**Project Title:**

**“Evaluating Retrieval-Augmented Generation (RAG) based Personalized Product Discovery and Automated Content Generation”**

**7COM1075: Data Science and Analytics Masters Project**

|  |  |
| --- | --- |
| **NAME** | **STUDENT ID** |
| **Anagha Santhosh** | **23001866** |

## AIM :

This project investigates how **Retrieval-Augmented Generation (RAG)** compares to **standalone Large Language Models (LLMs)** in the context of **personalized product recommendations** and **AI-driven content generation**. Specifically, we aim to:

* **Improve the relevance & personalisation** of product recommendations using RAG-powered retrieval.
* **Enhance explainability** by providing justifications for recommendations based on retrieved product reviews & metadata.
* **Automate marketing content generation** (e.g., product descriptions, ads, and blog content) using AI-enhanced retrieval.

This study will evaluate whether **RAG-based retrieval mechanisms outperform traditional LLMs** in delivering **accurate, explainable, and personalized content** for e-commerce applications.

Research Question:

“**How effective is RAG in generating personalized e-commerce recommendations and automated content compared to standalone LLMs?”**

Objectives:

* **Develop a RAG-based recommendation system** that retrieves customer reviews & product metadata to generate **contextually aware** recommendations.
* **Compare RAG vs. standalone LLMs** in terms of **recommendation relevance, explainability, and content quality.**
* **Automate AI-powered content generation** for product descriptions, email campaigns, and social media posts.
* **Incorporate explainability mechanisms** to justify why a product was recommended.
* **Benchmark RAG-based outputs against traditional LLMs** using objective evaluation metrics.

Description of the idea:

The proposed system will utilise **RAG with a vector database** to enhance **personalised product recommendations and content generation**. Unlike **standard LLMs**, which generate responses solely based on pre-trained knowledge, **RAG retrieves relevant customer reviews & metadata before generating recommendations**.

The **retrieval component** fetches the **top-K most relevant reviews & metadata**, while the **generation component** synthesizes personalized recommendations & AI-generated marketing content.

The primary motivation for integrating product recommendation and content generation is the growing influence of social media influencers as key drivers of marketing and sales. By providing both personalized recommendations and AI-generated content, this system can enhance influencer-driven promotions, ultimately boosting engagement and sales conversions.

Research Methodology:

### Identify dataset :

Dataset has been collected from:

<https://huggingface.co/datasets/McAuley-Lab/Amazon-Reviews-2023>

Contains:

* + - **Amazon Reviews Dataset**
    - **Product Metadata**

**I am considering to use only certain categories from the vast available resource, keeping in mind the resource limitations and complexity.**

**Here is a snippet of data in review and metadata files:**

**1: Amazon Review dataset:**

A screenshot of a computer code

AI-generated content may be incorrect.

**Preview of Amazon Review dataset:**

A screenshot of a computer

AI-generated content may be incorrect.

**2: Product Metadata :**

A screenshot of a computer

AI-generated content may be incorrect.

**Preview of product Metadata:**

A screenshot of a phone number

AI-generated content may be incorrect.

A close-up of a box

AI-generated content may be incorrect.

### Data preprocessing

* Cleaning up the data
* Tokenize and vectorize **reviews, descriptions and metadata**.

( I may have to use a relational db for the metadata, for key word search. Further research needed )

* Generate **embeddings** using suitable **embedding models**.
* Store indexed data in **vector DB for fast retrieval**

(Identified Chroma DB and FAISS as potential options)

### Model Implementation:

* **Baseline Model (Standalone LLMs) without retrieval**.
* **Experimental Model (RAG-Based):** Retrieves **top-K product reviews & metadata** before generating responses.

RAG implementation diagram:

A diagram of a algorithm

AI-generated content may be incorrect.

Developing Rag pipeline:

A white rectangle with black text

AI-generated content may be incorrect.

### Evaluation Metrics:

These are the metrics that are currently identified to measure the performance.

1. **Compare Precision@K of RAG vs. LLM** (Higher = Better Recommendations).
2. **Evaluate Explainability with ROUGE Scores** (Higher = More Transparent).
3. **NDCG@K (Normalized Discounted Cumulative Gain)** to measure Ranking quality of retrieved recommendations.

### Expected Results

* RAG-based recommendations will outperform standalone LLMs in relevance & personalisation.
* AI-generated marketing content from RAG will be more context-aware & engaging.
* Explainability mechanisms in RAG will improve user trust & satisfaction.

### Citations:

1. Vaswani, A., 2017. Attention is all you need. *Advances in Neural Information Processing Systems*.
2. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.T., Rocktäschel, T. and Riedel, S., 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, *33*, pp.9459-9474.
3. Wang, Y., Hernandez, A.G., Kyslyi, R. and Kersting, N., 2024. Evaluating quality of answers for Retrieval-Augmented Generation: A strong LLM is all you need. arXiv preprint arXiv:2406.18064.
4. McAuley, J. and Yang, A., 2016, April. Addressing complex and subjective product-related queries with customer reviews. In Proceedings of the 25th International Conference on World Wide Web (pp. 625-635).
5. Roumeliotis, K.I., Tselikas, N.D. and Nasiopoulos, D.K., 2024. LLMs in e-commerce: a comparative analysis of GPT and LLaMA models in product review evaluation. *Natural Language Processing Journal*, *6*, p.100056.
6. Liu, J., Lin, J. and Liu, Y., 2024. How much can rag help the reasoning of llm?. arXiv preprint arXiv:2410.02338.