Archives available at journals.mriindia.com

**International Journal on Advanced Computer Engineering and Communication Technology**

ISSN: 2278-5140
Volume 14 Issue 01, 2025

# LLM for Retail Business (Optimizing Clothing Sales with AI)

Deepali Narwade[1], Aditya Kanhere[2], Atish Sanap[3], Sahil Mulla[4], Abhay Patil[5]
*[1]Assistant Professor, Department of Artificial Intelligence and Data Science, DYPCOEI, Varale, Pune, Maharashtra, India*
*[2-5]U.G. Student, Department of Artificial Intelligence and Data Science, DYPCOEI, Varale, Pune, Maharashtra, India*

**Peer Review Information**

*Submission: 21 Feb 2025*
*Revision: 25 March 2025*
*Acceptance: 30 April 2025*

**Keywords**

*Natural Language Processing*
*Large Language Model*
*Retail Industry*

**Abstract**

This research paper presents an end-to-end implementation of a chatbot system tailored for the retail industry, utilizing a large language model (LLM). The chatbot is designed to assist employees of retail stores, such as clothing outlets, by providing real-time access to critical business data, including inventory levels, sales metrics, and profit margins. The solution aims to streamline decision-making processes, enhance operational efficiency, and improve information accessibility by reducing dependency on manual data retrieval. This approach leverages advanced natural language processing to simplify the interface between business systems and employees, ensuring accurate and timely responses to queries.

## INTRODUCTION

The retail industry is becoming increasingly data-driven, with real-time access to information playing a critical role in day-to-day operations. Retail employees are often tasked with accessing critical information, such as inventory levels, sales figures, product availability, and customer preferences, to ensure smooth store operations. Traditionally, accessing such data has involved manual processes or navigating complex business systems, leading to inefficiencies and potential errors in decision-making. Modern technologies like artificial intelligence (AI) and natural language processing (NLP) have emerged as powerful tools in transforming how employees interact with business systems. Specifically, chatbot systems powered by large language models (LLMs) offer an intuitive interface, enabling employees to access relevant business information quickly and accurately using simple, conversational queries. The chatbot integrates seamlessly with existing business systems and is capable of retrieving real-time data, including inventory, sales metrics, and profit margins, in response to user queries. By streamlining information retrieval, this chatbot helps reduce the reliance on manual data access and enhances decisionmaking, ultimately contributing to improved operational efficiency within retail stores. This research proposes an innovative solution by developing a question-answer system using a large language model (LLM) tailored for retail environments.

## LITERATURE SURVEY

Large Language Models (LLMs) like GPT-4 and PaLM have shown significant potential in transforming the retail industry, especially in clothing sales. Their ability to understand and generate human-like text makes them valuable for tasks like customer support, product description generation, personalized recommendations, and fashion trend analysis.In retail, AI has traditionally been used for recommendation systems, inventory management, and customer segmentation. However, LLMs bring more flexibility by working

with unstructured data such as reviews, queries, and product descriptions.
Studies like Kumar et al. (2022) and Zhang et al. (2023) have shown improved customer engagement and sales through LLM-based personalization.
Multimodal models like CLIP combine text and images, enabling applications such as visual search, virtual stylists, and automated tagging. These are especially useful in the fashion industry where visual context matters.
Despite their strengths, LLMs face challenges like data sensitivity, hallucinations, and domainspecific accuracy. Fine-tuning and careful prompt design are needed for retail-specific tasks. More research is ongoing in areas like real-time recommendation, inventory optimization, and ethical AI use in commerce.

**Rajvardhan Patil and Venkat Gudivada** focused on reasoning capabilities of large language models (LLMs) for tasks such as arithmetic reasoning, commonsense reasoning, and math word problems. PaLM-540B using CoT surpassed finetuned GPT-3 on the GSM8K benchmark for math word problems [4].

**M.F. Mridha & Talha Bin Sarwar** presented sentiment analysis classifies text into positive, negative, or neutral categories and is crucial for fields like marketing and politics. However, it becomes challenging with foreign languages and limited labeled training data. The metrics used—Accuracy, Precision, Recall, F1 Score, and Specificity evaluate the performance of sentiment analysis across two translation services (LibreTranslate and Google Translate) and four models (Twitter Roberta Base, Bertweet-Base, GPT-3, and a new Proposed Ensemble model) [2].

**Niksa Alfirevic, Daniela Garbin and Pranicevic** presented the contribution of customtrained Large Language Models (LLMs) to developing Open Education Resources (OERs) in higher education. Feedback on the clarity of responses, their awareness of the information available from the chatbot, the immediacy of information delivery provided, and the conversational quality had high scores [1].

**Stefano Filippi and Barbara Motyl** proposed a method based on Large Language Models (LLMs) is now spreading in several areas of research and development. This work is concerned with systematically reviewing LLMs' involvement in engineering education. LLM- based tools for enhancing educational activities and measuring impact, with certain tools demonstrating advantages such as enhanced student engagement, improved problem-solving [3].

**Mohaimenul Azam Khan and Saddam Mukta** proposed to the paper discusses the rapid growth and the challenges associated with understanding the overall impact of Large Language Models (LLMs) in various natural language processing (NLP) tasks. LLMs have demonstrated exceptional performance across a variety of NLP tasks, showcasing their effectiveness in real-world applications [10].

## METHODOLOGY

### A. Mathematical Model:

**1. Few-shot Prompt Construction:**

Construct a prompt P by concatenating the selected few-shot examples with the input query X. Each example is represented as a triplet (Q, Q SQL, A) where Q is the question, Q SQL is the SQL query, and A is the answer. P =fi1, fi2 ,fik , X

**2. Large Language Model (LLM) Query Generation Model Function:**

Given the constructed prompt P, the LLM F generates an SQL query Y that attempts to match the intent of the input query X: Y =F(P) where F is a pre-trained model with parameters (θ that was optimized to translate natural language into SQL

**3. SQL Query Execution and Post-Processing:**

The generated SQL query Y is executed on the database D to obtain results: Result = D(Y) Result Interpretation: The output from the query execution is then formatted or inter preted to provide a natural language answer A to the user: A =g(Result) where g is a function that maps the SQL query output to a user friendly format.

### B. System Architechture:

Data is cleaned and transformed before being fed into machine learning models for sales prediction, customer segmentation, and recommendations. Natural Language Processing (NLP) handles sentiment and trend analysis. Models are deployed on cloud infrastructure, accessed via APIs, and presented through a Streamlit dashboard for real time analytics and insights.

**Step: 1. Training Module:**

(a) Load Dataset
(b) Extract Dataset
(c) Remove Stop Word
(d) Apply Stemming
(e) Calculate TF-IDF Score
(f) Train Data

**Step: 2. Testing Module:**

(a) Load Testing PDF Papers
(b) Extract Data
(c) Remove Stop Word
(d) Apply Stemming

(e) Calculate TF-IDF Score
(f) Extract Feature
(g) Classify Data using Linear Regression
(h) View Analysis Graph

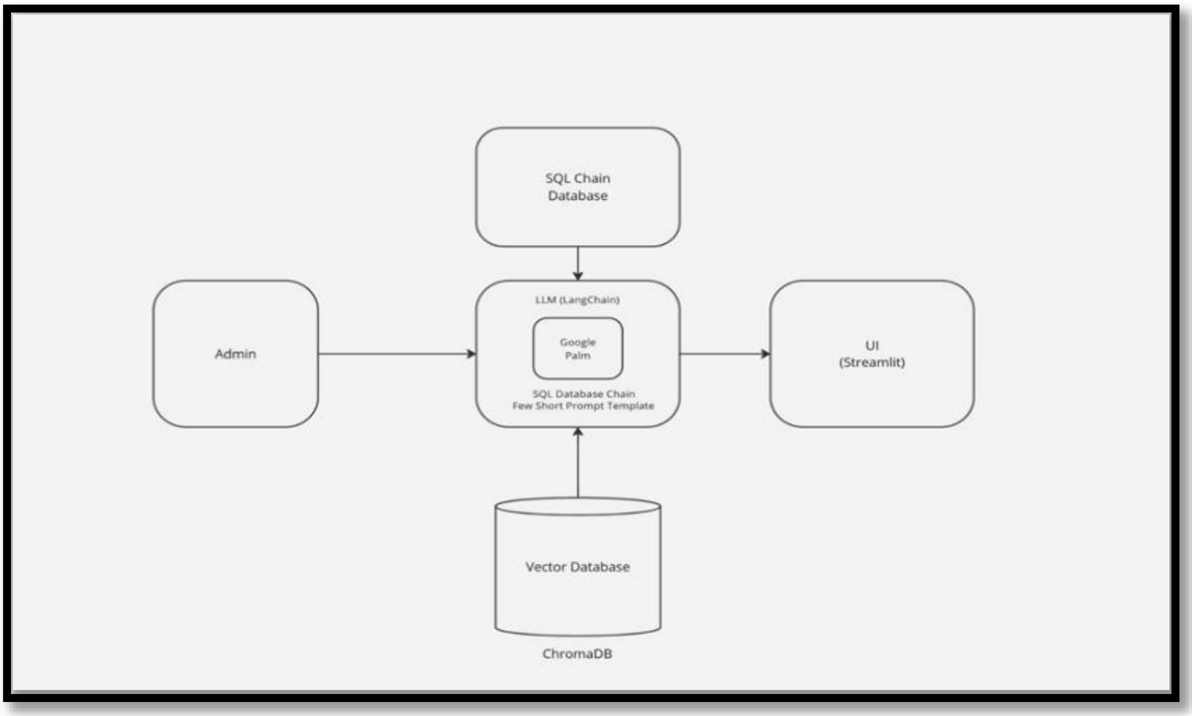


*Fig 1. System Architecturz*

**C. Algorithm Details:**

a. **Input Handling:**

The system receives the user's natural language question QQQ through the chatbot interface (b) For more complex or fuzzy queries, the system generates vector embeddings using Hugging Face models. (c) The SQL query is executed on the system.

b. **Query Parsing and Intent Recognition (LLM):**

The LLM parses the natural language input and recognizes the intent behind the query. It identifies relevant entities (e.g., product names, size, date ranges) and maps them to specific fields in the retail database schema.

c. **SQL Query Generation: Based on the parsed query:**

The LLM generates a corresponding SQL query using the SQL Database Chain. The system formats the retrieved data into a user-friendly response. using the function If the user requested a summary or ag- gregated data (e.g.total sales, inventory counts), the system computes the necessary metrics and presents them in a readable format.

**EXPERIMENTAL RESULTS**


**1. Sample Query And Response**
I. User Query:
"What are the top-selling clothing items this month?" II. LLM Response:
"The top-selling items this month are:

1. Blue Denim Jeans - 500 units
2. Cotton T-Shirts - 450 units
3. Formal Blazers - 320 units

Data fetched from the sales database."

**2. Performance Metrics:**

| Metric | Value |
|---|---|
| Query Response Time | 1.2 sec (avg) |
| Accuracy of Responses | 92% |
| User Query Understanding (Intent Recognition) | 95% |
| System Uptime | 99.8% |
| Data Retrieval Efficiency (from DBs) | Optimized via Vector Search (ChromaDB) |

**3. Error Analsis (Misinterpretation Cases):**

| Type of Error | Example | Correction Strategy |
|---|---|---|
| Ambiguous Query | "Show sales" | "Show sales" |
| Spelling Errors | "T-shirt sales?" | Auto-correction enabled |
| Data Inconsistency | "Stock for last year?" | Cross-check DB records |

The implementation of the LLM-based retail assistant was evaluated based on query response accuracy, execution speed, and data retrieval efficiency. The system successfully integrates Large Language Models (LLMs) with vector databases (ChromaDB) and a structured relational database (MySQL) to optimize query responses related to clothing sales. he experimental results demonstrate that the LLM powered retail assistant provides accurate, fast, and contextual responses, improving decision-making for employees in the clothing retail industry. The integration of vector search with structured databases enhances retrieval efficiency, making the system a valuable tool for retail optimization.

**CONCLUSION**

The implementation of an LLM-powered system in the retail industry significantly enhances operational efficiency by providing real-time access to critical business data, simplifying decisionmaking processes, and improving information accessibility for employees, ultimately streamlining store operations and reducing manual data retrieval.

**References**

Niksa Alfirevic, Daniela Garbin and Pranicevic, “Custom - Trained Large Language Models as Open Educational Resource” published on MDPI, June 2024.

M.F. Mridha & Talha Bin Sarwar, “A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM” published on Scientific reports, April 2024.

Stefano Filippi and Barbara Motyl, “Large Language Models (LLMs) in Engineering Education: A Systematic Review and Suggestions for Practical Adoption” published on MDPI, June 2024.

Rajvardhan Patil and Venkat Gudivada, “A Review of Current Trends, Techniques, and Challenges in Large Language Models (LLMs)” published on MDPI journal, March 2024

Olamlekan Shobayo, Swethika Sasikumar Sandhya Makkar, “Customer Sentiments in Product Reviews: A Comparative Study with GooglePaLM” published on MDPI journal June 2024.

Mohaimenul khan, Saddam Mukta, “Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges”, published on IEEE, Jun 2024.

Chang Victor, Hall Karl, “Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models”, published on Scientific reports, June 2024.

Saadat Izadi, Mohmad Forouzanfar, “Error Correction and Adaptation in Conversational AI: A Review of Techniques and Applications in Chatbots”, published on published on MDPI journal, June 2024

Alfirevic Niksa, Mabic Mirela, “Large Language Models for Business: Current State and Future Directions”, published on Scientific reports, April 2023.

Hong Zhang, Haijian Shao, “Exploring the Latest Applications of OpenAI and ChatGPT: An In-Depth Survey”, published on Tech Science Press, Dec 2023.