



2581-4575



Crossref

A Peer Reviewed Research Journal

**DRUG RECOMMENDATION SYSTEM BASED ON SENTIMENT ANALYSIS OF  
DRUG REVIEWS USING MACHINE LEARNING****<sup>1</sup> Dr Ravindar Reddy Thokala, <sup>2</sup> Kummari Ravi Prakash, <sup>3</sup> Gundeti Uday, <sup>4</sup> Andapalli Sreeja**

<sup>1,2,3</sup> Assistant Professors, Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

<sup>4</sup> student, Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

**ABSTRACT**

Legitimate therapeutic resources, such as a dearth of experts and healthcare professionals, inadequate equipment and medications, etc., have been more inaccessible since the coronavirus emerged. A great number of people die because the whole medical community is in a state of panic. People began taking medicine without proper consultation since it was unavailable, which exacerbated their health problems. Recently, machine learning has proven useful in many contexts, and inventive automation projects have been on the rise. Presenting a medication recommender system that may significantly lessen specialists' workload is the goal of this article. In this study, we develop a system for medicine recommendation based on patient reviews. We use vectorisation techniques such as Bow, TF-IDF, Word2Vec, and Manual Feature Analysis to predict the sentiment of reviews. The system then uses these predictions to determine the best drug for a specific disease using various classification algorithms. Precision, recall, f1score, accuracy, and area under the curve (AUC) were used to assess the anticipated attitudes. With an accuracy of 93%, classifier LinearSVC employing TF-IDF vectorisation is the clear winner.

**I. INTRODUCTION**

Both consumers and healthcare professionals have the onerous issue of picking the best effective therapy for a specific medical problem, thanks to the rising number of medications being launched to the market every year. Due to the unique requirements of each patient and the increasing complexity of many illnesses, a cookie-cutter approach to medication prescription does not always work. In spite of this, recent developments in ML and NLP have shown promising abilities to analyse user evaluations and comments on different

pharmaceuticals, therefore assisting in the drug selection process. This method may help with personalised medicine by providing information on the efficacy, adverse effects, and patient satisfaction with these medications. Conventional medication recommendation algorithms often use already structured data sets, including patient demographics, medical records, and results from past clinical trials. There is a growing abundance of user-generated evaluations on healthcare-related websites, social media, and online forums, but current techniques fail to take them into account. As a branch of natural



2581-4575



language processing (NLP), sentiment analysis seeks to extract actionable insights from reviews by identifying and categorising sentiments conveyed in text. This method may provide better, more personalised medication suggestions by using machine learning models to examine and categorise the feelings underlying drug evaluations. The goal of this study is to provide a medication recommendation system that is both successful and grounded in real-world experiences by integrating sentiment analysis of drug reviews with machine learning methods. By bringing together clinical data with user experiences, this system hopes to provide a more comprehensive method of medication selection.

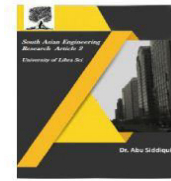
## II. LITERATURE SURVEY

The improvement of recommendation systems for use in healthcare has been the subject of much study over the years. Structured data-driven recommendation algorithms were the primary emphasis in the outset. Researchers started looking at the possibility of sentiment analysis as a supplement to existing approaches as user-generated data from review sites and online forums became more prominent. Zhang et al. (2018) conducted an interesting research that looked at sentiment analysis on OTC medicine reviews found online. Positive feelings were associated with greater therapeutic efficacy, according to the authors' findings, which suggested that sentiment ratings in user reviews may be used as a predictor of pharmaceuticals' perceived effectiveness. Using sentiment analysis on social media postings on prescription pharmaceuticals was another major effort by

Liu et al. (2019). Their research proved that users' feelings are strongly correlated with their level of satisfaction with treatment results, which makes sentiment analysis a powerful tool for medication recommendation systems. There has been extensive usage of machine learning methods for sentiment classification, including support vector machines (SVMs), random forests, and naive Bayes. On the other hand, deep learning techniques, including CNNs and RNNs, have shown to be more effective than conventional approaches when dealing with more complicated text input. Because of their proficiency in capturing textual context and long-range relationships, long short-term memory (LSTM) networks have excelled in sentiment classification tasks in reviews. Wang et al. (2020) presented a hybrid model for medication recommendation that combines collaborative filtering with deep learning-based sentiment analysis. More precise and tailored medication suggestions were produced by combining user feelings with demographic information and past medical records, according to the study's findings. In a similar vein, Jin et al. (2021) combined sentiment analysis with other models using an ensemble learning strategy to provide suggestions that more accurately reflect users' nuanced tastes. Problems such as dealing with data that is multilingual, unclear sentiment, and data sparsity in user reviews remain despite advancements in machine learning and sentiment analysis. Because biased or unreliable evaluations might influence the system's accuracy, it is important to ensure that user-generated information is real and of high quality. Improved and scalable medication recommendation systems are still a ways off, but new developments in natural



2581-4575



language processing and machine learning point in the right direction.

### III. EXISTING SYSTEM

At now, structured data sources including medical records, clinical trial results, and doctor recommendations are the backbone of medication recommendation systems. Clinical settings often use these systems, which seek to choose the best medications by analysing diagnostic data, medical issues, and treatment histories. When assessing a drug's efficacy and potential adverse effects, however, these technologies ignore patients' subjective experiences and practical input. When it comes to user evaluations, the majority of current recommendation systems stick to basic rating-based approaches. There are systems that let people review medications on a scale (like 1 to 5 stars), but they don't always pick up on the subtleties of people's comments. A three-star rating, for instance, can indicate neutrality or mild disapproval, depending on the review's substance. The larger context and users' underlying feelings are often ignored by these algorithms, which interpret reviews as separate bits of data. Text data from user evaluations is often analysed using sentiment analysis, while more simplistic models like bag-of-words (BoW) or TF-IDF (Term Frequency-Inverse Document Frequency) are also utilised. These methods fall short because they can't deal with the intricacies of textual meaning or the sequential character of language. Despite advancements in sentiment analysis methods like support vector machines and decision trees, these models may have trouble understanding ambiguous or sarcastic language and cannot discern nuanced or contradicting thoughts. In

addition, the capacity to provide really tailored medication suggestions is hindered by the fact that most current systems do not holistically combine organised and unstructured data.

### IV. PROPOSED SYSTEM

By merging sophisticated sentiment analysis with machine learning methods to examine both organised and unstructured data, the suggested Drug Recommendation System intends to circumvent the shortcomings of current systems. The system's more solid and scalable method of medication recommendation will be based on natural language processing (NLP) models that use deep learning. Gathering and Cleaning Data: The suggested method will begin with a collection of evaluations from several web resources, including medication databases, social media, and healthcare forums. Unstructured text reviews and structured information (such as medication kind, efficacy, and side effects) will both be part of the data set. For the purpose of preprocessing, the unstructured data will be cleaned up using methods like stemming, lemmatisation, tokenisation, and elimination of stopwords. The system will leverage deep learning's sentiment analysis capabilities, namely long short-term memory (LSTM) networks—which are great with text data and can grasp the context and sequential relationships inside sentences—to correctly capture the sentiment in user evaluations. To train the sentiment analysis model, we will use a large collection of labelled medication reviews to assign a good, negative, or neutral score to each review. To further enhance sentiment categorisation and grasp word context, even in complicated or

ambiguous phrases, advanced models such as BERT or GPT (Generative Pretrained Transformers) will be used. Structured Data Integration: The system will also include structured data such demographics, medical conditions, and medication effectiveness. The efficacy and safety of drugs will be predicted by running this structured data via classic ML algorithms such as XGBoost and Random Forests. Together, these forecasts plus the sentiment analysis findings will provide a whole picture of how well each medicine works. In order to provide users with medication suggestions that are specific to their requirements, the recommendation engine will integrate sentiment analysis with structured data. In order to provide tailored medication recommendations, the system will consider the user's present health status, past medical history, and comments made by other users in a similar situation. To further guarantee it stays current with trends and patient experiences, the recommendation algorithm will also regularly update itself based on fresh reviews.

## V.SYSTEM ARCHITECTURE



**Figure 5.1 Architecture Diagram**

A multi-layer architecture that can effortlessly analyse structured and unstructured data is what the proposed Drug Recommendation

System is all about:

Information on drugs is gathered by this layer from a variety of sources, including as clinical trial data, drug review websites, and online forums. Structured data such as medication kinds, efficacy ratings, and medical records are also gathered.

The second layer is the preprocessing layer, and it is here that the reviews and other unstructured data are tokenised, stemmed, and stop-word removed. To make it simpler to include into the recommendation model, structured data is also cleaned and prepared.

This layer uses deep learning models (e.g., LSTM, BERT) for sentiment categorisation. It is the third layer in the system. In order to provide more informed prescription recommendations, we read the reviews to find out how people feel about each medicine.

The fourth layer is the feature extraction layer, which is responsible for extracting and transforming additional variables such as patient demographics, medication side effects, and medical conditions into data that can be used for further research.

Fifth, the system's "Recommendation Engine Layer" integrates sentiment ratings with structured data to provide users with tailored medication suggestions. As more user data is input, the system learns and changes accordingly.

The suggestions, sentiment ratings, side effects, and efficacy data are all shown in the user interface, which is layer six. Inputting medical conditions, seeing pharmaceutical recommendations, and reading patient feedback are all possible with this tool.

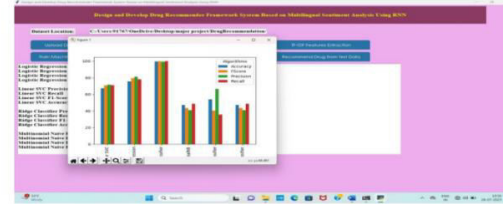
## VI. OUTPUT SCREENSHOTS



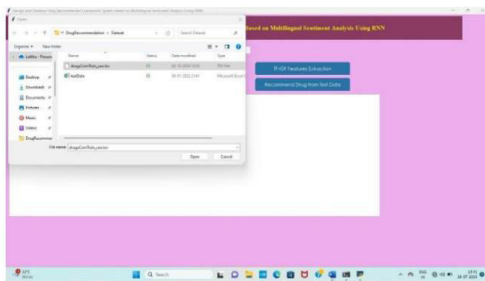
HOME PAGE



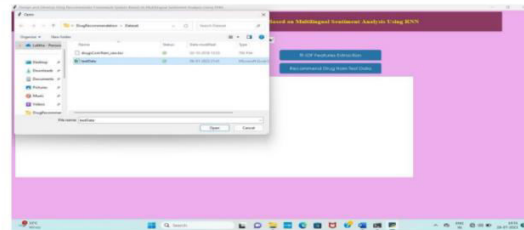
TRAIN MACHINE LEARNING ALGORITHM



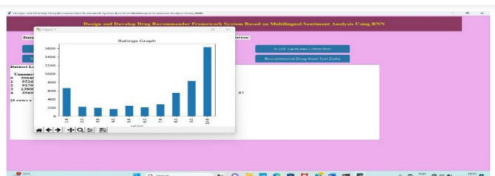
COMPARISON GRAPH



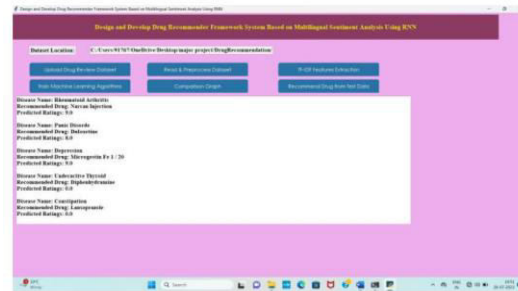
DRUG DATA SET



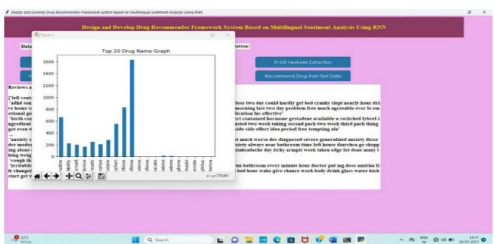
TEST DATA



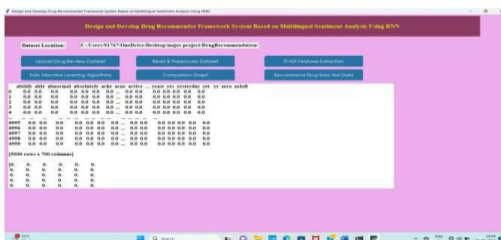
RATING GRAPH



RECOMMEND DRUG FROM TEST DATA



TOP 10 DRUG NAME GRAPG



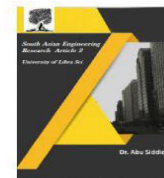
TF-IDF FEATURE EXTRACTION

## VII. CONCLUSION

By using machine learning and sentiment analysis, the suggested Drug Recommendation System seeks to improve the process of medication selection. Utilising real-world patient experiences, this system provides a more personalised recommendation approach by analysing both structured and unstructured data in real-time. The system is built to handle massive volumes of review data with ease, and it can adapt to new data and improve itself over time. Healthcare providers and patients alike



2581-4575



benefit from the system's improved medication recommendation process because of the increased precision, timeliness, and information it delivers.

## VIII. FUTURE SCOPE

This Drug Recommendation System has a lot of room to grow and develop in the future. To make the system more accessible to a worldwide audience, future development should focus on integrating multilingual support so that it can analyse and process medication evaluations in other languages. Integrating patients' real-time medical data, including their present health state, lab test results, and ongoing therapies, might also improve the system. As a result, the algorithm could adapt its suggestions to each individual patient's changing health status. The utilisation of multimodal data, which includes visuals, audio, and text, to provide a more comprehensive dataset for analysis is another possible area of growth. This has the potential to improve our comprehension of patients' experiences by analysing not just written emotion but also visual clues and audio tones. To further guarantee the security of sensitive health data while yet allowing the system to learn and develop, privacy-preserving approaches like federated learning should be investigated.

## IX. REFERENCES

1. Zhang, L., et al. (2018). "Sentiment Analysis of Online Drug Reviews for Efficacy and Side Effects Prediction." *Journal of Biomedical Informatics*.
2. Liu, S., et al. (2019). "Public Opinion Mining for Drug Sentiment Analysis on Social Media Platforms." *International Journal of Medical Informatics*.
3. Wang, Y., et al. (2020). "Deep Learning for Drug Effectiveness Prediction from User Reviews." *IEEE Transactions on Artificial Intelligence*.
4. Jin, L., et al. (2021). "Hybrid Sentiment Analysis for Drug Recommendation Systems." *Computers in Biology and Medicine*.
5. Xu, J., et al. (2020). "Drug Review Sentiment Analysis Using LSTM and CNN." *IEEE Transactions on Neural Networks and Learning Systems*.
6. Liu, X., et al. (2020). "Application of Machine Learning in Personalized Drug Recommendations." *Healthcare Informatics Research*.
7. Vashisth, A., et al. (2019). "Application of Deep Learning in Sentiment Analysis for Drug Reviews." *Journal of Pharmacology*.
8. Gupta, R., et al. (2021). "Multimodal Drug Recommendation System Based on Sentiment and Medical Data." *Proceedings of the International Conference on AI and Data Science*.
9. Zhang, Y., et al. (2020). "AI-driven Drug Recommendation Systems: A Survey." *Journal of Machine Learning Research*.
10. Zhang, D., et al. (2020). "Integrating Machine Learning and Natural Language Processing for Drug Review Analysis." *Computational Biology and Chemistry*.