

Demo Session: Price Discovery



Overview

- What has been completed
- Current Progress
- Next Steps

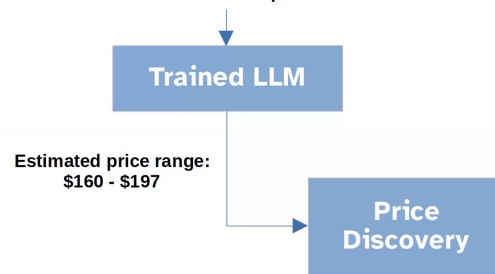


Objective

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 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

- Migrating the Vector Database from local machine to Cloud.
- Experimented with Knowledge graph database to improve model.
- Developed UI for product MVP

Database Migration



Database: Qdrant

- Efficient similarity search
- Vector Indexing
- REST API and Client Libraries

Progress on deployment

- Completed collection of ~50,000 products with embedded vectors.
 - Text collection - 49,982 products
 - Image collection - 50,000 products
- Ready to be implemented and used in production.

Proposed improvements.

- Using different embedding functions.
 - Dual-Path Convolutional Image-Text Embeddings

Resources used


- Sentence transformers in hugging face.
- Trung's 50,000 apparel dataset.

```
embeddings = OpenCLIPEmbeddings(model_name="ViT-B-32", checkpoint="laion2b_s34b_b79k")
```

img-collection

POINTS INFO SNAPSHOTS VISUALIZE

Collection Info

| | |
|-----------------------|---|
| status | green  |
| optimizer_status | ok |
| vectors_count | 50064 |
| indexed_vectors_count | 49352 |
| points_count | 50000 |
| segments_count | 2 |
| config | {...} 5 Items |
| payload_schema | { } 0 Items |

Collection Cluster Info

| Shard ID | Location | Status |
|----------|--------------------------|--------|
| 0 | Local (8318655547226364) | Active |

Point 00029ad1-a7cb-4084-85de-76b9771b58b1

Payload:

| | |
|--------|---|
| Price | 79.98 |
| asin | B0CF66FT9G |
| imgUrl | https://m.media-amazon.com/images/I/91SJMTwSM2L__AC_UL320_.jpg |
| title | Overwatch Black Steel Toe Sneakers for Men, Waterproof Comfortable Slip Resistant Mens Steel Toe Shoes, Men's Work & Safety Footwear - Medium and Extra Wide Width Invincible Mens Work Shoes |

Vectors:

Default vector


Length: 512

FIND SIMILAR

apparel-collection

POINTS INFO SNAPSHOTS VISUALIZE

Collection Info

| | |
|-----------------------|---|
| status | green  |
| optimizer_status | ok |
| vectors_count | 50001 |
| indexed_vectors_count | 40192 |
| points_count | 50001 |
| segments_count | 2 |
| config | { "_": 5 items |
| payload_schema | { } 0 items |

Collection Cluster Info



drant

Text Vector Collection

apparel-collection

POINTS INFO SNAPSHOTS VISUALIZE

Collection Info

| | |
|-----------------------|-------------------|
| status | green ● |
| optimizer_status | ok |
| vectors_count | 50046 |
| indexed_vectors_count | 48116 |
| points_count | 49982 |
| segments_count | 2 |
| config | > { ... } 5 Items |
| payload_schema | { } 0 Items |


Example point:

Point 00016b38-34aa-4228-ab2e-2df7f0ea2f5a

Payload:

| | |
|-------|---|
| Price | 41.5 |
| asin | B0771L565H |
| title | Women Hooded Knit Cardigans Button Cable Sweater Coat |

Vectors:

Default vector 

Length: 512

2,500 products visualized:



drant

Image Vector Collection

img-collection

POINTS INFO SNAPSHOTS VISUALIZE

Collection Info


| | |
|-----------------------|-------------------|
| status | green ● |
| optimizer_status | ok |
| vectors_count | 50064 |
| indexed_vectors_count | 49352 |
| points_count | 50000 |
| segments_count | 2 |
| config | > { ... } 5 Items |
| payload_schema | { } 0 Items |

Example point:


Point 000c8c21-2a24-4e5a-b2db-195181195380

Payload:

| | |
|--------|--|
| Price | 13.99 |
| asin | B07H6JVVVP |
| imgUrl | https://m.media-amazon.com/images/I/71TA025YE3L_AC_UL320_.jpg |
| title | Men's Winter Hat Thick Knit Beanie Cap with Visor Newsboy Beanie Hat |



Vectors:

Default vector  Length: 512 FIND SIM

2,500 products visualized:



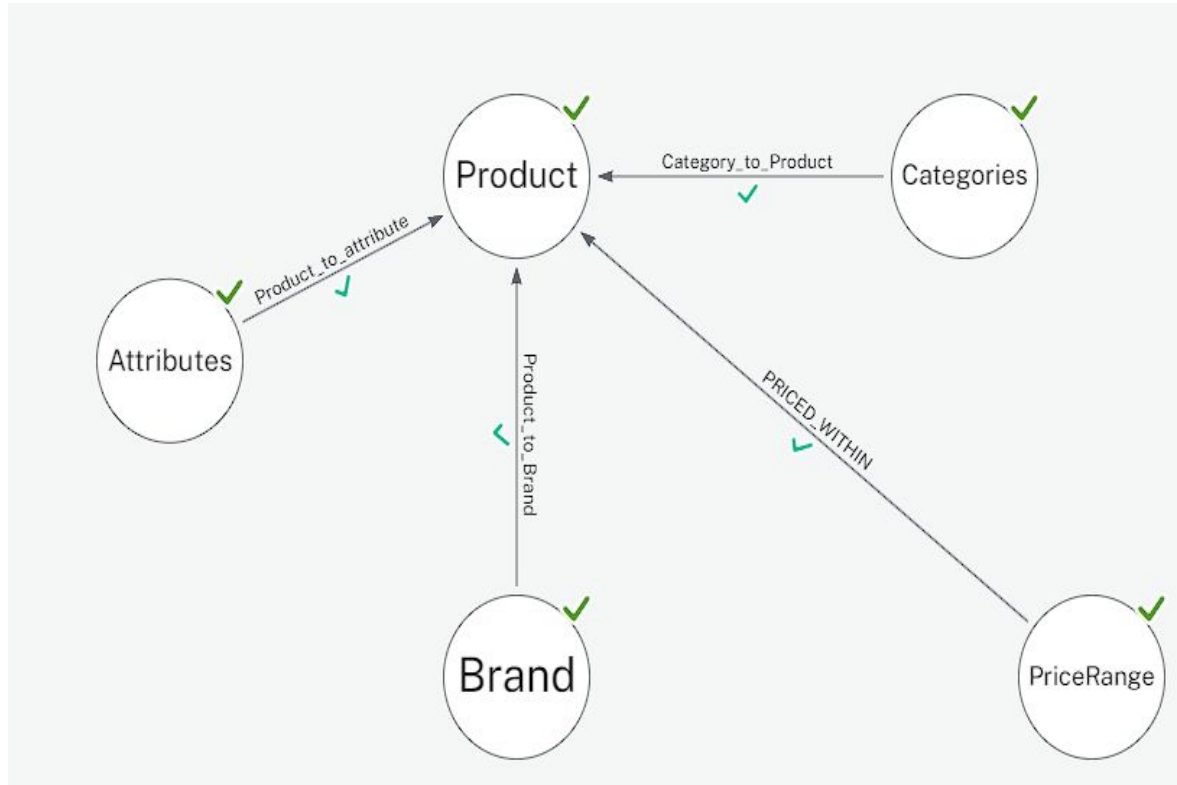
Knowledge Graph Database



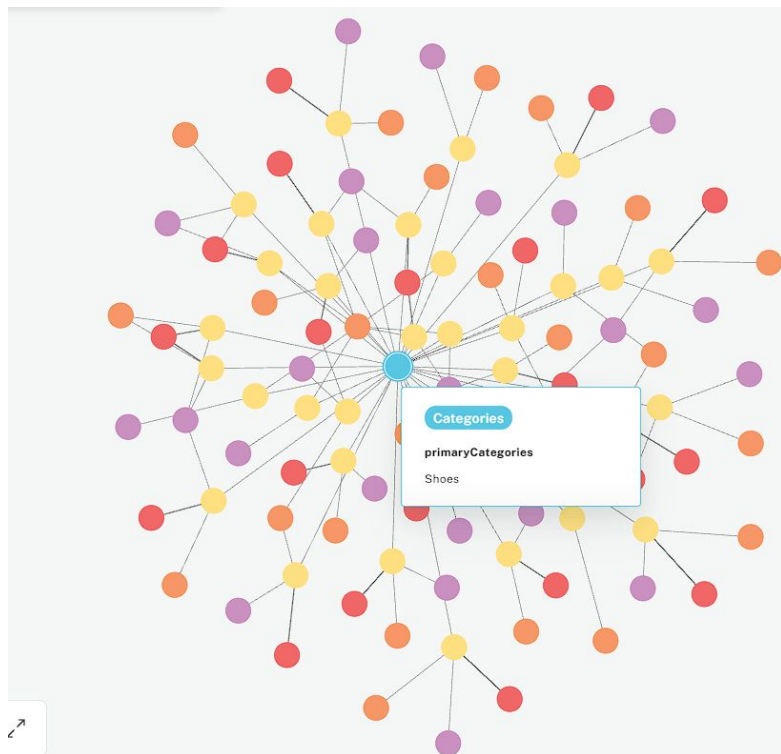
Why Knowledge graphs?






- **A knowledge graph is a way to store and organize information that makes the relationships between pieces of data clear and easy to understand.**
- **Follow up from last week; we used a data scraper to extract products from amazon in 10 different categories .**






Experimentation process



Results



| Nodes | | Relationships |
|---|------------|---------------|
| Filter categories | | |
| <input checked="" type="radio"/> All <input type="radio"/> In Scene <input type="radio"/> Off Scene | | |
|  | Attributes | 25 |
|  | Brand | 24 |
|  | Categories | 1 |
|  | PriceRange | 21 |
|  | Product | 34 |

| Nodes | | Relationships |
|---|----------------------|---------------|
| Filter relationships | | |
| <input checked="" type="radio"/> All <input type="radio"/> In Scene <input type="radio"/> Off Scene | | |
|  | PRICED_WITHIN | 20 |
|  | Price_to_Product | 25 |
|  | Product_to_Brand | 33 |
|  | Product_to_attribute | 34 |
|  | Product_to_category | 33 |

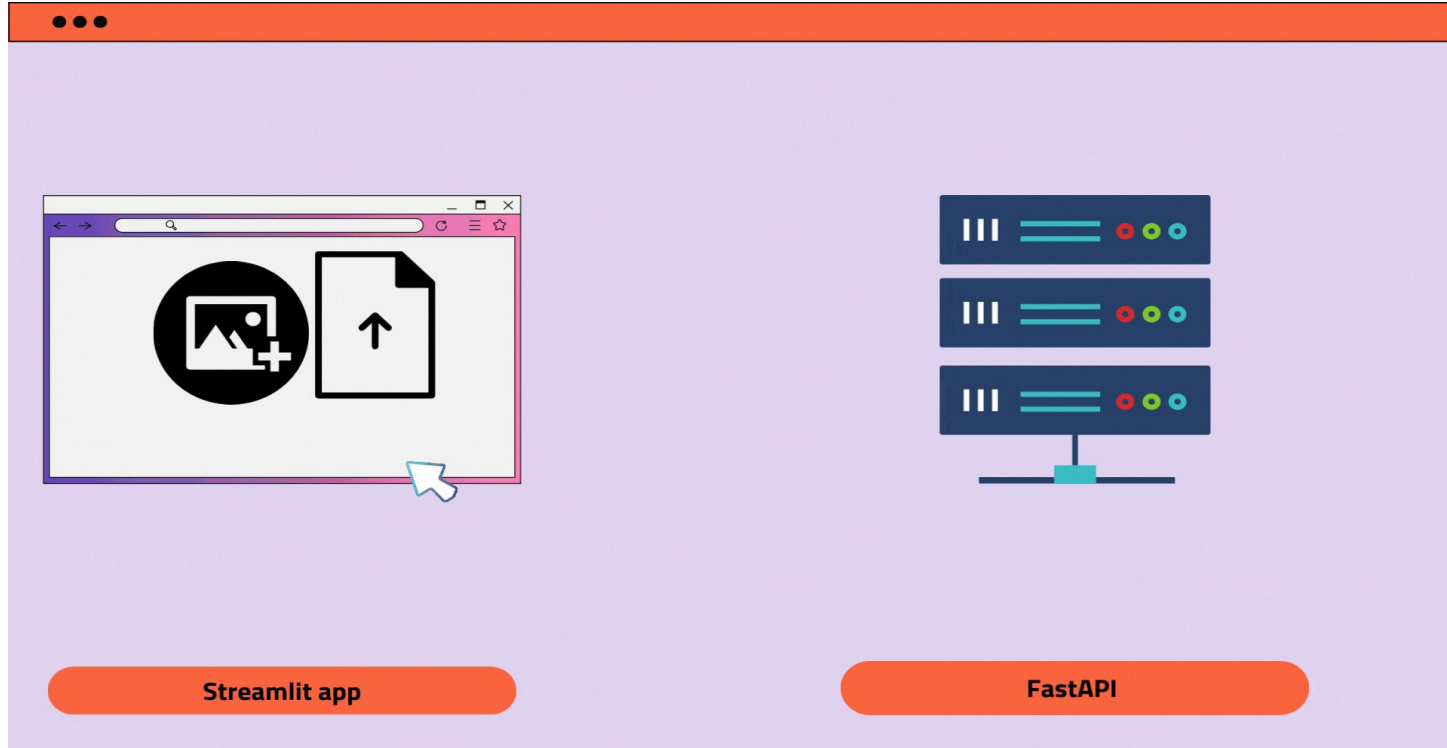
Limitations

- Missing data reduced inconsistency
- Browser was unable to support large range of products at a time
- Insufficient data, inconsistent data and Irregular values were the key blockers

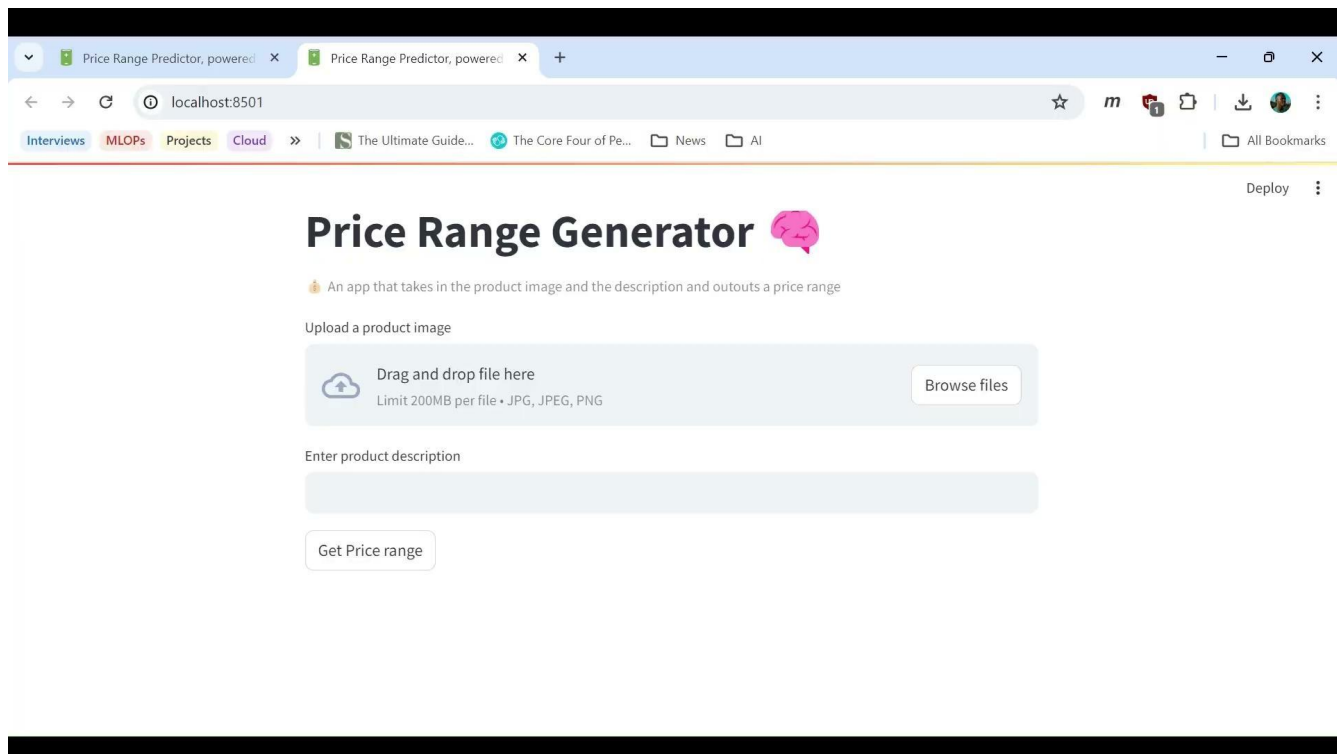
Product Demonstration



Basic Flow



Demo



Next Steps

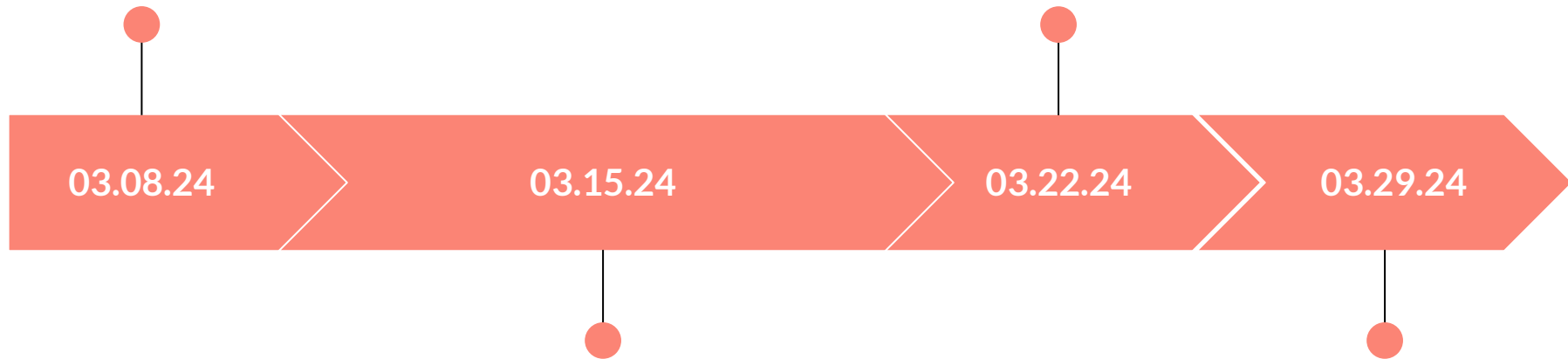


Next Steps

- UI
 - Prompting and Output formatting
 - Improve on the attractiveness of design
- Evaluation metric finalization
 - MAPE
 - Rouge & BLU
 - Ragas
- Pipeline Implementation
 - Show them that data is in database
 - Show images that show that product is in the database

Product Fine Tuning

Product & Evaluation
Metric finalization



Product Demonstration &
Cloud Deployment

Final Product
Demonstration &
Documentation



Feedback?

