Personalized Magazine Recommendation System Based on User Reviews and Social Network Dynamics

Richard R Lee
MEng, Computer Science
Virginia Tech, NVC
Fairfax, VA
richardva@vt.edu

I-An Cho
MEng, Computer Science
Virginia Tech, NVC
Fairfax, VA
cianvt@vt.edu

Pei-Hsuan Sung
MEng, Computer Science
Virginia Tech, NVC
Fairfax, VA
sungpeihsuan@vt.edu

***Abstract*—This study investigates the magazine subscription sector in the digital age. Amazon Review Dataset of 2023 from the 'Magazine Subscriptions' category is utilized. As Consumer preferences shift towards digital media and subscription models, traditional magazine publishers need to adapt to stay relevant and competitive. By employing data analytics and social media insights, this research provides a detailed analysis of consumers' reviews. This research employs sentiment analysis, topic modeling, and network analysis on reviews to provide a detailed analysis of consumer feedback. These methods aim to uncover key drivers of customer satisfaction and preferences within the magazine subscription domain. Additionally, the study incorporates a collaborative-approach recommendation system utilizing the reviews and interaction network to predict user preferences for magazines, further aiding publishers in understanding and enhancing their marketing strategies. Besides the obvious benefits of incorporating topic modeling and sentiment analysis into a recommendation system, our experiments with different weight configurations revealed that network factors can influence the results. This finding suggests that network factors are valuable and should be considered in the recommendation system. The findings offer practical strategies for magazine publishers to navigate digital disruptions and engage customers, contributing valuable insights to academic and industry practices.**

I. Introduction

The traditional magazine sector is transforming this age as industries across the board adapt to keep pace with evolving consumer behaviors and preferences. The rise of digital platforms and subscription-based models has put magazine publishers and sellers in a challenging position. They must reevaluate their marketing strategies and offerings to stay competitive.

This project examines magazine subscriptions, using the Amazon Review Dataset (2023) 'Magazine Subscriptions' section as a source of consumer feedback and insights. We use advanced analytics techniques to understand the factors that drive consumer satisfaction and preferences in the magazine subscription industry.

The Amazon Review Dataset is a comprehensive repository of consumer reviews that provides a unique opportunity to understand consumer sentiments, opinions, and experiences related to magazine subscriptions. We use social media analytics to extract insights that inform decision-making for magazine publishers and sellers.

This study aims to empower industry stakeholders with valuable insights into consumer preferences and behaviors. This will enable them to tailor their offerings, improve customer satisfaction, and optimize their marketing strategies to engage with their target audience in the digital era. This analysis aims to help the magazine industry navigate the challenges and opportunities presented by the digital landscape by identifying the main drivers of consumer satisfaction and preferences.

II. Background / Related Work

In the wave of digital transformation, the subscription model of the magazine publishing industry is facing unprecedented challenges and opportunities. This study's related work involves using big data analytics technologies, such as sentiment analysis and topic modeling, to deeply understand consumer feedback and preferences regarding magazine subscriptions. These technologies not only help identify positive or negative trends in consumer emotions but also reveal the main themes discussed, thereby guiding the development and optimization of recommendation systems.

Research shows that consumers' emotional responses to magazine content and thematic interests can significantly influence their subscription decisions. For instance, Abdellatif et al. (2023) demonstrated a framework for multilingual sentiment analysis, which helps to better understand consumer emotions across cultural contexts. [1] Additionally, topic modeling techniques have been widely applied to personalized recommendation systems that recommend content based on users' past behaviors and preferences.

In the magazine publishing field, the use of sentiment analysis and topic modeling is particularly important for developing personalized recommendation systems. These systems use weighted ranking to optimize the accuracy of recommendations based on the sentiment scores and topic relevance of reviews. Through intelligent matching techniques, such recommendation systems evaluate users' historical interactions and preferences to provide more relevant magazine choices.

However, despite these technologies providing new insights, existing methods often fail to fully consider the interaction between emotions and topics, or to adjust flexibly in a rapidly changing market environment. Korenčić et al. (2021) explored how topic coverage can be used to evaluate the effectiveness of topic models, which is crucial for improving topic modeling methods to more accurately capture the focus of consumer discussions. [2] These research findings not only enrich our understanding of the application of data analytics in magazine subscription services but also provide valuable references and insights for this study.

This study will build on these pioneering studies by incorporating state-of-the-art machine learning techniques and algorithms to develop an efficient personalized magazine recommendation system that aims to enhance consumer satisfaction and optimize publishers' marketing strategies. We will particularly focus on improving the adaptability of the technology to meet the challenges of rapid digitalization, thereby better satisfying consumer and market needs.

III. Approach

This study utilizes the 'Magazine Subscriptions' section of the Amazon Review Dataset from 2023, applying multi-stage data analysis and machine learning techniques to explore key factors influencing customer satisfaction and preferences.

A. Data Ingestion

We start by loading customer reviews and metadata from two separate JSON Lines files: "Magazine Subscriptions" and "Meta Magazine Subscriptions". These files contain structured data in the form of dictionaries representing individual reviews and metadata respectively. Review Data Attributes include ratings, text, user details, and timestamps. Metadata Attributes include descriptions, category information, and imagery details.

Utilizing the 'Magazine Subscriptions' section of the Amazon Review Dataset (2023), which includes thousands of consumer reviews. This large-scale Amazon Reviews dataset was collected in 2023 by McAuley Lab at UCSD. We focus on the Magazine Subscriptions data. The dataset is in JSON Lines format, with each JSON object representing a single data record. There are 2 types of files in the dataset:

- **Review data (`Magazine_Subscriptions`):** Size: 71,494 instances (Reviews) x 10 features. This file contains individual reviews for magazine subscriptions. Each record includes information such as the rating, title of the review, review text, images (if any), ASIN (Amazon Standard Identification Number), parent ASIN, user ID, timestamp of the review, votes for helpfulness, and whether the purchase was verified. Figure 1 shows the samples of data.
- **Metadata (`meta_Magazine_Subscriptions`):** Size: 3,391 instances (Magazines) × 16 features. This file contains metadata for the magazine subscriptions themselves. Each record includes details like the main category, title of the magazine, average rating, number of ratings, features, description, price, images, videos, store or publisher information, categories, date first available or manufacturer, parent ASIN, and related products bought together. Figure 2 shows the samples of metadata.

	rating	title	text	images	asin	parent_asin	user_id	timestamp	helpful_vote	verified_purchase
0	5	Great pictures and recipes	Wonderful recipes in this magazine.	[]	B00HLSSQKK	B00HLSSQKK	AE7Y5RLYIKHOZB5NKKOEKYG2SPSQ	2020-12-15 15:29:02.433	0	True
1	4	great for kids who love sports!	Great sports magazine that's on my 9 year olds...	[]	B002PXW04Y	B002PXW04Y	AHORTSSMI6ZZFUFWMPT4UFR2ISUQ	2016-08-12 03:28:32.000	1	True
2	5	A great look at what's new on the kosher scene...	"Joy of Kosher" magazine fills a much-needed n...	[]	B00HCR5090	B00HCR5090	AFW2PDT3AMT4X3PYQG7FJZH5FXFA	2014-04-06 23:32:53.000	4	False

Fig. 1. Sample Data

	main_category	title	average_rating	rating_number	features	description	price	images	videos	store	categories	details	parent_asin	bought_together	subtitle	author
0	Magazine Subscriptions	GQ Print Access Print Magazine	3.6	10	[]	[Product Description, Dive into, GQ, 's cultur...	None	[{'thumb': 'https://m.media-amazon.com/images/...	[]	Conde Nast Publications	[]	{'Date First Available': 'June 2, 2020', 'Manu...	B00FA7T630	NaN	NaN	NaN
1	Magazine Subscriptions	Hi-Fi + Print Magazine	4.4	44	[]	[Hi-Fi+ is Europe's premier English-language h...	None	[{'thumb': 'https://m.media-amazon.com/images/...	[]	Absolute Media Ltd/Hi-Fi+	[Magazine Subscriptions, Arts, Music & Photogr...	{'Date First Available': 'September 19, 2013',...	B00F8P62PO	NaN	NaN	NaN
2	Magazine Subscriptions	Paper Crafts	3.0	3	[]	[]	None	[{'thumb': 'https://m.media-amazon.com/images/...	[]	None	[]	{'Date First Available': 'May 12, 2021'}	B003F1W9T6	NaN	NaN	NaN

Fig. 2. Sample Metadata

Additional types of information included in the dataset are:

1) **Network information:** Users review the same magazines; interaction through helpful votes on reviews.
2) **Text information:** Review texts and magazine descriptions.
3) **Endorsement/recommendation information:** Ratings and helpful votes, average rating.

For more information, visit the dataset website at Amazon Review Data 2023.

B. Data Preprocessing

We start by cleaning and structuring the magazine subscription review data from Amazon. This includes addressing missing values and standardizing text data to establish a clean foundation for analysis. Our preprocessing steps involve:

- Tokenization: We use TweetTokenizer from NLTK for tokenizing the text, which is adapted to social media text and can handle informal language. Tokens=TweetTokenizer().tokenize(text.lower())
- Removing Stopwords: Our functions are designed to refine the text data by stripping away punctuation and removing stopwords—words that typically do not carry significant meaning.
- Stemming: The 'stem sentence' function is prepared to reduce words to their root forms. Although it is not applied universally within the primary analysis loop, it is available for use when normalization of vocabulary is necessary.

C. Sentiment Analysis

We employ Natural Language Processing (NLP) techniques to analyze the sentiment within the reviews using NLTK's

‘SentimentIntensityAnalyzer‘. This step calculates sentiment scores for each review, categorizing them into negative, neutral, or positive based on the compound score calculated. Implementing this step helps reveal the emotional tone and intensity of customer feedback towards products. [3], [4]

The sentiment scores are then added back to the data frame, enriching it with additional columns for negative, neutral, positive, and compound values. Beyond general sentiment metrics, our analysis also dives into sentiments associated with specific topics identified during topic modeling. This deeper insight into thematic sentiment distribution is invaluable, as it helps understand customer emotions in the context of specific content discussed in the reviews.

The table I demonstrates the results of sentiment analysis performed on the first three magazine reviews. Each row represents a review’s sentiment breakdown, which includes measurements of negative, neutral, positive, and compound scores. Below is an explanation of each column and the implications of the sentiment scores:

1) **Cleaned Text:** This shows the text after preprocessing. This cleaned text is used for analysis.
2) **Neg (Negative):** Represents the proportion of the text that was classified as negative. A higher score indicates more negative sentiment within the text.
3) **Neu (Neutral):** Indicates the proportion of text considered emotionally neutral. This score is particularly high when the text is factual or lacks emotional language.
4) **Pos (Positive):** Reflects the proportion of the text deemed to be positive. Higher scores suggest a positive response from the user.
5) **Compound:** A cumulative score that measures the overall sentiment of the text. Scores close to 1 indicate a strong positive sentiment, scores around -1 suggest a strong negative sentiment, and scores near zero imply a neutral sentiment.

From the data , the following could be observed. Review 0: The analysis shows a very high positive sentiment (0.481) with no detected negativity and a compound score of 0.5719, suggesting that the user had a positive experience with the magazine. Review 1: The positive score (0.354) and a compound score of 0.6588 indicate that the review is generally positive, likely appreciating the magazine’s suitability for children. Review 2: Despite a tiny negative sentiment (0.002), the overall sentiment is highly positive with a very high compound score of 0.9987, indicating satisfaction with the magazine’s content.

D. Topic Modeling

We apply Latent Dirichlet Allocation (LDA) to uncover the main topics discussed in consumer reviews. Using the Gensim library, an LDA model is trained to identify potential topic distributions. We select an appropriate number of topics to ensure model coherence and clarity, guided by measures such as topic coherence and separation. [2], [5]

As shown in Table II, topic modeling of magazine reviews identified several topics, each with a set of keywords. These topics provide insights into the diverse content reflected in the reviews. Below is an analysis of each topic from the topic modeling result:

TABLE I
EXAMPLE OF SENTIMENT ANALYSIS RESULTS

ID	Cleaned Text	Neg	Neu	Pos	Compound
0	[wonderful, recipes, magazine]	0.000	0.519	0.481	0.5719
1	[great, sports, magazine, that’s, year, olds]	0.000	0.646	0.354	0.6588
2	[joy, kosher, magazine, fills, much-needed, niche]	0.002	0.821	0.177	0.9987

TABLE II
TOPIC MODELING RESULTS

Topic	Keywords and Weights
0	0.025*”print” + 0.017*”guide” + 0.015*”version”
1	0.037*”like” + 0.020*”good” + 0.019*”ads”
2	0.012*”one” + 0.008*”reviews” + 0.006*”issue”
3	0.016*”fashion” + 0.012*”people” + 0.009*”women”
4	0.033*”great” + 0.026*”ideas” + 0.019*”articles”
5	0.051*”issue” + 0.041*”received” + 0.034*”first”
6	0.037*”articles” + 0.026*”cover” + 0.023*”read”
7	0.062*”subscription” + 0.033*”year” + 0.031*”gift”
8	0.038*”every” + 0.029*”read” + 0.028*”love”
9	0.184*”great” + 0.108*”love” + 0.069*”recipes”

1) **Topic 0 (Print and Digital Formats):** Keywords such as ”print,” ”guide,” and ”version” suggest discussions related to different formats of the magazines, possibly comparing print and digital versions.
2) **Topic 1 (General Appraisal):** This topic, represented by words like ”like,” ”good,” and ”ads,” indicates general opinions about the magazines, including both positive sentiments and mentions of advertisements.
3) **Topic 2 (Critical Reviews):** Keywords such as ”one,” ”reviews,” and ”issue” likely reflect more critical or detailed discussions about specific issues or editions of magazines.
4) **Topic 3 (Fashion and Lifestyle):** Dominated by ”fashion,” ”people,” and ”women,” this topic suggests content focused on fashion and lifestyle, potentially targeting a female audience.
5) **Topic 4 (Content Quality):** Words like ”great,” ”ideas,” and ”articles” indicate positive feedback on the content quality and the ideas presented within the magazines.
6) **Topic 5 (Subscription Experiences):** Featuring terms like ”issue,” ”received,” and ”first,” this topic likely covers user experiences with subscription management and receiving initial copies.
7) **Topic 6 (Engagement with Content):** With ”articles,” ”cover,” and ”read,” this topic suggests engagement with specific articles and features prominently on magazine covers.

8) **Topic 7 (Gift Subscriptions):** Highlighted by ”subscription,” ”year,” and ”gift,” it appears to focus on magazines given as gifts and annual subscriptions.
9) **Topic 8 (Regular Readership):** The presence of ”every,” ”read,” and ”love” indicates a strong affinity and regular readership, with expressions of enjoyment and routine reading habits.
10) **Topic 9 (High Engagement and Recipes):** Dominated by ”great,” ”love,” and ”recipes,” this topic stands out with high positive sentiment and likely pertains to magazines that feature cooking or culinary content.

Overall, topic modeling serves as a powerful tool for understanding reader preferences to align with consumer interests.

E. Visualization

We use Matplotlib and Seaborn to create intuitive charts and graphics that display the results of our data analysis. This includes sentiment distribution plots and topic prevalence bar charts, which provide a visual representation of the data and help identify trends and insights in customer preferences. We also use Gephi to visualize our network.

F. Sentiment Analysis + Topic Modeling

After assigning all sentiment scores and topic modeling scores to each review, we can start investigating the relationship between these scores and other features in review data. The first one to investigate is the average sentiment score for the reviews on each topic.

The table III presents the sentiment analysis results by topic derived from magazine reviews. Most topics show a high percentage of neutral sentiments, indicating that many reviews contain neutral statements without strong emotional expressions. There’s a notable variation in positive sentiment across topics, suggesting different levels of satisfaction associated with the content of various themes. Negative sentiments are relatively low across all topics.

With the keywords ”great”, ”love” and ”recipes”, Topic 9 (High Positivity and Engagement) stands out with the highest positive sentiment (0.621) and compound score (0.661), indicating very favorable reviews. The keywords ”great”, ”ideas” and ”articles” show positive sentiment in reviews, therefore Topic 4 (Content Appreciation) has higher positive sentiment (0.308) and the highest compound score (0.699).

In Figure 3, the box plot categorizes sentiment scores by topics identified in magazine content, ranging from topics related to magazine formats (Topic 0) to those discussing specific content like recipes (Topic 9). The analysis shows that different topics engage readers’ emotions to varying degrees. For instance, Topic 9, which likely includes lifestyle and culinary content, has a higher median compound sentiment score, indicating very positive feedback from readers. Topics like Topic 3 (Fashion and Lifestyle) and Topic 8 (Regular Readership) also show predominantly positive sentiments, suggesting these topics resonate well with the audience. However, Topics 0 and 2, which may involve more critical discussions, exhibit lower median sentiment scores and wider distributions, indicating a mixed response among readers. This plot reveals how sentiments in reviews can provide insights into which topics are popular and which may be unpopular.

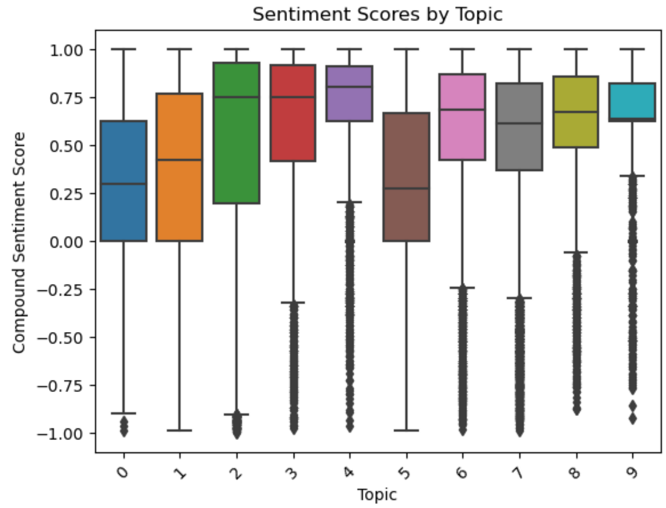

Fig. 3. Sentiment Scores by Topic

The figure 3 illustrates the sentiment score by topic in a visual way.

TABLE III
SENTIMENT ANALYSIS SCORES BY TOPIC

Topic	Neg	Neu	Pos	Compound
0	0.046763	0.651686	0.271078	0.279460
1	0.067486	0.765086	0.167428	0.318390
2	0.047285	0.797845	0.154873	0.507449
3	0.034885	0.741851	0.223269	0.583315
4	0.013417	0.678539	0.308045	0.699584
5	0.049036	0.818178	0.132775	0.243286
6	0.026994	0.692532	0.280481	0.575522
7	0.038426	0.668357	0.293214	0.485078
8	0.012864	0.679036	0.308100	0.616122
9	0.006579	0.372175	0.621245	0.661061

The second one to investigate is whether the rating affects the sentiment score. In Figure 4, the box plot illustrates the distribution of compound sentiment scores across different user ratings from 1 to 5 stars. Notably, higher ratings (4 and 5 stars) show a strong positive sentiment, with median values well above zero, indicating general satisfaction among reviewers. The boxes for these ratings are narrower. This implies less variability and a consensus of positive sentiments among users. Conversely, lower ratings (1 and 2 stars) exhibit a broader spread and medians closer to zero, highlighting a mix of sentiment that leans towards dissatisfaction. In particular, the 1-star rating shows some outliers with extremely negative sentiment scores, indicating strong dissatisfaction with certain reviews. This visual encapsulates the clear trend that as the rating increases. As ratings go up, so does customer satisfaction. Ratings are a reliable indicator of customer satisfaction.

The third one to investigate is whether different topics affect the number of helpful votes. In Figure 5, this scatter plot illustrates the distribution of helpful votes across different

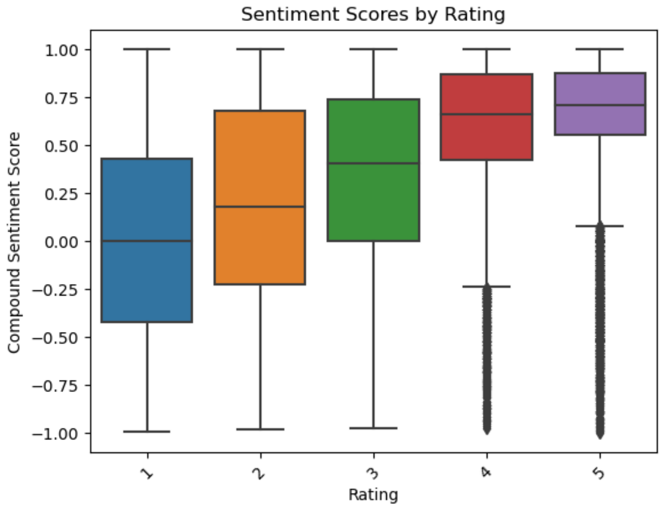

Fig. 4. Sentiment Scores by Rating

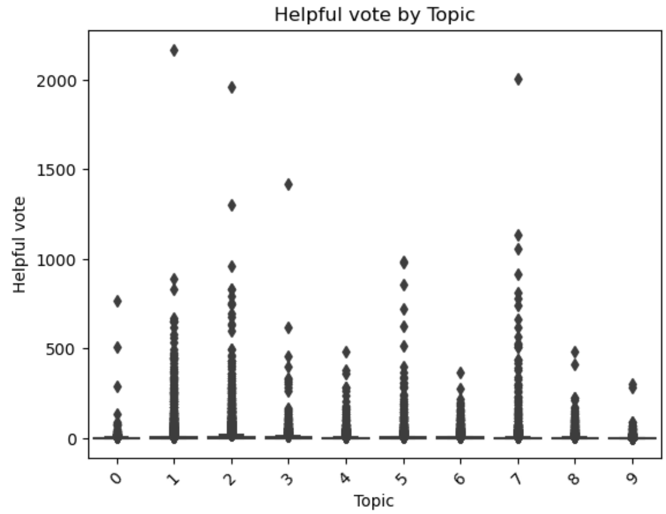

Fig. 5. Helpful vote by Topic

topics extracted from magazine reviews. The graph shows that while most topics receive a consistent level of participation, indicated by the concentration of votes near the bottom of the graph, certain topics stand out with higher peaks and a few spikes indicating exceptional participation.

For example, Topic 1, Topic 2, and Topic 7 show higher and more spread-out helpful vote counts, suggesting that reviews on these topics are perceived as particularly useful by other readers.

On the other hand, Topics like Topic 4, Topic 6, Topic 8, and Topic 9, though they have their share of helpful votes, do not reach the extremes seen in more specialized topics. This could mean that while these general topics are relevant, they do not always provide unique content that attracts significant helpful votes.

This pattern of helpful votes shows which topics engage more readers. It also shows which topics lead to more meaningful discussions, which makes the reviews on those topics more useful.

G. Network

To enhance our collaborative approach to the recommendation system, we utilize the network between users' interactions with the magazine through their reviews. In the content of our network analysis for our project, there are two types of nodes and one type of edge: users and magazines being nodes, and the reviews being the edge. We are interested in the relationships between these entities through edges as interactions: reviews. To represent the structures of our network, we are building a Bipartite graph. A bipartite graph is a specific type of graph where nodes are divided into two sets, and all the nodes can only connect to the different sets. The edges can only create the link between various types of nodes, meaning they can't connect directly if the nodes are the same types. In our case, since users are not communicating with different users directly, they build the connection by leaving the reviews

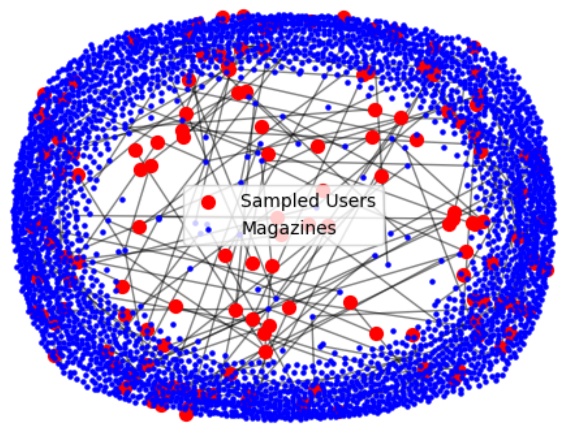

Fig. 6. A small portion of our Bipartite network

under the same magazine. Hence, the magazine also needs to be a node to create connections between users.

1) **Users as Nodes:** Each user is represented by a unique user ID, signifying their individuality in the network. Users engage with magazines by reviewing them. A user can leave multiple reviews under a magazine, indicating their active participation in the network.
2) **Magazines as Nodes:** Magazines are the products under review, serving as the focal point of user interactions. A single magazine can be reviewed by many users, highlighting its importance in the network.
3) **Reviews as Edges as interaction:** The reviews themselves create the edges in the network. When a user reviews a magazine, it creates an edge linking to the magazine.

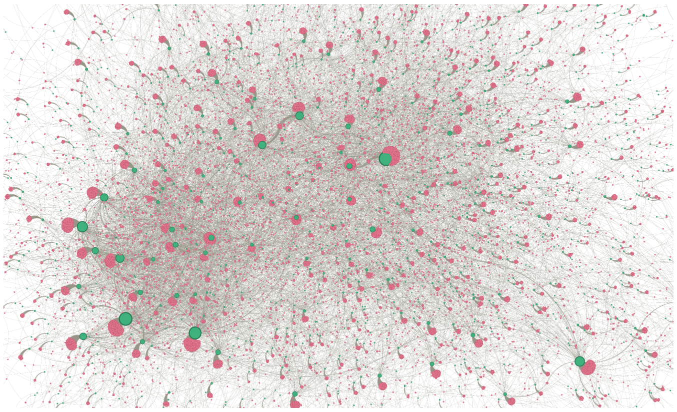

Fig. 7. The Bipartite network drawn by Gephi (green nodes = magazines, red nodes = users, edges(links) = reviews

H. Recommendation System

Integrate results from network analysis, sentiment analysis and topic modeling to recommend personalized magazines, using weighted ranking to optimize recommendation accuracy. The recommendation system evaluates users' historical interactions and preferences, implementing intelligent matching through machine learning technologies. [6] [7]

Parameters used in the recommendation system:

- Rating: The importance of the magazine's overall user rating in influencing the recommendation.
- Helpful votes: The importance of the number of helpful votes a review has received.
- Sentiment: The weight given to the sentiment analysis of the reviews (how positive or negative they are).
- Network centrality: We calculate centrality using degree centrality from our network analysis. This measure indicates how central or significant a particular review or reviewer is within the network of users and magazines. Degree centrality is particularly useful for identifying influential users or key reviewed items in our dataset. It is computed as the number of links incident upon a node (i.e., the number of ties that a node has).
- Topic: The importance of the thematic content of the reviews as determined by topic modeling.
- Netowrk community: To identify clusters of users with similar preferences or magazines that frequently get reviewed together.

These configurations are likely used to tailor how the recommendation system prioritizes different aspects of the data when making recommendations. Adjusting these weights allows the system to be fine-tuned to emphasize specific features more than others, potentially leading to different sets of recommended magazines, as shown in Test results. Each test uses a different weight configuration to see how changing these priorities affects the recommendations generated by the system. This method helps optimize the performance of recommendation algorithms to align with specific user preferences or business goals.

IV. Experiment

We are interested in investigating how different factors affect the recommendation result. Since we are focusing on building the recommendation system in a collaborative approach, which means the indicators are not as obvious and straightforward as the content-based approach(such as using categories or the title of the magazine), it is crucial to do a series of experiments to see how different factors affect the result. The biggest challenge in evaluating the result is that there is no way to assess the accuracy of the recommendation result, and the dataset needs to include the data for the recommendation. Also, we need more resources to implement feedback experiments to test whether the user likes the recommended items. For still evaluating the recommendation, we focus on the result similarity levels and the frequency of magazine recommendations across all configurations.

A. Experiment Setup

For finding the overall effectiveness of our collaborative-approach recommendation system, a comprehensive score will be calculated for each magazine candidate which considering various aspects of user interaction. Besides the rating and the number of helpful votes, we are also including the factors from our topic modeling, sentiment analysis, and network construction. Let's define the weights and scores for our model:

- w_r - weight for ratings
- w_h - weight for helpful votes
- w_s - weight for sentiment analysis
- w_c - weight for centrality in the network
- w_t - weight for topic relevance
- w_{com} - weight for community alignment
- $\bar{r}$ - average rating
- h - total helpful votes
- $\bar{s}$ - average sentiment score
- c - centrality score
- t - topic score
- com - community score

The score for a magazine can be calculated as follows:

$$\text{Score} = w_r \cdot \bar{r} + w_h \cdot h + w_s \cdot \bar{s} + w_c \cdot c + w_t \cdot t + w_{com} \cdot com$$

The table IV-A shows the configurations of weights used to calculate the recommendation scores based on different criteria:

TABLE IV
RECOMMENDATION SYSTEM WEIGHTS CONFIGURATION

Configuration	**Rating**	**Votes**	**Sentiment**	**Centrality**	**Topic**	**Community**
Review and Rating Centric	0.4	0.2	0.2	0.0	0.2	0.0
Only Network	0.0	0.0	0.0	0.6	0.0	0.4
Network and Rating Centric	0.4	0.0	0.0	0.3	0.0	0.3
All Balanced	0.167	0.167	0.167	0.167	0.167	0.167
High Emphasis on Review Itself	0.2	0.2	0.2	0.1	0.2	0.1
High Emphasis on Network	0.1	0.1	0.1	0.4	0.1	0.2

B. Experiment Results

For each time, the recommendation generator generates the top 8 magazines to be recommended based on the magazine's score. All six different sets of configurations were tested, and the results are listed below:

- Test 1 Recommendations: Review and rating centric
  - National Geographic Kids
  - National Geographic Magazine
  - Highlights For Children
  - Family Handyman
  - Food Network Magazine
  - AllRecipes
  - Real Simple
  - Family Handyman
- Test 2 Recommendations: Only network
  - National Geographic Magazine
  - Family Handyman
  - Reader's Digest
  - Food Network Magazine
  - Real Simple
  - National Geographic Kids
  - Popular Science
  - Popular Mechanics
- Test 3 Recommendations: Network and rating centric
  - National Geographic Magazine
  - Family Handyman
  - Reader's Digest
  - Food Network Magazine
  - Real Simple
  - National Geographic Kids
  - Food Network Magazine Print Edition
  - Numismatic News (1-year auto-renewal)
- Test 4 Recommendations: All balanced
  - National Geographic Kids
  - National Geographic Magazine
  - Highlights For Children
  - Family Handyman
  - Food Network Magazine
  - Real Simple
  - AllRecipes
  - Popular Mechanics
- Test 5 Recommendations: High emphasis on the review itself
  - National Geographic Kids
  - National Geographic Magazine
  - Highlights For Children
  - Family Handyman
  - Food Network Magazine
  - Real Simple
  - AllRecipes
  - Popular Mechanics
- Test 6 Recommendations: High emphasis on network
  - National Geographic Magazine
  - National Geographic Kids
  - Family Handyman
  - Reader's Digest
  - Real Simple
  - Food Network Magazine
  - Highlights For Children
  - Popular Mechanics

From the result, we can observe several things. First, we found that, in general, some magazines, such as National Geographic Kids and National Geographic Magazine, appear in all test results. This may be because those magazines are trendy, so the reviews are active and positive, and there are many interactions between reviewers (users). Second, different configurations affect the small to medium portion of the results. Not all the recommendations are the same; this proves our approaches affect the results, whether the sentiment and topic for the review or the network centrality and community detection. Third, the results from test 4 and test 5 are identical, at least from the top 8 recommendations. This may be because the weight distribution numbers do not have a noticeable difference, so the result is not affected much.

C. Similarity Analysis

To better represent the similarity level of each recommendation result, we adapted the Jaccard similarity level and built a heatmap in Figure 8. Jaccard similarity is a statistic used to gauge the similarity and diversity of sample sets. It's beneficial when comparing the similarity and diversity of the recommendation system's results.

The Jaccard similarity index J between two sets A and B is defined as the ratio of the number of elements in the intersection of the sets to the number of elements in the union of the sets. The formula is given by:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

where $|A \cap B|$ represents the number of elements common to both sets, and $|A \cup B|$ represents the total number of unique elements in both sets.

From the heatmap, it is evident that different weight configurations affect the outcome of the recommendation to some degree depending on the different parameter settings. The Jaccard similarity shows that adjusting the weights between each collaborative method can result in generating different results. The Jaccard score ranges from as low as 0.45 to a perfect 1.0 indicating that the results are similar in general, some magazines appear all the time due to it being already popular among the users, so it is active in all categories in parameters. For example, the results from test 4(all balanced) and test 5 (high emphasis on review itself) are identical with a score of 1.0, and they are also almost identical to the result from test 1 (review and rating-centric) with a high score of 0.88. This is reasonable since they all have a higher focus on the review itself, with a focus on high votes, ratings, and underlying text information for both the topic modeling score and sentiment score.

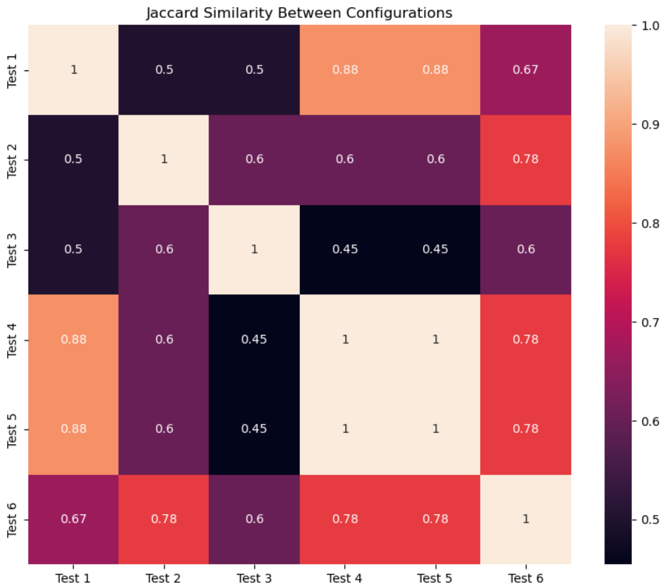

Fig. 8. Jaccard Similarity Between Configurations

On the other hand, the results from test 2 (only network), test 3 (network and rating centric), and test 6 (High emphasis on network) tell another story. They have a lower similarity score compared to the other tests, such as having a score of 0.45 between test 3 and test 4 (and test 5). From the difference, the results tend to be changed when it comes to involving the network parameters. This proves that the recommendation system values the magazines based on the user interactions inside the network, which may differ based on the review itself. We still can see some similar magazines appear in those tests, such as National Geographic Magazine and National Geographic Kids. Again, these magazines may appear everywhere since they are both some of the hottest magazines so they can have many famous and positive reviews and, simultaneously, are very active inside the interaction network.

The comparison shows that reviews and networks can impact the recommendation system differently. The recommendation system designer needs to decide whether to focus on the review itself or more on the network. Suppose we can access user purchasing history or browsing history. In that case, we can compare those activity data with the results we have and tune the weights of the recommendation system based on the comparison, but that is out of the scope of our current project.

V. Conclusion

In our endeavor to create a personalized magazine recommendation system, we adopted a collaborative approach that leverages sentiment analysis, topic modeling, and network analysis. This comprehensive strategy involved calculating the sentiment score and assigning topics to each review based on the results of topic modeling. We also constructed a Bipartite network, with users and magazines as nodes and reviews as edges, to represent the interactions between users and magazines. Subsequently, we developed the recommendation system and conducted experiments with various weight configurations to gauge their impact on the results. We found that when the network is involved in the system, the results tend to have some differences, which proves that they are valuable and should be considered in the recommendation system. The designer of the recommendation system needs to consider whether to value the network more; further investigation of users' purchasing or browsing habits may be useful for optimizing weight configurations to better match user interests.

Two future ideas could be investigated to evaluate the recommendation system further. First, we can try to find the data for each customer's browsing or purchasing history. We can utilize customers' history data on the magazines and compare them with the recommendation results. If the recommended magazine appeared in history, it may prove that it can successfully recommend the exciting magazine to the customer. Second, we can have a feedback experiment. We can gather these users and let them evaluate whether they like the recommendation results and are willing to browse or purchase the recommended magazine. We learned that the recommendation system is difficult to assess since it has no defined correct answer for the results. Each situation and entity is different; we can only do our best to suggest items based on our resources, and our collaborative approach can effectively adapt to user preferences and trends, enhancing both recommendation relevance and user satisfaction.

References

[1] A. Abdellatif, S. Sahmoud, and A. Nizam, "A unified framework for multi-language sentiment analysis," in *2023 3rd International Conference on Computing and Information Technology (ICCIT)*, 2023, pp. 280–284.

[2] D. Korenčić, S. Ristov, J. Repar, and J. Šnajder, "A topic coverage approach to evaluation of topic models," *IEEE Access*, vol. 9, pp. 123 280–123 312, 2021.

[3] X. Wu, Y. Linghu, T. Wang, and Y. Fan, "Sentiment analysis of weak-ruletext based on the combination of sentiment lexicon and neural network," in *2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, 2021, pp. 205–209.

[4] L. Wang, J. Niu, and S. Yu, "Sentidiff: Combining textual information and sentiment diffusion patterns for twitter sentiment analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 10, pp. 2026–2039, 2020.

[5] D. Yamunathangam, C. Priya, G. Shobana, and L. Latha, "An overview of topic representation and topic modelling methods for short texts and long corpus," in *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 2021, pp. 1–6.

[6] Q. Pu and B. Hu, "Intelligent movie recommendation system based on hybrid recommendation algorithms," in *2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE)*, 2023, pp. 1–5.

[7] P. Mathew, B. Kuriakose, and V. Hegde, "Book recommendation system through content based and collaborative filtering method," in *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*, 2016, pp. 47–52.