Utilizing Language Models and Machine Learning for Perfume Recommendations

Ashhar Anique - IIT2020174, Mohit Rathour - IIT2020194, Inder Sonu - IIT2020195
under the supervision of Prof. Pritish Kumar Varadwaj

Abstract - The recommendations presented in this report pertain to perfume datasets, but the methodology can be applied to other dataset types as well. Large Language Models (LLMs) such as GPT4, LLAMA, BERT, and GEMINI have demonstrated encouraging capabilities in the domain of natural language comprehension. While LLM has a wide range of applications, this paper focuses on two primary ones: text generation and the extraction of significant contextual features from a given text.

This study presents an innovative approach to perfume recommendation that utilizes Large Language Models (LLMs) in order to improve user contentment and experience during the fragrance selection process. Data Preprocessing and Feature Extraction, Data Generation, Fine-Tuning LLM, Bag of ML Models, and Ranking and Recommendation are the five primary phases of the strategy.

By employing Fragrantica as the source of data, the dataset is preprocessed in order to extract pertinent characteristics including age, gender, time, season, and fragrance notes. Following this, text paragraphs are generated by pretrained LLMs in response to customized queries extracted from the dataset. The fine-tuning LLM methodology utilizes methods such as parameter efficient fine-tuning (PEFT) in conjunction with quantized load sensitivity analysis (QLoRA) in order to enhance model performance and contextual sensitivity.

After this phase, perfume recommendations are generated using TF-IDF and Sentence Transformers in accordance with input parameters; the resulting recommendations are then combined to enhance accuracy. In the end, the ranking and recommendation system assigns priorities to perfumes according to parameters and fragrance characteristics provided by the user, including accord composition and note hierarchy.

The objective of this approach is to optimize the perfume selection process by providing individualized recommendations via interaction with natural language. The system endeavors to offer users more intuitive recommender systems through the utilization of LLMs and sophisticated techniques. This will enable users to engage in unrestricted communication with the system and acquire more precise recommendations through the implementation of natural language commands.

Keywords - LLM, ML, LLAMA, GPT, Recommendation, Perfume, Fragrantica, Notes, Accords, Perfume, Bag Of Models, Bagging, Cosine Similarity.

I. INTRODUCTION

Recommendation systems are becoming essential instruments in directing people in a variety of fields towards the best decisions. From personalized shopping recommendations to movie recommendations on streaming services, these systems examine user preferences and actions to offer customized recommendations that improve the decision-making process.

But among all the uses for recommendations, perfume recommendation—the process of recommending customized perfumes according to specific tastes—remains comparatively untapped. Perfume choice is very arbitrary and impacted by things like occasion, cultural background, and skin chemistry, unlike other recommendation areas.

In order to provide consumers recommendations, recommendation systems have historically used a variety of methods including content-based and collaborative filtering. These techniques find trends and similarities by analyzing user interactions and item properties, therefore providing pertinent recommendations. But these conventional methods frequently fall short in capturing the subtle preferences that are part of choosing a scent, which results in less than ideal recommendations for consumers.

With the development of sophisticated machine learning methods and the emergence of Large Language Models (LLMs), the field of recommendation systems has experienced tremendous change recently. These advanced models—like BERT and GPT-3—have shown amazing capacity for producing text that is appropriate to the context and for comprehending natural language. LLMs and ML algorithms combined allow recommendation systems to provide more precise and customized recommendations.

In spite of this, there are promising chances to completely transform the perfume suggestion process thanks to recent technological developments, especially the emergence of Large Language Models (LLMs) and machine learning methods. The complexity of perfume recommendation can be addressed and customized fragrance recommendations that speak to each person's particular tastes and traits can be provided by using the potential of LLMs and ML algorithms.

Through the use of these new technological developments, we can completely transform the process of perfume recommendation and offer consumers customized fragrance recommendations that suit their individual tastes and traits. Large-scale data analysis capabilities of LLMs and ML algorithms enable the identification of patterns and correlations that guide the recommendation process, including perfume descriptions, user ratings, and fragrance properties. Furthermore, we may further improve the relevance and accuracy of recommendations, hence improving the user experience and happiness, by customizing these models to certain areas like scent recommendation.

We introduce in this work a five-step method for perfume recommendation that leverages Machine Learning (ML) methods and Large Language Models (LLMs). Unlike conventional techniques, our method extends its applicability to several kinds of datasets depending on utilization, going beyond the constraints of perfume datasets. We present a flexible architecture that can transform recommendation systems across many domains, including but not limited to scent selection, by exploiting the strengths of LLMs and ML. By means of our creative approach, we hope to improve the recommendation experience by providing for the unmatched precision and customization of individual preferences and traits.

II. LITERATURE SURVEY

Perfume recommendation systems are like personal fragrance guides, helping people find scents that match their style and life. Surprisingly though, not much study has been done on this subject. Through their paper "Personalized Quiz-Based Perfume Recommender System Using Social Data," Elena-Ruxandra Lutan and Costin Badica attempted to close this gap. They devised a novel concept: they took into account the users' lifestyle and surrounding events in addition to what they liked. In this way, their technology seeks to provide customized scent recommendations that truly suit the preferences and circumstances of each individual. [1]

Large Language Models (LLMs) and other technological developments provide us new approaches to improving recommendation systems. Using LLMs, Arkadeep Acharya, Brijraj Singh, and Naoyuki Onoe attempted to get into greater depth in describing novels and movies. They looked at if LLM descriptions were as excellent as human-written ones using data from sites like Goodreads and MovieLens. LLMs come in useful since they can automatically produce better descriptions by understanding and writing in complex ways. [2]

Another great product is called Alpaca-Lora, an elegant name for an open-source LLM. It is highly proficient in text authoring, hence item descriptions for recommendation systems are created with it. This increases user interest in and utility of the recommendations. Though LLMs are excellent, they are not without flaws, such as fabricating information or misrepresenting the facts. Relax, though; intelligent individuals are trying to resolve these problems.

Junjie Zhang, Ruobing Xie, Yupeng Hou, Wayne Xin Zhao, Leyu Lin, and Ji-Rong Wen came up with a novel concept that modifies the way recommendation systems function. Using ordinary English, they let you tell the system exactly what you want, rather than assuming it. It resembles expressing your search needs to a friend. The advice thus seems more understandable and personal. Their approach, dubbed "InstructRec," facilitates people's ability to express their preferences, which improves suggestions generally. [3],[4]

We still need to look into a few aspects despite all these interesting research:

- Regional & Cultural Preferences: Depending on their origins or cultural affiliation, people may have varied tastes. To make recommendation systems more appropriate, we must investigate how these factors influence the perfumes that individuals enjoy.
- Making Better Use of LLMs Though they are quite useful, LLMs are not being used in scent recommendation systems enough. To offer more precise perfume recommendations, we must figure out how to improve them even more.
- Too Little Perfume Data: There isn't much solid information available on scents and personal preferences. Recommendation systems that truly grasp what people enjoy are difficult to create without reliable data. Better recommendations will entail gathering more information about scents and how individuals feel about them.

We can create more individualized, culturally sensitive, and simply better perfume suggestion algorithms if we concentrate on these areas. [5]

III. PROPOSED METHODOLOGY

Perfume Recommendation has always been an untapped field of research. Our methodology includes 5 major steps which helps to recommend perfumes - the identity and self esteem of a person with utmost correctness. Five steps involved in the methodology is as follows :

1. Data Preprocessing and Feature Extraction
2. Data Generation
3. Fine Tuning LLM
4. Bag of ML Models
5. Ranking and Recommendation

1. Data Preprocessing and Feature Extraction : There were numerous characteristics in the dataset that may not have been useful for generalizing the data. Instead of generalizing, the entire recommender will be overfit. Therefore, we had to begin by cleaning and preprocessing the dataset in order to extract crucial characteristics that would be useful when recommending various fragrances to users. As one examines the dataset, it becomes indisputable that there are five overarching categories that are universally comprehensible and utilized by all users:
fragrance which is the notes or type of smell they will be referring to,
age - This will denote the age group of our user whether youth (under 25) or adult (over 25)
gender - This denotes the gender of our user
time - time of the day (day or night) for which they are looking to use the perfume.
season - season when the user wants the perfume to be used.

Although age, gender, time, and season may appear to be categorical in nature, this is not the case. Because a user may be attracted to a fragrance that is suitable for both genders and can be worn throughout the year, including both summer and winter, day and night. In general, this significantly expands the range of these parameters.

Additionally, there are columns that are highly interconnected, which adds to the intricacy of data cleansing and feature extraction, both of which are crucial for the development of the perfume recommendation system. We began by extracting the essential attributes, including the ID, title, accords, date, and rating. The remaining columns must now be extracted or cleansed in order to obtain the data in the correct format.

It was necessary to extract the top, middle, and base notes from the corresponding columns (i.e., 0-64). As a result, these notes were extracted and stored in distinct columns as a string separated by commas. Following the notes extraction process, we required categorical user-specific data such as age, gender, time, and season, which were also highly interconnected. Therefore, the categorical variables were normalized as follows:

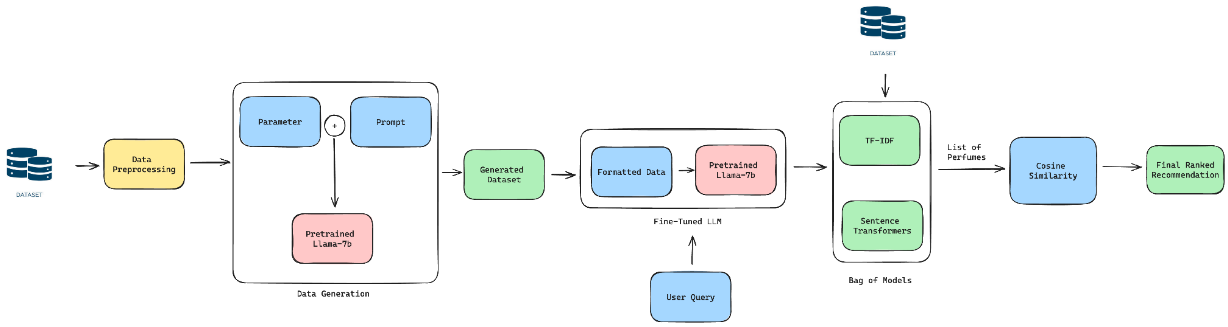

Fig 1 : Proposed Methodology

(Summer/(Summer+Spring+Autumn+Winter)). This attributed a value to each of the categorical values. But in our case, a perfume that spans two seasons may be appropriate. As a result, we established a threshold of 50% and obtained the season's name, which possessed a value exceeding 50%. The process was replicated for additional categorical columns, resulting in refined parameters and cleansed data.

Columns remaining in our dataset following preprocessing:
id : id of the perfume
title: title of the perfume
accords: accords associated with the perfume
date: date of formation of the perfume
rating: rating of the perfume
top_notes: top notes of the perfume
middle_notes: middle notes of the perfume
base_notes: base notes of the perfume
love: This is a normalized integer value which denotes how many people love the perfume.
like: This is a normalized integer value which denotes how many people like the perfume.
dislike: This is a normalized integer value which denotes how many people dislike the perfume.
season: This denotes the season in which perfume is mostly used
time: This denotes the time of the day in which perfume is mostly used
age: Age group which mostly uses it
gender: Gender which mostly uses it.

2. Data Generation : Navigating the domain of perfumery in the present-day consumer environment can be a formidable challenge, especially for those who lack familiarity with the complex terminology and subtleties of fragrance. Consequently, when inquiring about perfume recommendations, numerous users frequently abstain from employing precise terminology associated with perfumes. They have a preference for results that are concise and require little effort and research. Furthermore, in light of the swift progressions in technology, consumers are increasingly demanding immediate and effortless accessibility to information and services. In light of these dynamic inclinations, our objective is to optimize the perfume recommendation procedure through the direct extraction of contextual information from the textual input supplied by users.

In order to accomplish this goal, we initiated the process of training a Large Language Model (LLM) to discern pertinent contextual information from paragraphs supplied by users. Our attention was directed towards the Llama-2-7b-chat-hf model, which is a member of the LLAMA 2 family of LLMs. This iteration of LLAMA-2 is specifically engineered for chat-based applications due to its vast parameter count of 7 billion. As a result, it excels at natural language processing tasks, including those that entail extensive agent-to-agent interactions like conversational AI systems or chatbots.

To produce text paragraphs concerning personal fragrance preferences, we implemented a methodical methodology. Initially, we meticulously designed prompts specifically designed to assist the LLM in comprehending the intended structure and substance of the produced text. The prompts were carefully constructed to encompass a wide range of perfume preferences, such as age, gender, season, time of use, and fragrance notes. Following that, we utilized the pre-existing llama2-7b-chat-hf model and inputted customized prompts and randomly sampled values from our dataset for each parameter.

By means of this iterative procedure, the LLM was instructed to produce text paragraphs that effectively communicate user preferences regarding perfume in a logical and enlightening fashion. The generated paragraphs succinctly summarize essential information, including the season preferred, gender, age group, time of usage, and specific fragrance notes.

3. Fine-Tuning LLM : During the third stage of our methodology, our attention is directed towards refining our Large Language Model (LLM) in order to efficiently extract features from textual data. In order to accomplish this objective, the Llama-3-8B-Instruct model was selected on account of its notable sensitivity to context and effectiveness in information retrieval. For fine-tuning, Llama-3-8B-Instruct adheres to a particular format, which is denoted as

<|begin_of_text|><|start_header_id|>system<|end_header_id|>
{ System Message }
<|eot_id|><|start_header_id|>user<|end_header_id|>
{Users Query }
<|eot_id|><|start_header_id|>assistant<|end_header_id|>
{ Assistant Answer }
<|eot_id|><|start_header_id|>user<|end_header_id|>
{User Query 2 }
<|eot_id|><|start_header_id|>assistant<|end_header_id|>

Special Tokens used have following meaning :
<|begin_of_text|>: This is equivalent to the BOS token
<|eot_id|>: This signifies the end of the message in a turn.
<|start_header_id|>{role}<|end_header_id|>: These tokens enclose the role for a particular message. The possible roles can be: system, user, assistant.
<|end_of_text|>: This is equivalent to the EOS token. On generating this token, Llama 3 will cease to generate more tokens.
A prompt can optionally contain a single system message, or multiple alternating user and assistant messages, but always ends with the last user message followed by the assistant header.

The process of fine-tuning is predicated on the specification of the system prompt, user prompt, and expected model response within this format.

Following the reformatting of our dataset to conform to this structure, we proceed with the improvement of the LLM's fine-tuning procedure. A noteworthy enhancement is the integration of Quantized LoRA (QLoRA), a method that quantizes model parameters in order to minimize memory and processing demands without compromising precision. QLoRA utilizes a base model loading with a precision of 4 bits and provides the ability to specify the compute data type and quantization type. This guarantees effective utilization of resources, which is especially advantageous for computers that have restricted memory or processing capabilities.

In addition, the quantization process is optimized by QLoRA via the BitsAndBytes (BNB) configuration, which permits precise management of quantization type, compute data type, and nested quantization, among other parameters. In order to enhance the efficiency of the fine-tuning process and implement automated procedures such as batch processing, optimization, and recording, we utilize the Hugging Face Trainer framework. In order to mitigate concerns such as overfitting and vanishing gradients, this framework offers an assortment of training parameters to be utilized. These parameters comprise batch size, learning rate, optimization technique, and gradient accumulation stages.

After undergoing fine-tuning with the designated dataset and parameters, the model is retained for subsequent implementation. By doing so, it guarantees that the improved model can be employed for a multitude of subsequent Natural Language Processing (NLP) endeavors, thus optimizing its practicality and efficacy.

4. Bag of ML Models : After parameter extraction from the text, our recommendation system utilizes two discrete models—Sentence Transformers and TF-IDF—each of which is an essential element in our recommendation pipeline.

Vectors have been pre-generated for each perfume in our dataset using the TF-IDF model. Upon the reception of parameters, the text undergoes preprocessing through the elimination of stop words and the consolidation of pertinent information into a unified string. The TF-IDF model then calculates the cosine similarity between the input vector and each vector in our dataset that represents a perfume, using the processed string as input. The TF-IDF model determines the ten perfume recommendations that possess the greatest similarity scores via this computation.

Simultaneously, we utilize Sentence Transformers to encode descriptions of perfumes into embeddings, which are then entered into our database. The embeddings of the parameters are computed upon input, and a k-nearest neighbors algorithm is implemented to identify neighboring embeddings in the dataset that are also present in the input. This procedure enables the identification of the ten fragrances whose embeddings most closely resemble the user's preferences.

In addition to the two methods mentioned, we also incorporated ChromaDB into our system. ChromaDB leverages sentence transformers to identify and retrieve documents with similar content, making it highly effective for both storing and retrieving vector-based documents. This capability not only streamlined our document management but also enabled us to generate a curated list of perfume recommendations.

Although all models provide insightful information, they each have distinct drawbacks. In order to tackle this issue, we combine the suggestions produced by all algorithms and apply a bespoke cosine similarity function to them in order to determine their ranking. This process ensures that the capabilities of each model are utilized to generate more reliable and precise aroma suggestions for users.

By incorporating these models into our recommendation pipeline, we augment the system's overall performance and efficacy. By leveraging the embedding-based technique of Sentence Transformers in conjunction with the vector-based approach of TF-IDF, we can effectively capture a broader spectrum of user preferences and provide more individualized recommendations.

5. Ranking and Recommendation : A list of fragrances and input parameters are processed during this phase of our recommendation system in order to determine their ranking for user recommendation. By incorporating comprehensive data, including top notes and base notes, which offer significant insights into the unique qualities of each perfume, we optimize the ranking methodology in order to augment the pertinence of our suggestions. More specifically, we designate a higher rank to those perfumes whose accords list correspond to the user's requested fragrance. Furthermore, fragrances featuring top notes are given precedence over those featuring middle notes, and vice versa, with middle notes being given precedence over base notes. Through the incorporation of these variables into our ranking algorithm, we guarantee that the ultimate suggestions propelled to the user are more individualized and precisely synchronized with their inclinations. By implementing this strategic methodology, we not only elevate the standard of our recommendations but also augment the overall user experience through the provision of more pertinent and individualized suggestions.

Overall Flow : The recommendation system receives a text paragraph from the user in which they articulate their personal preferences regarding perfumes. The parameters of this text are extracted by our fine-tuned Large Language Model (LLM), which discerns significant attributes including gender preference, preferred fragrance notes, and seasonality, among others. These parameters are subsequently transmitted to the system's Bag of Models component.

Each sub-model in The Bag of Models possesses expertise in a distinct facet of perfume recommendation. The aforementioned models comprise TF-IDF, Sentence Transformers, ChromaDb Sentence Transformer, among others, all of which provide distinct perspectives for the recommendation procedure. On the basis of the extracted parameters, these models generate a catalog of perfumes that most closely correspond to the user's preferences.

Ultimately, the ranking algorithm is tasked with arranging the fragrances in a particular sequence in accordance with their pertinence to the user's specifications. This classification takes into account a multitude of factors, including the extent to which the perfume's attributes correspond with the user's personal preferences, in addition to any further criteria that the user specifies or that the system predefines.

By adhering to this efficient workflow, our recommendation system guarantees that users are provided with pertinent and personalized perfume suggestions that are customized to their specific inclinations and preferences. By implementing this strategy, not only is user contentment increased, but it also proves that sophisticated natural language processing methods can be utilized to provide superior recommendations.

IV. DATASET

Fragrantica is a well-known website devoted to colognes, perfumes, and scents. Since its founding in 2007, it has grown into a complete hub for fragrance lovers, providing a plethora of knowledge, opinions, and conversations on a diverse selection of scents. Fragrantica is an essential resource for fragrance enthusiasts looking to explore, discover, and engage with the community in the ever-evolving world of fragrances thanks to its extensive resources and captivating features. [6]

The dataset has in total 51213 unique perfumes which has been scraped from this site. All useful information like notes, accords, title, user preferences etc has been scraped and loaded into the dataset for a more refined recommendation.

Central to the dataset are a multitude of attributes meticulously gathered to paint a comprehensive picture of each perfume's profile. Starting with essential details such as the URL of the source, the perfume's title, and its release date, the dataset facilitates easy referencing and contextualization of the data within the Fragrantica ecosystem. Moreover, it captures vital aspects of a perfume's essence, including its rating, accord composition, longevity, and sillage, all contributing to a holistic understanding of its character.

The rating attribute encapsulates the collective evaluation of a perfume's quality, distilled from the wealth of user reviews and ratings found on Fragrantica. Accords, meanwhile, offer a glimpse into the fragrance's olfactory makeup, providing insights into its predominant scent profiles. Attributes detailing longevity and sillage further enrich the dataset, offering nuanced assessments of a perfume's endurance and projection across different user segments.

User preferences and interactions constitute another crucial facet of the dataset. Demographic insights such as age and gender-based preferences, alongside seasonal recommendations and suitability for various occasions and weather conditions, are meticulously documented. Metrics gauging user sentiment, including the number of users who loved, liked, or disliked a perfume, as well as current, past, and future user intentions, provide valuable engagement and sentiment analysis.

Moreover, the dataset delves into the intricate world of fragrance notes, cataloging the various scent components into top, middle, and base notes. This granular information enriches our understanding of each perfume's olfactory profile, potentially influencing user preferences and purchase decisions.

All in all the final features extracted from the website includes various types of data related to perfume. For eg : id, title, rating, date, accords, top_notes, middle_notes, base_notes. Apart from these it has various votes by people in various columns such as love_female_under25, like_female_under25, love_female_25under, love_male_25under, love_female_25older, love_male_25older, like_female_25under, like_male_25under, like_female_25older, like_male_25older, dislike_female_25under, dislike_male_25under, dislike_female_25older, dislike_male_25older, female_25under_winter etc same for summer, winter, autumn, spring, day and night. They are having integer value - number of people who voted them for that category.

V. EXPERIMENTAL SETUP

We have utilized the Fragrantica dataset to enhance the functionality of our recommendation system, allowing it to propose scents based on user input. Initially, we diligently processed the dataset by performing thorough cleaning and extracting pertinent features, adhering to our established methodology.

After preparing our dataset, we proceeded to train our Language Model (LLM) by utilizing a meticulously designed prompt. We chose the Llama-7b-chat-hf model, which is well-known for its astounding 7 billion parameters. To do this, we utilized the HuggingFace platform to load the pretrained model that was given by meta. We employed a text-generation pipeline that utilized the identical tokenizer as the initial training of the Language Model (LLM). To refine the model, we customized the prompt by incorporating sample input, intended output, and diverse parameters for generating paragraphs. Because the LLM's GPU utilization is demanding, we were only able to process 1000 scents at once, which resulted in a limited dataset. Upon generating the results, we isolated the text paragraphs and put them in a distinct dataframe for subsequent analysis.

Subsequently, we employed this dataset to train our second LLM model, selecting the Llama-3-8B-Instruct due to its proficiency in capturing contextual nuances in text. Prior to fine-tuning, it was necessary for us to restructure our dataset to align with the comprehension of the LLM. Llama-3-8B-Instruct processes training prompts in a particular structure, so we modified our 1000 generated words to adhere to this framework. Next, we readied the LLM for fine-tuning by importing the pretrained model and tokenizer using HuggingFace. We meticulously adjusted the QLoRA parameters, such as the LoRA attention dimension, alpha parameter, and dropout probability. Similarly, we configure the BitsandBytes parameters to enable the loading of a base model with 4-bit precision, using specific quantization settings. Subsequently, we established TrainingArguments, explicitly setting parameters such as the output directory, number of epochs, batch size, optimizer, and learning rate schedule. In order to validate the trained model, we loaded it and assessed its performance by supplying input prompts and verifying that it produced precise values.

Next, we analyzed the perfume dataset, which consisted of several parameters such as accords, notes, gender, age, time, and season. The NLTK library was used to apply text preprocessing techniques, including tokenization, lowercase conversion, and punctuation removal. Subsequently, we employed the TF-IDF technique to compute similarity scores between the user's input and the descriptions of perfumes, thereby facilitating individualized suggestions that align with individual preferences. In addition, we employed the Sentence Transformers library to produce embeddings for both fragrance notes and user preferences. We utilized a k-nearest neighbors (k-NN) approach with cosine similarity to train a model that can provide high-performing perfume suggestions. In addition to these we also used ChromaDB to find similar documents in the corpora of perfumes. Ultimately, we combined the results from all of these models.

Finally to provide a correct recommendation we took the merged list and compared it with the query giving it custom scores based on notes being similar to top note or middle note. This enhanced the ranking of the final recommendation produced finally allowing us to provide recommendations based on the text provided by the user.

VI. RESULTS

Experiments were conducted using the above dataset and methodology and results were analyzed.

- Successful Methodology: Our methodology effectively generated recommendations based on user input, providing balanced and comprehensive results. The model accurately extracted parameters and produced recommendations that aligned with user preferences.
- Validation of Text Generation: The text generation process was validated, demonstrating a balanced use of synonyms and thorough coverage of parameters. This ensured that the

generated text accurately reflected the input and provided relevant recommendations.

- Parameter Extraction by LLM Model: The LLM model successfully extracted parameters from shorter texts containing fewer parameters, demonstrating its capability in handling simpler input scenarios.
- Challenges with Complex Inputs: However, the system faced challenges when presented with texts containing numerous parameters. In such cases, the model struggled to extract all relevant parameters, leading to incomplete recommendations.
- Limitations of TF-IDF: TF-IDF, while useful for calculating similarity between two perfumes, proved insufficient for comprehensive perfume recommendations. Its reliance on similarity scores alone limited its effectiveness in capturing diverse user preferences.

Overall, while our methodology showed promise in generating recommendations, there were limitations in handling complex input scenarios and relying solely on TF-IDF for recommendation generation.

VII. LIMITATIONS

This project faced several limitations that impacted its development and effectiveness.

- Lack of research: The field of personalized perfume recommendation systems is still relatively underexplored, with limited research available. To enhance the system's capabilities, further research is needed to uncover new insights and approaches.
- Lack of dataset: A significant challenge was the scarcity of suitable datasets for training and testing the recommendation model. Existing datasets are often limited in size and may not adequately represent diverse preferences. Efforts to scrape data from perfume websites, particularly those catering to Western markets, are necessary to address this limitation.
- Lack of resources: The computational resources required for training and fine-tuning Language Model (LLM) models, particularly GPUs, were a major constraint. The high GPU demand of LLM models made it challenging to conduct extensive fine-tuning, limiting the model's optimization and performance. Acquiring adequate resources, such as GPUs, is essential to fully leverage the capabilities of these models and improve recommendation accuracy.

VIII. FUTURE WORKS

- Scalability Testing: We plan to execute the entire model on a high-end GPU to accommodate the entire dataset comprising 50,000 perfumes. This will allow us to scale up our operations and handle larger volumes of data efficiently.
- Enhanced Parameter Extraction: We intend to further refine our parameter extraction process to ensure comprehensive coverage of all relevant attributes in perfume descriptions. This includes exploring advanced natural language processing techniques to extract subtle nuances and characteristics from text data.
- Dataset Expansion: Additionally, we will focus on expanding our dataset by scraping data from diverse sources, including non-Indic perfume websites tailored primarily for Western audiences. This will enrich our dataset with a wider variety of perfumes, enhancing the diversity and representativeness of our recommendation system.
- User Feedback Integration: Incorporating user feedback into our recommendation system will be a key area of focus. By collecting and analyzing user preferences and satisfaction with recommended perfumes, we can iteratively improve the accuracy and relevance of our recommendations over time.

IX. CONCLUSIONS

To summarize, our research efforts in creating a scent recommendation system have produced encouraging outcomes, while also revealing several aspects that require additional improvement and investigation. We have proven the effectiveness of using complex natural language processing models, such as Llama-3 and Llama-2, to provide individualized scent suggestions based on user input. This was achieved through careful approach and experimentation.

The significance of thorough data preparation, feature extraction, and model fine-tuning in attaining precise and contextually appropriate recommendations is emphasized by our findings. By tackling challenges such as a lack of data, limited resources, and the ability to scale models, we have established a strong basis for future progress in this field.

In the future, we will prioritize improving the scalability, originality, and diversity of our recommendation system by incorporating cutting-edge language models and growing our dataset. In addition, we will give priority to incorporating user feedback, conducting ongoing evaluations, and deploying our system in real-world settings to guarantee that it is practical and effective in improving the experience of shopping for perfumes.

Our research is a significant advancement in creating intelligent suggestion systems that are customized to the specific preferences of individual users in the field of fragrance picking. Our goal is to use advanced technology and methods to give consumers individualized choices and enhance their satisfaction and delight while selecting perfumes.

X. REFERENCES

[1] Luţan, Elena-Ruxandra, and Costin Bădică. "Personalized Quiz-Based Perfume Recommender System Using Social Data." In International Conference on Computational Collective Intelligence, pp. 30-43. Cham: Springer Nature Switzerland, 2023.

[2] Acharya, Arkadeep, Brijraj Singh, and Naoyuki Onoe. "Llm based generation of item-description for recommendation system." In Proceedings of the 17th ACM Conference on Recommender Systems, pp. 1204-1207. 2023.

[3] Zhang, Junjie, Ruobing Xie, Yupeng Hou, Wayne Xin Zhao, Leyu Lin, and Ji-Rong Wen. "Recommendation as instruction following: A large language model empowered recommendation approach." *arXiv preprint arXiv:2305.07001* (2023).

[4] Liu, Dairui, Boming Yang, Honghui Du, Derek Greene, Aonghus Lawlor, Ruihai Dong, and Irene Li. "RecPrompt: A Prompt Tuning Framework for News Recommendation Using Large Language Models." arXiv preprint arXiv:2312.10463 (2023).

[5] Jingya Cao, Siuyan Chen, Xin Shen, Shixuan Wan. “The Perfume Recommendation System” December 2019

[6] Datasource : https://www.fragrantica.com/