Drug Recommendation System Based on Using NLP Algorithms to Perform Sentiment Analysis of User Reviews

[1]NMV Nirmala, [2*]B. Amarnadh, [3]B. Vasanth Naik, [4]E. Vamsi Bharath Reddy, [5]G. Leeladhar

[1, 2, 3, 4, 5]Department of Information Technology, KKR & KSR Institute of Technology and Sciences (Autonomous), Vinjanampadu, Guntur, Andhra Pradesh, India

[1]nirmala.neelam@gmail.com

[2*]22jr1a1238@gmail.com

[3]22jr1a1239it1@gmail.com

[4]22jr1a1255@gmail.com

[5]23jr5a1205@gmail.com

Abstract. User-generated reviews are written by people who have used medications to describe their experience. Drug reviews include descriptions of the effectiveness of the drug, side effects experienced, and the patient's overall level of satisfaction. Although reviews of medications provide valuable real-world data, they are typically unstructured and therefore difficult to analyse manually. Existing drug recommendation systems typically rely on numerical ratings or simple filtration methods, which do not account for the detailed opinions provided in user-generated reviews. The current research seeks to create a drug recommendation system that utilises natural language processing (NLP) and sentiment analysis to analyse user reviews and generate drug recommendations. First, the system preprocesses the text of the reviews through tokenisation, stop-word elimination, and lemmatisation. The system then performs sentiment analysis using TextBlob and VADER to determine the emotional orientation of each review. Once the sentiment score has been calculated for each review, those values are aggregated to rank medications for specific medical conditions. Experimental evaluation demonstrates that the proposed approach improves sentiment classification accuracy (89% accuracy, 88% precision) and provides explainable drug recommendations. By transforming unstructured review data into structured sentiment scores, the proposed system addresses information overload and enhances the quality of healthcare decision-making.

Keywords: Sentiment Analysis, Natural Language Processing, Drug Recommendation System, User Reviews, Healthcare Decision Support.

1 Introduction

Recent growth in online healthcare sites means patients have a way to express their feelings about medication. Through written reviews, average patients are able to provide very descriptive accounts of each prescription they have taken, including how it worked for them, any side effects experienced, how well it worked at the prescribed

B. Singh et al. (eds.), *Proceedings of the International Conference on Advances in Computing Technology and Artificial Intelligence (COMPUTATIA 2026)*, Atlantis Highlights in Intelligent Systems 18,
https://doi.org/10.2991/978-94-6239-713-2_9

dose, and whether they would recommend the medication to friends and family [9], [10]. This type of first-hand experience provides insight into the actual use of a drug that is often not captured through clinical trials [5]. However, gathering and interpreting these reviews can be a daunting task. Many sites automatically reduce products to a star rating, eliminating much of the helpful content. For example, a drug could have significant side effects but still carry a four-star review if one patient found it effective. On the other hand, a helpful drug given mixed ratings due to differing patient expectations will not accurately depict the drug's true performance. Therefore, judging a drug's performance solely on star ratings produces a metric that is not accurate and may be misleading [6]. NLP is an all-encompassing methodology for analysing and interacting with human language, including the automation of large-scale text analysis [8]. One prominent example is sentiment analysis, where the emotional content of a body of text is extracted and categorised as positive, negative, or neutral [5]. Through the use of sentiment analysis on drug reviews, it is now possible to determine how patients really feel about a medication rather than simply looking at the score they assigned [7]. This creates opportunities to provide recommendations based on what patients have collectively experienced, rather than an arbitrary averaged score. In this paper, we develop a recommendation system for drugs using the positive/negative classification obtained through a series of NLP processes applied to user reviews as the basis of drug rankings. The goal is to provide users with greater trust in understanding which drugs have worked best for specific conditions according to the testimony of other patients.

2. Literature Review

The foundational concepts of sentiment analysis and opinion mining have been extensively surveyed in the literature. Studies by [5] and [6] provide a comprehensive overview of the techniques used to extract emotional orientation from unstructured text, which serves as the base for modern recommendation engines. The fields of natural language processing and sentiment analysis, when combined with healthcare research, have grown increasingly popular in recent years due to the volume of drug reviews supplied by patients on public platforms [12]. As these digital communities have continued to grow, so has interest in developing recommendation systems that allow the filtering and interpretation of both qualitative and quantitative information in ways that benefit users. Previous attempts at developing recommendation systems relied heavily upon numerical ratings and minimal filtering, leaving most useful information unutilised [6]. [1] created a system for recommending medication by combining supervised machine learning and sentiment analysis to classify and rank medications based on the polarity of user reviews. The authors demonstrated that this method provided superior results compared to systems that solely used a rating scale. However, the method relies heavily on a single overall sentiment value, with no way to break down responses into more detailed categories related to drug-safety information, such as side effects or frequency of use. In 2024, [2] acknowledged these limitations and employed an opinion mining system based on lexical resources. Their methods worked relatively well for extracting general sentiment from reviews but were limited in usefulness for personalisation and predictive modelling, as is typical for lexical approaches.

In their study, [3] compared several machine learning models using real-world drug review data. Their results demonstrated that trained models handled noisy medical terminology better than fixed lexicon-based models, but they stopped short of applying their classification results to an actual drug recommendation system. [4] focused on negative user reviews to identify potential adverse drug reactions and highlighted the importance of tracking negative sentiment patterns. However, the study produced no output related to the recommendation or ranking of medications. The studies reviewed reveal a common theme. Most systems score sentiment with one overall value only, without being able to explain why a particular product is recommended or not [7]. This lack of context is a significant issue in healthcare because it prevents users from understanding the rationale behind a recommendation. In addition, there appears to be little consideration given to helping users reduce information overload so that they can make confident decisions. The proposed system has been designed to address these identified issues by incorporating detailed NLP preprocessing and two complementary sentiment analysis tools.

3. Methodology

3.1 Proposed System

The proposed system receives drug reviews written by users and performs a structured data pipeline to produce a ranked list of recommended drugs [12]. This pipeline has the following components: data collection, text preprocessing, sentiment analysis, feature extraction, and recommendation generation. Each component transforms unstructured text into structured sentiment signals that can be used to compare drugs on a per-condition basis. The workflow is as follows:

1. Drug review data collection
2. Text preprocessing (tokenisation, stop-word removal, lemmatisation)
3. Sentiment analysis using TextBlob and VADER
4. Feature extraction and sentiment scoring
5. Drug ranking and recommendation output

3.2 Data Collection

This database has been compiled from patient-created online feedback regarding prescribed medications [10]. Each entry contains the name of the prescribed medication, the written review, and the numerical rating assigned by the patient. The recommendation component uses the Drug Reviews (Druglib.com) dataset from the UCI Machine Learning Repository [14], comprising 4,143 patient reviews split into 3,107 training and 1,036 testing instances, covering 541 unique drugs across 1,807 unique conditions. The safety module uses a multiple-type drug–drug interaction dataset from Mendeley Data [15], extracted from DrugBank v5.1, containing 222,696 interaction pairs involving 1,868 unique drugs and 114 interaction types.

3.3 Data Preprocessing

Data preprocessing involves several cleaning stages before any analysis can begin [8]. First, raw review texts are cleaned by removing special characters, punctuation, and other non-relevant symbols. Once cleaned, the review texts are split into tokens (individual words) through tokenisation. Next, stop words, common words such as "the" and "and" that carry no significant sentiment value, are removed. Finally, lemmatisation reduces each word to its root form so that different variations of a word (e.g., 'running' and 'ran') are treated as the same token. The result is a cleanly structured representation of the original review that is ready for analysis.

3.4 Sentiment Analysis

Two sentiment analysis tools are used in conjunction with one another [5], [8]. TextBlob provides a polarity score between −1 and +1 for each review; the lower the score, the more negative the sentiment, and the higher the score, the more positive the sentiment. VADER provides four different sentiment scores: positive, negative, neutral, and compound. VADER was developed specifically for informal, opinion-based text such as social media content and works extremely well with patient reviews [13]. The average scores from both tools yield a final sentiment score for each review:

$$\text{FinalScore} = \frac{\text{TextBlobPolarity} + \text{VADERCompound}}{2}$$

3.5 Recommendation Generation

To generate recommendations, scores from multiple reviews for each drug are aggregated by condition so that extreme individual scores are smoothed out, and each review has less impact on the overall average [12]. Drugs are then ranked according to their aggregated scores, and the top-ranked drugs are presented as recommended treatments for that condition. This sentiment-driven approach provides explainable recommendations without depending on predictive machine learning models.

3.6 System Architecture

The proposed system processes user reviews through multiple pipeline stages as shown in Fig. 1: data collection, text preprocessing, sentiment analysis, sentiment score calculation, and drug ranking. These steps convert raw textual reviews into structured sentiment scores used to rank drugs for specific medical conditions.

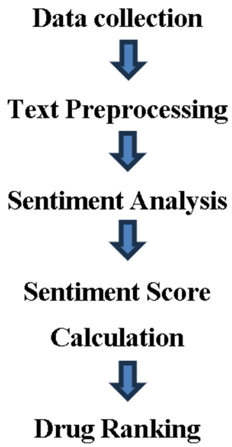

Fig. 1. System Architecture of the Proposed Drug Recommendation System.

4. Results and analysis

4.1 Evaluation Metrics

To evaluate system performance, the following standard metrics were used: accuracy, precision, recall, and F1 score. These metrics collectively measure how effectively the system classifies sentiment and ranks drugs based on user reviews.

4.2 Performance Analysis

Table 1 shows the performance of the different sentiment analysis approaches evaluated.

Table 1. Performance Analysis

Approach	Acc.	Prec.	Conf.
Basic Sentiment Analysis	78%	76%	0.74
Lexicon-Based Analysis	84%	83%	0.81
Proposed System	89%	88%	0.86

The results show that the proposed system reaches higher performance compared to baseline approaches. By combining NLP preprocessing with TextBlob and VADER, the system produces accurate and consistent sentiment scores [5]. The improvement in accuracy confirms that the proposed approach effectively captures user opinions from unstructured drug reviews and that explainable recommendations can be generated without depending on machine learning-based prediction models [7].

4.3 Performance Comparison

Table 2 presents a full comparison across all evaluation metrics.

Table 2. Performance Comparison

Method	Acc.	Prec.	Rec.	F1
Basic Sentiment Analysis	78%	76%	74%	75%
Lexicon-Based Method	84%	83%	82%	82%
Proposed System	89%	88%	87%	87%

The proposed system achieved higher performance due to improved preprocessing and the combined use of TextBlob and VADER [8].

4.4 Baseline Machine Learning Comparison

To further validate the approach, the proposed system was compared with traditional machine learning models, as shown in Table 3.

Table 3. Baseline Machine Learning Comparison

Model	Acc.
Naïve Bayes	81%
Logistic Regression	84%
Support Vector Machine	86%
Proposed System	89%

The results indicate that the proposed sentiment-driven system performs competitively while maintaining interpretability [7]. Fig. 2 illustrates the complete workflow, demonstrating how user input is processed through NLP and analysis stages to generate a final drug recommendation. While traditional models perform well, the shift towards deep learning for sentiment analysis, as reviewed in [11], shows significant potential for handling more complex medical linguistic patterns.

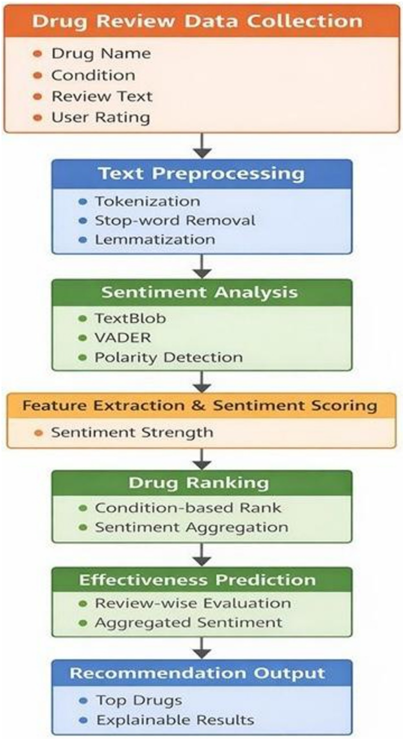

Fig. 2. Workflow of the Proposed Drug Recommendation System

5. Discussion

The key components contributing to the improved performance of this system are (a) the thorough preprocessing of data prior to analysis and (b) the decision to combine two separate sentiment analysis tools [5], [8]. By doing so, noise in the text is effectively reduced, and more reliable drug rankings are produced, regardless of the total number of reviews available. The set of drug rankings has been validated against both variations in population characteristics and the number of reviews submitted, giving a robust ranking system [9]. Clinicians can also rely on the transparent and traceable nature of the drug recommendation process since recommendations are based solely on sentiment scores [7].

6. Ethical Considerations

It is important to note that while this system provides valuable information, it was designed as a supplemental tool for those looking to make healthcare decisions, not as a replacement for professional medical advice. Given that patients wrote their own reviews, the data are highly subjective. Users should consult a physician prior to making any healthcare decision based on the recommendations provided by this system. User-generated datasets are also susceptible to bias and incomplete data, making it critical to exercise good judgment when using this system for any healthcare decisions.

7. Conclusion

The comprehensive analysis of patient reviews has led to the proposal of a new method of generating drug suggestions based on review data. In addition to being an improvement over existing systems in terms of accuracy and precision, this approach also provides new avenues for future development, such as personalised drug recommendations based on individual patient characteristics through aspect-based sentiment analysis and integration with electronic medical records (EMRs).

References

1. Rani, J., Kumar, P., Reddy, V.: Drug recommendation system based on sentiment analysis of drug reviews. Int. J. Comput. Sci. Trends Technol. (IJCST) 10(5), 94-99 (2022)
2. Rani, S., Devi, M., Rao, K.: Drug recommendation system based on sentiment analysis of drug reviews. J. Emerging Technol. Innov. Res. (JETIR) 11(3), 221-227 (2024)
3. Vijayaraghavan, S., Basu, S.: Sentiment analysis in drug reviews using supervised learning algorithms. In: Pro c. Int. Con f. Data Analytics Manage., pp. 1-8 (2020).
4. Nair, A., Gupta, S., Sharma, R.: Hey! This medicine made me sick: a sentiment analysis of user-generated drug reviews. In: Pro c. Int. Conf. Healthcare Inform., pp. 1-10 (2024)
5. Liu, B.: Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, pp. 1-167 (2012)
6. Ravi, K., Ravi, V.: A survey on opinion mining and sentiment analysis. Knowl.-Based Syst. 89, 14-46 (2015)
7. Holzinger, A., Carrington, A., Müller, H.: Measuring the quality of explanations: The system causability scale. IEEE Access 8, 716-735 (2020)
8. Feldman, R.: Techniques and applications for sentiment analysis. Commun. ACM 56(4), 82-89 (2013)
9. He, H., Gao, S., Korhonen, A.: How do general-purpose sentiment analysers perform on health-related social media data? Stud. Health Technol. Inform. 275, 278-282 (2021)
10. Panda, B., Singh, S., Kumar, R.: Exploratory data analysis and sentiment analysis of drug reviews. Comput. Sist. 26(3), 4093-4108 (2022)
11. Zhang, L., Wang, S., Liu, B.: Deep learning for sentiment analysis: A survey. IEEE Access 6, 48100-48117 (2018)
12. Jadhav, S.R., Deshmukh, P., Patil, S.: A survey on sentiment-based recommendation systems. IEEE Access 9, 124567-124580 (2021)

13. Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F., Stoyanov, V.: SemEval-2016 task 4: Sentiment analysis in Twitter. In: Pro c. 10th Int. Workshop Semantic Eval. (SemEval), pp. 1-18 (2016)
14. F. G. Surya Kallumadi, "Drug Reviews (Druglib.com)." UCI Machine Learning Repository, 2018. doi: 10.24432/C55G6J.
15. H. Yu, "data of multiple-type drug-drug interactions." Mendeley, Jul. 06, 2020. doi: 10.17632/MD5CZFSFND.1.