RELvogue AI (Relevant vogue)

Fashion Product Information Retrieval System.

Project Goals

- The primary goal of this project is to create an intelligent assistant capable of retrieving detailed information about fashion products based on user queries.
- The system integrates *semantic search, reranking* with *cross-encoders,* and *retrieval-augmented generation (RAG)* using OpenAI's GPT-3.5.

Data Sources

- The project utilizes a fashion dataset sourced from Myntra, containing attributes such as product names, descriptions, prices, colors, brands, ratings, and product attributes.
- This dataset serves as the foundation for building a robust information retrieval system.

Design Choices

1. Used Open AI's text embedding model:

- The *"text-embedding-ada-002"* model was used to set up embedding function.

2. Semantic Search with Cache:

- Implemented to quickly retrieve relevant documents based on user queries. Uses ChromaDB for efficient storage and retrieval of embeddings.
- Created two collections, one for the main storage and the other for the cache.

3. Reranking with Cross-Encoders:

- Utilized a pre-trained Cross-Encoder model to improve result relevance by scoring retrieved documents against the user query.
- The cross-encoder model used was *"cross-encoder/ms-macro-MiniLM-L-6-v2"*.

4. Retrieval-Augmented Generation (RAG):

- OpenAI's GPT-3.5 is employed to generate detailed responses based on both the user query and the retrieved document information.

5. Integration of Layers:

- The system seamlessly integrates these components to provide a user-friendly interface for querying fashion product details.

Challenges Faced

- **Data Preprocessing:**
  Handling missing data, text preprocessing, and ensuring uniformity in data formats across different attributes.
- **Integration Complexity:**
  Ensuring smooth integration between semantic search, reranking, and generation layers to provide coherent and accurate responses.
- **Performance Optimization:**
  Balancing performance with accuracy, especially in prompting.
- **Adding Documents to ChromaDB Collection:**
  Faced "rate limit reached" error many times from transformer model.

Conclusion

This project demonstrates the integration of advanced AI techniques to build a functional fashion product information retrieval system. By leveraging semantic search, cross-encoder reranking, and GPT-3.5 for generation, the system delivers accurate and contextually relevant responses to user queries in the fashion domain.

An Example of final Result:

User: I need a Kurta with palazzo and dupatta from Libas brand

Assisstant: I found a Libas product that matches your search query for a Kurta with palazzo and dupatta:

- Name: Libas Women Maroon Printed Kurta with Palazzo and Dupatta
- Details: This set includes a maroon printed kurta paired with matching palazzo pants and a dupatta.

This product should meet your requirements. If you need further assistance or have more specific preferences, feel free to ask!

The result from semantic search layer are as follows:

index	IDs	Documents	Distances	Metadatas	Reranked_scores
0	420	name: libas women black floral printed gotta patti pure cotton kurta with palazzos & with dupatta. products: kurta, palazzos, dupatta. price: 2999.0. colour: Black. brand: Libas. "image_url: http://assets.myntassets.com/assets/images/16857572/2022/2/2/3238da41-456a-4b7d-a5b9-3e7c9ca24e4b1643786560766-Libas-Women-Kurta-Sets-7241643786559765-1.jpg". rating count: 38.0. average rating: 3.789473684210526. description: our 3piece black printed cotton suit set features a floral print straightfit kurta with a notched neckline and gota patti accents on the yoke and sleeves flared palazzo pants with a partially elasticated waistband and roomy pockets and a complementing dupatta a perfect pick for celebratory eveningspure cottongentle machine washthe model height 58 is wearing a size s. Product Details: addons na body shape id 443333324424 body or garment size garment measurements in bottom closure slipon bottom fabric pure cotton bottom pattern printed bottom type palazzos character na dupatta with dupatta dupatta border printed dupatta fabric pure cotton dupatta pattern printed main trend na neck round neck number of pockets 2 occasion festive ornamentation gotta patti pattern coverage none print or pattern type floral sleeve length threequarter sleeves sleeve styling regular sleeves slit detail side slits stitch ready to wear sustainable regular technique na top design styling regular top fabric pure cotton top hemline straight top length calf length top pattern printed top shape straight top type kurta waistband partially elasticated wash care machine wash weave pattern regular weave type knitted and woven wedding guests gifting.	0.2508014440536499	{'Name': 'libas women black floral printed gotta patti pure cotton kurta with palazzos & with dupatta', 'Price': '2999.0', 'Product_ID': 16857572}	6.528972625732422