iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59247

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

Drug Recommendation System on Sentiment Analysis of Drug Reviews by Using Machine learning

B. Sivaiah[1], Grishma[2], S. Tharun[3], A. Shashi[4]
[1]*Associate Professor,* [2, 3, 4] *UG Student, Department of CSE, CMR College of Engineering & Technology, Hyderabad, Telangana*

Abstract: ***With the healthcare system facing increased challenges due to the COVID-19 pandemic, there's a big demand for new ideas to help doctors and nurses. This paper suggests a new way of using computers to help doctors decide which medicines to give to patients. By using smart computer programs, we can make it easier for healthcare workers to handle their workload and provide better care for patients. By analyzing patient reviews, we employ sentiment analysis using advanced vectorization methods like Bag of Words, Term Frequency-Inverse Document Frequency (TF-IDF), and Manual Feature Analysis. These techniques allow for the identification of subtle sentiment nuances crucial for tailoring personalized drug recommendations. We utilize various classification algorithms such as Naive Bayes, Support Vector Classifier (SVC), and Random Forest to predict sentiments. Among these, the LinearSVC classifier coupled with TF-IDF vectorization demonstrates superior performance in sentiment prediction. We tested our system using various measures like precision, recall, F1-score, and Area Under the Curve (AUC) to ensure its effectiveness. By incorporating advanced machine learning techniques, our system offers a strong foundation for enhancing the accuracy of drug prescriptions, which is a key issue in today's healthcare. Through better drug recommendations, our goal is to ultimately enhance patient health outcomes and play a part in advancing healthcare practices, especially in the face of complex and evolving healthcare challenges.***

Keywords: Healthcare, Machine learning, Dataset, Medicines, Individual's demise, Drug reviews, Sentiment analysis, Vectorization processes, Accuracy, Classification algorithms.

I. INTRODUCTION

As the number of coronavirus cases increases, the country faces a shortage of doctors, especially in rural areas where there are fewer specialists than in cities. It takes approximately 6 to 12 years for a doctor to gain the necessary qualifications. Therefore, doctors cannot expand rapidly in a short time. In these difficult times, the telemedicine framework should be energized wherever possible. Medical errors are common today. Improper inspections affect more than 200,000 people in China and 100,000 in the United States each year. In more than 40% of the drugs, experts make mistakes when prescribing them because they create solutions based on their very limited knowledge. It is important for patients to choose the best medicine as they need specialists with extensive knowledge about diseases, antibiotics and patients. New research emerges every day, including many medications and tests that healthcare professionals can use every day. For this reason, it becomes difficult for doctors to choose the treatment or medication to be given to the patient based on the teachings and past medical history. With the growth of the web and web-based business transactions, project review has become a necessary and important part of the treatment. Buy products worldwide. People all over the world check reviews and websites before making a purchasing decision. While past research has primarily centered on assessing expectations and recommendations within e-commerce contexts, there has been a noticeable lack of attention given to clinical settings. With a growing emphasis on health and well-being among the general population, the prevalence of online medical inquiries is on the rise. According to a 2013 survey conducted by the Pew Research Center, close to 60% of adults have sought out health-related information online, with approximately 35% specifically searching for medical diagnoses. This underscores the increasing reliance on internet resources for medical guidance and highlights the importance of further exploration in this area within academic and practical domains. Drug approval strategy is important because it helps professionals and patients develop knowledge about drugs based on specific conditions. A recommendation framework is a rule that recommends a product to the user based on their strengths and needs. In this process, customer research is used to analyze customers' opinions and make recommendations based on their actual needs.

©IJRASET: All Rights are Reserved

The drug approval process uses analytical thinking and engineering to provide drugs in specific situations based on patient reviews. Sentiment analysis is a set of ideas, methods, and tools used to distinguish and extract emotional information, such as emotions and attitudes, from language. Feature engineering, on the other hand, is the process of creating more of existing features; improves the performance of the model related work

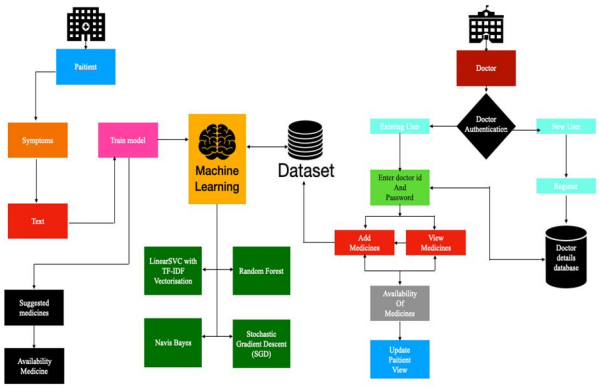
II. RELATED WORK

In the quest for innovation and efficiency, modern projects frequently rely on existing solutions as fundamental building blocks for development. This approach not only recognizes the expertise and advancements of those who came before us but also nurtures a collaborative ecosystem where ideas can evolve and confront new challenges. In our project, we wholeheartedly embrace this ethos, conscientiously integrating elements from existing solutions to enrich our endeavor. These existing solutions serve as guiding lights, offering insights and frameworks that shape the direction of our project.

1) *Drug Recommendations for Various Diseases in Medicine Using Machine Learning:* The best knowledge in the medical industry has been developed recently to improve patients' lives and enable better medical decisions. The application of machine learning and data mining can transform existing data into valuable information that can be used to identify symptoms and recommend appropriate medications. In this study, a machine learning method for multi disease drug recommendation is proposed to provide drug recommendations to patients with multiple diseases.
2) *Machine learning-based recommendation system for disease-drug material and adverse drug reaction:* Comparative review: In recent times, social media platforms have become significant channels for sharing patient medical content, leading to the extraction and analysis of this data for pharmacovigilance, drug repositioning, and advancements in the healthcare sector. Paramedical advertisers leverage this platform to gather trusted health information and correlate it with user profiles to recommend appropriate treatments. The exploration of patients' personal and medical details on social media has heightened awareness about recent advancements in drug treatments. To facilitate predictive computational medical services, there's a pressing need to establish a centralized virtual platform for pharmacovigilance activities using social media data. This platform would focus on identifying drugs and associated adverse drug reactions (ADRs). Additionally, it aims to foster an informational community of similar e-patients for more effective information sharing and emotional support.
   This paper provides a thorough examination of recommendation systems and highlights potential research gaps concerning drug repositories and ADRs based on recent studies. Moreover, it introduces a framework comprising clinical post pre-processing, clinical entity recognition, tweet pooling, and expert rating to classify the intersection of diseases, drugs, and ADRs, thus facilitating recommendations for drugs and ADRs. Furthermore, the paper conducts a comparative analysis of benchmark classifiers trained on labeled tweet corpora to automate drug repositioning and ADR detection, offering insights into the efficacy of the proposed framework.
3) *Machine Learning Based Health Drug Recommendation System:* In today's digital age, health is one of the important areas of health. People try to find health information they care about. The internet can be a great source of this information, but you need to be careful not to get bad information. Nowadays, a lot of medical information is scattered across different websites on the internet, making it impossible for users to find important information that can help improve their health. Misuse of medications is one of the most serious medical conditions and can be life threatening. These issues have increased the need for consensus systems in healthcare to help users make more informed and informed health decisions. In this paper, a drug recommendation system is designed to help end users choose specific and appropriate drugs for certain conditions based on the analysis of different drugs provided by other end users for many diseases. The purpose of this agreement is to use data mining concepts, visualization, reasoning, analytics to analyze the dataset and recommend medications based on the condition, providing assessment and treatment control for each condition of the patient using machine learning, context and collaborative filtering
4) *Computerized disease prediction and drug recommendation using machine learning:* The war between viruses and humans has been going on since ancient times. However, according to the theory of evolution, every part of the world, including small organisms, is constantly striving to survive. As a result, infectious diseases are increasing rapidly and burdening people with morbidity and mortality. Although we have many technologies for diagnosis, prevention and treatment of infectious diseases in this era, the emergence of new diseases worldwide continues to create serious problems for the public. A recent example is the novel coronavirus COVID-19, which was first discovered in Wuhan, China, and quickly evolved into a global pandemic. There are no drugs that can cure patients with this new disease.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

In this difficult situation, doctors and medical professionals recommend available medications based on the symptoms experienced by the patient. In this process, many patients lose their lives due to lack of suitable medicine. Therefore, in this study, we use disease prediction based on various symptoms of the disease. In addition, we have come up with an idea that could help the pharmaceutical industry develop drugs for all infectious diseases using machine learning. Essentially, the technology identifies symptoms and predicts the best medicine for any new disease. Additionally, the method can predict the content of drugs needed by pharmaceutical companies to develop new drugs under the supervision of pharmaceutical experts. Smart pharmaceutical industry and recommendations based on blockchain and machine learning for the management of medicines.



III. METHODS AND EXPERIMENTAL DETAILS

A. High- level methodology

Building a drug recommendation system is like putting together a puzzle with different pieces. The first piece is having good-quality information about drugs, like reviews from people who have used them. Then, there's understanding how these people feel about the drugs, which is where sentiment analysis comes in. It's like figuring out if they liked or disliked the drugs based on what they wrote. Next, we need smart algorithms that can use this information to suggest the right drugs to people based on their needs and preferences. But it's not just about making the system work; we also need to make sure it's doing its job well and following the rules. That means regularly checking and improving it to make sure it's giving helpful and trustworthy advice while also keeping people's information safe and following all the ethical and legal guidelines. In simpler terms, building a drug recommendation system is about collecting good information, understanding how people feel about it, using smart tools to make useful suggestions, and always making sure everything is done right.

B. Proposed work

Theoretically, this study is based on the fact that the recommended medication should depend on the patient's ability. For example, if the patient has been vaccinated, a reliable drug should be recommended at this time. A risk stratification method has been proposed to determine patient resistance. For example, more than 60 factors, such as blood pressure and alcohol consumption, are used to determine a patient's ability to protect themselves from infection. A web-based model using decision support has also been developed to help physicians choose their first medication. Three different algorithms are examined: decision tree algorithms, support vector machines (SVMs), and backpropagation neural networks that process the data. SVM was chosen as medical recommendation because it performed well on three parameters (model accuracy, model information, model versatility). Error control is also planned to ensure analysis, accuracy and quality control.

User: Users can register first. Registration requires the user's email address and mobile number for further communication. Once the user is registered, the administrator can activate the user. When the administrator is activated for the user, the user can log in to our system. Users can download files according to our matching system. The data must be in floating mode for the algorithm to execute. We use the recommended medication usage information here. Users can add new files to existing files based on our Django application.

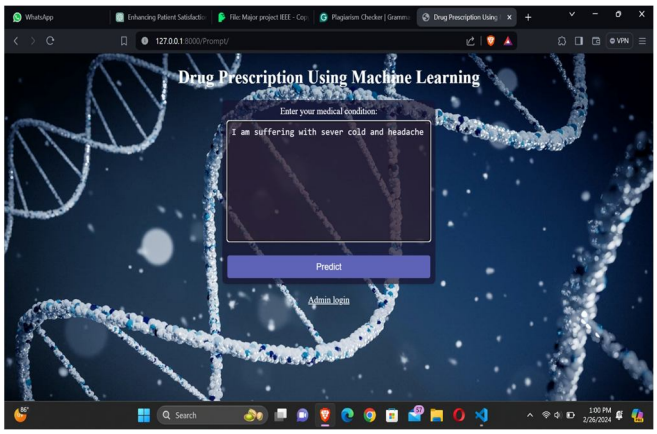


Figure 1 - User Interface

Administrator: Administrator are the doctors who can log in using login credentials. Administrators can initiate user registration. When activated, only the user can log in to our system. Administrators can view the entire profile in the browser. Administrators can click the "Remove" button on the page to see the actual, correct, reverse and F1 scores calculated according to the system. Once all the algorithms are executed the administrator can see the whole truth of the web page.

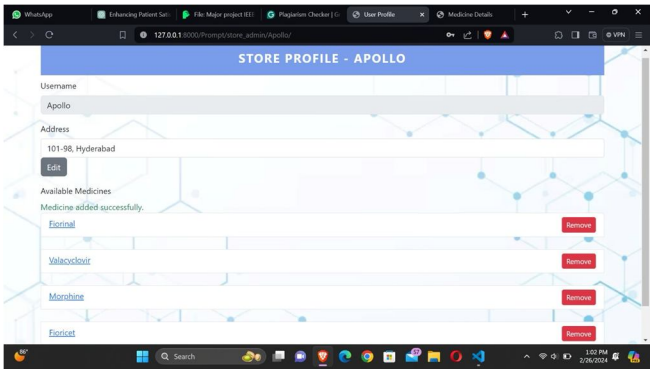


Figure 2 - Management interface

IV. RESULTS AND DISCUSSION

Following the split criterion, the preprocessed data is divided into 60% for training and 40% for testing. Subsequently, the dataset undergoes evaluation using five distinct machine learning classifiers: Decision Tree (DT), Random Forest (RF), Light Gradient Boosting Machine (LGBM), CatBoost, and Naive Bayes (NB). The accuracy of each classifier is computed and presented in the results. The classifier exhibiting the highest accuracy is identified as the optimal choice.

Prediction Integration

The presented system is a Django web application designed to facilitate disease prediction based on symptoms provided by patients. Upon receiving a submission through a form, the system processes the symptoms by converting them into a binary vector representation where each element corresponds to a specific symptom. This vector is then fed into a pre-trained machine learning model, likely trained on a dataset associating symptoms with various diseases. The model predicts the most probable disease based on the input symptoms and calculates a confidence score for the prediction using predict_proba(). Additionally, the system recommends a type of doctor for consultation based on the predicted disease, autonomy, and trust in their interactions and transaction.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

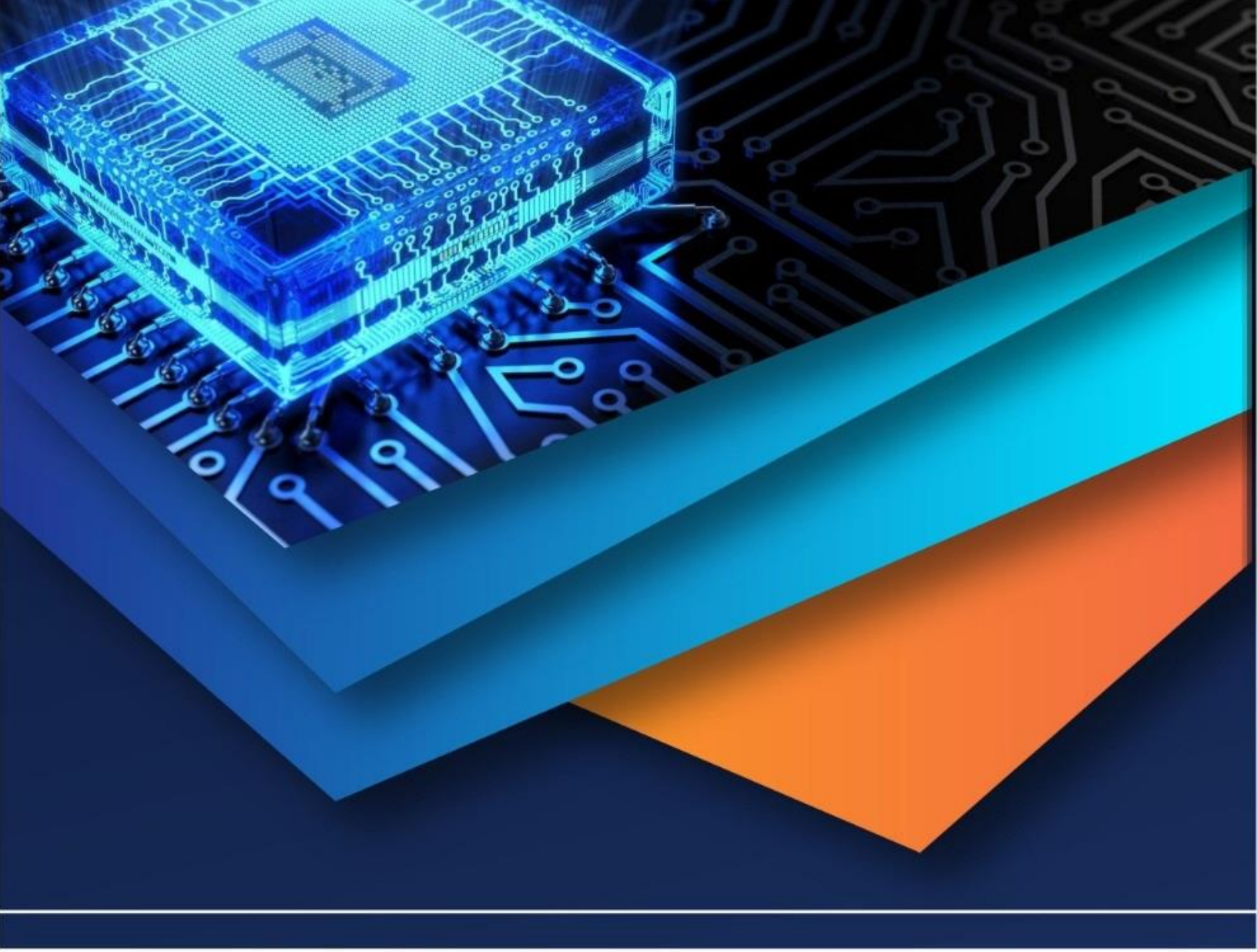
ALGORITHMS	ACCURACY %
Linear SVC with TF-IDF vectorization	87.492%
Naive Bayes	85.103%
RANDOM FOREST	89.004%
Stochastic Gradient Descent (SGD)	93.006%

V. CONCLUSION

Comments play an essential role in our daily decision-making processes, influencing choices such as online shopping or dining out. To assist in these decisions, we often turn to reviews for guidance. This study employs multiple machine learning classifiers, including logistic regression, perceptron, polynomial naive Bayes, ridge classifier, stochastic gradient descent, linear SVC, Arc reference, TF-IDF, trimmed trees, and random forests, to develop a consensus model.Word2Vec and manual methods are utilized, with LGBM and CatBoost serving as the classifiers for these techniques. Evaluation of these models is conducted using precision, recall, F1-score, accuracy, and AUC score metrics. Results indicate that TF-IDF paired with linear SVC achieves the highest accuracy at 91.5% compared to other models tested. Future research will focus on comparing different methodologies, exploring alternative n-gram approaches, and devising strategies to further enhance recommendation system performance. This study aims to contribute to the advancement of recommendation systems, ultimately aiding users in making informed decisions across various domains.

REFERENCES

[1] Wittich, C.M., Burkle, C.M., & Lanier, W.L. (2014). Medication errors: An overview for clinicians. Mayo Clinic Proceedings, 89(8), 1116-1125.
[2] Chen, M.R., & Wang, H.F. (2013). The reason and prevention of hospital medication errors. Practical Journal of Clinical Medicine.
Bartlett, J.G., Dowell, S.F., Mandell, L.A., File, T.M. Jr., Musher, D.M., & Fine, M.J. (2000). Practice guidelines for the management of community-acquired pneumonia in adults. Infectious Diseases Society of America. Clinical Infectious Diseases, 31(2), 347-382. DOI: 10.1086/313954.
[3] Fox, Susannah & Duggan, Maeve. (2012). Health Online 2013. Pew Research Internet Project Report.
[4] IRNSS “Signal In Space Interface Control Document for Standard Positioning Service”, version 1.0, ISRO, June 2014.
[5] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” IEEE J. Sel. Areas Commun., Vol. 18, No. 3, pp. 535–547, Mar. 2000.
[6] Venkatesan, K.G.S. "Comparison of CDMA & GSM Mobile Technology." Middle-East Journal of Scientific Research, vol. 13, no. 12, 2013, pp. 1590–1594.
[7] Priya, P. Indira, and K.G.S. Venkatesan. "Finding the K-Edge connectivity in MANET using DLTRT." International Journal of Applied Engineering Research, vol. 9, no. 22, 2014, pp. 5898–5904.
[8] Praveena, Ms. J., and K.G.S. Venkatesan. "Advanced Auto Adaptive edge-detection algorithm for flame monitoring & fire image processing." International Journal of Applied Engineering Research, vol. 9, no. 22, 2014, pp. 5797–5802.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)