

# Demo Session: Price Discovery



# Overview

- What has been completed
- Current Progress
- Next Steps



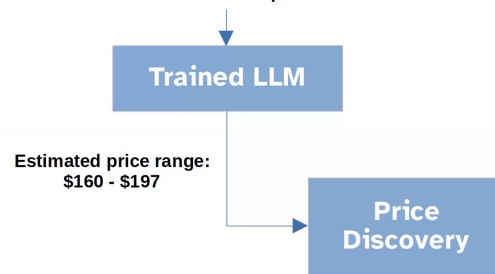
---

# Objective

Develop LLM that sees product image and reads product description to establish price range.



Silver necklace with semiprecious stone



# Recap

## Week 1

- Collected Amazon 2023 product dataset ~ 1.4m entries
- Feature engineering on dataset
- Data scraper tested

## Week 2

- CLIP w/ statistical Architecture designed and tested
- Text and image embeddings created
- Evaluation metrics selected and tested preliminarily

## Week 3

- vit-large CLIP model used
- Evaluation metrics improved
- RAG implementation
- Price range outputting for products

## Week 6

- Documentation
- Evaluation Metric

## Week 5

- Migrating DB to the cloud
- Experiment Knowledge graph db
- Product demonstration

## Week 4

- Improvement of prompts
- Agent and prompt evaluations
- Fine tuning LLM



# This Week

- Implemented databases into pipeline
- Finalize evaluation metric

---

# Database Implementation



# Outside Dataset

## Price Range Generator 🧠

🔗 An app that takes in the product image and the description and outputs a price range

Upload a product image



Drag and drop file here

Limit 200MB per file • JPG, JPEG, PNG

Browse files



910USMr6wL\_AC\_SX342\_.jpg 40.3KB



Enter product description

Women's Lightweight Open-Front Cardigan Sweater



Women's Lightweight Open-Front Cardigan Sweater

Get Price range

The Price Range (in Dollars) is 20.30-28.35

The price range is derived from the prices of similar Women's Lightweight Open-Front Cardigan Sweaters found on Amazon.com, ranging from 20.30 to 28.35. These prices are based on user ratings and are considered accurate and compact estimates.

**Actual Price: \$22**

# In Dataset

## Price Range Generator 🧠

🔗 An app that takes in the product image and the description and outputs a price range

Upload a product image



Drag and drop file here

Limit 200MB per file • JPG, JPEG, PNG

Browse files



813cdih06L\_AC\_SL1500\_.jpg 133.3KB



Enter product description

34-Inch Curved Computer Monitor



34-Inch Curved Computer Monitor

Get Price range

The Price Range (in Dollars) is 249.99-454.99

The price range for a 34inch curved ultrawide LG monitor with 3440 x 1440 resolution, HDR 10, and black color is estimated to be between 249.99 and 454.99, based on the prices of similar LG monitors found on the internet.

**Actual Price: \$427**

Action: search text db

Action Input: LG 34-inch UltraWide Curved QHD IPS Monitor with HDR10 and FreeSync (3440 x 1440) Black

Observation2024-03-21 15:09:09,289 - httpx - INFO - HTTP Request: POST https://1883d972-3d1a-4e6d-a846-86b9952741cd.us-east4-0.gcp.cloud.qdrant.io:6333/collections/apparel-collection/points/search "HTTP/1.1 200 OK"  
[ScoredPoint(id='0a5b4f55-5646-49de-bb91-837e0ec20962', version=1317, score=0.13870408, payload={'Price': 12.99, 'asin': 'B01NAPP8G0', 'title': 'White Wedding Handkerchief Pure Cotton, Soft Ladies Hanky with Crochet Lace -None}), ScoredPoint(id='1bf2f1db-dd88-411c-8650-2601222c13cd', version=1398, score=0.12775262, payload={'Price': 15.99, 'asin': 'B089756CJQ', 'title': 'Dreams@Bharatanatyam White & Orange Single Color Jasmine Flower symbol for Gorgeous and Stunning Hairstyles'}, vector=None, shard\_key=None), ScoredPoint(id='33beb874-eee1-4af5-983e-86493415123e', version=1190, score=0.12543732, payload={'Price': 29.66, 'asin': 'B08FDR06SS', 'title': 'Gorgeous Long Cathedral Veils for Brides Soft Tulle Bridal Veils with Comb 118" (Ivory)'}, vector=None, shard\_key=None)]The text database search did not yield any relevant results for the LG 34-inch UltraWide Curved QHD IPS Monitor (3440 x 1440) Black. I will now search for the product in the image database.

Action: search image db

Action Input: LG 34-inch UltraWide Curved QHD IPS Monitor with HDR10 and FreeSync (3440 x 1440) Black

Observation2024-03-21 15:09:10,318 - httpx - INFO - HTTP Request: POST https://1883d972-3d1a-4e6d-a846-86b9952741cd.us-east4-0.gcp.cloud.qdrant.io:6333/collections/img-collection/points/search "HTTP/1.1 200 OK"  
[ScoredPoint(id='c29ff088-10e4-47de-bbaa-59621db548ac', version=1295, score=0.14220831, payload={'Price': 10.99, 'asin': 'B09DG6DSY2', 'imageUrl': 'https://m.media-amazon.com/images/I/810Vimu85tL.AC\_UL320\_.jpg', 'title': 'Classic Movie Characters Acrylic Keyring Accessories for Keys Wallets Backpack Scary Movies Figure Pendant Hanging Key Chain Decorations with Rotatable Link for Gift'}, vector=None, shard\_key=None), ScoredPoint(id='6857a08854', version=1859, score=0.13978866, payload={'Price': 16.99, 'asin': 'B0CC91NKK6', 'imageUrl': 'https://m.media-amazon.com/images/I/81Nb4SMQQL.AC\_UL320\_.jpg', 'title': '20 Pieces Cute Ghost Enamel Pin Kawaii Pins Halloween Brooches Badges Aesthetic Gift Cartoon Boo Pumpkin Flying Spirit Enamel Pins for Women Men Backpack Jacket Clothes'}, vector=None, shard\_key=None), ScoredPoint(id='9213e398-6fa6-49e0-9bb7-21fbb7fd1146', version=1208, score=0.1199, 'asin': 'B0B5QH925T', 'imageUrl': 'https://m.media-amazon.com/images/I/71T95bFjOGL.AC\_UL320\_.jpg', 'title': '10 Packs Horror Classic Movie Characters Keychains Halloween Horror Keychains for Keys Wallets Backpack Pendant Hanging KeyChain Decorations Horror Movie Gifts Merch for Men Women'}, vector=None, shard\_key=None)]Thought: The image database search did not yield any relevant results for the LG 34-inch UltraWide Curved QHD IPS Monitor (3440 x 1440) Black. I will now search for the product on the internet.

Action: search internet

Action Input: LG 34-inch UltraWide Curved QHD IPS Monitor with HDR10 and FreeSync (3440 x 1440) Black



---

# Evaluation Metric



# Evaluation metrics

Evaluate the final answer: Price Range + Reason

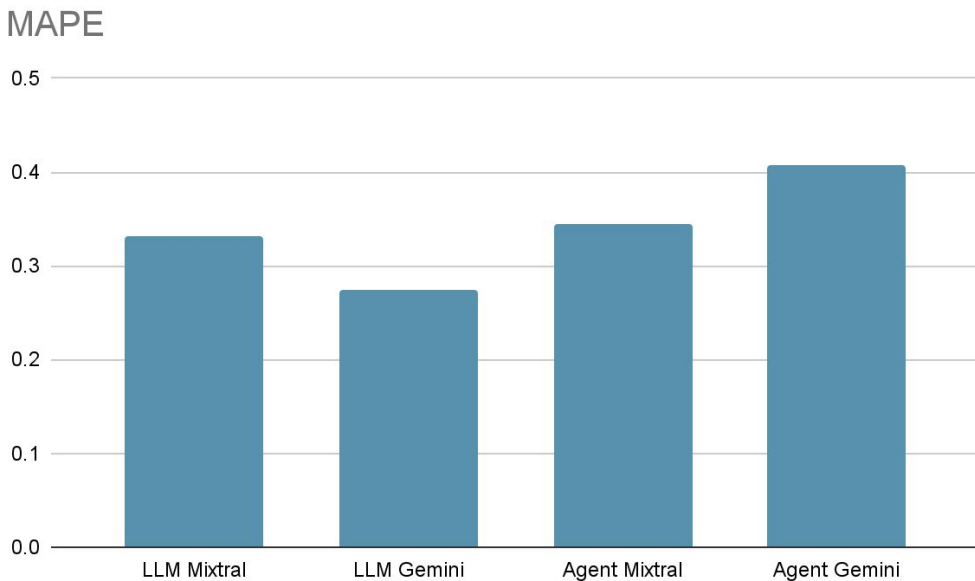
- Mean Absolute Percentage Error (MAPE)
- Reasoning score (LLM-based)

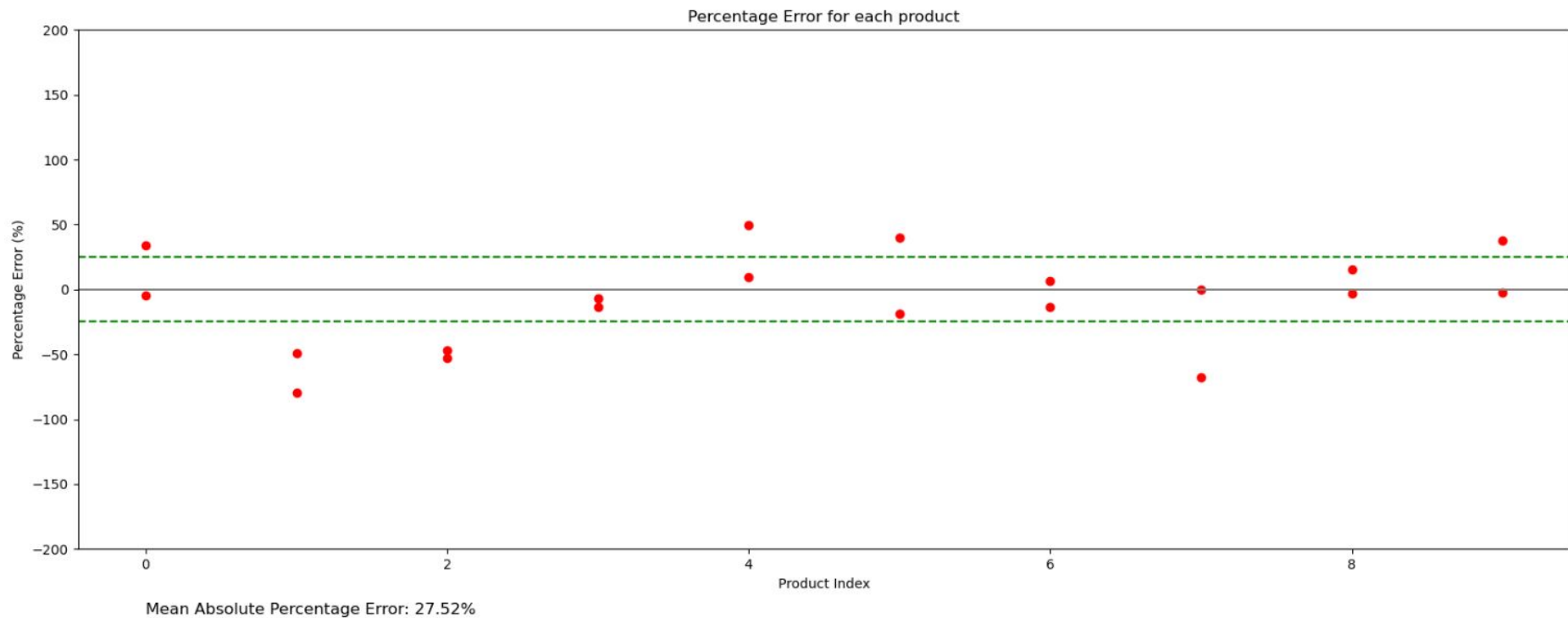
Evaluate the context: All products retrieved from internet or database

- BLEU and ROUGE score
- Context Relevance score (LLM-based)

# MAPE

Calculate the percentage error of the min and max of test products then take the average





Note: ylim is set to (-200, 200) but some errors may be outside this range

# Reasoning Score

Use an LLM to evaluate the reason produced

```
reasoning_with_cot_reasons_prompt_template = ""
```

```
CRITERIA: PRICE RANGE EVALUATION
```

You are a PRICE RANGE EVALUATOR; assessing the appropriateness and reasoning behind a given price range for a product or service.

Rate the quality of the PRICE RANGE RESPONSE as well as the PRICE RANGE and INTERMEDIATE STEPS on a scale from 0 to 10, where 0 indicates an extremely poor response \ and 100 indicates an excellent response. Ignore INTERMEDIATE STEPS if they are not provided.

Consider the following factors when evaluating:

Clarity: Is the provided price range an actual price range?

Justification: Are the reasons given for the price range logical and well-explained?

Completeness: Does the response cover all relevant factors that could influence pricing?

Accuracy: Does the price range seem reasonable and align with market rates based on your knowledge?

Level of Detail: Is the reasoning sufficiently detailed or does it oversimplify the pricing factors?

Step Validity: Are the intermediate steps taken to arrive at the price range valid and logical?

Step Completeness: Do the steps cover all the necessary considerations for determining an appropriate price?

A low score (0-3) should be given if the response is unclear, lacks justification, the steps are invalid, or the reasoning is blatantly inaccurate.

A mid-range score (4-7) is appropriate if the response provides a reasonable price range with some justification, but the steps are incomplete, \ lack important details or have minor inaccuracies.

A high score (8-10) should be awarded if the response gives a well-justified and accurate price range, the intermediate steps are valid and complete, \ covering all major pricing factors in a clear and detailed manner.

```
PRICE_RANGE_RESPONSE: {price_range_response}
```

```
PRICE_RANGE: {price_range}
```

```
INTERMEDIATE_STEPS: {intermediate_steps}
```

Please answer using the entire template below.

TEMPLATE:

Score: <The score 0-10 based on the soundness of the reasoning>

Criteria: <Criteria is always Price Range Evaluation>

Supporting Evidence: <Provide a detailed explanation justifying the score based on the criteria listed above, including an evaluation of the \ intermediate steps taken and how they factor into the overall scoring>

```
""
```

Score: 9

Criteria: Price Range Evaluation

Supporting Evidence:

Clarity: The provided price range is clear and unambiguous, falling between \$77 and \$150.

Justification: The response provides a valid justification for the price range by referencing similar products found in a text database search. The response also includes a search of the internet, providing further evidence for the price range.

Completeness: The response considers relevant factors that could influence pricing, such as the product's features (rose gold case, black band, water resistance) and comparable market prices.

Accuracy: The price range appears reasonable based on the information provided and aligns with market rates for similar products.

Level of Detail: The reasoning is sufficiently detailed, providing specific examples of comparable products and their prices.

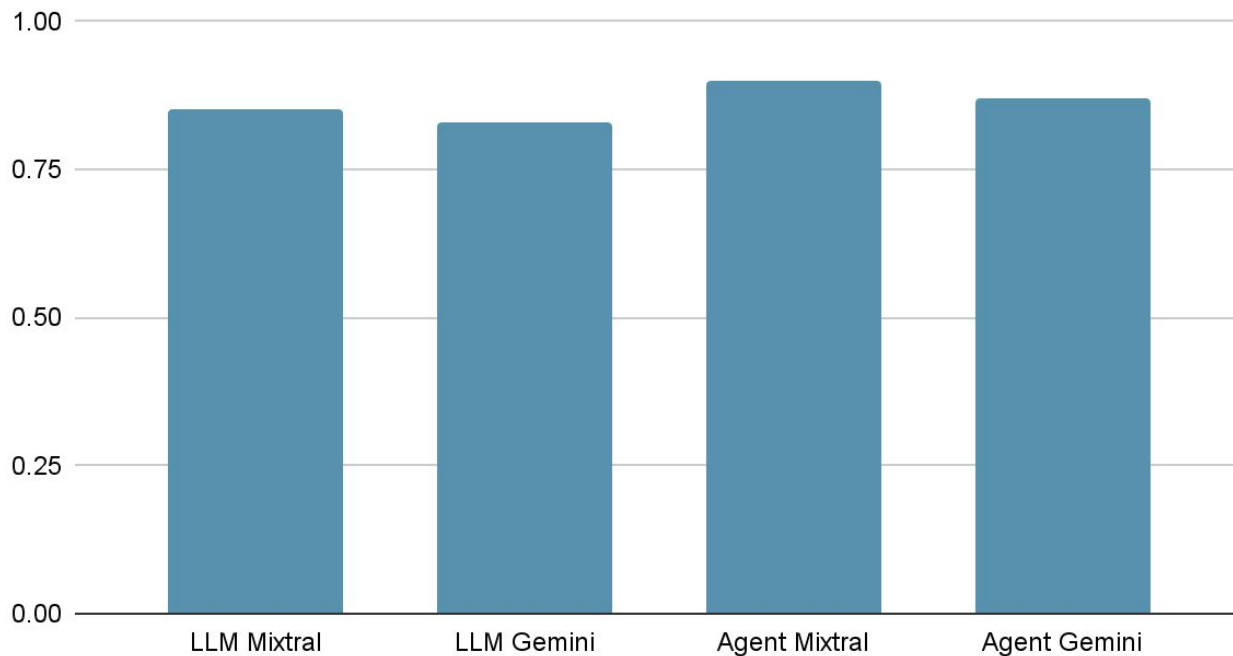
Step Validity: The intermediate steps are valid and logical, involving a search for similar products in both a text database and the internet.

Step Completeness: The steps cover the necessary considerations for determining an appropriate price, including identifying comparable products and their pricing.

Overall, the response demonstrates a clear understanding of the factors that influence pricing and provides a well-justified and accurate price range. The intermediate steps are valid and complete, supporting the overall reasoning.

# Reasoning Score

Reasoning score

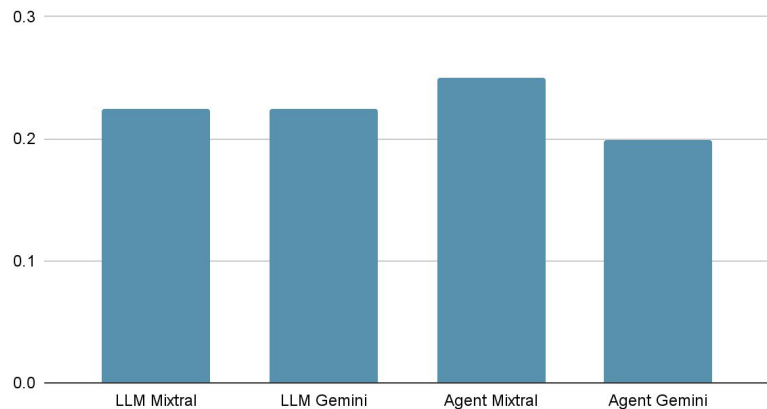




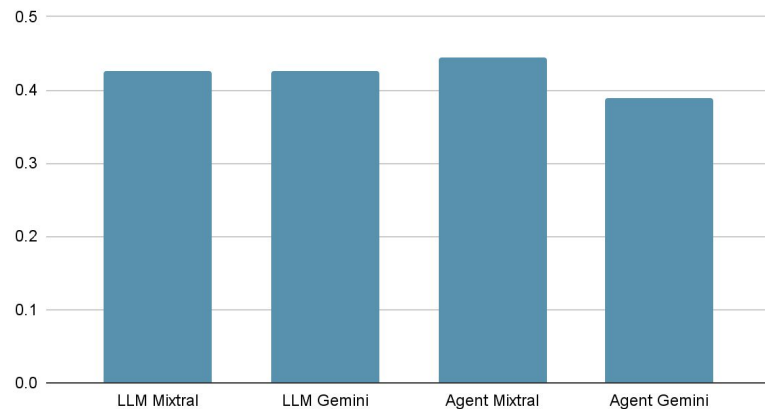
# BLEU and ROUGE score

Evaluate the retrieved products based on word match

BLEU



ROUGE



# Context Relevance score

Use an LLM to evaluate the relevance of the products retrieved

```
context_relevance_with_cot_reasons_prompt_template = """
CRITERIA: RELEVANCE
```

You are a PRODUCT RELEVANCE evaluator; assessing the relevance of the RETRIEVED PRODUCT to the ORIGINAL PRODUCT based on detailed criteria.

Rate the relevance on a scale from 0 to 10, where 0 signifies no relevance and 10 indicates perfect relevance.

A few additional evaluation guidelines:

- The length of the PRODUCT DESCRIPTIONS should not affect the score.
- The RELEVANCE score should increase as the RETRIEVED PRODUCT shares more similarities with the ORIGINAL PRODUCT in terms of:
  - Product Type: Both products belong to the same category (e.g., both are T-shirts, both are watches).
  - Functionality: Both products serve the primary use in a similar way (e.g., running shoes vs. casual shoes).
  - Target Users: Both products are intended for the same gender, age group, etc.
  - Style: Both products share a similar design, color scheme, etc.
  - Material: Both products are made from similar materials.
  - Brand: Both products are from the same manufacturer or brand.
- A RETRIEVED PRODUCT that matches only a few aspects of the ORIGINAL PRODUCT should receive a score of 2, 3, or 4. A higher score within this range indicates a greater degree of relevance.
- A RETRIEVED PRODUCT that matches many aspects of the ORIGINAL PRODUCT should receive a score of 5, 6, 7, or 8. A higher score within this range indicates a greater degree of relevance.
- A RETRIEVED PRODUCT that matches almost all aspects of the ORIGINAL PRODUCT should receive a score of 9 or 10, with 10 reserved for products that are highly relevant and comparable to the ORIGINAL
- Never elaborate.

ORIGINAL PRODUCT: {original\_product}

RETRIEVED PRODUCT: {context\_product}

RELEVANCE:

Please answer using the entire template below.

TEMPLATE:

Score: <The score 0-10 based on the relevance>

Criteria: <Criteria is always Relevance>

Supporting Evidence: <Provide your reasons for scoring based on the listed criteria step by step. Tie it back to the comparison being completed>

"""

Score: 9

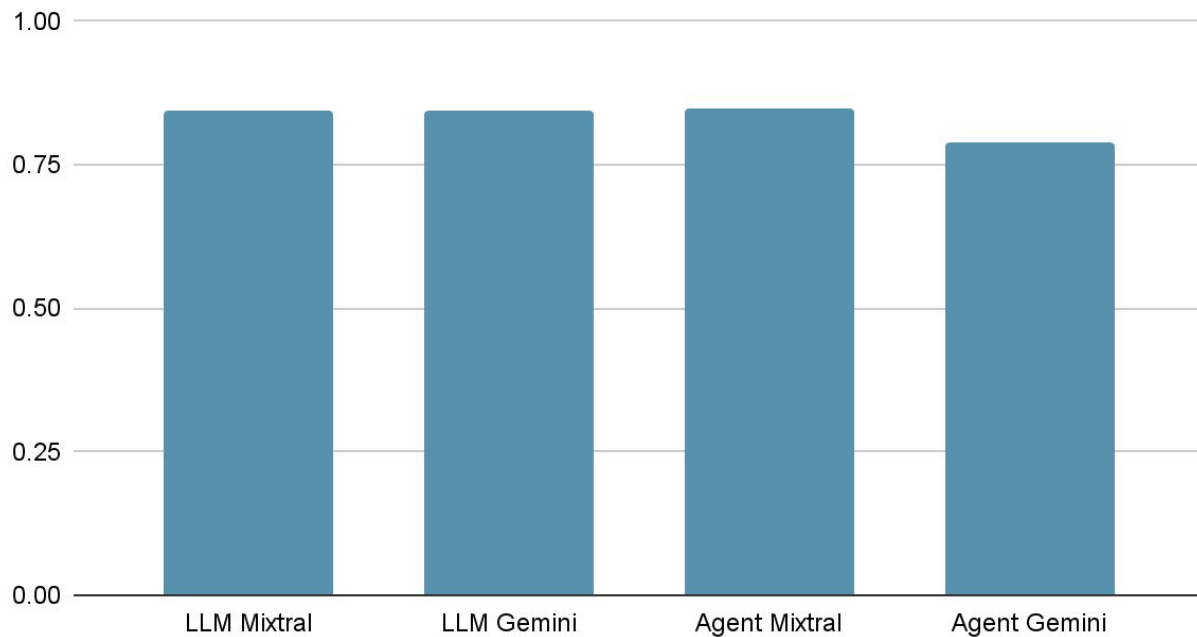
Criteria: Relevance

Supporting Evidence:

- Product Type: Both products are crossbody bags.
- Functionality: Both products provide anti-theft features and are designed for carrying personal belongings.
- Target Users: Both products are likely intended for adults.
- Style: Both products have a similar design and color scheme.
- Material: Although the material is not specified in the retrieved product description, both products are likely made from durable materials to withstand everyday use.
- Brand: Both products are from the same brand, Travelon, which specializes in anti-theft travel gear.

# Context Relevance score

Context Relevance



---

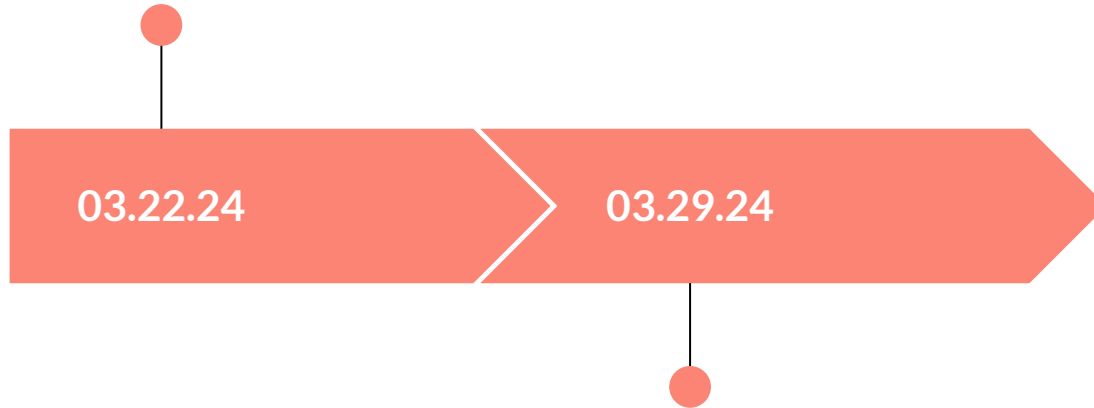
# Next Steps



# Next Steps

- Evaluation metric finalization
  - Testing out dataset of 100
- Pipeline Implementation
- Complete documentation
- UI
- This Week
  - Finalize documentation
    - Due by Wednesday
  - Create **docker**
    - **Kimani can do and finish by Tuesday**

Product & Evaluation  
Metric finalization



Product Demonstration,  
Product Deployment, &  
Documentation





---

# Feedback?

