

TEAM DUMPLINGS

# Ý TƯỞNG GIẢI QUYẾT BÀI TOÁN

**DATASET 1:** CHATBOT GỌI Ý SẢN PHẨM THEO  
YÊU CẦU CỦA KHÁCH HÀNG.

**DATASET 2:** CHATBOT AI HỖ TRỢ DỰ BÁO VÀ  
QUẢN LÝ HÀNG TỒN KHO

# **DATASET 1**

FASHION STORE



# GIỚI THIỆU VẤN ĐỀ

*Việc hạn chế trong quá trình mua hàng online là khách hàng không nhận được sự tư vấn cần thiết để chọn lựa sản phẩm phù hợp với nhu cầu cá nhân.*



# Mô tả vấn đề

Hệ thống tư vấn bán hàng tích hợp chatbot sử dụng trí tuệ nhân tạo góp phần nâng cao trải nghiệm người dùng và nâng cao uy tín thương hiệu.

1

**RECOMMAND CÁC SẢN PHẨM LIÊN QUAN ĐẾN  
MIÊU TẢ CỦA NGƯỜI DÙNG**

2

**ĐÓNG VAI TRÒ LÀ NGƯỜI CHĂM SÓC  
KHÁCH HÀNG**

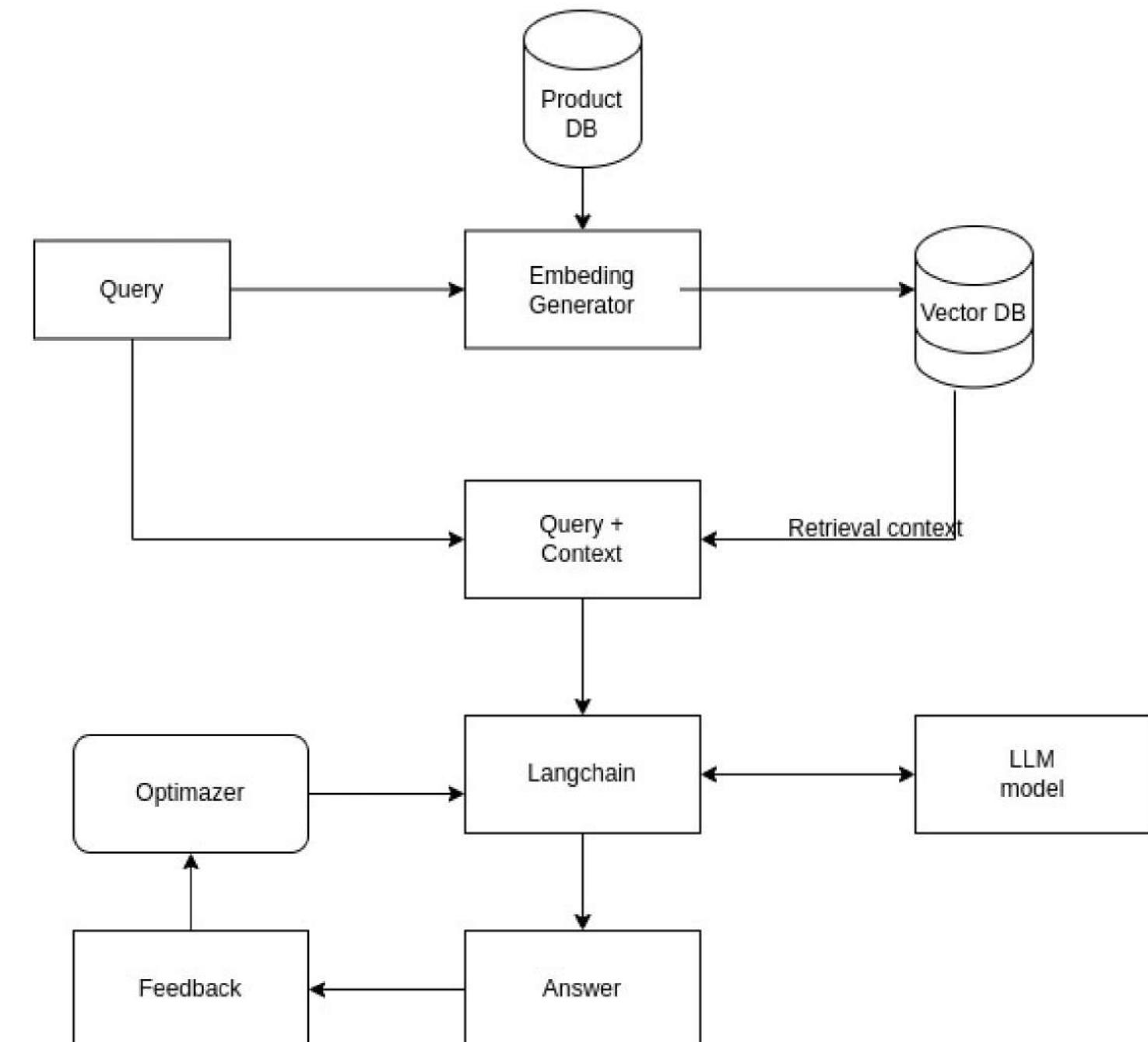
# TỔNG QUAN GIẢI PHÁP

Chatbot:

- Sử dụng framework langchain kết hợp với LLM để đưa ra sản phẩm phù hợp với miêu tả của khách hàng. Sau đó nhận feedback của khách hàng để cải thiện mô hình.

Example:

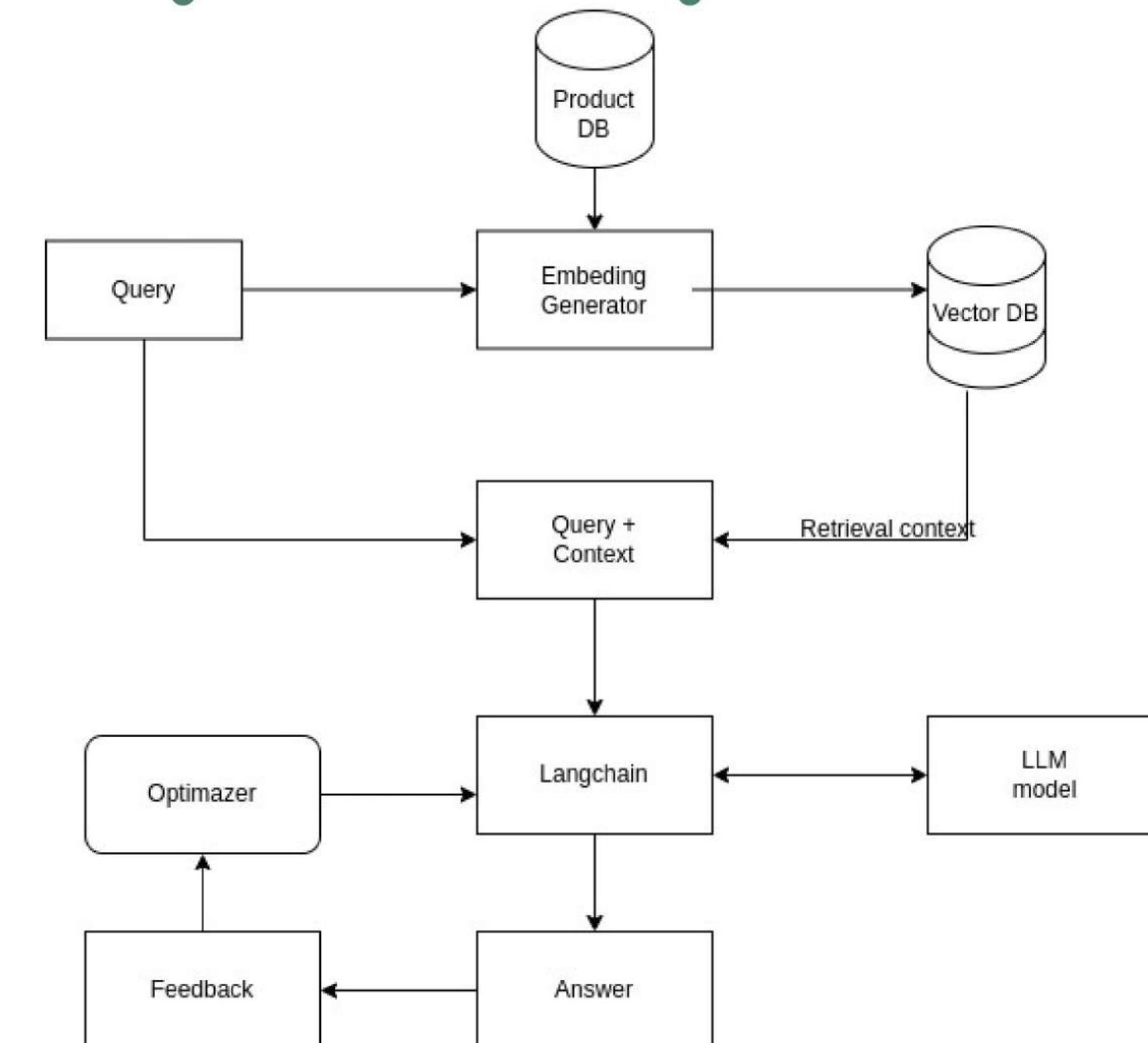
- Customer: “I need a pair of shoes to give my child a fun adventure.”
- Bot: “Based on your description, I recommend the green ZX 2K Shoes from adidas. This is the option...”
- 



# MÔ HÌNH THỰC HIỆN

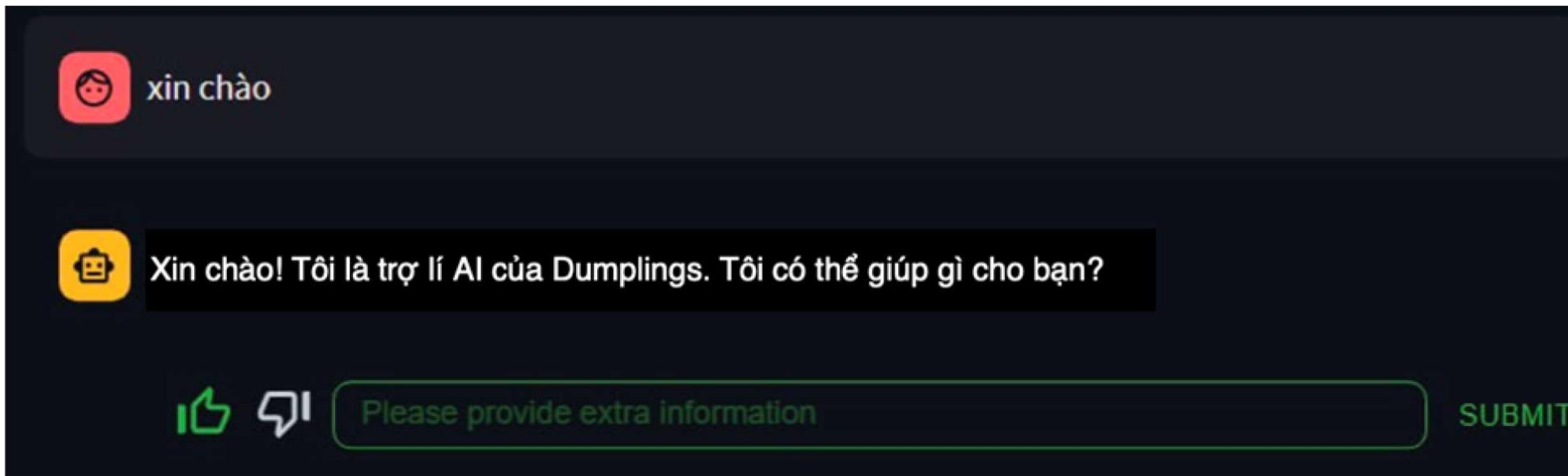
*Chatbot:*

- Embedding: RoBERTa, PhoBERT-v2
- LLM: Llama 7B, GPT 3.5 (API), GPT-J (finetune)
- Vector DB sử dụng FAISS.
- Lưu trữ thông tin sử dụng ZEPmemory, PostgreSQL (id, session\_id, question, answer, datetime, react, comment)



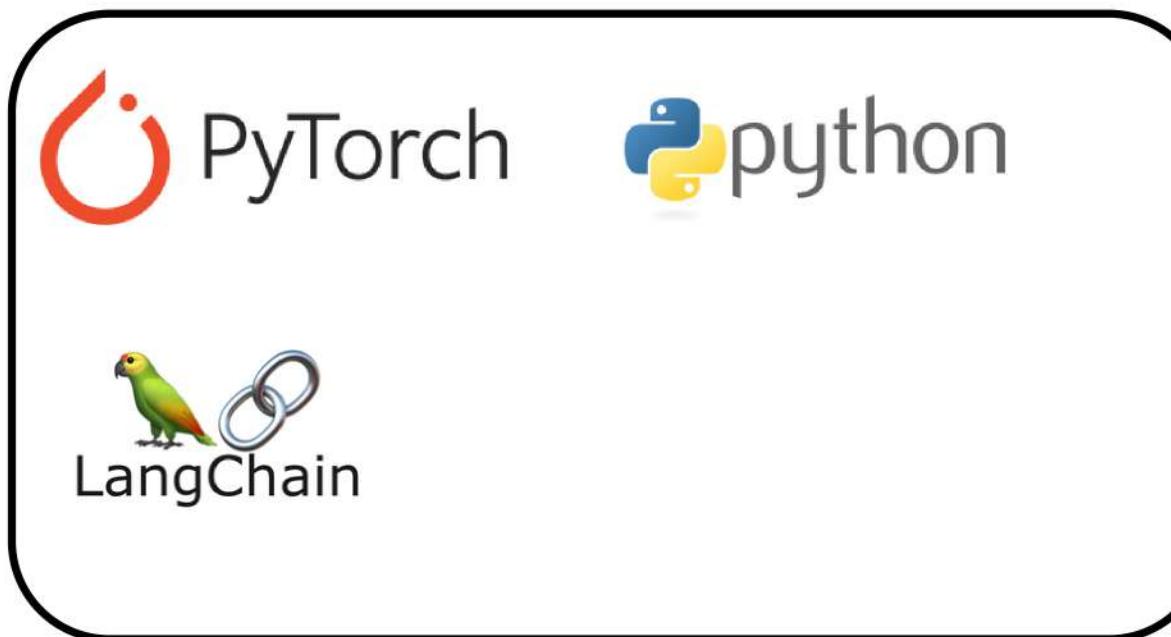
# PHƯƠNG PHÁP ĐÁNH GIÁ

- *Đánh giá chatbot:* Nhóm sẽ đánh giá dựa trên độ hài lòng của khách hàng bằng cách thêm tính năng like/dislike của mỗi câu phản hồi.

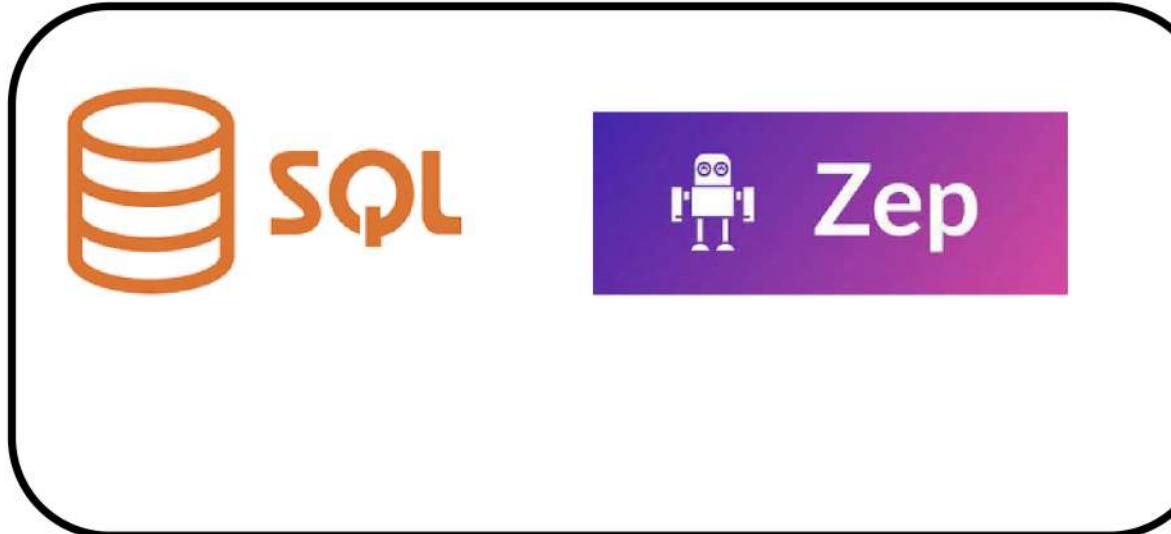


Minh họa độ đo dựa trên đánh giá của người dùng.

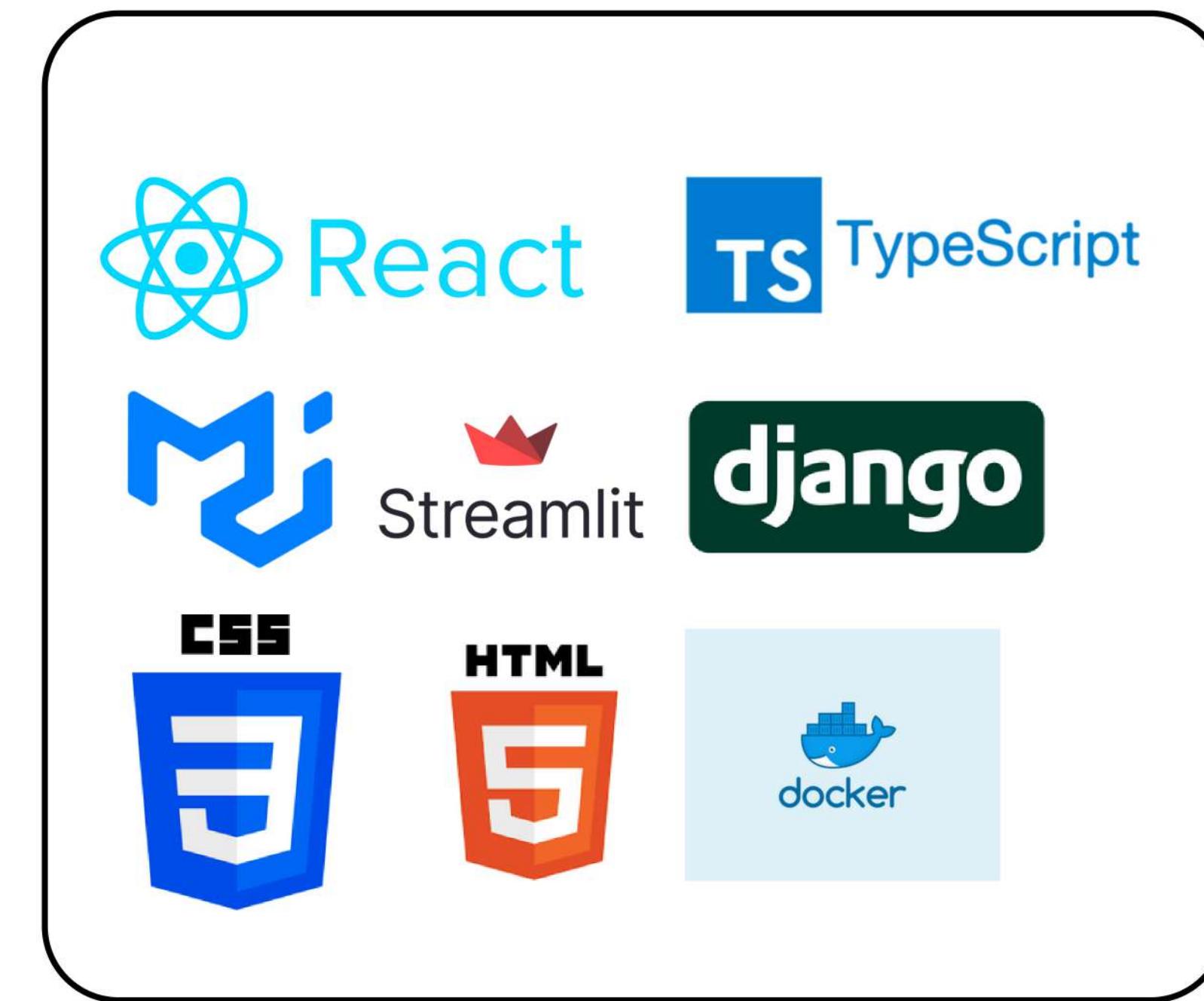
# CÔNG NGHỆ SỬ DỤNG



Model



Database



web-app

# TIMELINE VÀ ROAD-MAP

		Tháng 11							Tháng 12							Tháng 12								
		13	14	15	16	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Công việc																								
1	Lên ý tưởng	Dữ liệu và các chủ đề cuộc thi Thị trường và các vấn đề liên quan đến tài Hướng tiếp cận, độ khả thi																						
3	Tiền xử lý dữ liệu	Làm sạch dữ liệu Insert dữ liệu vào database																						
4	Chatbot AI	Chuẩn bị dữ liệu cho chatbot Config langchain Optimize hiệu suất và prompt engineer																						
5	Báo cáo	Viết báo cáo Deploy ý tưởng																						
Dataset		Adidas and Nike																						

DEPLOY Ý TƯỞNG

# **DATASET 2**

*SALES AND INVENTORY DATA  
OF VIETNAM RETAILERS*



# 1. GIỚI THIỆU VẤN ĐỀ

Theo McKinsey Digital, hệ thống dự báo nhu cầu tích hợp AI có thể giúp giảm 50% lỗi dự báo và 65% doanh thu bị mất so với các phương pháp truyền thống.[1] Thông qua cuộc thi Datathon team 112 muốn đưa đến một ý tưởng bao gồm:

- MVP: “AI hỗ trợ dự báo và quản lý kho cho các nhà bán lẻ”.
- MVP bao gồm **công cụ phân tích trực quan hóa dữ liệu, chatbot hỗ trợ dự báo nhu cầu và đề xuất phương án quản lý hàng tồn kho tối ưu.**  
=> Chủ động **điều chỉnh kế hoạch kinh doanh** dựa trên các báo cáo để tối đa hóa doanh số và hiệu quả hoạt động.  
=> Đưa ra các quyết định nhanh chóng hơn qua trợ lý AI góp phần **giảm thời gian tìm kiếm và tự phân tích**.

[1] Top 10 xu hướng ứng dụng AI trong ngành bán lẻ năm 2023

## 2. Mô tả vấn đề

Các vấn đề:

1

### CHƯA THỂ ĐO LƯỜNG

- “If you can't measure it, you can't improve it”.

2

### HÀNG TỒN

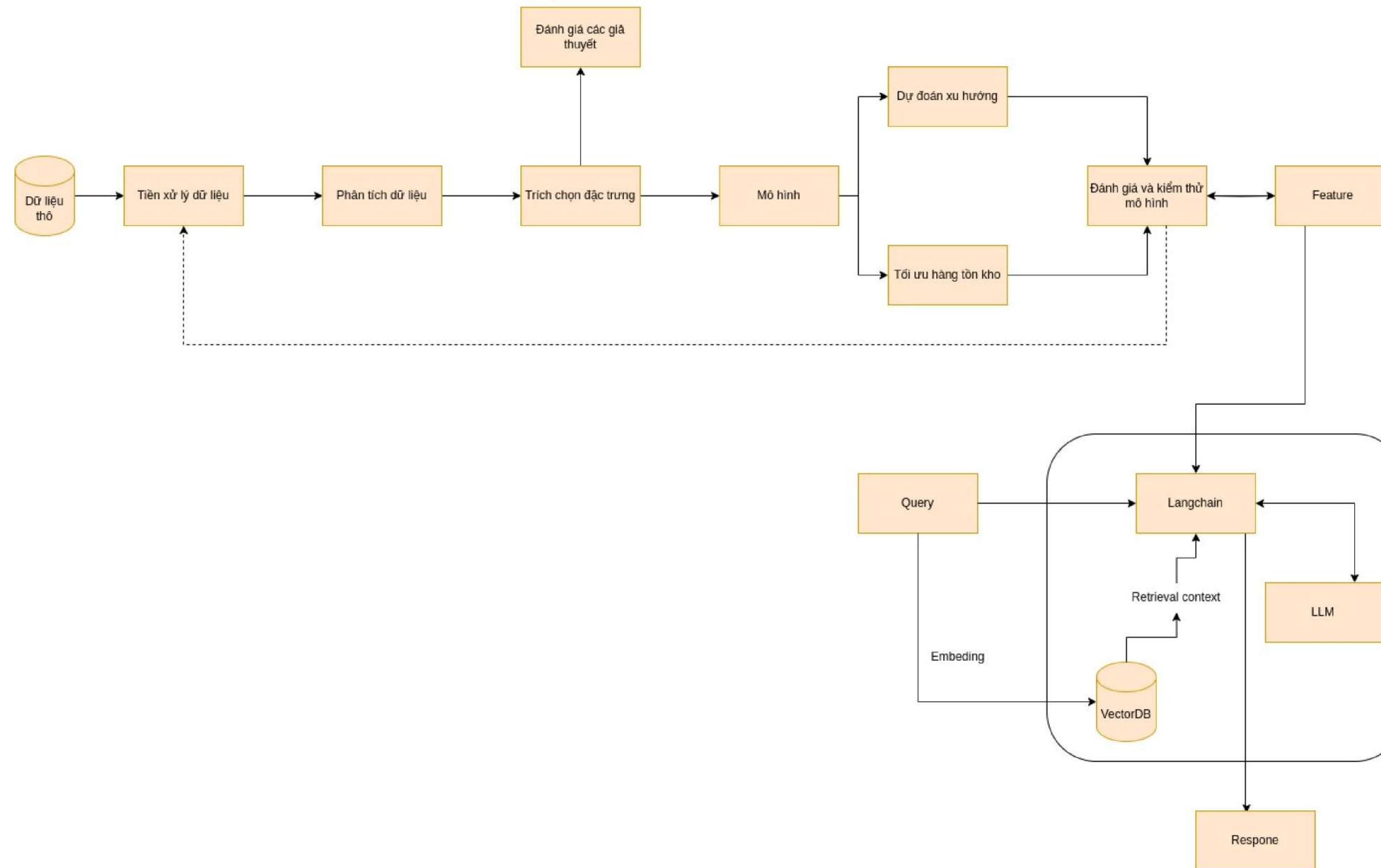
- Các cửa hàng thời trang thường xuyên phải đổi mặt với tình trạng hàng tồn kho lỗi thời do dự báo sai xu hướng thị trường.

3

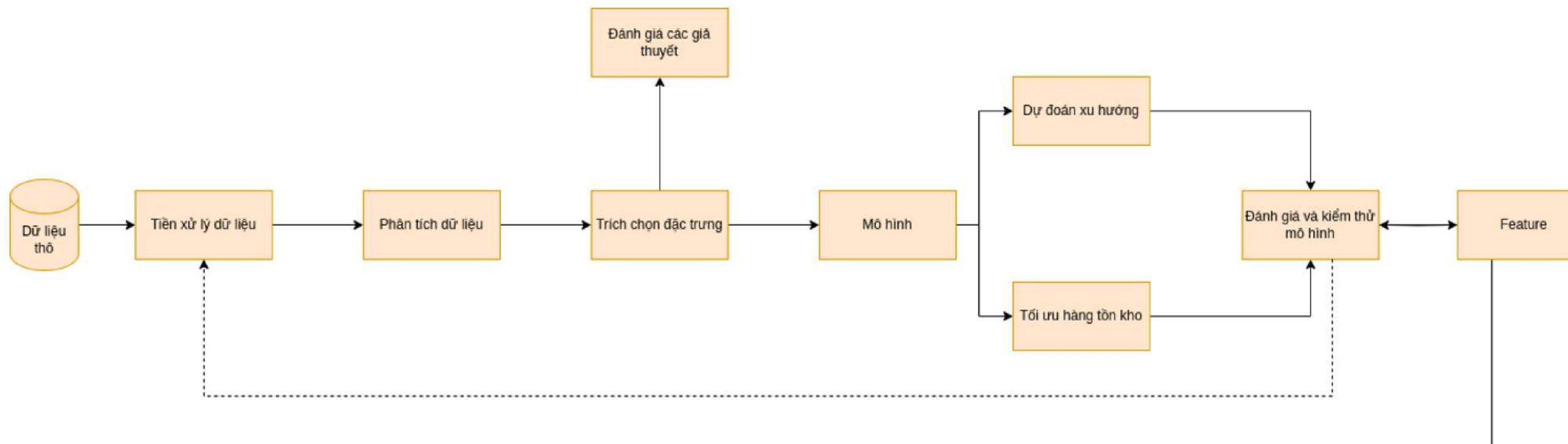
### CÔNG CỤ QUẢN LÍ HỖ TRỢ KINH DOANH

- Các cửa hàng thiếu công cụ hỗ trợ dự báo chính xác nhu cầu và lập kế hoạch kinh doanh.
- Việc theo dõi và quản lý hàng tồn kho phức tạp do có nhiều mẫu mã, kích cỡ khác nhau

# 3. TỔNG QUAN GIẢI PHÁP

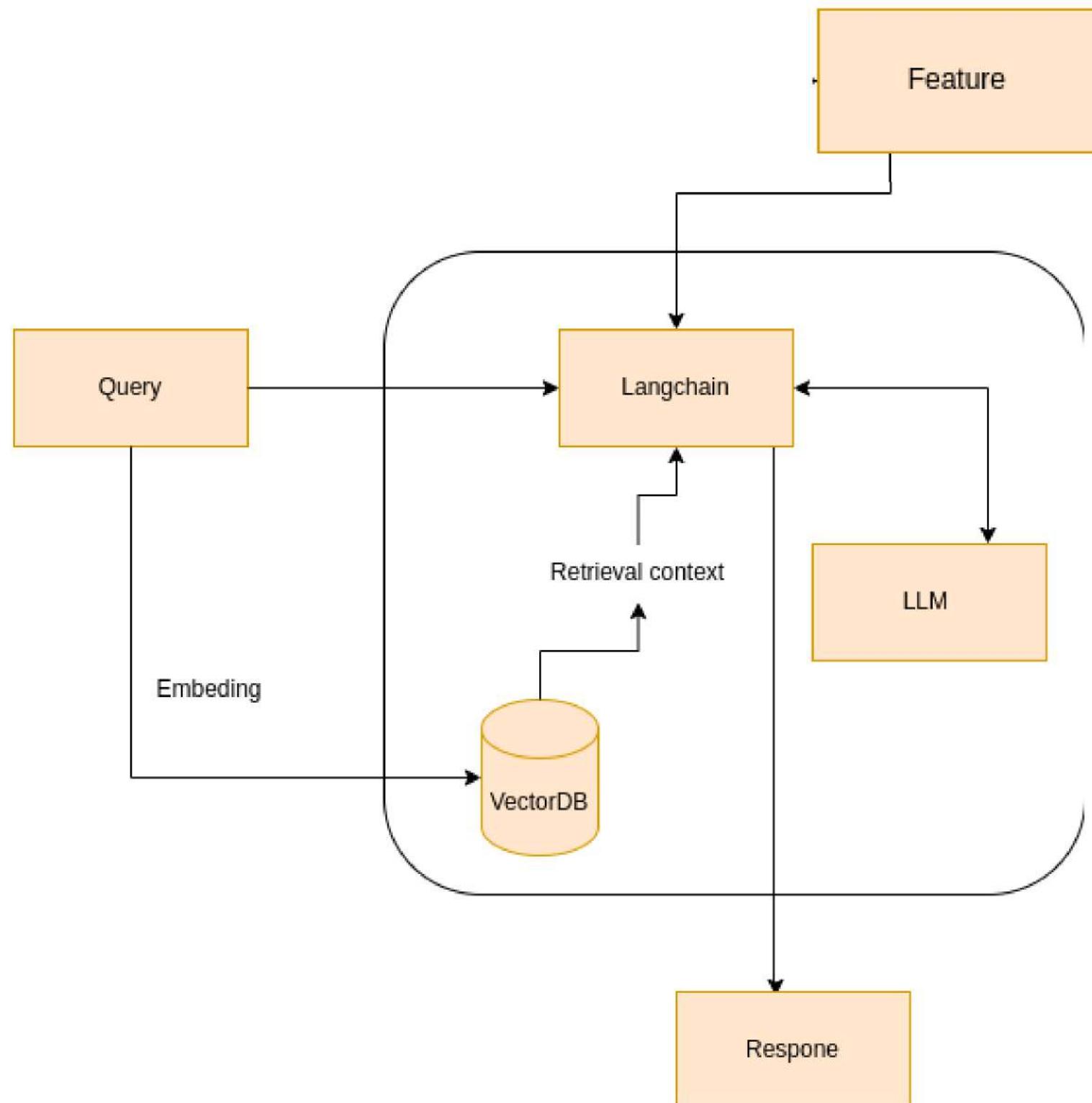


# 3. TỔNG QUAN GIẢI PHÁP



- Phase 1:
  - Phân tích dữ liệu và trích chọn các đặc trưng có độ tương quan cao với lợi nhuận, các mặt hàng và xu hướng. Sau đó, dùng các kỹ thuật thống kê như Causal Impact, T-test, Anova,.. để đánh giá kiểm định các giả thuyết về hàng tồn kho. Từ đó, đưa vào mô hình học máy như Prophet, ARIMA, Linear Regression, XGBoost, LightGBM để dự đoán xu hướng, nhu cầu, tối ưu hóa hàng tồn.
  - Input: dữ liệu thô, output: Dự đoán xu hướng, chiến lược xử lý hàng tồn kho dựa vào các phân tích đặc trưng trên.

# 3. TỔNG QUAN GIẢI PHÁP



- Phase 2:

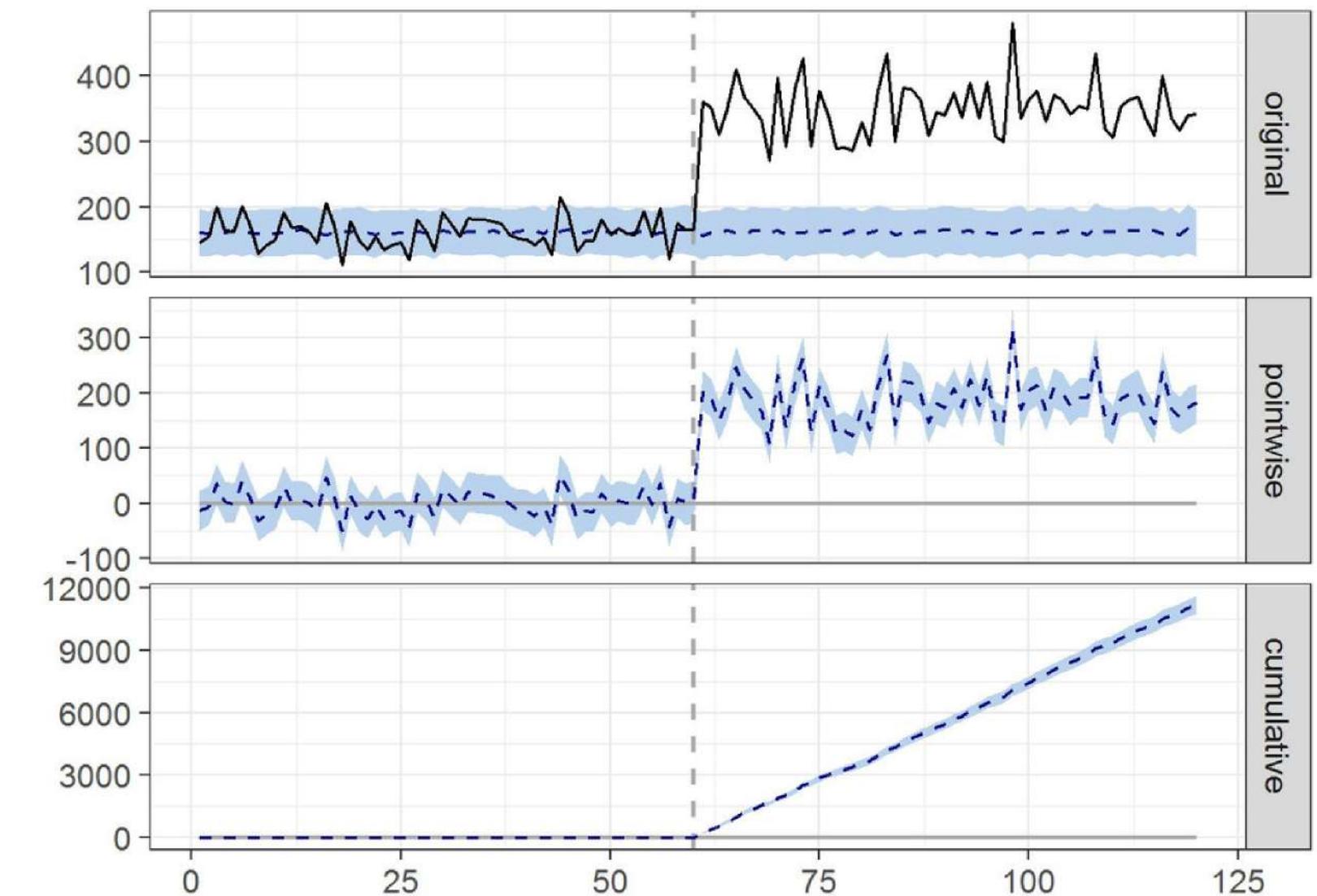
- Chatbot sẽ được huấn luyện để cung cấp các khuyến nghị và báo cáo cá nhân hóa cho người dùng dựa trên kết quả dự báo và tối ưu hóa.
- Input: [Query, Feature], output: text trả lời câu hỏi liên quan đến dự đoán xu hướng, chiến lược xử lý hàng tồn kho .

# 4. PHƯƠNG PHÁP LUẬN

1

## CAUSAL IMPACT

Phân tích tác động nhân quả là mô hình thống kê ước tính tác động của một sự kiện hoặc can thiệp nhất định lên một chuỗi kết quả sau can thiệp.



## 4. PHƯƠNG PHÁP LUẬN

2

### THỐNG KÊ T-TEST VÀ ANOVA

-Kiểm định T được dùng để so sánh trung bình của hai nhóm mẫu độc lập. Ví dụ, so sánh trung bình doanh số của mặt hàng tồn kho A và B để xem mặt hàng nào có khả năng bán chạy hơn,...

-ANOVA được dùng để so sánh trung bình của nhiều hơn 2 nhóm mẫu. Ví dụ, so sánh trung bình doanh số của các nhóm mặt hàng để xem nhóm mặt hàng nào có tiềm năng tăng trưởng cao nhất.

## 4. PHƯƠNG PHÁP LUẬN

### 3 MÔ HÌNH PROPHET

Mô hình Prophet là một mô hình dự báo tự động được phát triển bởi Facebook. Được sử dụng để dự báo các chuỗi thời gian có xu hướng, chu kỳ và các sự kiện đặc biệt, chẳng hạn như ngày lễ.

$$y_t = g(t) + s(t) + h(t) + \varepsilon_t,$$

- $y_t$  là giá trị dự báo tại thời điểm  $t$
- $g(t)$  là hàm xu hướng, mô tả xu hướng tổng thể của chuỗi thời gian
- $s(t)$  là hàm chu kỳ, mô tả các chu kỳ lặp đi lặp lại trong chuỗi thời gian
- $h(t)$  là hàm sự kiện, mô tả tác động của các sự kiện đặc biệt
- $\varepsilon_t$  là sai số

# 4. PHƯƠNG PHÁP LUẬN

## 4 MÔ HÌNH ARIMA VÀ SARIMA

-Mô hình ARIMA:

- Dựa trên giả thuyết chuỗi dừng và phương sai sai số không đổi. Mô hình sử dụng đầu vào chính là những tín hiệu quá khứ của chuỗi được dự báo để dự báo nó.
- Các tín hiệu đó bao gồm: chuỗi tự hồi qui AR (auto regression) và chuỗi trung bình trượt MA (moving average).
- Hầu hết các chuỗi thời gian sẽ có xu hướng tăng hoặc giảm theo thời gian, do đó yếu tố chuỗi dừng thường không đạt được.
- Trong trường hợp chuỗi không dừng thì ta sẽ cần biến đổi sang chuỗi dừng bằng sai phân. Khi đó tham số đặc trưng của mô hình sẽ có thêm thành phần bậc của sai phân  $d$  và mô hình được đặc tả bởi 3 tham số ARIMA( $p, d, q$ ).

# 4. PHƯƠNG PHÁP LUẬN

## 4 MÔ HÌNH ARIMA VÀ SARIMA

-Mô hình SARIMA:

- Về bản chất đây là mô hình ARIMA nhưng được điều chỉnh đặc biệt để áp dụng cho những chuỗi thời gian có yếu tố mùa vụ.
- Bằng cách tìm ra chu kỳ của qui luật mùa vụ và loại bỏ nó khỏi chuỗi ta sẽ dễ dàng hồi qui mô hình theo phương pháp ARIMA.

## 4. PHƯƠNG PHÁP LUẬN

### 5 MÔ HÌNH XGBOOST

XGBoost (*Extreme Gradient Boosting*):

- **Là một giải thuật được dựa trên gradient boosting**, tuy nhiên kèm theo đó là những cải tiến to lớn về mặt tối ưu thuật toán, về sự kết hợp hoàn hảo giữa sức mạnh phần mềm và phần cứng, giúp đạt được những kết quả vượt trội cả về thời gian training cũng như bộ nhớ sử dụng.
- **Hoạt động bằng cách tạo một chuỗi các cây quyết định** (decision trees), trong đó mỗi cây tiếp theo được tạo để sửa lỗi của các cây trước đó. Thuật toán có nhiều siêu tham số (*hyperparameters*) có thể được điều chỉnh để tối ưu hóa hiệu suất của thuật toán, bao gồm tốc độ học, số lượng cây và độ sâu tối đa của mỗi cây.

# 4. PHƯƠNG PHÁP LUẬN

## 6 MÔ HÌNH LIGHTGBM

Mô hình LightGBM:

- Dựa trên thuật toán cây quyết định tăng cường độ dốc (gradient boosting decision tree - GBDT), thuật toán này xây dựng một tập hợp các cây quyết định dự đoán biến mục tiêu bằng cách kết hợp các dự đoán của nhiều mô hình yếu hơn.
- Mỗi cây quyết định được xây dựng bằng cách lặp đi lặp lại quá trình thêm các nhánh mới vào cây hiện có với thuật toán tối ưu hàm mất mát.
- Thuật toán cũng kết hợp một kỹ thuật gọi là lấy mẫu một phía dựa trên độ dốc (Gradient-based one-side sampling - GOSS), GOSS lấy mẫu dữ liệu bằng cách ưu tiên các trường hợp có độ dốc lớn hơn, giúp giảm chi phí tính toán và cải thiện độ chính xác của mô hình.

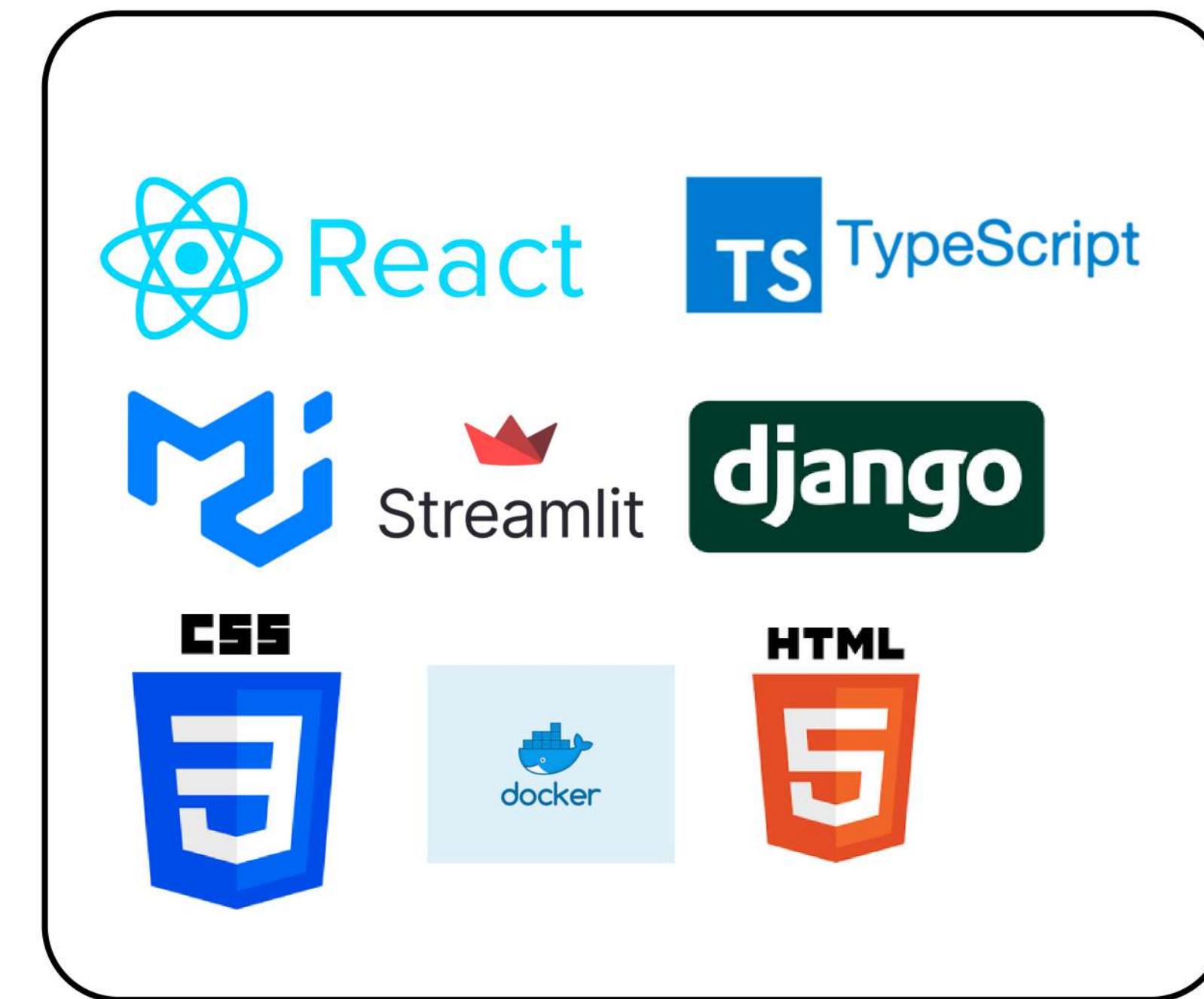
# CÔNG NGHỆ SỬ DỤNG



Model



Database



web-app

# PHÁC THẢO CÁC TÍNH NĂNG VÀ CHỨC NĂNG CHÍNH CỦA MVP

Tính năng chính:

- Dự đoán xu hướng, nhu cầu của khách hàng như: ở mặt hàng giày dép thì đế dẹp hay cao 3 phân thì thường chạy hàng vào tháng nào,..
- Trực quan hóa dữ liệu: Bằng cách tạo dashboard các số liệu về hàng tồn theo thời gian và dự đoán, xu hướng, doanh thu..
- Chatbot hỗ trợ tư vấn, trả lời tương tác với người dùng về vấn đề hàng tồn kho.

# PHÁC THẢO CÁC TÍNH NĂNG VÀ CHỨC NĂNG CHÍNH CỦA MVP

*Chức năng chính:*

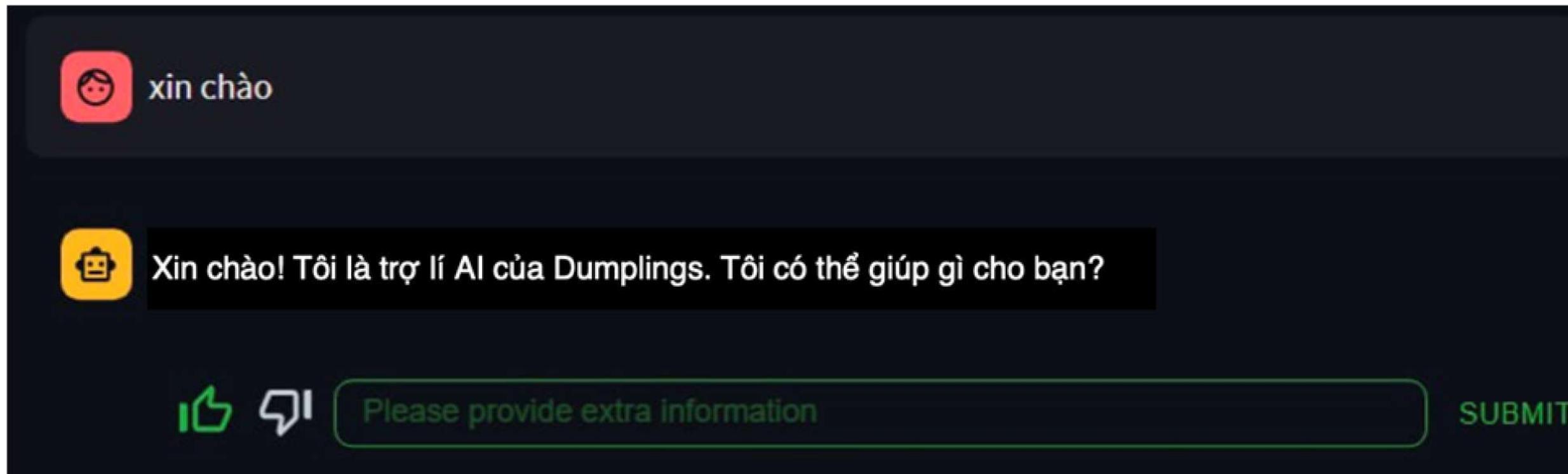
- *Báo cáo và thống kê:* tạo báo cáo tổng quan như xem chi tiết top những mặt hàng bán chạy, bán kém của từng chi nhánh cụ thể để có kế hoạch đẩy hàng nhanh chóng,..
- *Cảnh báo hàng tồn:* Khi các hàng tồn quá lâu nếu không xử lý sẽ gây thiệt hại cần ưu tiên để xử lý.
- *Dự đoán:* Dự đoán xu hướng tồn kho trong tương lai, mặt hàng nào có xu hướng tồn kho giảm, sẽ bán chạy ở tương lai.
- *Đánh giá hiệu quả:* của mặt hàng cho từng kênh và hướng xử lý hàng tồn dựa trên các kênh.
- *Chatbot:* Trả lời các câu hỏi về hàng tồn, các chiến lược cho người dùng.

# PHƯƠNG PHÁP ĐÁNH GIÁ

- Đánh giá mô hình dự đoán: **Accuracy, F1-score, R2 score, Precision, Recall.**
- Đánh giá chiến lược tối ưu hóa hàng tồn kho:
  - **Inventory Turnover:** cho biết số lần hàng tồn kho đã được bán và thay thế trong một khoảng thời gian cụ thể.
  - **Carrying Cost of Inventory:** Chi phí lưu kho trong một khoảng thời gian cụ thể. Hàng tồn trong kho càng lâu thì chi phí này càng cao
  - **Inventory-to-Sales Ratio:** đo lường số lượng hàng tồn kho đang thực hiện so với các đơn đặt hàng đã lên kế hoạch.
  - **GMROI:** là một tỷ lệ tài chính đo lường lợi nhuận của một công ty từ các khoản đầu tư hàng tồn kho của mình.
- Độ chính xác của dự báo nhu cầu: Đo lường mức độ tồn kho phù hợp với nhu cầu thực tế như thế nào. Độ chính xác cao có thể dẫn đến mức tồn kho tốt hơn và giảm chi phí.

# PHƯƠNG PHÁP ĐÁNH GIÁ

- Đánh giá các giả thuyết cho việc chọn mặt hàng tồn kho: *P value*.
- Đánh giá chatbot: Nhóm sẽ đánh giá dựa trên độ hài lòng của khách hàng bằng cách thêm tính năng like/dislike của mỗi câu phản hồi.



Minh họa độ đo dựa trên đánh giá của người dùng.

# TIMELINE VÀ ROAD-MAP

## Chatbot AI hỗ trợ dự báo và quản lý hàng tồn kho

Người thực hiện

Team Dumplings

DATE: 19/11/2023

Bước	Chi tiết	W1	W2		W3		D E P L O Y Ý T Ủ R Ọ N G
		Tháng 11	Tháng 12		Tháng 12		
	Công việc	13 14 15 16 17	1	2 3 4 5	6 7 8	9 10 11 12 13	14 15 16 17
1	Lên ý tưởng	Dữ liệu và các chủ đề cuộc thi Thị trường và các vấn đề liên quan đến tài Hướng tiếp cận, độ khả thi					
2	Tiền xử lý dữ liệu	Làm sạch dữ liệu Insert dữ liệu vào database					
3	Mô hình dự đoán	Kiểm định T-test, ANOVA, Causal Impact và Prophet Mô hình XGBoost, LightGBM,.. Huấn luyện và optimize Review lại kết quả					
4	Chatbot AI	Chuẩn bị dữ liệu cho chatbot Text to SQL Huấn luyện và optimize					
5	Báo cáo	Viết báo cáo Deploy ý tưởng					

Dataset

Sales and inventory Snapshot data

# CƠ HỘI

- AI có thể xử lý lượng lớn dữ liệu, phân tích và dự báo với độ chính xác cao nhờ máy học.
- Cải thiện độ chính xác dự báo nhu cầu và giảm lỗi so với các phương pháp truyền thống.
- Giúp tối ưu hóa việc nhập hàng, quản lý tồn kho và lên kế hoạch khuyến mãi.

# THÁCH THỨC

- Đòi hỏi lượng lớn dữ liệu để huấn luyện mô hình AI.
- Tích hợp và vận hành AI đòi hỏi kỹ năng chuyên môn cao.
- Mô hình AI cần được cập nhật và huấn luyện liên tục để duy trì độ chính xác.