**Alex Nguyen – LLM Project**

**Proposal – Cologne Trend Using LLMs**

**Evaluating LLM Effectiveness in Consumer Fragrance Insights**

**Abstract**

This project will evaluate how effectively large language models (LLMs) can extract, categorize, and analyze data related to men's cologne trends, specifically identifying top-selling fragrances each year, their scent profiles, and the evolution of public sentiment. The project will use web-sourced product descriptions, customer reviews, and sales rankings, along with historical fragrance databases. I will build a baseline evaluation pipeline using structured prompts and parsing techniques to extract scent categories (e.g. woody, citrus, spicy), temporal sentiment changes, and popularity trends. Evaluation metrics will include accuracy of scent categorization, trend consistency with known rankings, and sentiment alignment with verified reviews.

**Introduction**

Cologne is a highly subjective product, with consumer interest driven by scent profile, marketing, celebrity endorsement, and cultural trends. While datasets on fragrances exist, they are often fragmented across e-commerce sites, blogs, and databases. This project will explore whether LLMs can consolidate and interpret this scattered information — identifying top colognes by year, categorizing them by scent families, and summarizing how public reception has changed over time. This provides a real-world test of an LLM’s ability to structure unstructured product and consumer data.

**Motivation**

Consumer behavior analysis is a growing use case for LLMs, particularly in industries like fashion, beauty, and fragrance, where subjective interpretation plays a large role. Men's cologne is both culturally and commercially significant, with global sales in the billions. LLMs may help reveal hidden insights by analyzing historical trends, scent descriptors, and sentiment trajectories. This project aims to evaluate the reliability and utility of LLMs in this lifestyle-focused application, contributing to both consumer research and digital marketing insights.

**Problem Formulation**

The central question is: **Can a large language model accurately identify top-selling colognes over time, classify them by scent category?** To answer this, I will extract data using an LLM from publicly available sources such as Reddit.

Key tasks include:

* Identifying the top 5 men's colognes for a given year (e.g. 2014–2024)
* Classifying each fragrance by scent family (e.g., woody, aromatic, spicy)
* Extracting and evaluating review sentiment over time
* Comparing LLM outputs to known benchmarks and datasets

The dataset will be semi-structured and drawn from multiple web-based sources. Sentiment will be assessed by combining LLM summaries with keyword-based analysis and time-stamped user reviews.

**Workplan**

1. **Baseline Prompting & Extraction**  
   Use structured prompts to have the LLM extract top-selling colognes by year and their scent profiles.
2. **Data Collection & Preprocessing**
   * Gather sales rankings and user review data
   * Use web scraping tools if needed for review aggregation
   * Normalize scent descriptors (e.g., standardizing “citrus”, “woody-aromatic”, etc.)
3. **Scent Categorization Validation**
   * Compare LLM-extracted categories to expert classification databases (e.g., Fragrantica)
   * Track accuracy and consistency of LLM-generated scent families
4. **Sentiment Tracking Over Time**
   * Analyze reviews for recurring themes and descriptors
   * Use the LLM to summarize year-by-year changes in public reception
5. **Evaluation & Metrics**
   * **Coverage**: How many top colognes are accurately retrieved?
   * **Categorization Accuracy**: Does the LLM match known scent profiles?
   * **Sentiment Consistency**: Are model summaries aligned with verified user reviews?
   * **Trend Accuracy**: Do trends match actual commercial performance?
6. **Final Analysis & Reporting**
   * Visualize trendlines in fragrance popularity and sentiment
   * Discuss LLM strengths and limitations in lifestyle-based analysis
   * Recommend future applications (e.g., personalized scent recommendations)