Optimizing Product Recommendations on Amazon with Advanced NLP and ML Techniques

Yumin Wang
Department of Computer Science
Virginia Tech
Falls Church, VA, USA
yuminw22@vt.edu

Jiayi Zhang
Department of Computer Science
Virginia Tech
Falls Church, VA, USA
jiayizhang@vt.edu

Weiting Li
Department of Computer Science
Virginia Tech
Falls Church, VA, USA
weitil6@vt.edu

Abstract

Our project aims to recommend Amazon products to users, based on Amazon product co-purchasing network. The goal of our project is to address the problem of designing and developing a recommender system that is accurate and efficient. We conducted network analysis and used Vadar to perform sentiment analysis on product reviews to understand users' perceptions of products better. What is more, we implemented topic modeling to gain a more detailed understanding of user preferences. Our work is detailed and it covers many areas. We also analyze the strengths and weaknesses of our methods and suggest future work.

Keywords: Datasets, Network Analysis, Sentiment Analysis, Topic Modeling, Product Recommendation, Amazon, Social Media Analytic

ACM Reference Format:
Yumin Wang, Jiayi Zhang, and Weiting Li. 2023. Optimizing Product Recommendations on Amazon with Advanced NLP and ML Techniques. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Amazon, founded by Jeff Bezos in 1994, has revolutionized the way people shop. This is because Amazon not only offers a convenient and accessible online shopping platform with competitive prices, but also uses innovative technology to provide users with a user-friendly interface, including features like customer reviews and personalized product recommendations. Amazon's co-purchasing network and ground-truth communities show the interdependent relationship between Amazon products and users who have purchased them together. This dataset holds value for researchers, e-commerce platforms, and participants of online marketplaces, including both sellers and buyers. Our aim in conducting a research study on Amazon's co-purchasing network is to assist e-commerce platforms in enhancing their services and product offerings, thereby tackling social issues like polarization and the dissemination of false information.

Studying social media is crucial as it has a significant impact on society, both positively and negatively. With the integration of features like buying and selling on social media platforms such as Instagram and TikTok, the importance of studying social media has grown. While social media can serve as an effective means of networking, sharing, and publishing, it can also have detrimental effects such as spreading hate and misinformation that can harm people's mental health and the social environment. Therefore, improving the social media environment is necessary to ensure its positive effects are maximized while minimizing its negative impacts.

Subsequently, we can identify frequent purchasing patterns, related products, and user behavior through the analysis of co-purchasing network. So, the findings of this study will provide valuable insights to Amazon and other e-commerce platforms to enhance their product recommendations, marketing strategies, and overall user experience. The results can be used to improve recommendation algorithms and suggest relevant products to customers, leading to an improved user experience and increased sales. Additionally, this research can be used to develop product bundles and promote related products, resulting in better marketing strategies and higher sales. Furthermore, the study of user behavior and preferences can assist companies in creating better products, improving customer service, and enhancing the overall user experience.

Hence, our research on Amazon's co-purchasing network will be divided into three main areas: network analytics, text analytics, and action analytics. Network analytics will enable us to identify patterns, relationships, and trends between Amazon products and their buyers. Text analytics will help us comprehend the meaning and sentiment of text data using techniques like natural language processing (NLP). Finally,

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06...$15.00
https://doi.org/XXXXXXX.XXXXXXX

action analytics will involve analyzing user actions such as clicks and purchases to gain a deeper understanding of how users interact with products and what factors influence their purchasing decisions.

The overall plan for approaching the problem includes but is not limited to:

1) Pre-process the raw data in order to improve the precision of the outcomes.
2) Conduct network analysis to gain an understanding of the network's structure.
3) Apply sentiment analysis to obtain insights about the general sentiment expressed in the reviews and comments related to the products.
4) Use topic modeling to identify primary topics discussed in Amazon co-purchasing network review and comments.
5) Utilize Machine Learning models to make a practical product recommendation.

2 Related Work

Gutiérrez [4]suggests the creation of a machine learnin model to examine customer evaluations of goods sold on Amazon and to glean insightful data regarding the user's viewpoint. The model seeks to find patterns and trends in the data to provide sales people useful information. The study emphasizes the significance of user feedback analysis for businesses selling goods on online marketplaces and suggests a model that makes use of a number of machine learning approaches to ascertain whether a user gave a product a favorable or negative rating based on their review.

The study by Haque [5] emphasizes the growing significance of customer evaluations in the increasingly digitalized e-commerce world and suggests using a machine learning algorithm to detailize and examine huge Amazon datasets. In order to train the model and obtain adequate accuracy in reviews, the article uses machine learning techniques. This renders it easier for customers to make purchasing choices based on the opinions of others.

Sharma [7]'s work provides an overview of sentiment analysis techniques for large sided data produced by social media. They highlighted the importance of sentiment analysis in decision-making in the study, as well as the difficulties in reading and evaluating the enormous volumes of text data. The study does a performance comparison analysis and gives a overview of sevral sentiment analysis approaches and algorithms. The survey shed lights on information on the potential of sentiment analysis to enhance digital world decision-making processes.

Table 1. Group Distribution

Group	Number
Baby Product	1
Book	393561
CE	4
DVD	19828
Music	103144
Software	5
Sports	1
Toy	8
Video	26131
Video Games	1

3 Approach

We thoroughly analyzed three different datasets including Amazon product co-purchasing network, product reviews of music, and product reviews of videos. We first collected the data, then pre-processed them, conducted basic network analysis and performed topic modeling for category of the data. Next, we did sentiment analysis on both dataset of product reviews. And last step involves using a machine learning model to make product recommendations. Finally, we refined the recommender system based on the results we obtained from our experiment.

3.1 Data Collection

We obtained Amazon product co-purchasing network metadata on snap website and amazon review data on the computer department of USCD. Amazon product co-purchasing network metadata was collected by crawling Amazon.com which contains product metadata and review information about 548,552 different products (as shown in the Table.1 below).

Among them, we will use the data related to music and video. We chose these two data types because they are representative of the products and the percentage of the four main groups is shown in the Figure.1 below.

3.2 Data Pre-processing

In order to parse the Amazon product co-purchasing network metadata, we first read the data in the dataset line by line, since each line represents a different type of information. After reading, we extracted the information corresponding to each product based on the keywords at the beginning of each line. We removed spaces in the lines and discarded spaces in the data, as well as punctuation and stop words. This is simply because punctuation and spaces do not have any effect on the results, but they can cause a lot of trouble in our follow-up work. Finally, we stored the processed data in a dictionary and converted them into a Dataframe. We

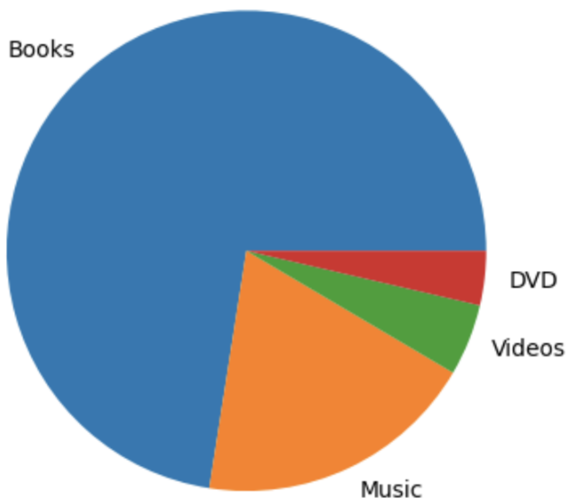

Figure 1. The percentage of the four main groups.

refered to Gounder [3] for guidance.

We also provided sentiment analysis using Amazon review data for music and videos. To maximize the accuracy of our results, we took several steps to parse the data. Since sentiment analysis is applied to each word of the review, we first tokenized each review into words. Then, we removed punctuation, URLs, stopwords, and other symbols from each tokenized word. It is important to note that we removed commonly spoken English words, as well as some music and video related words that are unnecessary for sentiment analysis. Finally, we regrouped the parsed words in each review into sentences for the convenience of the analysis process.

3.3 Network Analysis

Initially, we generate the csv file of the nodelist based on the similar products, ratings, and reviews of the product. After that, we generated edgelist csv files based on similar products. According to the generated nodelist and edgelist files, the corresponding graphs were drawn. The last step was to use the networkx package to calculate the degree centrality, average degree, average density, and maximum connected images for each product.

3.4 Topic Modeling

Topic modeling is used to process natural language techniques for analyzing hidden topics in texts. The most commonly used approach is LDA, a model that will help us find the most frequently occurring combinations of topic words. Topic modeling is performed on the category of the product, which is used to determine the most popular types of products. Each product may belong to many categories, so finding the most popular product category can help us in further product recommendation. We also use Gensim, a library used to process datasets with very large amounts of data. Finally, we used pyLDAvis researched by Chuang et al. [2] to visualize our model.

3.5 Sentiment Analysis

Sentiment Analysis is a critical aspect of our attempt to refine the Amazon recommender system, as it enables us to gain a comprehensive understanding of how different users view each product. By utilizing this technique, we can learn the emotions behind each review of each product, allowing us to determine whether customers like or dislike a product. For this experiment, we utilized the SentimentIntensityAnalyzer provided by VADER. This dictionary-based sentiment intensity analyzer is an NLTK module that can label the sentiment of each text. We can gain insights into the popularity of the product based on whether the majority of users responded positively or negatively to each product, and thereby determine if we should recommend a product.

3.6 Product Recommendation

For our product recommendation system, we decided to go with machine learning based collaborative filtering. Collaborative filtering based on machine learning is a technique used in recommendation systems to make recommendations to consumers based on their shared preferences. In order to discover patterns and similarities between users and objects, it entails training a machine learning algorithm on a dataset of user-item interactions. After researching, famous recommender machine learning algorithms are k-Nearest-Neighbor, and singular value decomposition.

4 Experiment

Using the methods mentioned in Section 3, we implemented a detailed evaluation of the data using Python and finally performed a comprehensive analysis on product recommendation. As the videos and music categories were the most symbolic to our experiment, we only utilized data from these categories for topic modeling, sentiment analysis, and product recommendation studies.

4.1 Basic Analysis

By sorting the data parsed by Amazon product co-purchasing network based on the number of reviews, we got the three most popular music and video (Table.3 and Table.4).

4.2 Network Analysis

We carried out network analysis on the generated network graph. First of all, this graph has 548,589 points and 4798,509

Table 2. Top 3 reviewed music

ASIN	Reviewed Times
B000084T18	993
B000088E6Q	993
B00005RGNI	990

Table 3. Top 3 reviewed video

ASIN	Reviewed Times
0783113145	99
6303514723	99
0767819543	99

Table 4. Network Analysis

Mean Degree	17.4940
Max Degree Centrality	0.6882
Min Degree Centrality	0
Mean Degree Centrality	3.188914760748274e-05
Std Degree Centrality	0.0028
Density	3.1889147607482734e-05

edges, we calculate that the average degree length is 17.494. We figured out that it is not a connected graph, thus, it does not have diameter and shortest path. Next, the degree centrality of the graph is found, and the maximum, minimum, mean and standard deviation of the degree centrality are calculated. Lastly, this graph has the largest connected graph with 379130 nodes and 4798509 edges, which we can find that the density of the connected graph is 6.676689089843261e-05 is larger than the density of network graph. We refered to vishalsha [8] for guidance.

4.3 Topic Modeling

We first retrieved the products with group as music and video, and generated Dataframe respectively. Foremost, we divided the category of all the products into three topics and modeled the topics, the results of which are shown in the following Table. After that, we modeled and visualized the topics for music products and video products for each of them. The most popular categories of music are often related to rock, pop, classical, and so on. Similarly, top video categories may be related to drama or comedy, and so forth. By using topic modeling, we came up with the most relevant words to the datasets (Figure.n and Figure. n) and the most relevant words for each topic (Table.n and Table. n).

4.4 Sentiment Analysis

We conducted sentiment analysis on customer reviews of Amazon's video and music products using NLTK's VADER module, as described in Section 3. We applied sentiment analysis to 1,584,082 music reviews. Due to the RAM limit of our laptops, we performed sentiment analysis on 16,010 video

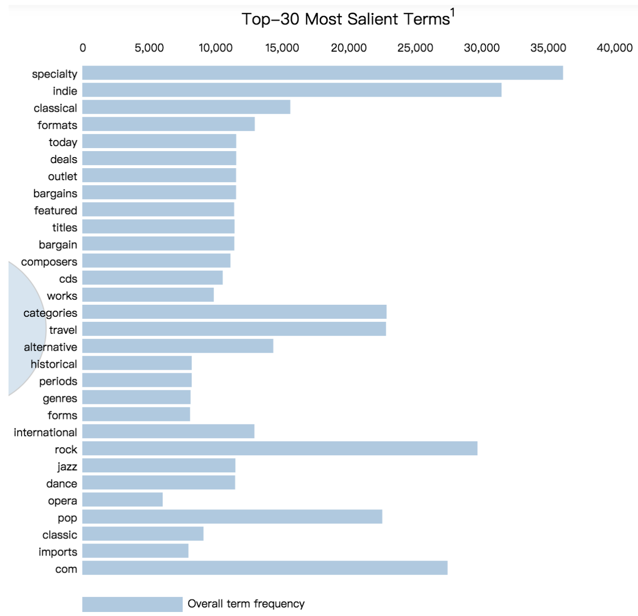

Figure 2. Top-30 Most Salient Terms in music.

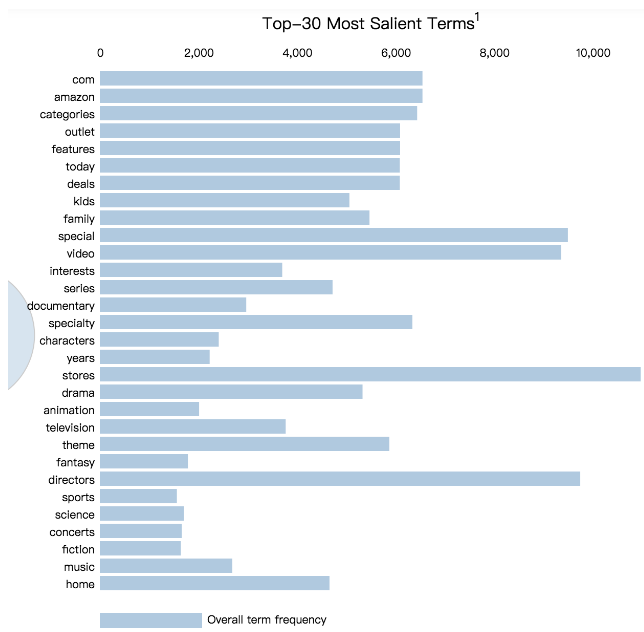

Figure 3. Top-30 Most Salient Terms in video.

Table 5. Topic modeling for all products

Topic 1	specialty, indie, rock, pop, alternative
Topic 2	categories, travel, formats, amazon,com
Topic 3	classical, featured, composers, works, historical

Table 6. Topic modeling for Music and Video

	Music	Video
Topic 1	specialty indie rock pop alternative	directors theme drama comedy specials
Topic 2	categories travel formats amazon com	video amazon com categories outlet
Topic 3	classical featured composers works historical	kids family special interests series

reviews out of a total of 160k video reviews.

Sentiment Analysis on Amazon's Music Products

By applying sentiment analysis to the parsed reviews of Amazon's music products, we have learned that the majority of customers who have purchased these products have expressed positive opinions. Out of a total of 1,584,082 reviews, 1,161,501 or 73% showed positive emotions, while only 2% of the reviews expressed negative emotions.

We also calculated the average score for each sentiment type based on the compound score of each type. The average positive sentiment score for Amazon's music products is 0.79, which is consistent with the findings of the Kernel Density Estimation (KDE) graph in Figure 4. In addition, we found that the average sentiment score for all of Amazon's music products is 0.61, which is more than 0.05. This indicates that the average score is equivalent to the ratio of the sentiment of all reviews, where the majority exhibited positive opinions.

Table 7. Reviews on Amazon's Music Products

	Positive	Negative	Neutral
Number of Reviews	1,161,501	29,464	393,117
Ratio	73%	2%	25%
Average Score	0.79	-0.68	0.17

Another important aspect of our sentiment analysis is that we examined the top -10 most frequent words in each sentiment category among all the reviews [1], as shown in figure 5. Our analysis of Amazon's music products revealed that the top 10 most frequent words in positive reviews [1] (as

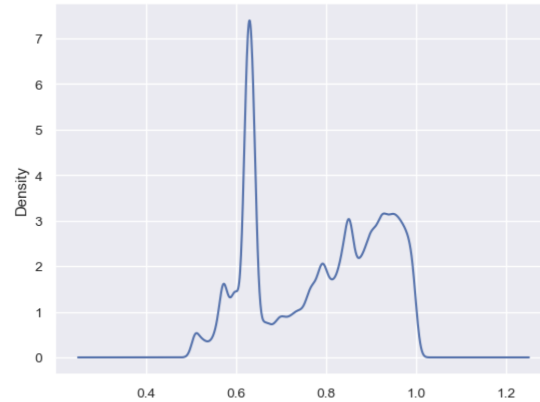

Figure 4. Kernel Density Estimation for Positive Sentiment

displayed in figure 6) share many similarities with the most frequent words in all of the music reviews. Interestingly, the top 4 most frequent words in positive reviews are identical to those in all reviews, which are "great", "love", "like", and "good". This finding is also consistent with the overall trend of the dataset, which consists mostly of positive reviews.

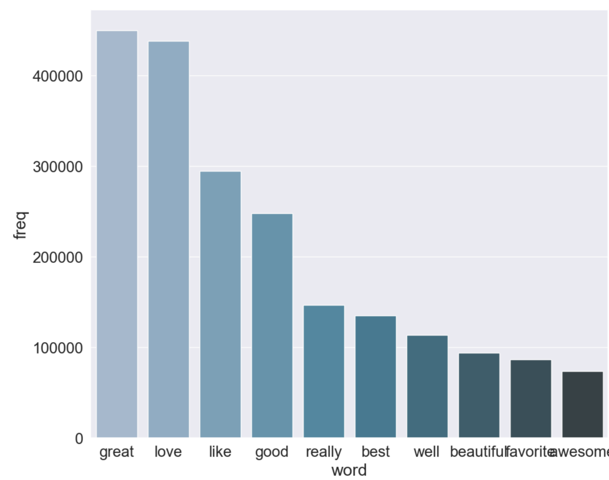

Figure 5. Top-10 Most Frequent Words in All Reviews

Overall, analyzing customer reviews is an important aspect of understanding Amazon's music products' quality and popularity. And we have found that most customers are satisfied with their music product purchases from Amazon.

Sentiment Analysis on Amazon's Video Products

In comparison, we also conducted sentiment analysis on all of the parsed video reviews using NLTK's VADER analyzer. Our analysis revealed that 68% of the customers who purchased Amazon's video products are satisfied with their

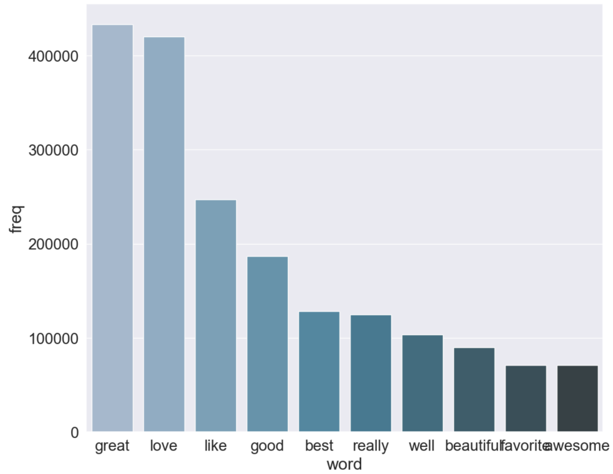

Figure 6. Top-10 Most Frequent Words in Positive Reviews

purchase, while only 7% of customers expressed negative emotions.

The percentage of customers who are satisfied with their video related purchase are slightly lower than those have who have purchased music related products. But the average positive sentiment score for the video products is high, which is 0.83. This finding is also consistent with the Kernel Density Estimation graph for positive video reviews as shown in figure 7. In addition, we have also discovered that the average sentiment score for all of the reviews on video products is 0.54, which matches our results that most customers convey positive opinions.

Table 8. Reviews on Amazon's Video Products

	Positive	Negative	Neutral
Number of Reviews	10,863	1,171	3,976
Ratio	68%	7%	25%
Average Score	0.83	-0.82	0.15

Furthermore, we also analyzed the top-10 most frequent words of each sentiment group and all of the reviews in the Amazon Video Review dataset [1]. The results are similar to the ones in Amazon's music reviews' top-10 most frequent words. The top-10 most frequent words in positive reviews (as shown in figure 9) are also very similar to the ones in the top-10 most frequent words in all reviews (as shown in figure 8). However, interestingly, the top-10 most frequent words in negative reviews (as displayed in figure 10) contain words like "evil" and "horror" [1]. Both of these words can generally convey negative emotions, but not in the case of

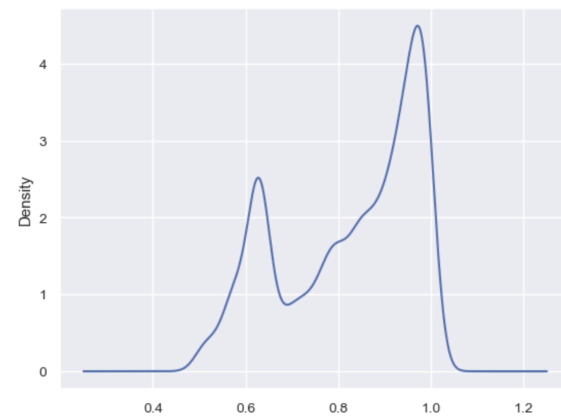

Figure 7. Kernel Density Estimation for Positive Sentiment

videos. Because the term "evil" can be used to describe a character, while the word "horror" can be referring to "horror movies".

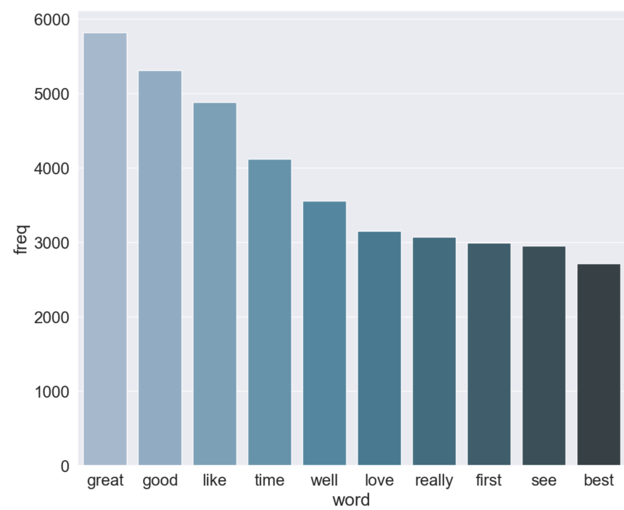

Figure 8. Top-10 Most Frequent Words in All Reviews

4.5 Product Recommendation

We proceed with experimenting both of them on the 'music' dataset and 'movie_tv' dataset based on the rating score. And we used root mean squared error as the metric to determine accuracy. We refered to Meda [6] and Venkata's recommender system for guidance.

Due to the large size of the music dataset and our machine RAM limitation, we cannot apply KNN strategy on the music video dataset. As a result, our model end up with a best score of 0.7 on video dataset shown in table n, and a best score of 0.48 on music dataset shown in table n+1.

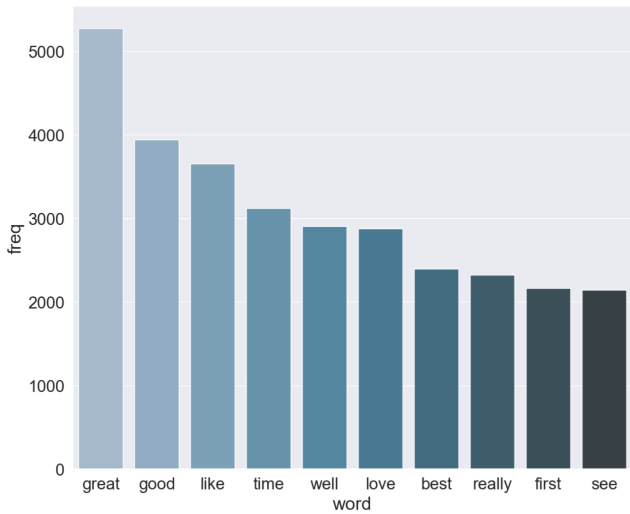

Figure 9. Top-10 Most Frequent Words in Positive Reviews

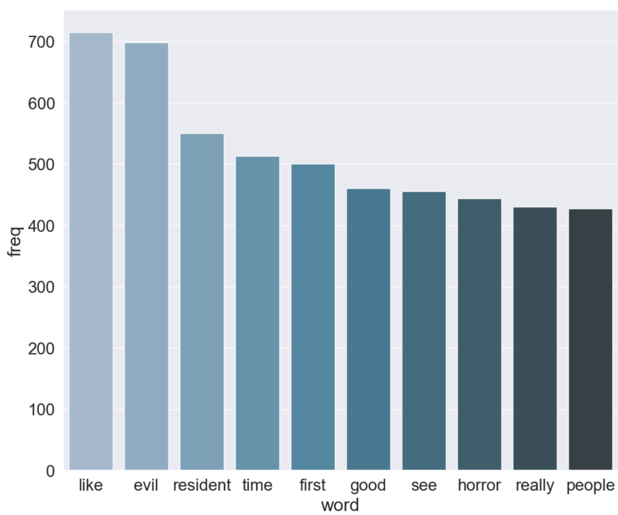

Figure 10. Top-10 Most Frequent Words in Negative Reviews

Table 9. Recommendation performance on video dataset

Model Type	RMSE
KNN k=10	1.133
KNN k=15	1.133
KNN k=20	1.133
SVD n=25	0.995
SVD n=50	0.999
SVD n=75	0.993
HyperTuned SVD n = 95	0.700

5 Limitations

For product recommendation, our system lacks the ability to train the model if the dataset were too large. In our case, it caused memory error due to lack of RAM to use. In the second place, because of the data provided by snap and UCSD, which are all available six years ago, the results we get are not up to date. Additionally, due to hardware problems, we have no way to determine whether the transformed network graph follows a power-law distribution. Moreover, there are no stopwords libraries for commonly used music-related and video-related words that are not very useful for sentiment analysis. So we had difficulties with eliminating all music-related and video-related stopwords during the pre-process stage.

Table 10. Recommendation performance on music dataset

Model Type	RMSE
SVD n=25	0.655
SVD n=50	0.654
SVD n=75	0.654
HyperTuned SVD n = 65	0.480

6 Future Work

For product recommendation algorithm, we should not just limit ourselves to supervised learning, but we should also take a look into deep learning field. Because deep learning as following advantages compared to machine learning:

1. non-linearity: deep learning algorithms like neural network is capable of handling non linear relationships that is not obvious in the dataset.
2. robustness on large dataset: deep learning algorithm is able to perform well on large dataset especially ones related to natural language processing.
3. ability to evolve: deep learning can continually improve its performance by feeding new data into the model and adjusting its parameters to optimize its accuracy.

Therefore, in the future, we can experiment with some of the popular deep learning algorithms like multi-layer neural network, convoluted neural network etc.

7 Conclusion

In conclusion, our approach initially turned the network of Amazon product co-purchasers into a network graph where items were nodes and related products were neighbors. When we performed a simple network analysis on this graph, we discovered that it was unconnected. Then we determined the top 30 terms relating to music and video using LDA and pyLDAvis package.Additionally, we discovered that Topic 1 accounted for the majority of topics—roughly half—in both datasets. These results indicate that Topic 1 is a substantial topic in both the music and video datasets and that the co-purchasing network is not well linked.

Furthermore, based on our sentiment analysis experiment, we found that a majority of customers who purchased music-related or video-related products are satisfied with their purchase. And we also have learned that eliminating prudct-related stopwords can enhance the results of sentiment analysis. According to our study, we have gained insights into which products customers responded to most positively and negatively. This will enable the personalized recommender system to recommend similar products to customers in the future while also avoiding those to which they responded negatively.

Finally, we successfully implemented a high accuracy product recommender system using svd and grid search to find the best parameters.

References

[1] Bert Carremans. 2018. Sentiment Analysis with Text Mining. (2018). https://towardsdatascience.com/sentiment-analysis-with-text-mining-13dd2b33de27

[2] Jason Chuang, Christopher D. Manning, and Jeffrey Heer. 2012. STermite: Visualization Techniques for Assessing Textual Topic Models. (2012).

[3] Susheel Gounder. 2023. *Co products recommender system using Amazon metadata Public*. https://github.com/susheelg1197/Co-products-recommender-system-using-Amazon-metadata

[4] Gina V. Acosta Gutiérrez. 2020. A Comparative Study of NLP and Machine Learning Techniques for Sentiment Analysis and Topic Modeling on Amazon Reviews. (April 2020).

[5] Tanjim Ul Haque. 2018. Sentiment analysis on large scale Amazon product reviews. (May 2018).

[6] Venkata Pinnimty Sai Meda. 2019. *Amazon Product Recommendation*. Department of Computer Science, Blacksburg, VA. https://github.com/ramapinnimty/CS5664-Amazon-Product-Recommendation/blob/main/Report.pdf

[7] Dipti Sharma. 2019. Sentiment Analysis Techniques for Social Media Data: A Review. (Nov. 2019).

[8] vishalsha. 2021. *Amazon Products Analysis*. https://github.com/vishalsha/amazon-products-analysis