Qualitative and Quantitative Methods in Libraries (QQML) 15,2:243-254, 2026

Applying Sentiment Analysis of Customer Reviews to Support Library Book Selection

Poulami Pal[1], Riya Karmakar[2], Rajashree Das[3] and Sibsankar Jana[4]

[1,2,4] University of Kalyani

[3] Bidhan Chandra College

ABSTRACT: The rapid growth of online customer reviews has created new opportunities for libraries to utilize user-generated feedback for informed collection development. This study applies sentiment analysis techniques to examine customer opinions and support effective library book selection. Reviews were collected from Amazon (www.amazon.in) for two popular self-help books: *The Power of Your Subconscious Mind* by Joseph Murphy and *The Power of a Positive Attitude: Your Road to Success* by Roger Fritz. A lexicon-based sentiment analysis approach was implemented using the R programming language. The dataset was pre-processed to improve data quality and analysed to identify emotional dimensions such as joy, trust, anger, fear, and overall sentiment polarity. The results indicate that both books exhibit strongly positive sentiment, reflecting high levels of user satisfaction and acceptance. These findings demonstrate that sentiment analysis can efficiently process large volumes of textual review data and serve as a valuable decision-support tool for librarians. The study highlights the practical application of text mining techniques in supporting evidence-based library collection development.

Keywords: Sentiment Analysis, Opinion Mining, Natural language processing, Customer feedback, Book selection process.

Received: 19.1.2026 / Accepted: 22.5.2026 ISSN 2241-1925

1 INTRODUCTION

In the digital communication environment, feedback plays an important role in understanding users' reactions to various types of content such as services, products, entertainment, news, and tutorials shared through mass media. This feedback may be collected through formal methods such as surveys or may appear as unsolicited comments, criticisms, and suggestions on social media platforms. Such feedback helps in evaluating the effectiveness of services and communication.

To understand feedback more effectively, it is necessary to consider two fundamental questions:

• What is the feedback about?

• Is the feedback positive or negative?

The first question can be addressed using text mining techniques, which help identify the subject or topic of the feedback. The second question relates to automatic sentiment classification, which enables the identification of the polarity of the feedback. This makes it possible to focus on areas requiring improvement when feedback is negative, or to identify areas of success when feedback is positive (Gamon, 2004).

Furthermore, information shared on social media often reflects users' opinions, emotions, and attitudes, which may be positive, negative, or neutral. Sentiment analysis, also known as opinion mining, is used to identify and extract such subjective information from textual data, combining Natural Language Processing (NLP), text mining, and statistical methods to classify sentiment into categories such as positive, negative, or neutral. In addition to polarity classification, sentiment analysis can also detect emotional states such as joy, sadness, anger, fear, and trust, thereby providing deeper insight into customer perceptions and satisfaction levels (Kumar & Zymbler, 2019; Shara, 2021).

Two commonly used approaches to sentiment analysis are machine learning-based methods and lexicon-based methods. Machine learning approaches often achieve high predictive accuracy but may be less transparent and interpretable due to their black-box nature. In contrast, lexicon-based approaches rely on predefined dictionaries of positive and negative words to calculate sentiment scores. In this study, a lexicon-based approach was adopted because of its transparency and suitability for analyzing review data.

In recent years, sentiment analysis has become an important tool for analyzing large volumes of online textual data. In the context of libraries, reviews available on e-commerce platforms can provide valuable support for collection

development. Incorporating user-generated feedback can complement traditional selection methods and offer a more user-centered perspective.
Therefore, the present study applies a lexicon-based sentiment analysis approach to evaluate customer reviews of books and determine their overall sentiment polarity and emotional characteristics. The findings of this study are expected to assist librarians in making more informed, data-driven, and user-centered book selection decisions.

2 LITERATURE REVIEW

In this section, various studies on sentiment analysis using Natural Language Processing and Deep Learning techniques are discussed. This review is conducted on the basis of numerous latest studies related to sentiment analysis of web content about users' opinions, emotions, reviews in the field of sentiment analysis. Here, we review the related articles and assess whether there is any gap between them or not.

Khanna et al (2017) described a very easy, cost-effective approach that exposed the opinions of much larger public which would seem impossible and also presented an exhaustive study of the efficiency of R language in opinion mining and how opinion data can be extracted from X (formerly Twitter) database (Khanna et al, 2017). In the paper "*An Application Example of Sentiment Analysis Using R*", Shara (2021) described an application of sentiment analysis using R. For analysis, he had used the AFINN, BING and NRC lexicons. Some text analysis, such as bigrams, trigrams and related methods, was also conducted (Shara, 2021). Vatambeti et al (2023) analysed customers' tweets about three different food delivery apps using R-Studio and a deep-learning based opinion examine approach by combining the Convolutional Neural Network (CNN) and Bi-directional Long Short-Term Memory (Bi-LSTM) models (Vatambeti et al, 2023). Hassan et al (2022) explored the netizen's opinions on cryptocurrency by emotion analysis theory and lexicon sentiments analysis via machine learning and implications of the cryptocurrency phenomenon under the emotion theory. Specifically, he showed how to supervise machine learning which could measure emotion using Twitter data to explain the implications of cryptocurrencies (Hassan et al, 2022). Kawade and Oza (2017) discussed tweets of Uri attack and found emotions and polarity of tweets were expressed by people regarding that particular topic. To mine emotions and polarity in tweets, text mining techniques are used in their study (Kawade & Oza, 2017).

In the paper of Jaichandran et al (2019) performed sentiment analysis on Twitter data by using Natural Language Processing for the movies titled Kabali, Bharath Ane Nenu Mersal, and Dangal (Jaichandran et al, 2019). Sivakumar and Reddy (2017) proposed a new method of analyzing online student feedback collected from Twitter API by measuring semantic relatedness between aspect words and student opinion sentences, which will help the students to improve their studied and help the instructors to improve their teaching skills. They used classification and clustering techniques to categorize the opinions (Sivakumar & Reddy, 2017). Tul et al (2017) highlighted latest studies regarding the implementation of deep learning models such as deep neural networks, convolutional neural networks and many more for solving different problems of sentiment analysis such as sentiment classification, cross lingual problems, textual and visual analysis and product review analysis, etc. (Tul et al, 2017). Gamon (2004) demonstrated that automatic sentiment classification in very noisy domain of customer feedback data by using large feature vectors in combination with feature reduction. Linear support vector machines achieved high classification accuracy on data that presents classification using feature reduction (Gamon, 2004).

We have seen some related work and discussed those articles. Here we observe that there are some articles on sentiment analysis using R-studio of reviews collected from Twitter of netizen's opinion about cryptocurrency, Uri attack, student feedback, food-delivery app review, application of sentiment analysis on book analysis etc. Henceforth, we decide to choose the platform Amazon.in from where datasets are collected and using R language and R-studio with so many built-in software packages for opinion mining.

3 METHODOLOGY

This study adopted a quantitative research design using a lexicon-based sentiment analysis approach. Customer review data were collected from the Amazon (*www.amazon.in*) platform for two purposively selected books. The collected review texts were pre-processed and analysed using the R programming language in the RStudio environment. Several R packages were used, including *tm* for text processing, *SnowballC* for stemming, *syuzhet* for sentiment scoring and emotion analysis, and *ggplot2* for data visualisation. The overall procedure involved dataset creation, text pre-processing, sentiment score calculation, and interpretation of results.

3.1 Dataset creation

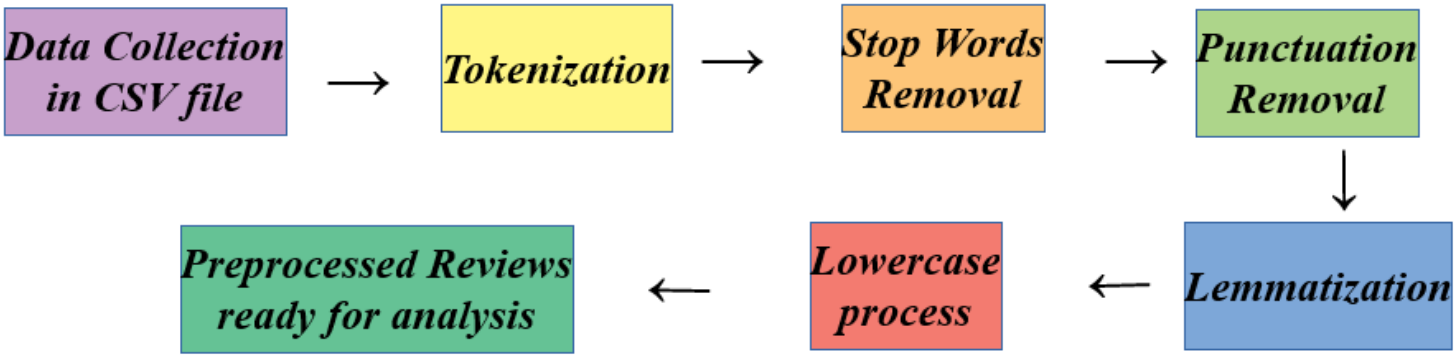
The dataset for this study was created by collecting customer reviews from the Amazon (*www.amazon.in*) platform. Reviews were extracted using the *Amazon Review Exporter* extension in Google Chrome and saved in CSV format for further analysis.

Two popular self-help books were selected as samples for this study: *The Power of Your Subconscious Mind* (ISBN: 978-8194790839) by Joseph Murphy and *The Power of a Positive Attitude: Your Road to Success* (ISBN: 978-9389432640) by Roger Fritz. These books were chosen due to their high number of customer reviews and popularity among readers.

The collected dataset consists of user-generated review texts and corresponding ratings, which were used for sentiment analysis.

3.2 Data pre-processing

In Natural Language Processing (NLP), text pre-processing is an essential step for improving the quality and consistency of textual data. The sequence of pre-processing steps applied in this study is illustrated in Figure 1.



The collected dataset consisted of raw review texts and ratings, which required cleaning before analysis. Initially, irrelevant elements such as usernames, special characters, and symbols were removed.

The text data was then tokenized into individual words to facilitate further processing. Stop words, including frequently occurring words such as “the”, “is”, and “of”, were removed as they do not contribute significantly to sentiment analysis. In addition, punctuation marks and extra whitespace were eliminated.

Stemming was applied to reduce words to their root forms, thereby minimizing variations of similar words. Lemmatization was also performed to convert words into their base or dictionary forms, improving linguistic consistency.

Finally, all text was converted to lowercase to ensure uniformity across the dataset. The resulting pre-processed data was then ready for subsequent sentiment analysis.

3.3 Sentiment Score Calculations

After text cleaning and corpus creation, a Term Document Matrix (TDM) was generated to represent the frequency of terms across the review documents. Based on the matrix, visualisations such as word clouds and bar charts were generated to identify prominent terms in the dataset.

Subsequently, sentiment scores were computed using the *syuzhet* package. The scores were classified into emotional categories including anger, anticipation, disgust, fear, joy, sadness, surprise, trust, as well as overall positive and negative polarity.

4 DATA ANALYSIS

The pre-processed customer review datasets were analysed using the *syuzhet* package in the R programming environment to determine both emotional dimensions and overall sentiment polarity. Summary tables and graphical visualisations were generated to present the sentiment patterns of the selected books. Two popular self-help books were taken as samples for analysis: *The Power of Your Subconscious Mind* by Joseph Murphy and *The Power of a Positive Attitude: Your Road to Success* by Roger Fritz.

Sample 1: *The Power of Your Subconscious Mind*

Figure 2 presents the distribution of customer ratings for the book. The ratings indicate that the book received favourable responses from readers. Table 1 shows the emotional sentiment scores extracted from the reviews. Among the

emotions, trust (102), joy (82), and anticipation (82) recorded the highest values, suggesting that readers viewed the book positively and found it encouraging and reliable. Lower scores for negative emotions such as anger, disgust, and sadness indicate limited dissatisfaction among reviewers.

Table 2 presents the overall polarity of sentiment. The positive score was 188 (87.44%), whereas the negative score was 27 (12.56%), resulting in a net sentiment score of 161. Figure 3 further illustrates the overall positive sentiment associated with the book. These findings suggest that *The Power of Your Subconscious Mind* is highly appreciated and widely accepted by readers.

Figure 2: Ratings for the book "*The Power of Your Subconscious Mind*" given by customers

Table 1: Sentiment Score for Each Emotion

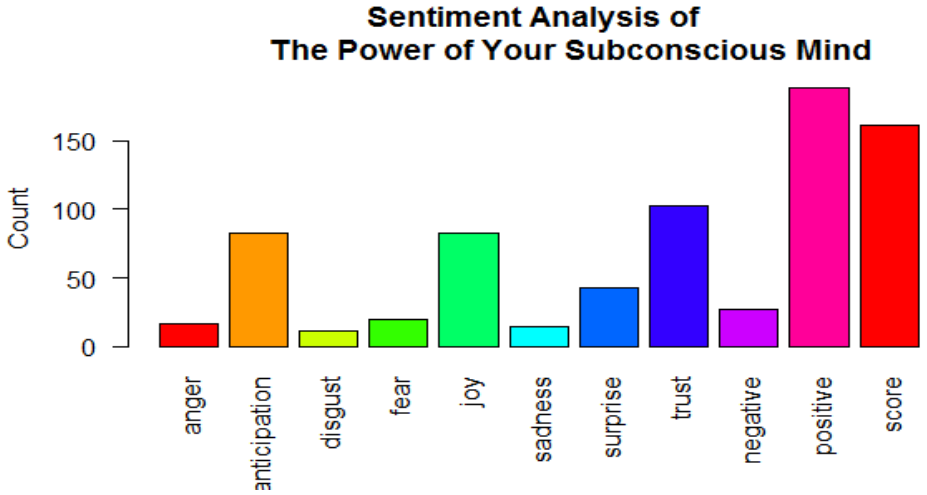
Emotions	**Anger**	**Anticipation**	**Disgust**	**Fear**	**Joy**	**Sadness**	**Surprise**	**Trust**
Score	17	82	11	20	82	14	43	102

Table 2: Polarity of Overall Sentiment

Polarity	**Positive**	**Negative**	**Net Sentiment**
Values	188	27	161

Percentage	87.44%	12.56%	74.88%

Figure 3: Sentiment Analysis of The Book
"*The Power of Your Subconscious Mind*"



Sample 2: ***The Power of a Positive Attitude: Your Road to Success***

Figure 4 shows the distribution of customer ratings for the second book, which also reflects generally favourable reader responses. Table 3 presents the emotional sentiment scores. Similar to the first sample, trust (83), joy (66), and anticipation (66) were the dominant emotions, indicating that readers perceived the book as motivating and beneficial. Negative emotions such as anger, disgust, and sadness were comparatively low.

Table 4 presents the polarity scores for this book. The positive score was 132 (86.84%), while the negative score was 20 (13.16%), resulting in a net sentiment score of 112. Figure 5 also demonstrates the overall positive nature of customer feedback. These results indicate that *The Power of a Positive Attitude: Your Road to Success* was well received by readers.

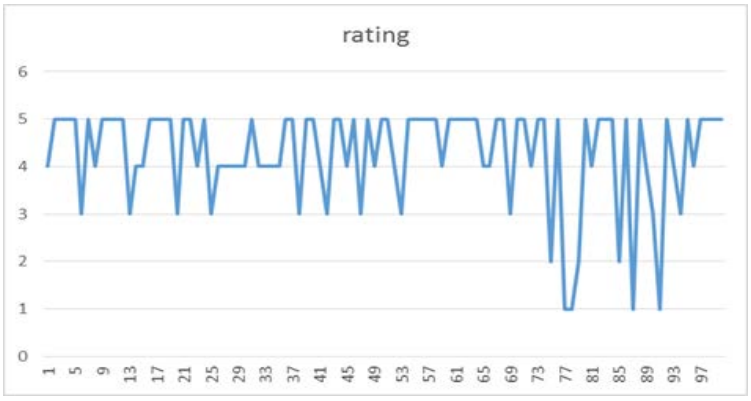


Figure 4: Ratings for the book "*The Power of a Positive Attitude: Your Road to Success*" given by customers

Table 3: Sentiment Score for Each Emotion

Emotions	**Anger**	**Anticipation**	**Disgust**	**Fear**	**Joy**	**Sadness**	**Surprise**	**Trust**
Score	11	66	09	14	66	12	47	83

Table 4: Polarity of Overall Sentiment

Polarity	**Positive**	**Negative**	**Net Sentiment**
Values	132	20	112
Percentage	86.84%	13.16%	73.68%

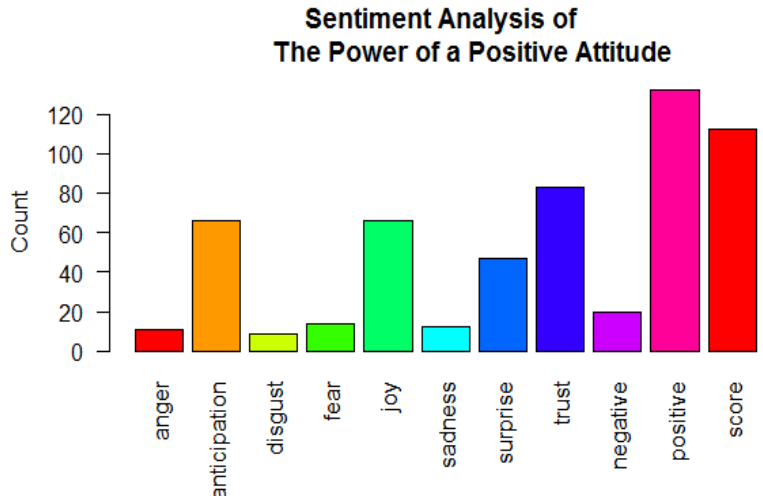


Figure 5: Sentiment Analysis of the Book "The Power of a Positive Attitude: Your Road to Success"

Comparative Observation

A comparison of both books shows that each received predominantly positive reviews. However, *The Power of Your Subconscious Mind* achieved a higher overall sentiment score (161) than *The Power of a Positive Attitude: Your Road to Success* (112). This suggests that the first book generated relatively stronger positive engagement among customers. Overall, the findings demonstrate that sentiment analysis can effectively interpret user reviews and provide useful evidence for book selection decisions in libraries.

5 CONCLUSIONS

This study shows how sentiment analysis can be used as a practical tool for understanding customer opinions in the context of library book selection. By applying a lexicon-based approach, it was possible to analyse user-generated reviews and identify both the overall sentiment and the emotional patterns expressed by readers. The results clearly indicate that both selected books received strongly positive responses, reflecting a high level of reader satisfaction.

These findings suggest that analysing online reviews can provide useful insights for librarians when making collection development decisions. Instead of relying only on traditional methods, incorporating user feedback allows for a more user-centred and evidence-based approach to selecting books. This can help libraries better align their collections with readers' interests and preferences.

However, the study is limited to a small number of books and a single data source. Future research can expand the dataset, include different genres, and explore more advanced techniques to further improve the accuracy and applicability of sentiment analysis in library decision-making.

References

Gamon, M. (2004). Sentiment classification on customer feedback data: Noisy data, large feature vectors, and the role of linguistic analysis. Proceedings of the 20th International Conference on Computational Linguistics, 841–847.

Hassan, M. K., Hudaefi, F. A., & Caraka, R. E. (2022). Mining netizen's opinion on cryptocurrency: Sentiment analysis of Twitter data. Studies in Economics and Finance, 39(3), 365–385. https://doi.org/10.1108/SEF-06-2021-0237

Jaichandran, R., Basha, C. B., Rajaprakash, K. L., & Raja, S. K. (2019). Sentiment analysis of movies on social media using R Studio. International Journal of Engineering and Advanced Technology, 8(6), 2171–2175. https://doi.org/10.35940/ijeat.F8586.088619

Kawade, D. R., & Oza, K. S. (2017). Sentiment analysis: Machine learning approach. International Journal of Engineering and Technology, 9(3), 2183–2186. https://doi.org/10.21817/ijet/2017/v9i3/1709030151

Khanna, P., Kumar, S., Mishra, S., & Sinha, A. (2017). Sentiment analysis: An approach to opinion mining from Twitter data using R. International Journal of Advanced Research in Computer Science, 8(8), 252–256. https://doi.org/10.26483/ijarcs.v8i8.4716

Kumar, S., & Zymbler, M. (2019). A machine learning approach to analyze customer satisfaction from airline tweets. Journal of Big Data, 6(1), Article 62. https://doi.org/10.1186/s40537-019-0224-1

AlQahtani, A. S. M. (2021). Product sentiment analysis for Amazon reviews. International Journal of Computer Science and Information Technology, 13(3), 15–30. https://doi.org/10.5121/ijcsit.2021.13302

Shara, J. (2021, September 22). An application example of sentiment analysis using R.

Sivakumar, M., & Reddy, U. S. (2017). Aspect based sentiment analysis of students' opinion using machine learning techniques. 2017 International Conference on Inventive Computing and Informatics (ICICI), 726–731. https://doi.org/10.1109/ICICI.2017.8365231

Tul, Q., Ali, M., Riaz, A., Noureen, A., Kamranz, M., Hayat, B., & Rehman, A. (2017). Sentiment analysis using deep learning techniques: A review. International Journal of Advanced Computer Science and Applications, 8(6). https://doi.org/10.14569/IJACSA.2017.080657

Vatambeti, R., Mantena, S. V., Kiran, K. V. D., Manohar, M., & Manjunath, C. (2023). Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique. Cluster Computing. https://doi.org/10.1007/s10586-023-03970-7