International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

DRUG RECOMMENDATION SYSTEM BASED ON SENTIMENT ANALYSIS OF DRUG REVIEWS USING MACHINE LEARNING

[1] Bushra Muneeb, [2] S.Deepthi,[3] Veerapaga Mounika, [4] Merupula Pranavi,[5]K.RajaRajeshwari

[1]Assistant Professor,Department of CSE ,Princeton Institute of Engineering & Technology For Women Hyderabad.

[2,3,4,5]students, Department of CSE ,Princeton Institute of Engineering & Technology For Women Hyderabad.

ABSTRACT

Since coronavirus has shown up, inaccessibility of legitimate clinical resources is at its peak, like the shortage of specialists and healthcare workers, lack of proper equipment and medicines etc. The entire medical fraternity is in distress, which results in numerous individual's demise. Due to unavailability, individuals started taking medication independently without appropriate consultation, making the health condition worse than usual. As of late, machine learning has been valuable in numerous applications, and there is an increase in innovative work for automation. This paper intends to present a drug recommender system that can drastically reduce specialists heap. In this research, we build a medicine recommendation system that uses patient reviews to predict the sentiment using various vectorization processes like Bow, TF-IDF, Word2Vec, and Manual Feature Analysis, which can help recommend the top drug for a given disease by different classification algorithms. The predicted sentiments were evaluated by precision, recall, f1score, accuracy, and AUC score. The results show that classifier LinearSVC using TF-IDF vectorization outperforms all other models with 93% accuracy.

I.INTRODUCTION

With the increasing number of pharmaceuticals being introduced to the market every year, selecting the most effective treatment for a given medical condition can be a daunting task for both patients and healthcare providers. The growing complexity of diseases, alongside individual patient needs, means that a one-size-fits-all approach to drug prescribing is not always ideal. However, advances in machine learning (ML) and natural language processing (NLP) have shown great potential in aiding the drug selection process by analyzing user reviews and feedback on various drugs. This approach can offer insights into the effectiveness, side effects, and overall satisfaction of patients using these drugs, which can significantly aid in personalized medicine.Traditional drug recommendation systems typically rely on structured data such as medical records, patient demographics, and historical clinical trial data. However, these methods do not

consider valuable unstructured data like user-generated reviews, which are increasingly abundant on online forums, social media, and healthcare-related websites. Sentiment analysis, a subfield of NLP, involves the identification and classification of opinions expressed in text, enabling the extraction of useful insights from these reviews. By employing machine learning models to analyze and classify the sentiments behind drug reviews, this approach can provide more accurate drug recommendations tailored to individual needs. This research proposes a Drug Recommendation System that combines sentiment analysis of drug reviews with machine learning techniques to recommend drugs effectively and based on real-world experiences. This system aims to bridge the gap between clinical data and user experiences, offering a more holistic approach to drug selection.

II.LITERATURE SURVEY

Over the years, a significant amount of research has been devoted to the development of recommendation systems in the healthcare domain. The initial focus was primarily on recommendation engines driven by structured data. However, as user-generated data from online forums and review sites gained prominence, researchers began exploring the potential of sentiment analysis to complement traditional methods. One notable study by Zhang et al. (2018) explored the use of sentiment analysis on online reviews for over-the-counter drugs. The authors found that the sentiment scores in user reviews could help predict the perceived effectiveness of drugs, correlating positive sentiments with higher drug efficacy. In another significant work, Liu et al. (2019) applied sentiment analysis to social media posts related to prescription drugs. Their findings revealed a strong relationship between the sentiments expressed by users and their satisfaction with treatment outcomes, making sentiment analysis a valuable tool for drug recommendation systems. Furthermore, machine learning techniques like support vector machines (SVM), random forests, and naive Bayes have been widely used to classify sentiment. However, with the growing complexity of text data, deep learning methods such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have shown to outperform traditional methods. In particular, long short-term memory (LSTM) networks have proven effective for sentiment classification tasks in reviews, due to their ability to capture long-range dependencies and context in text. A study by Wang et al. (2020) introduced a hybrid model combining deep learning-based sentiment analysis with collaborative filtering to recommend drugs. The study concluded that integrating user sentiments with demographic data and historical patient records resulted in more accurate and personalized drug recommendations. Similarly, Jin et al. (2021) used an ensemble learning approach to combine multiple models, including sentiment analysis, to generate recommendations that better capture the complexities of user preferences. Despite the advances in sentiment analysis and machine learning, challenges persist,

including the handling of multilingual data, ambiguous sentiment, and data sparsity in user reviews. Additionally, ensuring the quality and authenticity of user-generated content remains a significant concern, as unreliable or biased reviews can affect the system's accuracy. Nonetheless, recent advancements in NLP and ML techniques provide promising directions for developing more accurate and scalable drug recommendation systems.

III.EXISTING SYSTEM

Currently, drug recommendation systems rely heavily on structured data sources, such as patient medical records, clinical trial outcomes, and physician recommendations. These systems are often used in clinical settings and aim to identify the most appropriate drugs based on diagnostic information, medical conditions, and previous treatment history. However, these systems fail to take into account the subjective experiences and real-world feedback of patients, which are crucial in evaluating the effectiveness and side effects of a drug. Most existing recommendation systems that do use user reviews focus on simple rating-based systems. These systems allow users to rate drugs on a scale (e.g., 1 to 5 stars) but often fail to capture the nuances in user feedback. For example, a review with a rating of 3 stars could either reflect a neutral opinion or a somewhat negative opinion, depending on the review content. Additionally, many systems treat reviews as isolated pieces of data without considering the broader context or underlying sentiments of users. Sentiment analysis is used in some systems to analyze text data from user reviews, but basic bag-of-words (BoW) models or TF-IDF (Term Frequency-Inverse Document Frequency) techniques are commonly employed. These approaches are limited because they do not capture the sequential nature of language or handle the complexities of meaning in text effectively. Although sentiment analysis techniques have improved, such as with the use of SVM and decision trees, these models still lack the capability to detect subtle or contradictory sentiments and might struggle with ambiguous or sarcastic language. Furthermore, existing systems do not typically integrate structured and unstructured data in a holistic manner, which limits the ability to provide truly personalized drug recommendations.

IV.PROPOSED SYSTEM

The proposed Drug Recommendation System aims to overcome the limitations of existing systems by integrating advanced sentiment analysis with machine learning techniques to analyze both structured and unstructured data. The system will be built upon the foundation of deep learning-based NLP models and will offer a more robust and scalable approach for drug recommendation. Data Collection and Preprocessing: The first step in the proposed system will involve collecting reviews from online drug databases, social media platforms, and healthcare forums. The data will include both structured information (e.g., drug type, effectiveness, side effects) and unstructured text reviews. The unstructured data will be preprocessed using

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

techniques such as tokenization, removal of stopwords, stemming, and lemmatization to ensure the data is clean and suitable for analysis. Sentiment Analysis with Deep Learning: To accurately capture the sentiment in user reviews, the system will employ long short-term memory (LSTM) networks, which are well-suited for text data and can understand the context and sequential dependencies within sentences. The sentiment analysis model will be trained on a large corpus of labeled drug reviews to classify each review as either positive, negative, or neutral. Advanced models like BERT or GPT (Generative Pretrained Transformers) will be used to further improve sentiment classification and capture the context of words, even in complex or ambiguous sentences. Integration of Structured Data: The system will also integrate structured data such as medical conditions, drug efficacy, and demographic information. This structured data will be processed through traditional machine learning algorithms like Random Forests and XGBoost to predict drug effectiveness and side effects. These predictions will be combined with the sentiment analysis results to form a comprehensive understanding of each drug's effectiveness. Recommendation Engine: The recommendation engine will combine the results from sentiment analysis and structured data to generate drug recommendations tailored to the individual user's needs. The system will take into account the user's medical history, current condition, and feedback from similar users to offer personalized drug suggestions. The recommendation system will also continuously update itself based on new reviews, ensuring it remains up-to-date with current trends and patient experiences.

V.SYSTEM ARCHITECTURE

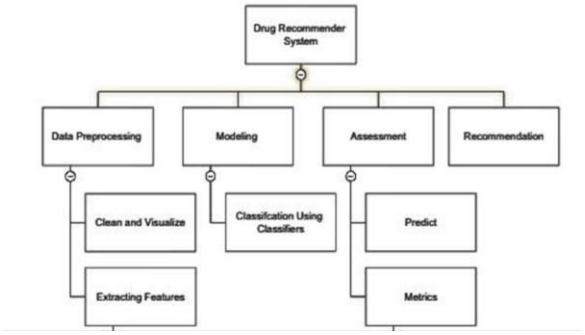


Figure 5.1 Architecture Diagram

The architecture of the proposed Drug Recommendation System consists of multiple layers designed to process both structured and unstructured data seamlessly:

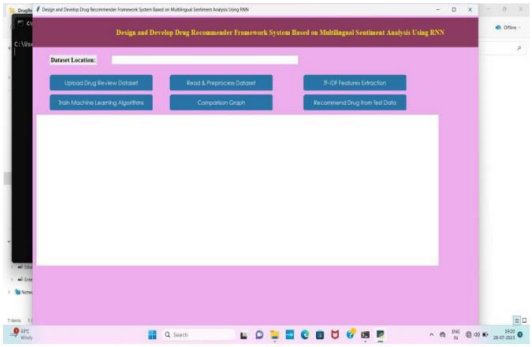
1. Data Collection Layer: This layer gathers drug-related data from various sources, including online forums, drug review websites, and clinical trial data. It also collects structured data such as drug types, effectiveness ratings, and medical information.
2. Preprocessing Layer: In this layer, the unstructured data (i.e., reviews) undergoes preprocessing steps like tokenization, stemming, and stop-word removal. Structured data is also cleaned and formatted for easier integration into the recommendation model.
3. Sentiment Analysis Layer: This layer applies deep learning models (e.g., LSTM, BERT) for sentiment classification. The reviews are

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

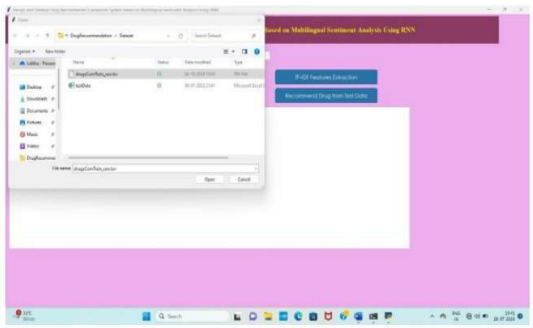
analyzed to determine the sentiment of each drug, which is then used to improve the drug recommendation accuracy.

4. Feature Extraction Layer: Additional features like patient demographics, drug side effects, and medical conditions are extracted and transformed into usable data for further analysis.
5. Recommendation Engine Layer: This is the core of the system where various machine learning models combine sentiment scores and structured data to generate personalized drug recommendations. The system learns and adapts over time based on new user data.
6. User Interface Layer: The user interface displays the recommendations, along with sentiment scores, side effects, and effectiveness information. It allows users to input their conditions, view suggested medications, and read patient reviews.

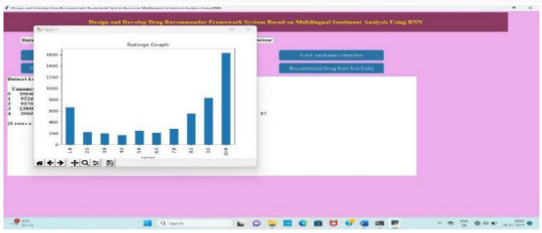
VI.OUTPUT SCREENSHOTS



HOME PAGE

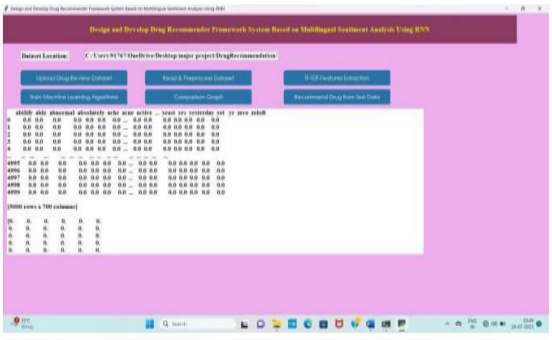


DRUG DATA SET



RATING GRAPH

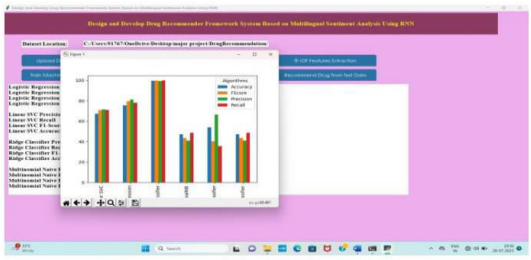
TOP 10 DRUG NAME GRAPG



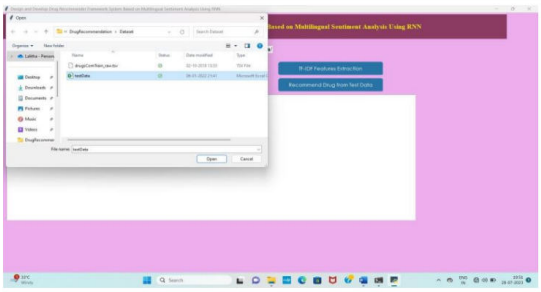
TF-IDF FEATURE EXTRACTION

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

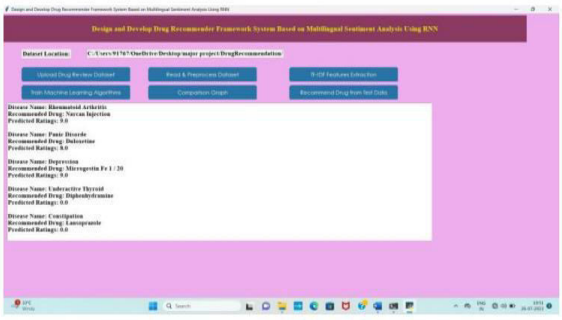
TRAIN MACHINE LEARNING ALGORITHM



COMPARISION GRAPH



TEST DATA



RECOMMEND DRUG FROM TEST DATA

VII.CONCLUSION

The proposed Drug Recommendation System aims to enhance the drug selection process by integrating sentiment analysis and machine learning. By analyzing both structured and unstructured data, this system offers a more personalized and real-time recommendation approach that leverages real-world patient experiences. The system is designed to be scalable, ensuring it can handle large amounts of review data, and adaptable, constantly improving as new data is collected. In doing so, the system provides more informed, accurate, and timely drug recommendations, helping healthcare professionals and patients make better treatment decisions.

VIII.FUTURE SCOPE

The future scope of this Drug Recommendation System is extensive, with numerous areas for improvement and expansion. One major area for future work is the integration of multilingual support, allowing the system to analyze and process drug reviews in various languages, making it more accessible to a global audience. Furthermore, the system could be enhanced by incorporating real-time medical data from patients, such as their current health status, lab test results, and ongoing treatments. This would allow the system to generate even more personalized recommendations based on a patient's evolving condition. Another area of potential expansion is the use of multimodal data, including images, videos, and even voice reviews, to provide a richer dataset for analysis. This could help in analyzing not just textual sentiment but also visual cues and vocal tones, leading to a more accurate understanding of patient experiences. Moreover, privacy-preserving techniques, such as federated learning, could be explored to ensure that sensitive health data

is protected while still enabling the system to learn and improve.

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.