2026 Volume: 02 Issue: 01 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-09

SENTIMENT ANALYSIS AND ITS APPLICATION IN DRUG RECOMMENDATION

***Dev Krishan, Dr. Vishal Shrivastava, Dr. Akhil Panday**

Computer Science & Engineering, Arya College of Engineering & I.T. Jaipur, India.

Received: 26 November 2025
Article Revised: 16 December 2025
Published on: 06 January 2026

***Corresponding Author: Dev Krishan**
Computer Science & Engineering, Arya College of Engineering & I.T. Jaipur, India. DOI: https://doi-doi.org/101555/ijrpa.4937

ABSTRACT

The COVID-19 pandemic exposed major gaps in healthcare systems, such as the shortage of doctors, nurses, medical equipment, and essential medicines. Because of these limitations, many people started taking medicines on their own without proper medical advice, which often made their health conditions worse. With the growth of technology and data-driven solutions, there is a growing need for systems that can assist both patients and healthcare professionals in making better treatment decisions. This work focuses on building a drug recommendation system that can help reduce the workload of medical specialists by suggesting appropriate medicines based on patient feedback. Using sentiment analysis, the system examines user reviews of various drugs to understand public opinion about their effectiveness. Machine learning algorithms are then applied to predict the overall sentiment and recommend the most suitable drug for a specific disease or condition.

KEYWORDS: Drug Recommendation, Sentiment Analysis, Drug Reviews, Machine Learning.

INTRODUCTION

In many situations, people start taking medicines on their own without consulting a medical professional, especially when proper healthcare services are not easily available. However, this self-medication often leads to complications and worsens the health condition. The main aim of this study is to develop a drug recommendation system that can support both patients and healthcare providers by reducing the burden on medical specialists.

During situations like the COVID-19 pandemic, the number of patients increased rapidly, while the availability of qualified doctors, particularly in remote and rural regions, remained very limited. Since becoming a medical specialist requires years of education and training, it is not possible to increase the number of doctors immediately. Therefore, telemedicine and intelligent healthcare assistance systems have become essential in dealing with such circumstances.

Medical prescription errors have also become a growing concern worldwide. Every year, thousands of people suffer due to incorrect prescriptions or dosage mistakes. One of the main reasons is that doctors often rely on their previous knowledge and experience, which may not always cover the continuously growing list of new medicines, treatments, and research findings. New drugs and diagnostic information are being introduced frequently, making it difficult for healthcare workers to stay updated all the time.

As the use of the internet continues to grow, many individuals now search online for health-related information before visiting a doctor. Survey reports, such as one from the Pew Research Center, indicate that a large percentage of adults look up symptoms or medical advice online. Because of this trend, there is a strong need for reliable systems that can provide guidance based on accurate medical knowledge rather than random online opinions.

RELATEDWORKS

The authors in study [9] introduced GalenOWL, a semantic-based online platform designed to support medical professionals in identifying appropriate drugs. The system provides recommendations by considering the patient’s disease, allergies, and possible drug interactions. To achieve this, clinical data are converted into standard medical ontologies such as ICD-10 and UNII before being integrated with patient information.

In another study, Leilei Sun [10] analyzed large-scale treatment records to determine the most suitable treatment plans for patients. The research made use of a semantic clustering algorithm to measure similarity between treatment histories. A recommendation framework was then developed to evaluate the effectiveness of the suggested treatments. Electronic Medical Records (EMR) collected from several hospitals were used for testing, and the results indicated an improvement in patient recovery rates.

In study [11], sentiment analysis was performed on multilingual medical reviews using two approaches—Naive Bayes and Recurrent Neural Networks (RNN). A translation API was used to convert non-English text to English before processing. The findings showed that

RNN achieved a higher accuracy rate of 95.34%, outperforming Naive Bayes, which achieved 77.21%.

The researchers in study [12] emphasized that drug recommendations should align with the patient's health condition and immunity level. More than sixty risk factors, such as hypertension and alcohol dependency, were considered to classify patients into different risk categories. A decision-support web application was also developed to help doctors choose the most suitable initial medications for each patient.

Xiaohong Jiang et al. [13] compared three machine learning models—Decision Tree, Support Vector Machine (SVM), and Backpropagation Neural Network—on treatment datasets. The SVM model showed the best performance in terms of accuracy, efficiency, and scalability, and was therefore selected for the recommendation component. An additional error-checking mechanism was included to ensure precise and reliable recommendations.

Mohammad Mehedi Hassan et al. [14] proposed a cloud-based drug recommendation system called CADRE. Initially, the system used collaborative filtering by grouping drugs based on their functional characteristics. However, issues such as high computation costs and data sparsity led the researchers to shift toward a cloud-assisted model using tensor decomposition, which improved the overall recommendation performance.

Lastly, Jiugang Li et al. [15] developed a hashtag recommendation model using the skip-gram method to capture contextual word meanings. Convolutional Neural Networks (CNN) were applied to create semantic sentence vectors, and Long Short-Term Memory (LSTM) networks were used for final classification. Their model showed better performance compared to traditional approaches like SVM and standard RNN, demonstrating the importance of semantic feature preservation in predictive tasks.

PROPOSED SYSTEM ARCHITECTURE

A recommender system is generally designed to suggest an item to a user based on their preferences and requirements. In the context of drug recommendation, the system analyzes patient reviews to understand experiences shared by users who have previously taken the medication. These reviews are processed to identify whether the opinions expressed are positive or negative, which helps in recommending the most suitable drug for a particular health condition.

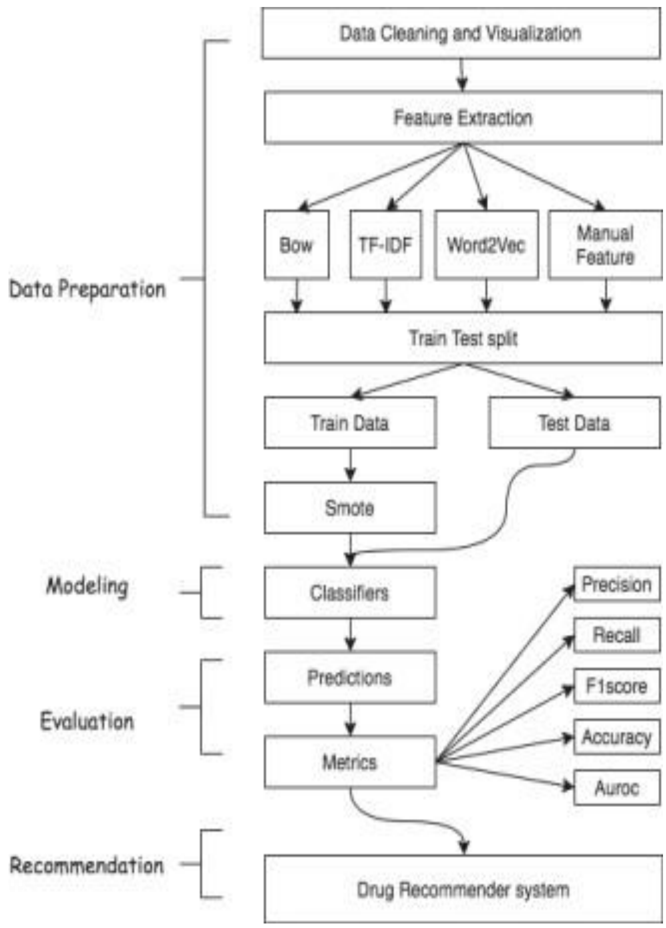

Fig.1 Proposed system process flow.

RESULTS AND DISCUSSION

In this study, each review in the dataset was labeled as either positive or negative based on the user's star rating. Reviews with ratings above five were considered positive, whereas those rated between one and five were treated as negative. Initially, the dataset contained 111,583 positive reviews and 47,522 negative reviews. To handle the imbalance between these two classes, the SMOTE oversampling technique was applied, increasing the number of negative samples to approximately 70% of the positive class. As a result, the updated dataset consisted of 111,583 positive reviews and 78,108 negative reviews.

To perform sentiment classification, four text representation techniques were used: Bag of Words (BoW), TF-IDF, Word2Vec, and manually designed features. These representations were then evaluated using ten different machine learning algorithms. After analyzing the individual model performances, the best-performing models from each representation method were selected: Perceptron for BoW, LinearSVC for TF-IDF, LGBM for Word2Vec, and Random Forest for manual features. Their prediction outputs were combined to create a final aggregated prediction score.

The main goal behind combining the results was to ensure reliability. If a drug was identified as effective by all models, it received a higher overall score. Conversely, if one model predicted it negatively, the score for that drug would decrease. To further refine the score, the combined prediction was multiplied by a normalized “useful count” measure, which reflects how many users found a review helpful. This step ensured that drugs with more credible and widely acknowledged reviews ranked higher. The final score for each drug was normalized on a per-condition basis to avoid score inflation for conditions with more review volume.

Although the results from all four approaches were satisfactory, the system still requires improvement before being adopted in real-world medical environments. The performance gap between positive and negative classes suggests that further balancing techniques such as ADASYN or SMOTETomek may yield better outcomes. Additionally, fine-tuning model hyperparameters may help improve classification precision and robustness.

The recommendation step in this study relied on a simple combination of the best individual models. A more advanced ensemble method could enhance predictive reliability and explanation clarity. Figure 2 illustrates the top recommended drugs for five health conditions: Acne, Birth Control, High Blood Pressure, Pain, and Depression.

Further analysis of the useful count distribution (as shown in Figure 3) revealed a large variance, with values ranging from very low to over 1300. This variance arises because popular drugs naturally accumulate more review interactions, regardless of review sentiment. To address this, the useful count was normalized within each condition before being used in the recommendation scoring process.

Several machine learning algorithms were evaluated for sentiment prediction. Logistic Regression, Multinomial Naive Bayes, Perceptron, LinearSVC, Ridge Classifier, and Stochastic Gradient Descent were used on the sparse BoW and TF-IDF representations due to their efficiency with high-dimensional data. For Word2Vec and manually engineered features, tree-based models such as Decision Tree, Random Forest, LGBM, and CatBoost were implemented. Given the dataset size of approximately 210,000 reviews, computational efficiency was an important consideration, and models were selected accordingly to minimize training time while maintaining strong predictive performance.

condition	drugName	Score
Acne	Retin-A	0.069334
Acne	Atralin	0.088545
Acne	Magnesium hydroxide	0.088545
Acne	Retin A Micro	0.097399
Birth Control	Mono-Linyah	0.005448
Birth Control	Gildess Fe 1.5 / 30	0.005987
Birth Control	Ortho Micronor	0.006149
Birth Control	Lybrel	0.027766
High Blood Pressure	Adalat CC	0.303191
High Blood Pressure	Zestril	0.305851
High Blood Pressure	Toprol-XL	0.362589
High Blood Pressure	Labetalol	0.367021
Pain	Neurontin	0.158466
Pain	Nortriptyline	0.171771
Pain	Pamelor	0.231829
Pain	Elavil	0.304513
Depression	Remeron	0.124601
Depression	Sinequan	0.146486
Depression	Provigil	0.240185
Depression	Methylin ER	0.328604

Fig.2 Recommendation of top four drugs on top five conditions.

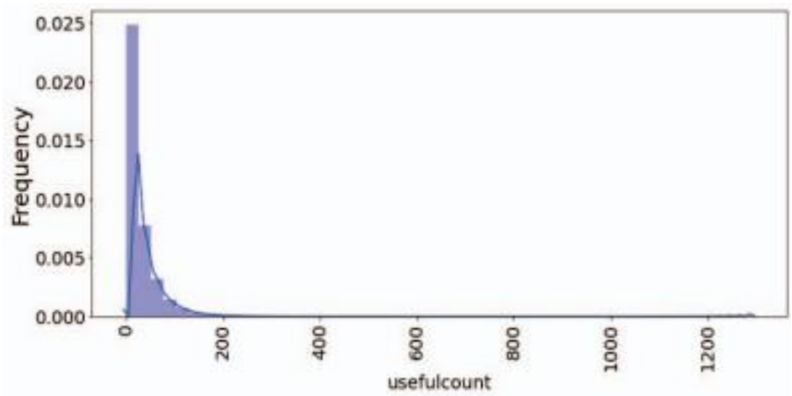

Fig.3 Distribution of useful count.

FUTURE SCOPE AND CONCLUSION

Reviews have become an essential part of decision-making in everyday life. Whether purchasing products online, choosing a doctor, or selecting a restaurant, people rely heavily on the experiences shared by others. Inspired by this behavior, the present work focused on analyzing sentiment in drug reviews to develop a drug recommendation system. Various machine learning models were implemented for sentiment classification. Logistic Regression, Perceptron, Multinomial Naive Bayes, Ridge Classifier, Stochastic Gradient Descent, and LinearSVC were applied to Bag of Words and TF-IDF representations, while Decision Tree, Random Forest, LGBM, and CatBoost were used with Word2Vec and manually engineered features.

REFERENCES

1. Wittich CM, Burkle CM, Lanier WL. Medication errors: an overview for clinicians. Mayo Clin Proc. 2014 Aug;89(8):1116-25.
   Telemedicine, https://www.mohfw.gov.in/pdf/Telemedicine.pdf
2. CHEN, M. R., & WANG, H. F. (2013). Thereason and prevention of hospital medication errors. Practical Journal of Clinical Medicine, 4.
3. Drug Review Dataset, https://archive.ics.uci.edu/ml/datasets/Drug% 2BReview%2BDataset%2B%2528Drugs.com% 2529#
4. Fox, Susannah, and Maeve Duggan. ”Health online 2013. 2013.” URL: http://pewinternet.org/Reports/2013/Health- online.aspx
5. Bartlett JG, Dowell SF, Mandell LA, File TM Jr, Musher DM, Fine MJ. Practice guidelines for the management of community-acquired pneumonia in adults. Infectious Diseases Society of America. Clin Infect Dis. 2000 Aug;31(2):347- 82. doi: 10.1086/313954. Epub 2000 Sep 7. PMID: 10987697; PMCID: PMC7109923.
6. Fox, Susannah & Duggan, Maeve. (2012). Health Online 2013. Pew Research Internet Project Report.
7. T. N. Tekade and M. Emmanuel, ”Probabilistic aspect mining approach for interpretation and evaluation of drug reviews,” 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, 2016, pp. 1471-1476, doi: 10.1109/SCOPES.2016.7955684.
8. Doulaverakis, C., Nikolaidis, G., Kleontas, A. et al. GalenOWL: Ontology-based drugrecommendations discovery. J Biomed Semant 3, 14 (2012). https://doi.org/10.1186/2041-1480-3-14
9. Leilei Sun, Chuanren Liu, Chonghui Guo, Hui Xiong, and Yanming Xie. 2016. Data-driven Automatic Treatment Regimen Development and Recommendation. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’16). Association for Computing Machinery, New York, NY, USA, 1865–1874. DOI:https://doi.org/10.1145/2939672.2939866
10. V. Goel, A. K. Gupta and N. Kumar, ”Sentiment Analysis of Multilingual Twitter Data using Natural Language Processing,” 2018 8th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2018, pp. 208-212, doi: 10.1109/CSNT.2018.8820254.
11. Shimada K, Takada H, Mitsuyama S, et al. Drug-recommendation system for patients with infectious diseases. AMIA Annu Symp Proc. 2005;2005:1112.

12. Y. Bao and X. Jiang, ”An intelligent medicine recommender system framework,” 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, 2016, pp. 1383-1388, doi: 10.1109/ICIEA.2016.7603801.
13. Zhang, Yin & Zhang, Dafang & Hassan, Mohammad & Alamri, Atif & Peng, Limei. (2014). CADRE: Cloud-Assisted Drug REcommendation Service for Online Pharmacies. Mobile Networks and Applications. 20. 348-355. 10.1007/s11036-014-0537-4.
14. J. Li, H. Xu, X. He, J. Deng and X. Sun, ”Tweet modeling with LSTM recurrent neural networks for hashtag recommendation,” 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 1570-1577, doi: 10.1109/IJCNN.2016.7727385.
15. Zhang, Yin & Jin, Rong & Zhou, Zhi-Hua. (2010). Understanding bag-of-words model: A statistical framework. International Journal of Machine Learning and Cybernetics. 1. 43-52. 10.1007/s13042-010-0001- 0.
16. J. Ramos et al., “Using tf-idf to determine word relevance in document queries,” in Proceedingsof the first instructional conference onmachinelearning, vol. 242, pp. 133–142, Piscataway, NJ, 2003.
17. Yoav Goldberg and Omer Levy. word2vec Explained: deriving Mikolov et al.’s negative- sampling word-embedding method, 2014; arXiv:1402.3722.
18. Danushka Bollegala, Takanori Maehara and Kenichi Kawarabayashi. Unsupervised Cross- Domain Word Representation Learning, 2015; arXiv:1505.07184. Textblob,https://textblob.readthedocs.io/en/dev/.
19. Van der Maaten, Laurens & Hinton, Geoffrey. (2008). Viualizing data using t-SNE. Journal of Machine Learning Research. 9. 2579-2605.
20. N. V. Chawla, K. W. Bowyer, L. O. Hall and W.P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique, 2011, Journal Of Artificial Intelligence Research, Volume 16, pages 321-357, 2002; arXiv:1106.1813. DOI: 10.1613/jair.953.
21. Powers, David & Ailab,. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. 2. 2229-3981. 10.9735/2229-3981
22. Haibo He, Yang Bai, E. A. Garcia and Shutao Li, ”ADASYN: Adaptive synthetic sampling approach for imbalanced learning,” 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, 2008, pp. 1322-1328, doi: 10.1109/IJCNN.2008.4633969.

23. Z. Wang, C. Wu, K. Zheng, X. Niu and X. Wang, ”SMOTETomekBased Resampling for Personality Recognition,” in IEEE Access, vol. 7, pp. 129678-129689, 2019, doi: 10.1109/ACCESS.2019.2940061.