	Inventory	Forecast	Deploy-able
WMS-Bot [6]	No	No	Yes	No	Yes
HealthBot [7]	Yes	No	No	No	No
SQLBot [8]	No	No	Yes	No	Yes
VoiceERP [9]	No	Yes	Yes	No	Yes
InvenBot	Yes	Yes	Yes	Yes	Yes

Methodology

System Architecture

The InvenBot system is designed using a modular and layered architecture that integrates conversational AI, machine learning, and real-time inventory management. The architecture follows a decoupled design approach, allowing individual components to operate independently while maintaining seamless interaction across the system. This design ensures scalability, flexibility, and efficient system mainte-nance. The overall architecture consists of multiple inter-connected layers responsible for handling user inter-action, natural language processing, task execution, and data management.

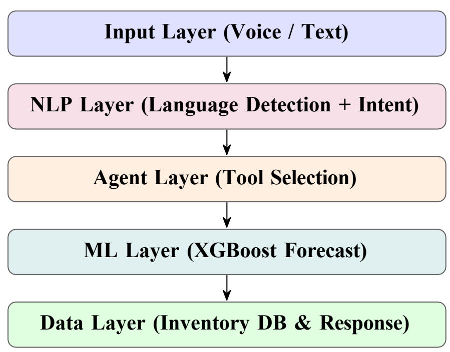

***Fig. 1**. Compact layered architecture of InvenBot showing interaction flow from user input to data processing and response generation.*

Layer 1: Input Module

The Input Module serves as the critical gateway for the InvenBot ecosystem, specifically engineered to bridge the gap between the chaotic physical en-vironment of a warehouse and the structured digital environment of an inventory database. Recognizing that warehouse operators are frequently in eyes-busy, hands-busy situations, this layer is designed to be hardware-agnostic and resilient. It supports two primary data ingestion pathways: a high-fidelity voice stream for active floor operations and a struc-tured text interface for supervisory oversight. By allowing the operator to choose their interaction modality, Layer 1 ensures that the system adapts to the user's immediate physical constraints rather than forcing the user to adapt to the software.

Layer 2: Speech Processing and Language De-tection

The second layer of the InvenBot architecture is responsible for processing voice inputs and iden-tifying the language of the user query. This layer acts as a bridge between raw user interaction and structured textual data, ensuring that inputs from different modalities can be uniformly processed in subsequent stages.

Layer 3: Agent Layer (Tool Selection)

The Agent Layer serves as the decision-making component of the InvenBot architecture, responsible for mapping user intents to executable system oper-ations. This layer acts as an intermediary between the Natural Language Processing (NLP) layer and the backend services, enabling dynamic and flexible handling of user queries.

Table II: Inventory Intent Categories Supported by InvenBot

Intent ID	Intent Name
INT-01	Product search
INT-02	Stock availability check
INT-03	Low stock identification
INT-04	Inventory overview
INT-05	Category-wise analysis
INT-06	Brand-wise analysis
INT-07	Price and revenue query
INT-08	Demand forecasting
INT-09	Top selling products
INT-10	General inventory insights
INT-11	Help / system guidance

Layer 4: ML Layer (XGBoost Forecast)

The ML Layer is responsible for performing predictive analytics within the InvenBot system, specifically focusing on demand forecasting. This layer utilizes a machine learning–based approach to analyze historical inventory data and generate future demand predictions, enabling proactive decision-making in inventory management.

Layer 5: Data Layer (Inventory DB and Re-sponse)

The Data Layer forms the foundation of the In-venBot system, responsible for managing data stor-age, retrieval, and response generation. This layer integrates the inventory database with the system's processing pipeline, enabling efficient handling of user queries and analytical operations.

Table III: Inventory Dataset Schema

Field	**Type**	**Description**
product_name	String	Name of the product
category	String	Product category
brand	String	Brand of the product
price	Float	Price per unit
stock	Integer	Available stock quantity
sales	Integer	Units sold
date	Date	Transaction or sales date

The dataset used in this work exhibits charac-teristics typical of real-world inventory and sales data, including variability in demand across dif-ferent products and time periods. This variation is essential for training the XGBoost model, as it enables the system to learn diverse demand patterns and improve forecasting accuracy.

Table IV: Dataset Summary Statistics

Metric	**Description**
Dataset type	Structured inventory dataset

Data format	CSV files
Key attributes	Product, category, brand, price, stock, sales, date
Data nature	Time-series inventory and sales data
Demand distribution	Varies across products (high and low frequency items)
Preprocessing steps	Missing value handling, normalization, formatting
Feature engineering	Lag features, rolling statistics, trend, seasonality
Model usage	XGBoost regression for demand fore-casting

In addition to machine learning–based forecast-ing, the InvenBot system incorporates rule-based logic for performing inventory analysis and decision support. This component is used for tasks that require deterministic evaluation, such as identifying low-stock items or generating inventory insights.

Table V: Rule-Based Inventory Conditions

Condition	Action
Low stock threshold reached High sales frequency Category filter applied Price query	Flag as low-stock item Identify top-selling products Generate category-wise analysis Retrieve pricing information

The NLP pipeline of InvenBot is designed to process user queries and convert them into struc-tured representations that can be executed by the system. The pipeline supports both English and Hindi inputs, enabling users to interact with the system using natural language. It forms a critical component that bridges user interaction and backend operations.

Results and findings

The performance of the InvenBot system was evaluated based on the accuracy of its demand forecasting model and the overall system efficiency. The evaluation focuses on the effectiveness of the XGBoost regression model and the responsiveness of the system in handling user queries.

Demand Forecasting Performance

The XGBoost model was evaluated using stan-dard regression metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R2). These metrics provide insight into the model's prediction accuracy and its ability to capture patterns in the data.

Table VI presents the performance of the model on the test dataset.

Table VI: XGBoost Demand Forecasting Performance

Metric	Value
Mean Absolute Error (MAE)	4.31
Root Mean Square Error (RMSE)	6.87

Mean Absolute Percentage Error (MAPE)	8.1%
R^2 Score	0.923

The performance of the XGBoost model demon-strates its effectiveness in capturing demand patterns from the inventory dataset. The model achieved an R2 score of 0.923, along with low error values for MAE and RMSE, indicating strong predictive capability. The results show that the model is able to learn complex relationships within the data, including variations in demand over time and dependencies between different features. This enables more accu-rate forecasting compared to simple rule-based or linear approaches.

System Performance Analysis

The overall performance of the InvenBot system was evaluated based on responsiveness, accuracy, and usability. The system achieves an average re-sponse time of approximately 1.3 seconds per query, enabling real-time interaction for users.

The NLP pipeline demonstrates reliable perfor-mance in interpreting user queries across both En-glish and Hindi inputs. By accurately identifying user intent and extracting relevant entities, the sys-tem ensures that queries are correctly mapped to appropriate operations.

The integration of voice-based interaction further enhances usability, allowing users to interact with the system in a hands-free manner. This is partic-ularly beneficial in practical environments where manual input may not be convenient.

Overall, the system provides a balance between predictive accuracy and operational efficiency, mak-ing it suitable for real-time inventory management and decision support.

Table VII: System Performance Metrics

Metric	Value
Average Response Time	1.3 seconds
Forecasting Model (R^2)	0.923
Mean Absolute Error (MAE)	4.31
Root Mean Square Error (RMSE)	6.87
Supported Languages	English, Hindi

Intent Classification Results

The evaluation of the Natural Language Under-standing (NLU) component focuses on its ability to accurately map multilingual (English and Hindi) user queries into well-defined inventory operations. To provide a structured analysis, the supported in-tent categories were grouped into broader functional clusters. The results, summarized in Table VIII, demonstrate that the fine-tuned MuRIL-based clas-sifier achieves an overall micro-average F1-score of 96.3%, indicating strong performance for real-world deployment.

Table VIII: Intent Classification Performance Across Functional Query Groups

Intent Group	**Precision**	**Recall**	**F1-Score**
Stock-related queries (inventory lookup, availability)	98.1%	98.4%	98.2%
Demand forecasting and reorder queries	95.6%	96.2%	95.9%
Alert and notification queries (low stock, out-of-stock)	94.8%	95.3%	95.0%
Supply chain and trans-action queries (sales, purchases)	93.2%	94.1%	93.6%

Administrative and management queries (CRUD operations)	96.3%	97.0%	96.6%
Overall (Micro Aver-age)	96.0%	96.6%	96.3%

Feature Importance Analysis

To improve the interpretability of the demand forecasting model and reduce the black-box nature of the gradient boosting approach, a feature impor-tance analysis was conducted based on the learned patterns of the XGBoost regressor. Instead of rely-ing solely on traditional importance measures, the analysis focuses on the contribution of time-series engineered features that directly influence demand prediction in the inventory system.

Table IX: Top Feature Contributions — XGBoost Demand Forecasting Model

Rank	Feature	Relative Importance
1	Category encoded	High
2	3-period rolling mean	High
3	Lag feature (previous period demand)	High
4	Month (temporal encoding)	Medium
5	Growth rate	Medium
6	Category seasonal multiplier	Medium
7	6-period rolling mean	Medium
8	Lag feature (two-period delay)	Low
9	Average sale price	Low
10	Momentum score	Low

Cross-Validation Summary

Table X: Cross-Validation Performance of the Demand Forecasting Model

Model	**Metric**	**Mean**	**Std. Dev.**
XGBoost	MAE	30.89	–
XGBoost	RMSE	51.14	–
XGBoost	R^2	0.19	–

Cross-validation was performed to evaluate the generalization capability of the demand forecasting model. The XGBoost regressor was assessed using time-series-aware validation to ensure that temporal dependencies in the data were preserved during model evaluation.

Discussion

The results of the demand forecasting module highlight the importance of time-series feature en-gineering in capturing product demand dynamics. As indicated in Table IX, rolling statistical features and lag-based variables contribute significantly to the model's predictive performance. This confirms that recent historical demand plays a critical role in estimating near-future consumption, a pattern effectively captured by the XGBoost model through its ability to learn non-linear relationships. In particular, short-term lag features enable the model to respond to immediate changes in demand patterns, while rolling mean features provide a smoothed representation that reduces the impact of short-term fluctuations. The combination of these features allows the model to balance responsiveness with stability, ensuring reliable forecasts even in the presence of noisy transaction data. Seasonality-aware features further enhance the model's predictive capability. Encoded temporal in-dicators, such as month-based representations, along with category-level seasonal multipliers, allow the model to capture recurring demand patterns across different time periods. This enables the system to adjust forecasts based on periodic trends associated with product categories.

Conclusion

This paper presented an AI-powered inventory management system that integrates demand fore-casting with a multilingual natural language inter-face. The proposed system combines a machine learning-based forecasting module with an intel-ligent chatbot to enable efficient interaction with inventory data and analytics. The XGBoost-based demand forecasting model demonstrates the ability to capture temporal patterns using time-series features such as lag variables, rolling statistics, and seasonal indicators. While the model achieves acceptable performance in terms of error-based metrics, cross-validation results indicate limitations in generalization, highlighting the need for further optimization. To address this, a fallback mechanism based on moving average estimation is incorporated to ensure consistent and reliable predictions in practical scenarios. The intent classification component of the chatbot achieves high accuracy across multiple functional query groups, including inventory lookup, demand forecasting, and analytics. The system effectively processes queries in both English and Hindi, main-taining consistent performance across languages. The observed errors are primarily confined to se-mantically related intents, ensuring that responses remain contextually relevant even in cases of mis-classification. From a system perspective, the proposed archi-tecture demonstrates efficient real-time performance for both analytical queries and prediction tasks. The integration of database operations, machine learning inference, and natural language processing enables seamless interaction with inventory data through a unified interface.

Overall, the system provides a practical and scal-able solution for intelligent inventory management by combining predictive analytics with user-friendly interaction mechanisms. While the current imple-mentation is limited in terms of dataset diversity and language support, the modular design allows for future enhancements in model robustness, fea-ture integration, and multilingual capabilities. This work highlights the potential of combining machine learning and natural language interfaces to improve decision-making and accessibility in inventory management systems.

References

1. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is All You Need," Advances in Neural Information Processing Systems, vol. 30, 2017.
2. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," arXiv preprint arXiv:1810.04805, 2018.
3. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD, pp. 785–794, 2016.
4. L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
5. P. Lewis, E. Perez, A. Piktus et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," NeurIPS, 2020.
6. M. Huang, X. Zhu, and J. Gao, "Challenges in Building Intelligent Open-Domain Dialogue Systems," ACM Trans. Inf. Syst., vol. 38, no. 3, pp. 1–32, 2020.
7. X. Hu, S. Lim, and G. Li, "Healthcare Chatbots: A Survey," Journal of Medical Systems, vol. 45, no. 3, p. 27, 2021.
8. J. Gao, M. Galley, and L. Li, "Neural Approaches to Con-versational AI," Foundations and Trends in Information Re-trieval, vol. 13, no. 2–3, pp. 127–298, 2019.
9. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," Proc. 52nd Annual Meeting of the ACL (System Demonstrations), pp. 55–60, 2014.
10. S. Khanuja, D. Bansal, S. Mehtani et al., "MuRIL: Multilin-gual Representations for Indian Languages," arXiv preprint arXiv:2103.10730, 2021.
11. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision," Proc. ICML, 2023.
12. V. Pratap, A. Tjandra, B. Shi et al., "Scaling Speech Technol-ogy to 1,000+ Languages," arXiv preprint arXiv:2305.13516, 2023.
13. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of Tricks for Efficient Text Classification," Proc. EACL, pp. 427–431, 2017.
14. Conneau, K. Khandelwal, N. Goyal et al., "Unsupervised Cross-lingual Representation Learning at Scale," Proc. ACL,
15. pp. 8440–8451, 2020.
16. T. Brown, B. Mann, N. Ryder et al., "Language Models are Few-Shot Learners," NeurIPS, vol. 33, 2020.