RESEARCH PAPER

Sentiment Analysis for Patient Satisfaction Assessment in Bilingual E-Pharmacy Platforms

Anand Kumar Gupta[1*], Indrajeet Kumar[2]

[1,2]*Associate Professor, Department of CE & IT, School of Engineering, P P Savani University, Surat, Gujarat, India*

ABSTRACT

Patient-generated content has grown significantly as e-pharmacy continues to grow quickly, especially in China, which has one of the biggest and most vibrant online pharmacy ecosystems. Product reviews are one of the most important categories of user-generated data since they offer authentic, experience-based insights straight from patient. These reviews have an impact on business-level planning, product development strategies, and general market trends in addition to individual purchasing decisions. Despite their significance, product reviews are difficult to analyze because of their unstructured nature and the linguistic diversity of customers around the world. Every day, millions of reviews are posted on Chinese e-commerce sites like Taobao, JD.com, Tmall, and Pinduoduo. Cross-lingual sentiment analysis is challenging for researchers from around the world because many of these reviews are written in Chinese. In order to get around this obstacle, the current study uses well-known Natural Language Processing (NLP) tools and machine learning techniques to conduct a comparative and standardized analysis of Chinese product reviews that have been translated into English. People still talk about how the BERT-based classifier really stood out among all the models. The results make it clear that BERT hits an impressive 98.10 percent accuracy and F1 score of 0.972. This NLP model outperforms the traditional models.

Keywords: NLP, Cross-Lingual Sentiment Analysis, E Pharmacy Analytics, BERT, Classifier, Deep Learning for Text Classification, Healthcare Data Analytics, Opinion Mining

How to cite this article: Gupta AK, Kumar I. Sentiment Analysis for Patient Satisfaction Assessment in Bilingual E-Pharmacy Platforms. Int J Drug Deliv Technol. 2026;16(63s):1859-1863. DOI: 10.25258/ijddt.16.63s.193

INTRODUCTION

E-commerce's quick-tempered rise has profoundly transformed how consumers shop and how companies' role globally [1]. User-generated information, including reviews, comments, and postings, is abundant on social media platforms and reveals consumer preferences, opinions, and trends. Natural Language Processing (NLP) is essential in this situation. NLP makes it possible for organizations to automatically analyze documented data, categorize content, identify sentiment, spot patterns, and even forecast consumer behavior [2] [6] [10]. Businesses may enhance consumer interaction, make smarter decisions, and modify their offerings to better meet user expectations by using precise sentiment analysis [7].

Sentiment analysis, which examines thoughts and feelings in text, has emerged as a crucial instrument for businesses to track client feedback [8]- [10]. This study uses a more granular methodology, known as feature-level sentiment analysis, to identify the exact terms that carry sentiment and how they connect to various product categories, whereas previous methods mostly focus on overall sentiment (positive, negative, and neutral) [11] - [17]. Nonetheless, there is still a lack of knowledge regarding the precise characteristics or terms of a product that cause gratification or discontent [18] - [25]. The goal of this research paper is to close that gap. It provides a more sophisticated approach to sentiment analysis in bi-lingual reviews by utilizing transformer-based NLP models such as BERT. In order to assess performance, it also contrasts these models with conventional machine learning [26] techniques like Random Forest and SVC follow.

RELATED WORKS

The several studies that examine the major variables affecting customer behavior and satisfaction in e-commerce settings are reviewed in this section. The influence of demographic characteristics, platform-specific features [14], trust-building strategies, and the use of sentiment analysis to comprehend user sentiments are all highlighted in the research.

Consumer Behavior in E-Pharmacy

Studies on the "her economy" reveal that issues including product safety, pricing transparency, deceptive advertising, and after-sales support have an impact on women's financial behaviors [25]. Another factor is age; younger people are typically more receptive to digital transactions and e-commerce [26]. Although expertise with online shopping doesn't always correlate with intention, it does influence perceived risk. Higher incomes are more willing to test new platforms and services. Satisfaction and trust have a favorable impact on customer loyalty, which enhances a business's competitive advantage [27] - [32].

B. External Platform Influences

It's interesting to note that reviews that are neutral or somewhat critical are typically seen as more genuine than those that are extremely enthusiastic, which may come across as biased [25] - [30] . There is a limit, though: evaluations that are too lengthy could overwhelm readers and diminish their perceived worth.

METHODOLOGY

A. Datasets

In this research, we utilized social media–based review data to analyze patient satisfaction and behavioral patterns in bilingual e-pharmacy platforms. These user-generated reviews encompass a wide range of sentiments, including complaints, suggestions, and overall experiences related to online medicine purchasing and healthcare services. These comments might range from complaints to suggestions. Two distinct datasets that were initially written in Chinese were used. We used the Google convert API to convert the text such that it was appropriate for English-based analysis.

A common benchmark in sentiment analysis research, particularly in Chinese-language research, is the first dataset, often known as Dataset. 61,000 items re-views in ten distinct product categories, such as:

1. *Books*
2. *Tablets*
3. *Fruits*
4. *Shampoos*
5. *Mobile Phones*
6. *Water Heaters*
7. *Clothing*
8. *Computers*
9. *Dairy Products (e.g., Mengniu)*
10. *Hotels*.

B. Dataset Analysis

This research employed the Sentiment Intensity Analyzer from the NLTK Python module to determine how consumers feel about various product categories. Com-pared to only utilizing positive or negative labels, this tool offered more precise sentiment scores (ranging from -1 for very negative to +1 for very positive).

Key statistics including the mean, median, and standard deviation of sentiment scores were calculated for each product category as part of the analysis.

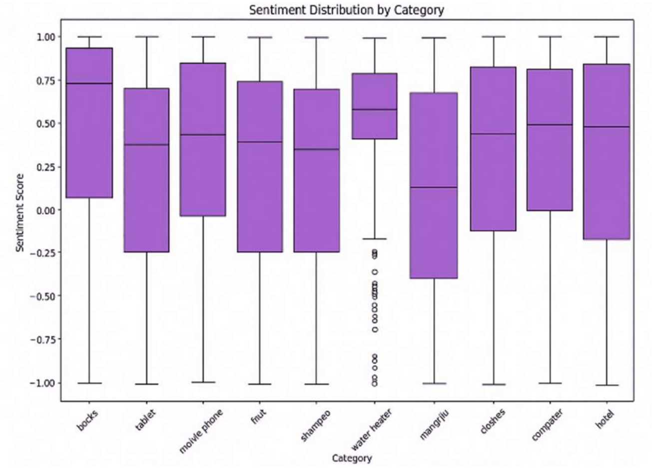

Fig. 1 Category-Based Sentiment Scores of various items.

Fig.1 shows the category- based sentiment analysis of various items like: Books, Tablets, Fruits, Shampoos, Mobile Phones, Water Heaters, Clothing, Computers, Dairy Products (e.g., Mengniu) and Hotels.

Cumulative Sentiment Analysis of All items Reviews.

Histograms with density curves were used to display sentiment distributions across all categories in Fig. 2, providing a general picture of how consumers felt.

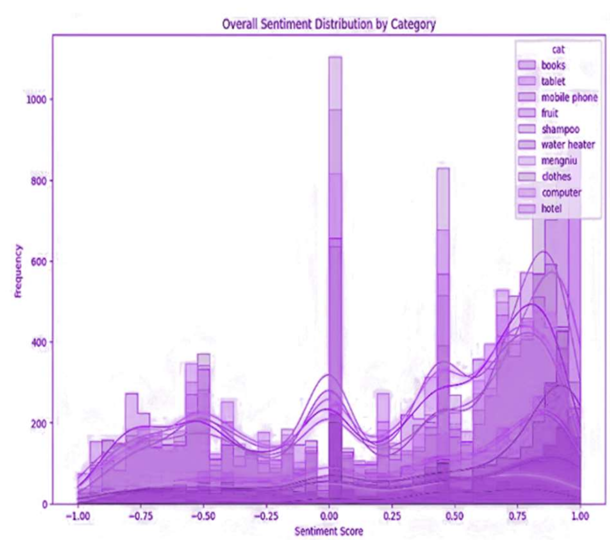

Fig. 2. Overview of Pseudo Label Generation and Question Localization Loss (LQL).

Histograms with density curves were used to display sentiment distributions across all categories in Fig. 2, providing a general picture of how consumers felt.

RESULT AND DISCUSSION

We have carried out two essential assessments to make sure that the datasets and our model predictions matched well. First one that compared the performance of various models. Second oner that used cosine similarity to gauge how closely the translated text matched with the original one.

A. Dataset Translation Accuracy

The main idea of the Chinese reviews is essentially captured by Google Trans-late, albeit occasionally it provides more of an interpretation than a literal translation. Despite these variations, the overall results on Datasets A and B indicate that the English translations mostly maintained the original meaning, with very sporadic semantic drift.

B. Model Performance Analysis

The model that performed the best out of all those evaluated was BERT. With an accuracy of 0.8243 and an F1-score of 0.8358, the Random Forest and XGBoost combination produced respectable results. Additionally, it attained recall of 0.8627 and precision of 0.8760, suggesting a small bias toward avoiding false positives. Although it doesn't match transformer-level performance, its 0.09-hour runtime makes it a quick and dependable choice for time-sensitive applications. Lastly, the baseline model was Linear Regression with SGD. With an accuracy of 0.3320, an F1-score of 0.2771, and a very high loss of 2.4564, it received the lowest score. Despite having the fastest runtime of only 0.0417 hours, it was not appropriate for complicated text classification tasks because to its in capacity to capture the non-linear patterns in the data.

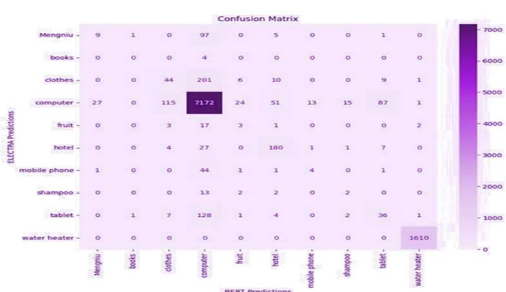

Fig. 3. Confusion matrix with BERT Prediction Vs ELECTRA Prediction.

Fig. 3 presents the confusion matrices obtained from the prediction results of the BERT and ELECTRA models. The figure illustrates the distribution of correctly and incorrectly classified samples for each model, enabling a comparative evaluation of their classification performance. By analyzing true positives, true negatives, false positives, and false negatives, the confusion matrices highlight the strengths and limitations of both models in handling the prediction task. Overall, the visualization helps in understanding how BERT and ELECTRA differ in terms of prediction accuracy and error patterns.

There were noticeably fewer matched predictions for categories like "books," "fruit," "shampoo," and "water heater," indicating that the algorithms had more difficulty with them. Common mismatches were also displayed in the confusion matrix, such as the 201 times in which "clothes" was mistaken for "computer," and the 128 instances in which "tablet" was mistaken for "computer." Semantic overlap in the language used in the evaluations is probably the cause of these kinds of errors, which make it more difficult for the models to discern between comparable product types.

CONCLUSION AND FUTURE WORK

This research paper validates how sentiment analysis, mainly when applied to multi-lingual e-commerce product reviews, can significantly improve our under-standing of consumer behavior and satisfaction. This Paper shows that trans-former-based designs outperform the conventional models by utilizing sophisticated machine learning models like BERT, ELECTRA, and BiLSTM. The trans-former-based designs with an F1 score of 0.9711 and the maximum accuracy up to 98.09%. In order to improve the efficiency of NLP model's capacity to fore-cast and comprehend the customer behavior more precisely. We have included the demographic Information, purchase history, and social media activity to improve the NLP model's capacity.

REFERENCES

1. Q. Guishen, ‘‘Analysis of the impact of e-commerce anchors on user consumption behavior–based on the perspective of gatekeeper the ory,’’ E-Commerce Lett., vol. 13, no. 2, pp. 3733–3738, 2024, doi: 10.12677/ecl.2024.132456.
2. I. O. Adeniyi, A. Akinkunmi, N. A. Sande, A. A. Author, and I. Oluwasegun, ‘‘Social media sentiment analysis: A comprehensive anal ysis,’’ 2024, doi: 10.13140/RG.2.2.31094.37441.
3. K. N. Manasa and M. C. Padma, ‘‘A study on sentiment analysis on social media data,’’ in Emerging Research in Electronics, Computer Science and Technology (Lecture Notes in Electrical Engineering), vol. 545, 2019, pp. 661–667, doi: 10.1007/978-981-13-5802-9_58.
4. K. R. Mabokela, T. Celik, and M. Raborife, ‘‘Multilingual senti ment analysis for under-resourced languages: A systematic review of the landscape,’’ IEEE Access, vol. 11, pp. 15996–16020, 2023, doi: 10.1109/ACCESS.2022.3224136.
5. N. Punetha and G. Jain, ‘‘Game theory and MCDM-based unsupervised sentiment analysis of restaurant reviews,’’

Int. J. Speech Technol., vol. 53, no. 17, pp. 20152–20173, Mar. 2023, doi: 10.1007/s10489-02304471-1.

6. N. M. Alharbi, N. S. Alghamdi, E. H. Alkhammash, and J. F. Al Amri, ''Evaluation of sentiment analysis via word embedding and RNN variants for Amazon online reviews,'' Math. Problems Eng., vol. 2021, pp. 1–10, May 2021, doi: 10.1155/2021/5536560.
7. T. Chen, P. Samaranayake, X. Cen, M. Qi, and Y.-C. Lan, ''The impact of online reviews on consumers' purchasing decisions: Evidence from an eyetracking study,'' Frontiers Psychol., vol. 13, pp. 661–667, Jun. 2022, doi: 10.3389/fpsyg.2022.865702.
8. R. Liu and J. Xiao, ''Factors affecting users' satisfaction with urban parks through online comments data: Evidence from shenzhen, China,'' Int. J. Environ. Res. Public Health, vol. 18, no. 1, p. 253, Dec. 2020, doi: 10.3390/ijerph18010253.
9. X. Tao and W. Hu, ''The influence of characteristics of e-commerce anchors information source on college students' purchase intention,'' Int. J. Bus. Manage., vol. 19, no. 5, p. 47, Aug. 2024, doi: 10.5539/ijbm.v19n5p47.
10. Gupta, Anand Kumar, Srinivasulu, Asadi, Hiran, Kamal Kant, Sreenivasulu, Goddindla, Rajeyyagari, Sivaram, Subramanyam, Madhusudhana, Prediction of Omicron Virus Using Combined Extended Convolutional and Recurrent Neural Networks Technique on CT-Scan Images, Interdisciplinary Perspectives on Infectious Diseases, 2022, 1525615, 11 pages, 2022. https://doi.org/10.1155/2022/1525615.
11. Shankar R., Kumar I., Mishra R.K. (2019). Outage probability analysis of MIMO-OSTBC relaying network over Nakagami-m fading channel conditions, Traitement du Signal, Vol. 36, No. 1, pp. 59-64. https://doi.org/10.18280/ts.360108
12. Kumar, I., Mishra, M.K., Mishra, R.K. (2021). Performance analysis of NOMA downlink for next-generation 5G network with statistical channel state information. Ingénierie des Systèmes d'Information, Vol. 26, No. 4, pp. 417-423. https://doi.org/10.18280/isi.260410
13. Shankar, R., Kumar, I., Mishra, R.K. (2019). Pairwise error probability analysis of dual hop relaying network over time selective Nakagami-m fading channel with imperfect CSI and node mobility. Traitement du Signal, Vol. 36, No. 3, pp. 281-295. https://doi.org/10.18280/ts.360312
14. Gowda, V. Dankan, Sharma, Avinash, Kumaraswamy, S., Sarma, Parismita , Hussain, Naziya, Dixit, Santosh Kumar & Gupta, Anand Kumar(2023) A novel approach of unsupervised feature selection using iterative shrinking and expansion algorithm, Journal of Interdisciplinary Mathematics, 26:3, 519-530, DOI: 10.47974/JIM-1678
15. Kumar I, Kumar A, Kumar Mishra R. Performance analysis of cooperative NOMA system for defense application with relay selection in a hostile environment. The Journal of Defense Modeling and Simulation. 2022;0(0). doi:10.1177/15485129221079721
16. Ashish, I. Kumar and R. K. Mishra, "Performance Analysis For Wireless Non-Orthogonal Multiple Access Downlink Systems," 2020 International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET), Patna, India, 2020, pp. 1-6, doi: 10.1109/ICEFEET49149.2020.9186987.
17. Kumar, I., Mishra, R.K. (2021). An investigation of spectral efficiency in linear MRC and MMSE detectors with perfect and imperfect CSI for massive MIMO systems. Traitement du Signal, Vol. 38, No. 2, pp. 495-501. https://doi.org/10.18280/ts.380229
18. Kumar, I., Mishra, R.K. (2020). An efficient ICI mitigation technique for MIMO-OFDM system in time-varying channels. Mathematical Modelling of Engineering Problems, Vol. 7, No. 1, pp. 79-86. https://doi.org/10.18280/mmep.070110
19. Kumar I., Sachan V., Shankar R., Mishra R.K. (2018). An investigation of wireless S-DF hybrid satellite terrestrial relaying network over time selective fading channel, Traitement du Signal, Vol. 35, No. 2, pp. 103-120. https://doi.org/10.3166/TS.35.103-120
20. A. K. Gupta, A. Sharma, A. Srinivasulu, T. Barua, S. Rajeyyagari and M. Subramanyam, "Early Prediction of Breast Cancer through Deep RNN Approach," 2022 International Conference on Trends in Quantum Computing and Emerging Business Technologies (TQCEBT), Pune, India, 2022, pp. 1-4, doi: 10.1109/TQCEBT54229.2022.10041634.
21. Sachan, V., Kumar, I., Shankar, R., Mishra, R.K. (2018). Analysis of transmit antenna selection based selective decode forward cooperative communication protocol. Traitement du Signal, Vol. 35, No. 1, pp. 47-60. https://doi.org/10.3166/TS.35.47-60
22. Indrajeet Kumar, Vikash Sachan, Ravi Shankar, Ritesh Kumar Mishra, Performance Analysis of Multi-User Massive MIMO Systems with Perfect and Imperfect CSI, Procedia Computer Science, Volume 167, 2020, Pages 1452-1461, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.356.
23. Vikash Sachan, Indrajeet Kumar, Lokesh Bhardwaj, Ritesh Kumar Mishra, Pairwise Error Probability Analysis of SM-MIMO system employing k – μ Fading Channel, Procedia Computer Science, Volume 167, 2020, Pages 2516-2523, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.304.
24. R. Shankar, I. Kumar, A. Kumari, K. N. Pandey and R. K. Mishra, "Pairwise error probability analysis and optimal power allocation for selective decode-forward protocol over Nakagami-m fading channels," 2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET), Chennai, India, 2017, pp. 1-6, doi: 10.1109/ICAMMAET.2017.8186700.

25. Gupta, Anand Kumar, Srinivasulu, Asadi, Oyerinde, Olutayo Oyeyemi, Pau, Giovanni, Ravikumar, C. V., COVID-19 Data Analytics Using Extended Convolutional Technique, Interdisciplinary Perspectives on Infectious Diseases, 2022, 4578838, 10 pages, 2022. https://doi.org/10.1155/2022/4578838.
26. Gupta, Y., Verma, R., Sharma, S.S.P.M.B., Kumar, I. (2024). An IoT Application Based Decentralized Electronic Voting System Using Blockchain. In: Pareek, P., Gupta, N., Reis, M.J.C.S. (eds) Cognitive Computing and Cyber Physical Systems. IC4S 2023. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 537. Springer, Cham. https://doi.org/10.1007/978-3-031-48891-7_12
27. Mishra, N., Raghuwanshi, R., Maurya, N.K., Kumar, I. (2024). Efficient Fuel Delivery at Your Fingertips: Developing a Seamless On-Demand Fuel Delivery App with Flutter. In: Pareek, P., Gupta, N., Reis, M.J.C.S. (eds) Cognitive Computing and Cyber Physical Systems. IC4S 2023. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 537. Springer, Cham. https://doi.org/10.1007/978-3-031-48891-7_11
28. Trivedi, D., Saxena, M., Sharma B, S.S.P.M., Kumar, I. (2024). Harmonizing Insights: Python-Based Data Analysis of Spotify's Musical Tapestry. In: Pareek, P., Gupta, N., Reis, M.J.C.S. (eds) Cognitive Computing and Cyber Physical Systems. IC4S 2023. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 536. Springer, Cham. https://doi.org/10.1007/978-3-031-48888-7_3.
29. S. Kim, S. Ham, H. Moon, B.-L. Chua, and H. Han, ‘‘Experience, brand prestige, perceived value (functional, hedonic, social, and financial), and loyalty among GROCERANT customers,’’ Int. J. Hospitality Manage., vol. 77, pp. 169–177, Jan. 2019, doi: 10.1016/j.ijhm.2018.06.026.
30. R. Shankar, I. Kumar, M. Kashyap, A. K. Jha, and B. P. Chaudhary, “A Review on NOMA scheme for emerging 6G wireless networks: State of the Art, Key Schemes, Future scope and Security Issues,” Radioelectronics and Communications Systems, vol. 68, no. 5, pp. 271–284, 2025. DOI: 10.3103/S0735272725010017.
31. Q. Zhang and M. Abisado, ‘‘A novel context-aware deep learn ing algorithm for enhanced movie recommendation systems,’’ Math. Model. Eng. Problems, vol. 10, no. 6, pp. 2031–2038, Dec. 2023, doi: 10.18280/mmep.100613.
32. Anand Kumar Gupta, Asadi Srinivasulu, Kamal Kant Hiran, Tarkeswar Barua, Goddindla Sreenivasulu, Sivaram Rajeyyagari and Madhusudhana Subramanyam. Early prediction and analysis of mammary glands cancer through deep learning approaches. World Journal of Advanced Engineering Technology and Sciences, 2022, 06(01), 018–024. Article DOI: https://doi.org/10.30574/wjaets.2022.6.1.0056.