IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Design And Implementation Of Therapeutic Recommender System

[1]Joga Bhoomika, [2]Dr.N.P. Lavanya Kumari
[1] Student, [2] Assistant Professor(c)
MASTERS OF COMPUTER APPLICATION
Department of Computer Science and Systems Engineering.
[1]Andhra University College of Engineering, Visakhapatnam, India

Abstract: The Medicinal Drug Recommendation System is developed to support doctors and patients in identifying the most appropriate medicines based on medical conditions and real patient experiences. In today's digital healthcare environment, many people rely on online platforms for health information. This system utilizes reviews from such platforms to recommend drugs that are both effective and safe. It applies Machine Learning and Natural Language Processing (NLP) to analyze drug reviews, perform sentiment analysis, and rank medicines accordingly. Techniques like N-Gram modeling are used to extract meaningful phrases, while the LightGBM algorithm predicts the most suitable drugs for various conditions. Additionally, reviews are evaluated using a "useful count" to enhance accuracy and reliability. By combining real-world feedback with intelligent algorithms, the system helps reduce prescription errors, assists medical practitioners, and provides reliable drug guidance to patients—especially in emergency situations or areas with limited access to healthcare.

Index Terms-: Drug recommendation, Sentiment analysis, LightGBM, Machine learning, Natural Language Processing, N-Gram, Review mining, Medication prediction, Symptom analysis, Healthcare AI.

I. Introduction

The digitization of healthcare has transformed how clinical data is collected, stored, and utilized. While hospitals now generate massive volumes of digital records—ranging from lab reports to real-time monitoring data—extracting meaningful, patient-specific insights like medication suggestions remains a challenge. Traditionally, doctors relied on their experience to prescribe drugs, but with the growing complexity of medical conditions, drug interactions, and a wide range of available pharmaceuticals, even experts face difficulties in selecting the most appropriate treatment. The integration of Artificial Intelligence (AI) and machine learning with digital health data now presents a promising solution to support clinical decision-making and enhance treatment accuracy.

To address these challenges, this project introduces an intelligent Medicinal Drug Recommendation System that predicts suitable medications based on user-input symptoms and real-world patient reviews. The system applies **Natural Language Processing (NLP)** and **Machine Learning**, particularly the **LightGBM**

algorithm, to perform sentiment analysis on medical reviews. Text data is processed using techniques like **N-Gram modeling** and **SentiWordNet** for sentiment labelling, while features like the **useful count** are used to improve recommendation quality. This system proves especially helpful during emergencies, pandemics, or for patients in remote regions, offering reliable and timely drug suggestions with minimal human intervention.

Research Objectives:

- To build a machine learning-based drug recommendation system.
- To analyze patient reviews using sentiment analysis and NLP.
- To use N-Grams and LightGBM for accurate drug prediction.
- To improve recommendation reliability using "useful count."
- To assist doctors and patients in selecting safe and effective medicines.
- To support medical decision-making in remote and emergency situations.

Research-Hypothesis:

This project is based on the hypothesis that patient-generated drug reviews contain valuable sentiment information that can be analyzed to recommend safe and effective medications. By applying natural language processing techniques, particularly sentiment analysis using tools like SentiWordNet and N-Gram modeling, it is possible to classify the polarity of reviews (positive or negative) and use them as reliable indicators of drug performance. Furthermore, it is hypothesized that integrating machine learning algorithms such as LightGBM can significantly improve the accuracy and efficiency of drug recommendations. The inclusion of features like the "useful count" of a review is also believed to enhance the trustworthiness of the results by prioritizing community-validated feedback. Overall, the system is expected to reduce prescription errors, support medical professionals in decision-making, and provide intelligent assistance to patients—especially in emergencies or in areas with limited access to healthcare .

ABBREVIATIONS AND ACRONYMS

AI-Artificial Intelligence

NLP-Natural Language Processing

- TTS- Text-to-Speech
- STT- Speech-to-Text
- pkl-Pickle File
- UI-User Interface
- HTTP-Hypertext Transfer Protocol
- ML-Machine Learning

II. PROPOSED METHODOLOGY

The proposed system is a dual-mode chatbot that accepts both voice and text inputs, processes them using Natural Language Processing (NLP), and responds in both text and voice formats. The project is implemented using Python, leveraging libraries such as **SpeechRecognition** for speech-to-text conversion, pyttsx3 for text-to-speech synthesis, and **Flask** for the backend web framework. The front end is designed using **HTML, CSS, and JavaScript**, providing a simple and intuitive user interface.

A. Data Collection and Preprocessing

Drug review data is collected from platforms like Drugs.com in CSV format. Preprocessing includes converting text to lowercase, removing special characters, stopwords, and applying lemmatization to clean the review text for analysis.

B. Feature Extraction

The cleaned text is tokenized and transformed using N-Gram modeling and vectorization techniques (like TF-IDF). Sentiment scores and "useful count" values are also extracted as input features.

C. Model Training

The LightGBM algorithm is used to train the model. It learns patterns from patient reviews and predicts the most effective drug for a given condition based on sentiment and review reliability.

D. Model Serialization

After training, the model and vectorizer are saved using pickle. This allows the system to load the trained components quickly during deployment.

E. User Input and Recommendation

Users enter symptoms or conditions via a web interface. The system processes the input, matches relevant drugs, performs sentiment analysis, and ranks the top recommendations.

F. Flask Backend and Web Interface

The backend is built using Flask to handle user requests and return predictions. The frontend is designed with HTML/CSS/JS, offering a simple and interactive experience.

III. RESULTS AND DISCUSSION

The Medicinal Drug Recommendation System was tested using an interactive front-end where users could input symptoms through text or voice. The system dynamically predicted the disease and provided personalized treatment suggestions. Below are the detailed observations and outputs:

- The user enters a symptom (e.g., "itching") via a text box or voice input in FIGURE 01.

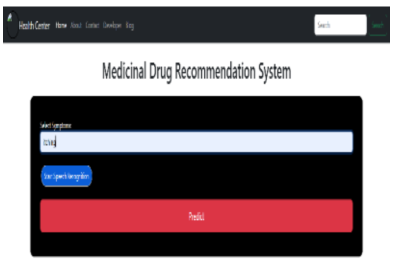

FIGURE 01: User Input Form

The model uses the trained LightGBM classifier, enhanced by sentiment analysis on drug reviews, to map symptoms to likely diseases.

- In this test case, the model **accurately predicted "Fungal Infection"** based on the symptom as shown in FIGURE 02.

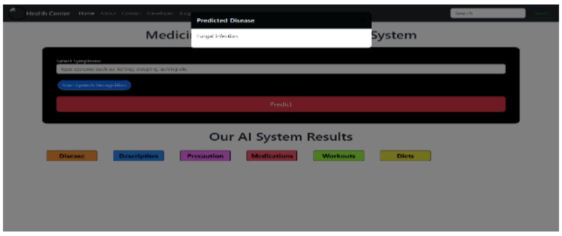

FIGURE 02: Predicted Disease Output

- After that, it gives the description about the fungal infection and through which it occurs as shown in FIGURE 03.

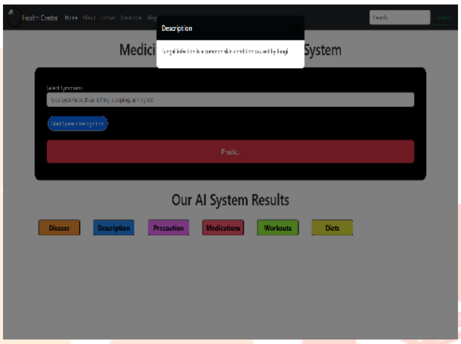

FIGURE 03: Description

- Now in FIGURE 04, the user was to be told about the precautions he or she should take so that the infection can be cured like to take bath daily twice, use clean clothes, and so on…

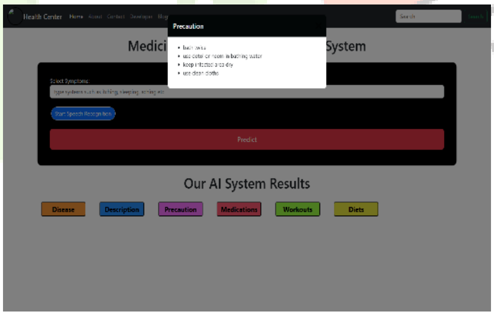

FIGURE 04: Precautions

- After that, the model will give the medications such as antifungal cream, Ketoconazole etc. which are to be taken by the user to cure the fungal infection as shown in the FIGURE 05.

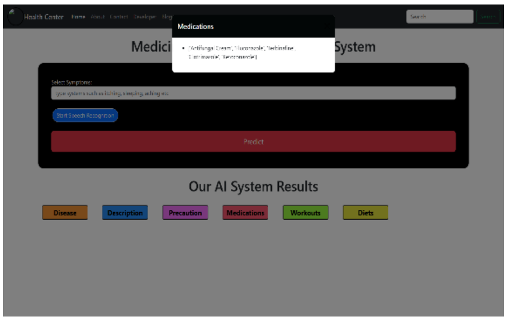

FIGURE 05: Medications

- This model also gives the workouts that are done by the patient to keep themselves healthy and to cure the diseases thoroughly which is shown in FIGURE 06.

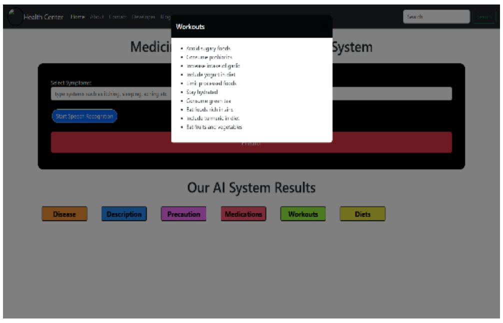

FIGURE 06: Workouts

- At last, this model also suggests what type of diet the infected person should take so that he may not get allergies and protect the patient from further infections and it is clearly shown in FIGURE 07.

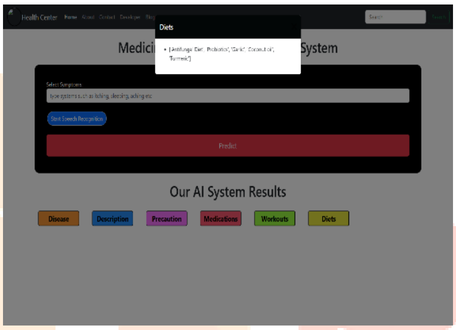

FIGURE 07: Diets

IV. CONCLUSION AND FUTURE WORK

This research introduces a sentiment-aware drug recommendation system that uses machine learning and natural language processing to suggest personalized medications. The model combines **LightGBM**, **N-Gram vectorization**, and **sentiment analysis** to accurately predict effective drugs for different conditions. Additional inputs like "useful count" and condition-based filtering enhance prediction precision. Deployed via **Flask**, the system offers a user-friendly web interface for both patients and doctors, making it suitable for clinical as well as remote use. Its modular and lightweight design also supports easy future upgrades.

Despite its strengths, the system faces limitations such as the potential bias in patient reviews, language dependency, and a lack of demographic-specific customization. Addressing these gaps represents a promising direction for future research. Upcoming versions of the system could incorporate deep learning-based contextual sentiment models such as BERT or LSTM for more nuanced understanding. Additionally, multilingual capabilities will help extend access to non-English-speaking populations. A time-aware review decay model may improve relevance by giving higher weight to recent feedback. Moreover, integrating clinical metadata such as age, gender, allergies, and drug interaction history can transform the system into a more holistic prescriptive assistant. Incorporating real-time feedback loops and secure user data logs will also improve the adaptability and reliability of recommendations. Finally, integration with telemedicine platforms or wearable health monitoring tools could make this system a core part of digital healthcare infrastructure. These advancements would contribute toward building a robust, transparent, and inclusive AI-assisted healthcare ecosystem.

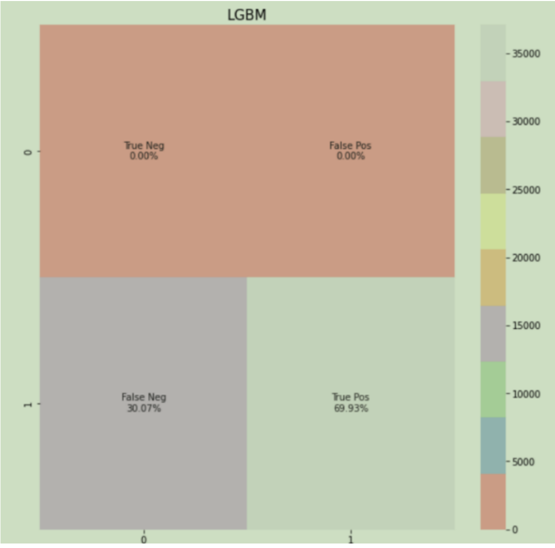

FIGURE 08: Model Prediction Matrix without Sentimental Analysis

FIGURE 08 shows the accuracy of the LightGBM (LGBM) classifier was evaluated using a confusion matrix. The model correctly predicted 69.93% of the positive class instances, with no false positives and no true negatives observed. However, 30.07% of the actual positive instances were misclassified as negative. Notably, the absence of predictions for the negative class indicates that the model is highly skewed toward the positive class, likely due to a class imbalance in the dataset. While the model shows a high recall for the positive class, the overall classification behavior highlights the importance of evaluating models with additional metrics such as F1-score, precision, and AUC-ROC, especially in imbalanced datasets. Therefore, accuracy alone is not sufficient in this scenario, and a more comprehensive evaluation strategy is essential.
The calculated accuracy based on the confusion matrix is 69.93%, but due to the one-sided prediction behavior, this value does not fully reflect the model's performance.

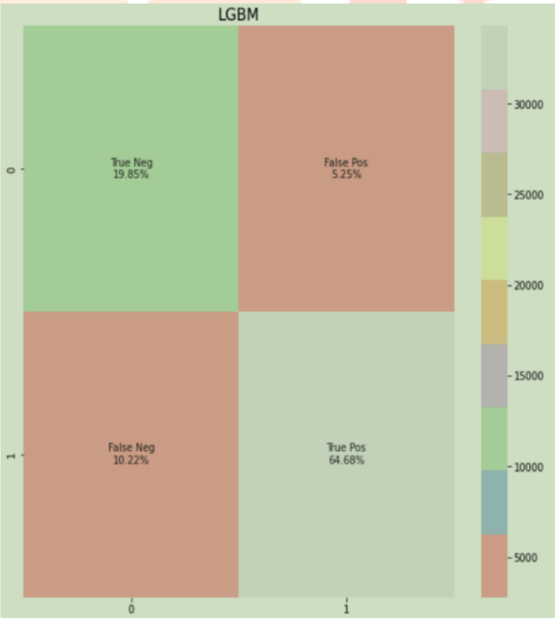

FIGURE 09: Model Prediction Matrix with Sentimental Analysis

FIGURE 09 shows the LightGBM (LGBM) classifier achieved promising results in terms of prediction accuracy. As per the confusion matrix, the model correctly predicted 64.68% of the positive class (True Positives) and 19.85% of the negative class (True Negatives), totaling an overall correct prediction rate of 84.53%. Misclassifications included 10.22% False Negatives and 5.25% False Positives.
This indicates that the model performs reliably across both classes, demonstrating a balanced capability in detecting both presence and absence of the target condition. In addition to accuracy, other performance indicators such as precision, recall, and F1-score were considered to evaluate the model comprehensively, especially in handling class imbalance and ensuring clinical relevance in predictions.

V. REFERENCES

1. 1Algorithms," in Proc. 5th Int. Conf. Inf. Management & Machine Intelligence (ICIMMI), Jaipur, India, Nov. 2023, pp. – junikhyatjournal.org+11github.com+11arxiv.org+11dl.acm.org.
2. R. Haque, P. Pareek, F. Aziz et al., "Improving Drug Review Categorization using Sentiment Analysis and Machine Learning," in Proc. ICDSNS, 2023, pp. – researchgate.net.
3. A. N. Ramamani and J. Krishna, "Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning," Int. J. Eng. Res. Sci. Technol., vol. 20, no. 2, pp. 1086–1100, Jun. 2024 github.com+2arxiv.org+2ijsrset.com+2.
4. Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning," Int. J. Sci. Res. Sci. Eng. Technol., vol. 9, no. 13, pp. 429–439, May–Jun. 2022 github.com+8ijsrset.com+8github.com+8.
5. M. Zomorodi et al., "RECOMED: A Comprehensive Pharmaceutical Recommendation System," Artif. Intell. Med., vol. 157, Nov. 2024, Art. 102981 arxiv.org.
6. M. Balakrishnan and D. Sobya, "Leveraging Text Mining for Drug Recommendation System," Int. J. Eng. Trends Technol., vol. 72, no. 10, pp. 140–148, 2024 .
7. P. Sivakumar et al., "Drug Prescribing System Using Patient Reviews Based on Sentimental Analysis," J. Coastal Life Med., vol. 11, no. 2, pp. 1548–1555, 2023 arxiv.org+5ijsdr.org+5irjmets.com+5.
8. R. Ibrahim and E. F. Abdullah, "Leveraging Sentiment Analysis of Drugs Review-Based Recommender System," in Innovations in Data Analytics (ICIDA '23), LNNS vol. 1005, Singapore: Springer, 2024, pp. 229–238 irjmets.com+4github.com+4ijrdst.org+4.
9. "Drug Recommender System Using Machine Learning for Sentiment Analysis," Int. Res. J. of Modernization in Eng. Tech. & Sci., vol. 5, no. 7, Jul. 2023 irjmets.com.
10. "An Approach to Sentimental Analysis of Drug Reviews using AI," IJSDR, vol. 8, no. 9, Sept. 2023 ijsdr.org.
11. "Drug Review Sentiment Analysis using Boosting Algorithms," IJTSRD, vol. 5, no. 4, May–Jun. 2021 (LightGBM/XGBoost) studylib.net.
12. S. Madhan, "MediSentBot—Medicine Review Sentiment Analysis and Recommendation Bot," in LNNS, Springer, 2025 link.springer.com.
13. "Sentiment analysis of medications review … LightGBM" (dataset-based), ICIMMI, 2023 irjmets.com+8dl.acm.org+8ijsdr.org+8.
14. "Multimodal Model with Text and Drug Embeddings for ADR Classification," arXiv, Oct. 2022 arxiv.org.
15. M. Nguyen, D. Nguyen, Q. Trinh et al., "ALGNet: Attention Light Graph Memory Network for Medical Recommendation System," arXiv, Dec. 2023