Decoding Amazon Reviews: Innovative Insights with Agentic Language Models

Data
My data is from Datafini available here on Kaggle.
It lists over 34,000 consumer reviews for Amazon products like the Kindle, Fire TV Stick, and more provided by Datafiniti's Product Database. The dataset includes each product's basic product information, rating, review text, and more.

Business Problem Description and Importance

In today's digital age, businesses and organizations are constantly bombarded with vast amounts of text data from various sources such as customer reviews, social media comments, and tweets. This data, if analyzed properly, can provide invaluable insights into customer sentiments, emerging trends, and key areas requiring attention. However, the sheer volume and diversity of this data make manual analysis impractical and ineffective. Businesses need an automated, intelligent solution that can process, analyze, and visualize this information in real-time to make data-driven decisions that enhance customer satisfaction, improve products, and optimize marketing strategies.

Using Data Analytics to Solve the Problem

To address this business problem, we propose developing a comprehensive data analytics solution that employs multiple specialized agentic language models (LLMs). Each agent will be trained to perform a specific task: topic modeling, sentiment analysis, text classification, and clustering. By leveraging advanced natural language processing (NLP) techniques, these agents will automatically analyze large datasets of customer reviews and social media comments to extract meaningful insights.

- **Topic Modeling Agent:** This agent will identify and extract key topics from the text data, helping businesses understand the main themes and issues discussed by their customers.
- **Sentiment Analysis Agent:** This agent will determine the sentiment expressed in the text, enabling businesses to gauge customer satisfaction and identify areas of concern.
- **Classification Agent:** This agent will categorize the text data into predefined classes, such as product categories or types of customer feedback, allowing for more organized and targeted analysis.
- **Clustering Agent:** This agent will group similar text entries together, uncovering patterns and relationships within the data that might not be immediately apparent.
- **Visualization Agent:** The results from each agent will be integrate and visualized in a possibly interactive, automatically updating dashboard or graphs. This dashboard will provide real-time insights through various visualizations such as bar charts, word clouds, sentiment trend lines, and cluster scatter plots.

By consolidating and presenting the data in an easily interpretable format, the dashboard will enable businesses to make informed decisions quickly and efficiently, ultimately driving better outcomes and fostering a deeper understanding of their customer base.

Steps taken to solve the problem

We experimented with several approaches to develop an automated data analytics solution for analyzing large datasets of customer reviews and social media comments. Initially, we attempted to have a single agent complete all tasks using LangChain integrated with OpenAI GPT-3.5. This approach centralized all functionalities but faced challenges in handling the diverse requirements of each task. To enhance performance, we then provided the agent with tools (Python classes) and specific instructions, including category labels for bucketing reviews and techniques for clustering and topic ideas.

Next, we explored a more specialized approach by employing individual agents for each task—topic modeling, sentiment analysis, text classification, and clustering—while a visualization agent integrated the results. We first implemented these agents with Python classes, ensuring each was finely tuned for its specific function. Following this, we used a multi-agent LangChain model, assigning one agent per task **(Fig 1)**. This setup, utilizing LangChain with GPT-3.5 Turbo through the OpenAI API, aimed to leverage the strengths of specialization while maintaining seamless integration through LangChain's framework. Each of these approaches provided unique insights into optimizing the solution for real-time, comprehensive data analysis and visualization.Ultimately, we chose the LangChain multi-agent framework because it provided better insights about the reviews. Below are the findings from that approach.

Findings

Sentiment Analysis:

The sentiment distribution pie chart **(Fig 2)** shows:88.7% Positive sentiment.This overwhelmingly positive sentiment suggests that customers are generally very satisfied with the product or service. However, the 7.67% negative sentiment shouldn't be ignored, as it represents areas for potential improvement.

The Average Sentiment Score per Topic chart (**Fig 3**) shows that all topics have positive sentiment scores, with some variation. This reinforces the overall positive sentiment while highlighting subtle differences between topics that could be further investigated.

The word clouds (Image 4) provide visual representations of frequently used words in positive, negative, and neutral reviews some common words in Positive reviews (**Fig 4)** : Tablet, use, great, love, kindle, Amazon, good, books, read, easy, Negative (**Fig 5**): Tablet, use, kindle, app, work, time, book, bought, disappointed, Neutral (**Fig 6**): Bought, tablet, use, kindle, gift, iPad, Amazon, books, hand, read. It seems that most reviews are about tablets and people have different opinions about them.

The heatmap of sentiment scores by topic and category shows a couple of key insights (**Fig 7**). Overall sentiment is generally positive across topics and categories.The "Suggestion" category shows consistently high sentiment scores, indicating customers are providing constructive feedback."Inquiry" and "Feedback" categories also show positive sentiment, suggesting customers are engaged and mostly satisfied."Complaint" category has lower but still positive scores, implying even critical feedback isn't overly negative.

Topic Modeling

The topic distribution bar chart (**Fig 8**) shows five main topics, each containing key words related to the product or service. Most reviews appear to be about a tablet, likely an Amazon Kindle Fire, used for reading and apps.Customers frequently mention buying the product as a gift, suggesting it's popular for gift-giving.Ease of use and reading capabilities are commonly discussed, indicating these are important features to users. The variety of topics suggests tablets serve multiple purposes (e.g., reading, apps, gift) which can be emphasized in marketing strategies.

The word frequency graph **(Fig 9)** shows that "Tablet" is by far the most frequently used word, confirming the product focus. Positive descriptors like "great", "good", "love", and "easy" are very common. Usage-related terms ("use", "apps", "screen", "games", "books") are prevalent, highlighting key features."Kindle" and "Amazon" appear frequently, indicating brand recognition. Words like "bought", "price", and "product" suggest discussion of purchasing decisions.

Classification

The review classifications bar chart (**Fig 10**) categorizes reviews into Complaint (highest number), Inquiry, Suggestion, Feedback (lowest number). The high number of complaints indicates areas needing improvement.The significant number of inquiries suggests customers may need more information or support. Amazon should pay attention to suggestions as they can provide valuable ideas for product improvements or new features. The lower number of general feedback entries might indicate a need to encourage more diverse types of customer feedback.

Clustering

The plot showing the clusters **(Fig 11**)shows that there is a cluster of reviews around the 0-1 rating range with varying helpfulness votes.Most reviews seem to fall in the positive range (3-5 stars) based on the clustering.Helpful votes don't necessarily correlate directly with high ratings, as some lower-rated reviews also received helpful votes. However, as seen on the figure, the clusters are not completely distinct from each other. This may be because a sample of 300 reviews were used due to the rate limit of the Open AI API.

Business Recommendations

Several key business recommendations are essential for Amazon based on the comprehensive analysis of the graphs and insights provided.Firstly, Amazon should capitalize on the overwhelmingly positive sentiment (88.7%) in its marketing campaigns by highlighting key positive aspects like ease of use, reading capabilities, and versatility of the Kindle tablets. This positive sentiment reflects strong customer satisfaction, which can be leveraged to attract new customers. Additionally, the analysis reveals that the product is frequently mentioned as a gift, suggesting a targeted marketing strategy around gift-giving. Amazon could develop special gift bundles or promotional offers during holiday seasons to cater to this market.

The analysis also highlights that features such as reading experience, app functionality, and user interface are highly appreciated. Therefore, Amazon should focus on maintaining and improving these features, and consider developing more family-friendly features, as indicated by the frequent mention of "daughter" in one of the topics. While the sentiment is largely positive, there are still complaints and negative sentiment (7.67%) that need to be addressed. Amazon should investigate the sources of these complaints and implement a systematic approach to address common issues, potentially through software updates or customer education.

Given the high number of inquiries identified in the analysis, enhancing product documentation, FAQs, and customer support resources could reduce these inquiries and improve overall customer satisfaction. The popularity of the Kindle Fire as both an e-reader and tablet suggests an opportunity to expand the ecosystem of compatible apps and content, increasing user engagement and loyalty.

Implementing a structured process to collect and utilize customer feedback can lead to product improvements and innovations that align closely with customer needs. The identified topics and clusters can be used to create segmented marketing campaigns for different user segments, such as avid readers, app users, and gift-givers. The mention of iPad in neutral reviews indicates a need to highlight Kindle's unique advantages over competitors in marketing materials and product descriptions.

To gain a more balanced view of customer experiences and needs, Amazon should implement strategies to encourage more varied types of feedback, not just complaints and inquiries. Addressing any recurring issues with device functionality or durability is essential to maintain the high positive sentiment observed in the analysis. Lastly, the popularity of Kindle tablets can be leveraged to promote other Amazon services like Prime, emphasizing the integrated ecosystem of products and services.

By implementing these recommendations, which are directly tied to the insights from the data analysis, Amazon can further enhance customer satisfaction, address pain points, and potentially increase its market share in the tablet and e-reader market. The focus should be on maintaining strengths while systematically addressing areas of improvement identified through this analysis.

Future Work

For future work, there are several avenues to explore to enhance the effectiveness and efficiency of the agentic language models (LLMs). One potential direction is to have the agents write the code to perform the tasks instead of executing the tasks directly. This approach, although initially seeming like it could be done by chatting with a regular chatbot like ChatGPT, might yield better results and offer new insights into the agents' capabilities. Additionally, finding ways to fine-tune the agents more precisely could be beneficial. While we allowed the agents significant freedom and included a field for them to output their thinking process, ensuring that the tools provided are of the highest quality might further improve performance. Another idea is to conduct a quick exploratory data analysis (EDA) before giving instructions to the agents. This preliminary step could provide a better understanding of the data, helping to tailor the agents' tasks more effectively. Although this was not done in the current project to see the pure capabilities of agentic LLMs, it may prove useful in future studies. Furthermore, exploring different configurations of multi-agent frameworks that have higher rate limits (making it possible to work on larger datasets) and experimenting with additional datasets could provide deeper insights and enhance the robustness of the solutions. It may also help make clusters more distinct by providing more data.

For the analysis of Amazon reviews, future work could involve deeper dives into specific product categories or customer segments to uncover more granular insights. Analyzing seasonal trends or the impact of marketing campaigns on customer sentiment could also provide valuable information for strategic decision-making. Additionally, integrating data from other sources, such as social media or competitor reviews, could offer a more comprehensive view of customer perceptions and market positioning. By pursuing these directions, the potential of agentic LLMs in automated data analytics can be more fully realized, and the understanding of customer feedback for Amazon can be significantly enhanced.

Appendix

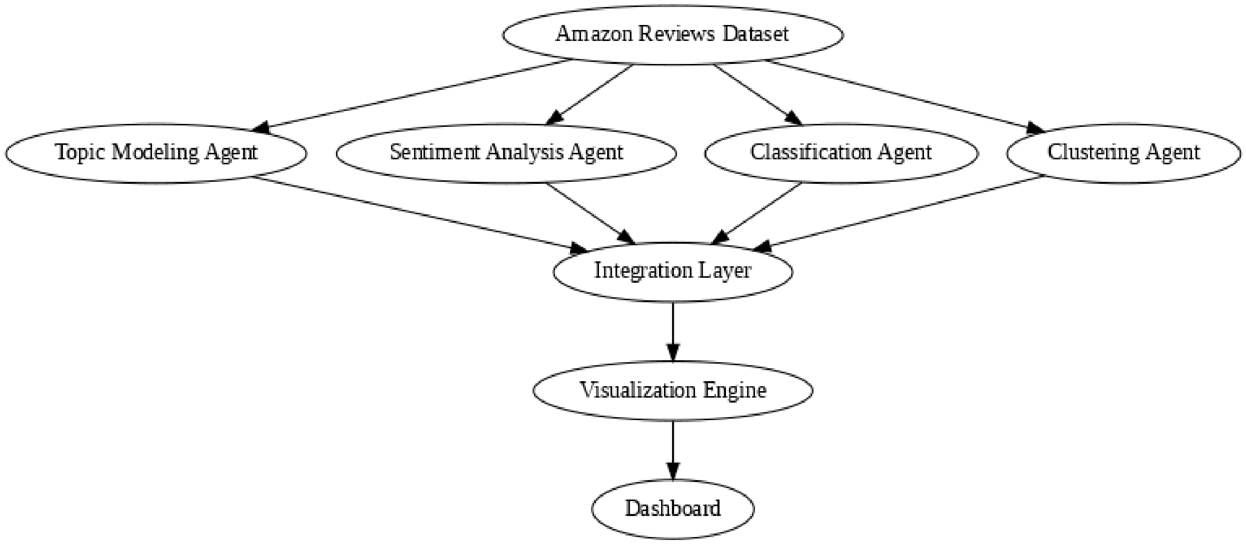

Figure 1: Multi Agent Network Structure

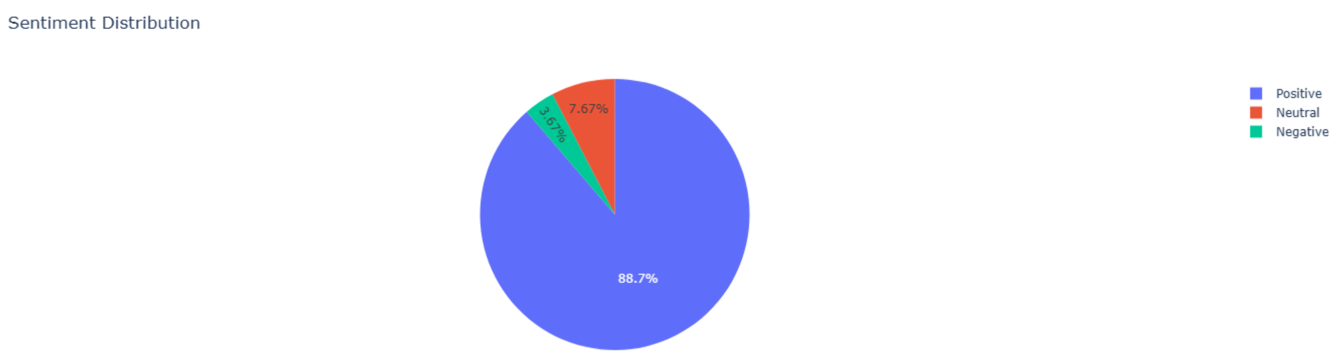

Figure 2: Sentiment Distribution

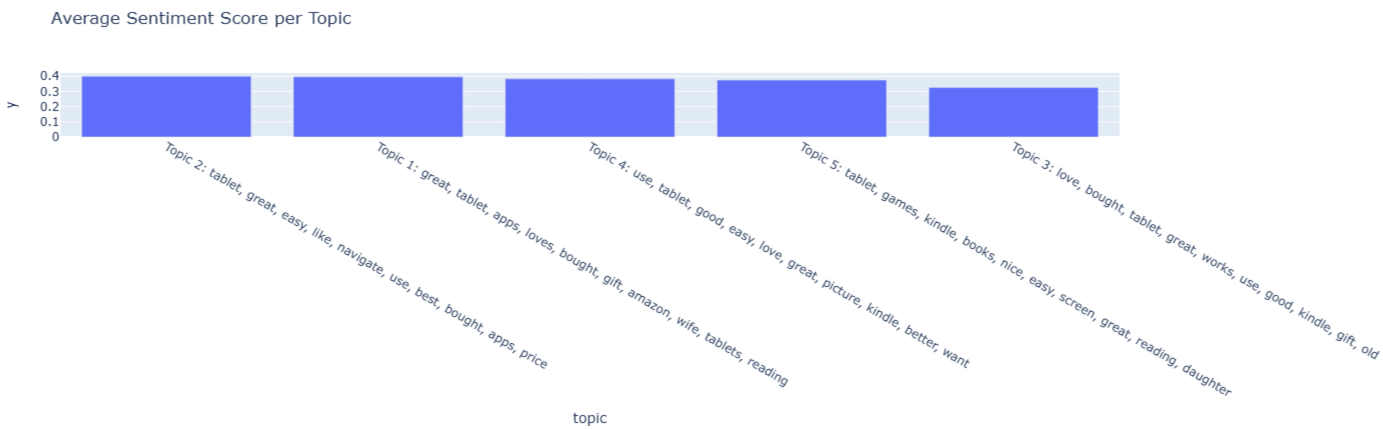

Figure 3: Average Sentiment Score per Topic

Figure 4: Positive Reviews Wordcloud

Figure 5: Negative Reviews Wordcloud

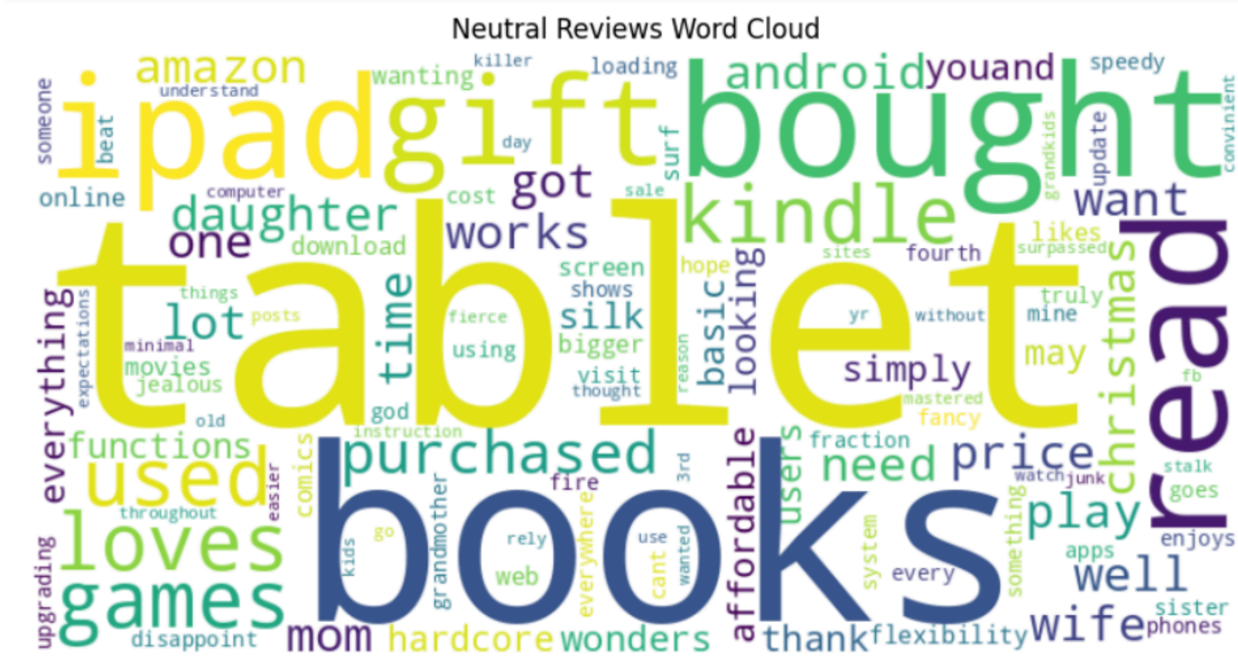

Figure 6: Neutral Reviews Wordcloud

Figure 7: Heatmap of Sentiment Scores by Topic and Category

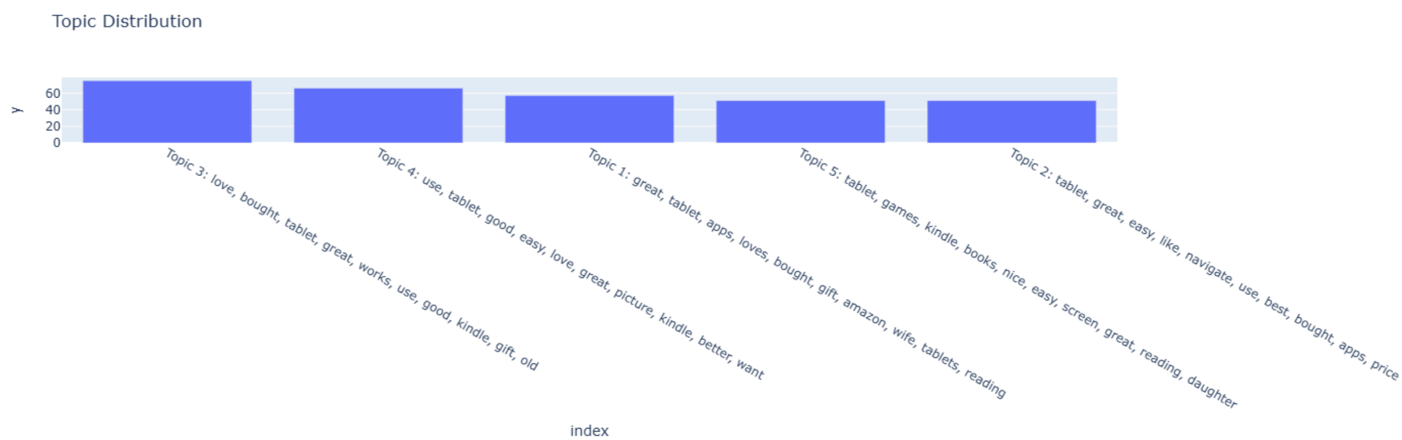

Figure 8: Topic Distribution

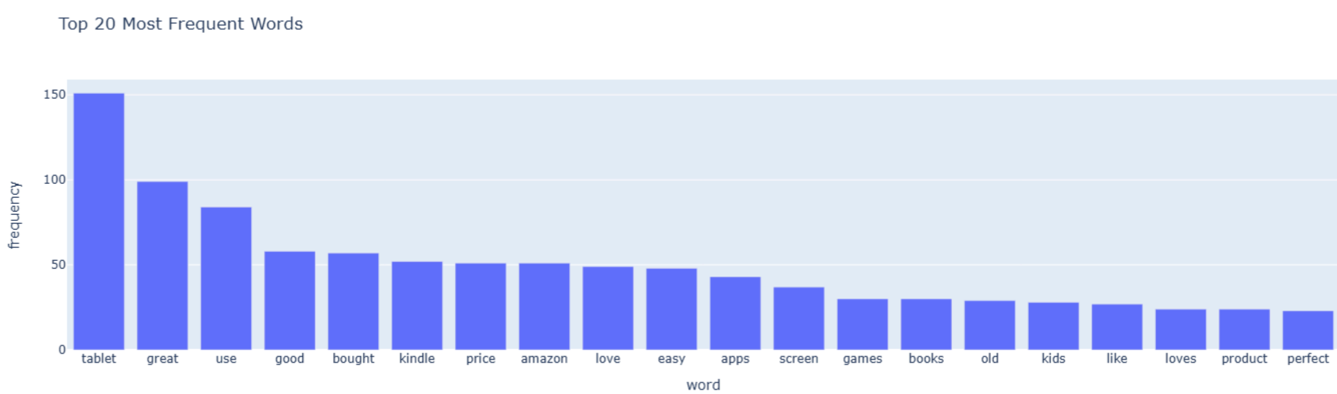

Figure 9: Word Frequency Plot

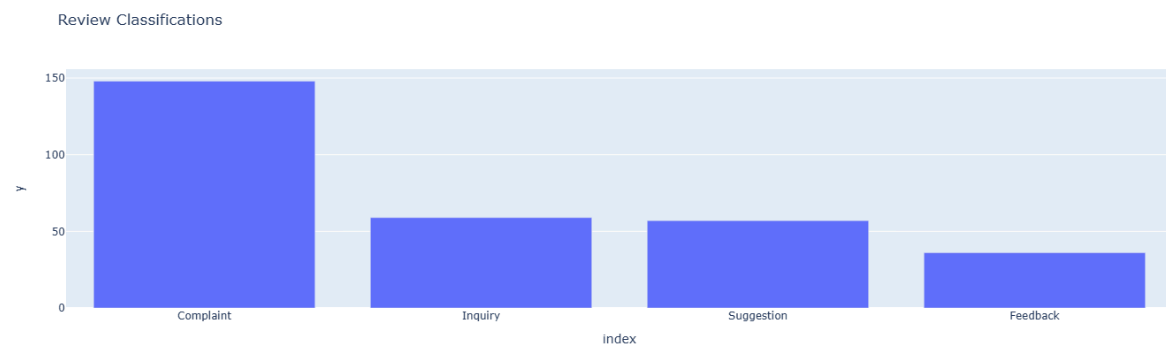

Figure 10: Reviews Classification

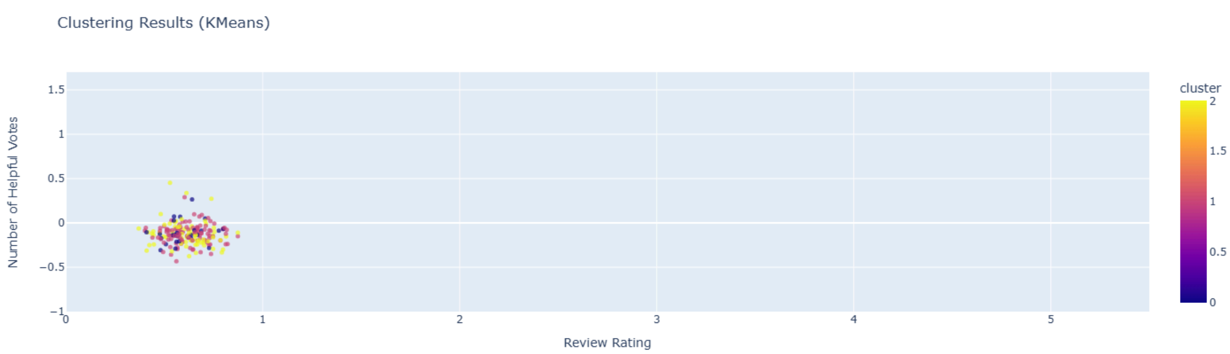

Figure 11: Kmeans Clusters

References

[How to Build the Ultimate AI Automation with Multi-Agent Collaboration (langchain.dev)]

[gpt-researcher/multi_agents at master · assafelovic/gpt-researcher · GitHub]