iPhone 14 Reviews Analysis: Uncovering User Sentiment and Emerging Themes

Executive summary

In this project, we analyze customer reviews on the iPhone 14 to uncover user sentiments and identify emerging patterns. Using a dataset from Kaggle, we combined LDA topic modeling and VADER sentiment analysis to explore major topics including brand, camera performance, battery health, and delivery services. The project aims to extract actionable insights that can support improvements in Apple's product development, marketing strategy, and customer support. By evaluating different areas and different time periods' user feedback, we found out some key advantages from the quality of the phone camera, and some issues about the battery and delivery delays-related complaints. Although certain limitations remain, including biased reviews and missing demographic information, our analysis offers constructive guidance to Apple and retailers, which can help them design information, enhance product function, and solve the issues that are commonly complained about by users.

Project objectives

Our project objective is to uncover customer insights from iPhone 14 user reviews, which is helpful on product development, marketing strategy, and customer service improvement. Through applying text mining techniques, including LDA topic modeling and VADER sentiment analysis, we aim to ensure the main theme from the user's review, and evaluate overall sentiment trend. Our goal is to help Apple and mobile retailers better understand what customers' demand and what they really care about, highlight the features that are being praised, detect recurring issues, and optimize strategy based on the sentiment region and time variations. Through the analysis, we are trying to provide actionable suggestions to enhance customer overall experience, reduce churn, attract new customers, and provide guidance for future product innovations.

Data description

The dataset used for the project is from Kaggle, which includes customer reviews from public review platforms. The data contains multiple geographic regions and covers reviews posted after the iPhone 14 release in late 2022. Each review entry includes a title, full text review, rating, customer name, review date, and location information, which provide qualitative feedback and structured data for data mining and modeling. The key columns are Review, a short headline summarizing the review, and Rating, a numeric score indicating satisfaction, where "1" means "poor" and "5" means "excellent". Those two parts are the parts we mainly focused on and used for sentiment analysis and topic modeling.

There are several limitations within the dataset that could affect the analysis outcomes. First, there is a significant over-representation of highly positive and highly negative reviews, which may introduce bias and reduce the representativeness of the overall customer sentiment. The distribution of ratings in the dataset is shown in chart 1. We found that most rating scores are between 4.25-5. No customer gave a rating below 3, which shows obvious selection bias in the data and ratings are right skewed. We thought two reasons could explain the bias. One is that collectors of the data delete some negative comments before posting it. They wanted the iPhone 14 to look more attractive and persuaded more customers to buy it. Another reason could be customers who are satisfied with iPhones have higher probability to purchase iPhones in the future and leave positive comments. Customers who were unhappy with iPhones wouldn't choose them after trying one time, so we couldn't collect negative comments from them. Because of the bias, we predicted that in our sentiment analysis, positive reviews are much more than negative reviews.

Second, some reviews lack substantial content, consisting mainly of emojis or very brief statements, which limits the depth of insights that can be extracted through text mining. Furthermore, important demographic information, such as the reviewer's gender and age, is missing, making it difficult to segment customer feedback across different population groups. These factors could be potential limitations to impact the reliability, generalizability, and richness of the analysis results.

Methodology

We first cleaned the dataset by removing non-English reviews and filtering out low-content entries such as emoji-only or blank reviews. This preprocessing ensured the input quality for our models, though it introduced the limitation of potentially excluding meaningful but brief feedback. For the text vectorization step, we used bag-of-word method for topic modeling and TF-IDF method for sentiment analysis. We also tried the n-grams method to do text vectorization, but it wasn't as efficient as the other two methods and produced some meaningless short word combinations. Considering the convenience of implementation of bag-of-word and the emphasis of words related to sentiment, we finally chose the two methods.

We used topic modeling and sentiment analysis to analyze customer reviews on iPhone 14 and filter the information with business insights. We tried different numbers of topics(3,4 and 5) and we wanted words in these topics to be useful and these topics were distinguished with each other. We finally decided four topics (K=4) based on clarity and interpretability, labeled them manually. While LDA allowed us to extract recurring themes efficiently, the manual labeling process introduced subjectivity, and some topics may have overlapped due to the lack of detail in some reviews.

Then, we applied the Valence Aware Dictionary and Sentiment Reasoner(VADER) method to do sentiment analysis because this method informs both about the polarity(positive/negative) and intensity of sentiment. We computed compound sentiment scores and categorized them into positive, neutral, and negative using established thresholds. This method was easy to implement and interpretable, though it was limited in detecting sarcasm and complex tone. We then combined the topic and sentiment results by mapping sentiment scores to each LDA-identified topic. This integration allowed us to detect sentiment patterns across different product features. For instance, we found that camera-related reviews were mostly positive, while battery and delivery topics had more negative sentiment. While this dual approach provided actionable insights, it treated topics and sentiment independently, which may overlook their potential interaction.

Overall, this methodology enabled us to uncover valuable customer insights despite some data limitations. Our combined use of LDA and VADER provided a structured way to surface both what customers are talking about and how they feel about it, directly supporting our project's objective of helping Apple and retailers respond to customer needs more effectively.

Results & Discussion

We applied LDA topic modeling (K=4) to reviews, extracted the top 10 keywords for each topic, and labeled the topics as camera performance, battery performance, iPhone brand, and delivery service. Table 2 reports the results of our topic modeling analysis and words in each topic. We then computed compound sentiment scores using VADER and classified reviews as positive (score $\geq$ 0.05), negative (score $\leq$ −0.05), or neutral otherwise. The results are shown in table 3. We found that 72.4% of camera performance reviews were positive, 20.1% neutral, and 7.5% negative; for battery performance, 78.6% positive, 13.7% neutral, and 7.9% negative; for iPhone brand, 85.4% positive, 9.2% neutral, and 5.4% negative; and for delivery service, 87.2% positive, 7.3% neutral, and 5.5% negative. We also observed that neutral reviews tended to be factual and descriptive, while negative feedback focused primarily on battery life and delivery delays, highlighting clear areas for improvement. Based on these findings, we recommend emphasizing delivery reliability and brand trust in marketing, prioritizing camera stabilization and battery-life enhancements in product development, and in future research integrating metadata such as star ratings and timestamps to analyze temporal and regional sentiment trends for more targeted operational and product strategies.

Conclusion

In our analysis, we discovered that iPhone 14 users had an overall positive sentiment toward the product. However, there were some potential growth areas that were revealed from our topic modeling and sentiment analysis. One of these areas is battery life, we recommend to Apple to continue to augment

not only the battery-life but also the performance of the battery over the lifetime of the device. We also recommend updating the delivery systems in place to increase reliability. This is a very important factor for brand trust, which had one of the highest negative review rates in our results. Our results also show that the camera performance had the lowest positive review percentage.We recommend that Apple upgrade the stability of the camera to increase positive reviews, which could help their brand trust among camera enthusiasts.

We considered that three potential shortcomings exist in our data and analysis. Considering little negative reviews exist in our data, it's probably that the dataset contains selection bias, which makes the results of sentiment analysis less useful for business instruction. To reduce the risk of bias, we should figure out the source of the dataset and find more customer reviews from multiple resources, such as the record from the online retailers and offline stores. But, during the process, we need to make sure these data resources are credible and unbiased.

Another one is lacking detail in a lot of reviews, with some of them just consisting of emojis or short statements. Finally, there was no demographic information about the reviewer, which could give us more of an ability to target our marketing efforts. One way to deal with the lack of demographics is to send a follow up survey to the reviewers. This way we could try to see if there is any demographic trend associated with the reviews. The way we could deal with emojis in the reviews is to assign each emoji to a sentiment, and if the emoji is unknown we could put it as neutral. We would weigh the emojis, along with the short statements, less in our model because there could be some ambiguity on what people are trying to convey with the emoji. For example, the emoji could be used in a sarcastic manner which would throw off our sentiment analysis. Taking in all of this information, our analysis offers Apple actionable insights to improve the areas that were highlighted in these reviews. These actions will strengthen Apple's product satisfaction, as well as their brand perception.

Appendix
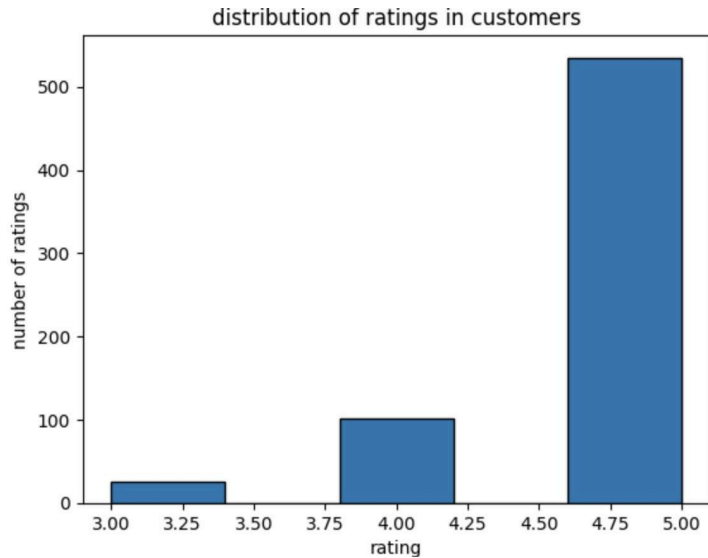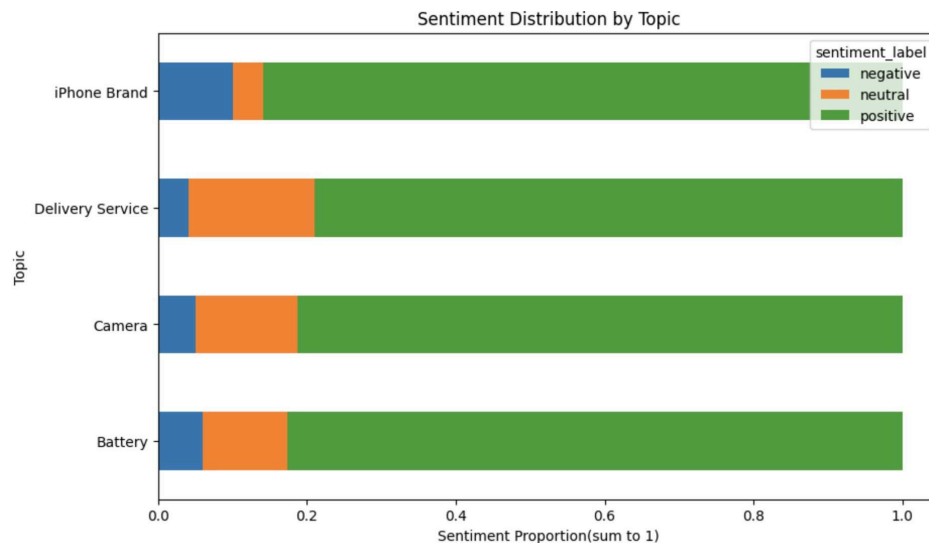
Table 1: distribution of ratings in customers

Table 2: Top 10 frequent words in each topic(4 topics)

```
Topic 0, Delivery Service: product flipkart delivery happy nice mobile thank
price exchange work
Topic 1, Camera: good camera phone awesome quality performance love nice
battery amazing
Topic 2, Battery: use iphone battery issue phone day video low heat user
Topic 3, iPhone Brand: iphone great best apple excellent display camera first
android device
```

Table 3: Sentiment Distribution by Topic

References

IPhone 14 Customer Reviews Dataset. (2024, March 2). Kaggle.
https://www.kaggle.com/datasets/shahriarkabir/iphone-14-customer-reviews-dataset-ratings