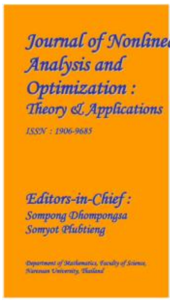
Journal of Nonlinear Analysis and Optimization
Vol. 15, Issue. 1, No.15 : 2024
ISSN : **1906-9685**



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

[1]*Reddi Amarr Kannthh Kosegi*,[2]Khaja Nazeeruddin,[3]Asha Kiran Annaladasu
[1,2,3]*Assistant Professor*
Department of CSE
Visvesvaraya College Of Engineering & Technology,Ibrahimpatnam,Telangana

ABSTRACT

The Drug Recommendation system based on sentiment analysis of drug reviews utilizes machine learning for sentiment analysis, a significant advancement in healthcare technology. This innovative approach helps medical professionals and patients make informed drug- related decisions by extracting and categorizing attitudes in medication evaluations. 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 specialist 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 Linear SVC using TF-IDF vectorization outperforms all other models with 93% accuracy. Index Terms—Drug, Recommender System, Machine Learning, NLP, Smote, Bow, TF-IDF, Word2Vec, Sentiment.

I. INTRODUCTION

With the number of coronavirus cases growing exponentially, the nations are facing a shortage of doctors, particularly in rural areas where the quantity of specialists is less compared to urban areas. A doctor takes roughly 6 to 12 years to procure the necessary qualifications. Thus, the number of doctors can't be expanded quickly in a short time frame. A Telemedicine framework ought to be energized as far as possible in this difficult time.

Clinical blunders are very regular nowadays. Over 200 thousand individuals in China and 100 thousand in the USA are affected every year because of prescription mistakes. Over 40% medicine, specialists make mistakes while prescribing since specialists compose the solution as referenced by their knowledge, which is very restricted. Choosing the top-level medication is significant for patients who need specialists that know wide-based information about microscopic organisms, antibacterial medications, and patients. Every day a new study comes up with accompanying more drugs, tests, accessible for clinical staff every day. Accordingly, it turns out to be progressively challenging for doctors to choose which treatment or medications to give to a patient based on indications, past clinical history.

With the exponential development of the web and the web-based business industry, item reviews have become an imperative and integral factor for acquiring items worldwide. Individuals worldwide become adjusted to analyse reviews and websites first before settling on a choice to buy a thing. While most of past exploration zeroed in on rating expectation and proposals on the E-Commerce field, the territory of medical care or clinical therapies has been infrequently taken care of. There has been an expansion in the number of individuals worried about their well-being and finding a diagnosis online. As demonstrated in a Pew American Research centre survey directed in 2013, roughly 60% of grown-ups searched online for health- related subjects, and around 35% of users looked for diagnosing health conditions on the web. A medication recommender framework is truly vital with the goal that it can assist specialists and help patients to build their knowledge of drugs on specific health conditions.

A recommender framework is a customary system that proposes an item to the user, dependent on their advantage and necessity. These frameworks employ the customers' surveys to break down their sentiment and suggest a recommendation for their exact need. In the drug recommender system, medicine is offered on a specific condition dependent on patient reviews using sentiment analysis and feature engineering. Sentiment analysis is a progression of strategies, methods, and tools for distinguishing and extracting emotional data, such as opinion and attitudes, from language. On the other hand, featuring engineering is the process of making more features from the existing ones; it improves the performance of models. This examination work separated into five segments: Introduction area which provides a short insight concerning the need of this research, Related works segment gives a concise insight regarding the previous examinations on this area of study, Methodology part includes the methods adopted in this research, The Result segment evaluates applied model results using various metrics, the Discussion section contains limitations of the framework, and lastly, the conclusion section.

II. LITERATURE SURVEY

With a sharp increment in AI advancement, there has been an exertion in applying machine learning and deep learning arXiv:2104.01113v2 [cs.IR] 5 Apr 2021 strategies to recommender frameworks. These days, recommender frameworks are very regular in the travel industry, e-commerce, restaurant, and so forth. Unfortunately, there are a limited number of studies available in the field of drug proposal framework utilizing sentiment analysis on the grounds that the medication reviews are substantially more intricate to analyse as it incorporates clinical wordings like infection names, reactions, a synthetic name that used in the production of the drug.

The study presents Galen-OWL, a semantic-empowered online framework, to help specialists discover details on the medications. The paper depicts a framework that suggests drugs for a patient based on the patient's infection, sensitivities, and drug interactions. For empowering Galen-OWL, clinical data and terminology first converted to ontological terms utilizing worldwide standards, such as ICD-10 and UNII, and then correctly combined with the clinical information.

Leilei Sun examined large scale treatment records to locate the best treatment prescription for patients. The idea was to use an efficient semantic clustering algorithm estimating the similarities between treatment records. Likewise, the author created a framework to assess the adequacy of the suggested treatment. This structure can prescribe the best treatment regimens to new patients as per their demographic locations and medical complications. An Electronic Medical Record (EMR) of patients gathered from numerous clinics for testing. The result shows that this framework improves the cure rate.

In this research, multilingual sentiment analysis was performed using Naive Bayes and Recurrent Neural Network (RNN). Google translator API was used to convert multilingual tweets into the English language. The results exhibit that RNN with 95.34% outperformed Naive Bayes, 77.21%.

The study is based on the fact that the recommended drug should depend upon the patient's capacity. For example, if the patient's immunity is low, at that point, reliable medicines ought to be recommended. Proposed a risk level classification method to identify the patient's immunity. For example, in excess of 60 risk factors, hypertension, liquor addiction, and so forth have been adopted, which decide the patient's capacity to shield himself from infection. A web- based prototype system was also created, which uses a decision support system that helps doctors select first-line drugs.

Xiaohong Jiang et al. examined three distinct algorithms, decision tree algorithm, support vector machine (SVM), and backpropagation neural network on treatment data. SVM was picked for the medication proposal module as it performed truly well in each of the three unique boundaries - model exactness, model proficiency, model versatility. Additionally, proposed the mistake check system to ensure analysis, precision and administration quality. Mohammad Mehedi Hassan et al. developed a cloud assisted drug proposal (CADRE). As per patients' side effects, CADRE can suggest drugs with top-N related prescriptions. This proposed framework was initially founded on collaborative filtering techniques in which the medications are initially bunched into clusters as indicated by the functional description data. However, after considering its weaknesses like computational, costly, cold start, and information sparsity, the model is shifted to a cloud-helped approach using tensor decomposition for advancing the quality of experience of medication suggestion.

Considering the significance of hashtags in sentiment analysis, Jiu-data. However, after considering its weaknesses like computational, costly, cold start, and information sparsity, the model is shifted to a cloud-helped approach using tensor decomposition for advancing the quality of experience of medication suggestion. Considering the significance of hashtags in sentiment analysis, Jiugang Li et al. constructed a hashtag recommender framework that utilizes the skip-gram model and applied convolutional neural networks (CNN) to learn semantic sentence vectors. These vectors use the features to classify hashtags using LSTM RNN. Results depict that this model beats the conventional models like SVM, Standard RNN. This exploration depends on the fact that it was undergoing regular AI methods like SVM and collaborative filtering techniques; the semantic features get lost, which has a vital influence in getting a decent expectation.

III. SYSTEM ANALYSIS

3.1 PROPOSED METHODOLOGY

The dataset used in this research is Drug Review Dataset (Drugs.com) taken from the UCI ML repository [4]. This dataset contains six attributes, name of drug used (text), review (text) of a patient, condition (text) of a patient, useful count (numerical) which suggest the number of individuals who found the review helpful, date (date) of review entry, and a 10-star patient rating (numerical) determining overall patient contentment. It contains a total of 215063 instances. Fig 3.1 shows the proposed model used to build a medicine recommender system. It contains four stages, specifically, Data preparation, classification, evaluation, and Recommendation.

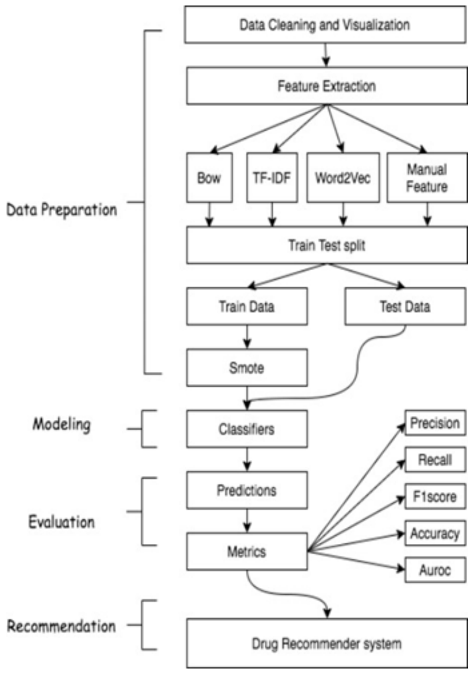


Fig .1 Flowchart of the proposed Fig model

3.1 Modules

3.1.1 Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Browse and Train & Test Data Sets , View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Prediction Of Drug Recommendation Type, View Drug Recommendation, Type Ratio Download Trained Data Sets, View Drug

Recommendation, Type Ratio Results, View All Remote Users.

3.1.2 View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

3.1.3 Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT DRUG RECOMMENDATION TYPE, VIEW YOUR PROFILE.

IV. OUTPUT SCREENS

Fig .2 HOME PAGE

Fig 3. ADMIN PAGE

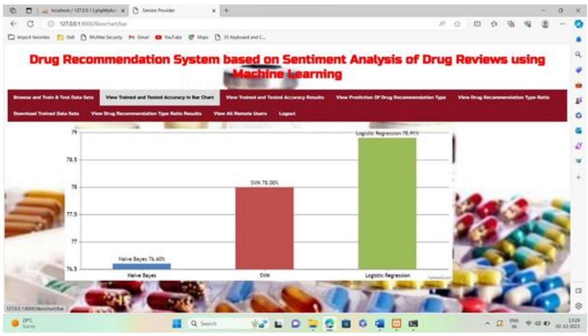


Fig 4. VIEW TRAINED AND ACCURACY IN BAR GRAPH

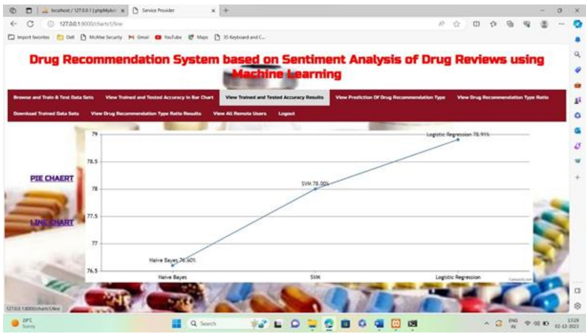


Fig 5. VIEW TRAINED AND TESTED ACCURACY RESULTS

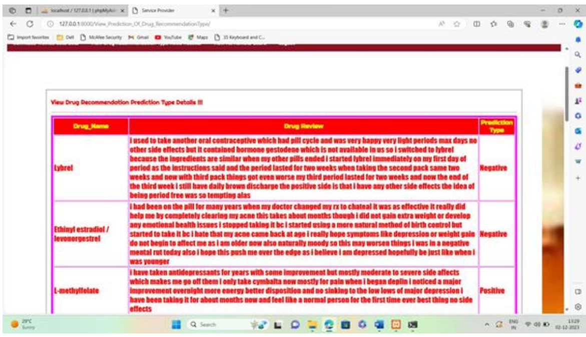


Fig 6. VIEW PREDICATION OF DRUG RECOMMENDATION TYPE

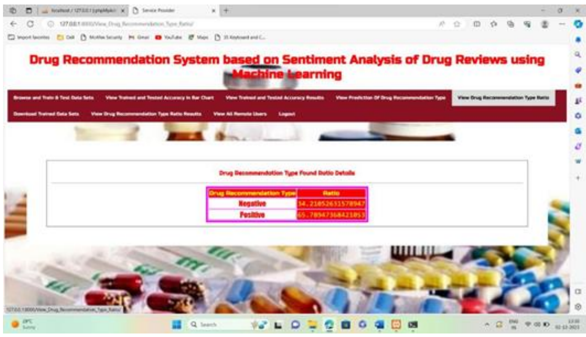


Fig 7. VIEW DRUG RECOMMENDATION TYPE RESULTS

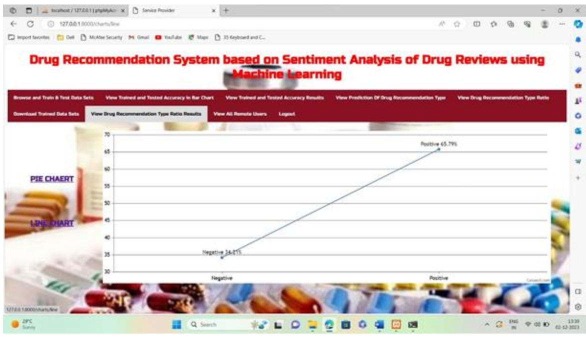


Fig 8. VIEW DRUG RECOMMENDATION TYPE RATIO RESULTS

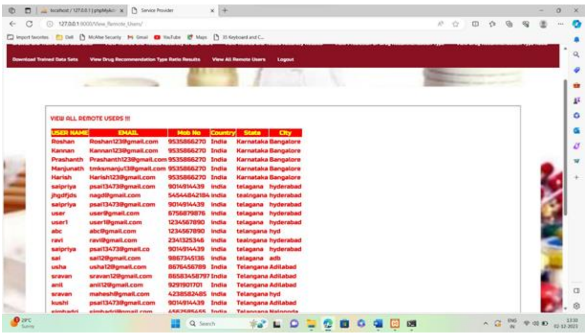


Fig 9. VIEW ALL REMOTE USERS

Fig 10. REGISTRATION PAGE

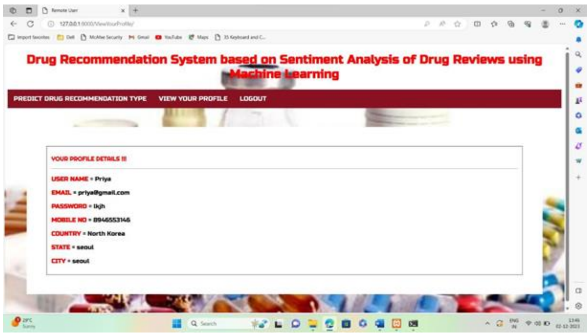


Fig 11. USER PROFILE

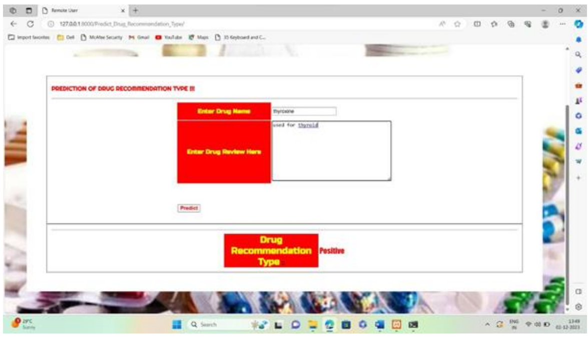


Fig 12. RECOMMENDATION OF DRUG WITH RESULTS

V. CONCLUSION

Reviews are becoming an integral part of our daily lives; whether go for shopping, purchase something online or go to some restaurant, we first check the reviews to make the right decisions. Motivated by this, in this research sentiment analysis of drug reviews was studied to build a recommender system using different types of machine learning classifiers, such as Logistic Regression, Perceptron, Multinomial Naive Bayes, Ridge classifier, Stochastic gradient descent, Linear SVC, applied on Bow, TF-IDF, and classifiers such as Decision Tree, Random Forest, Lgbm, and Cat boost were applied on Word2Vec and Manual features method. We evaluated them using five different metrics, precision, recall, f1score, accuracy, and AUC score, which reveal that the Linear SVC on TF-IDF outperforms all other models with 93% accuracy. On the other hand, the Decision tree classifier on Word2Vec showed the worst performance by achieving only 78% accuracy. We added best-predicted emotion values from each method, Perceptron on Bow (91%), Linear SVC on TF-IDF (93%), LGBM on Word2Vec (91%), Random Forest on manual features (88%), and multiply them by the normalized useful Count to get the overall score of the drug by condition to build a recommender system. Future work involves comparison of different oversampling techniques, using different values of n- grams, and optimization of algorithms to improve the performance of the recommender system.

FUTURE SCOPE

The goal of a recommender system is to generate meaningful recommendations to a collection of users for items or products that might interest them. Drug recommendation systems are one area being studied in this field. The goal of drug recommendation systems is to provide a list of relevant drugs based on the patient's disease conditions. Recommendation systems rely heavily on data, and the quality of the data can significantly impact the accuracy and effectiveness of the recommendations. Poor quality data, such as incomplete or inaccurate user data, can lead to irrelevant or even incorrect recommendations These frameworks employ the customers' surveys to break down their sentiment and suggest a recommendation for their exact need. In the drug recommender system, medicine is offered on a specific condition dependent on patient reviews using sentiment analysis and feature engineering.

REFERENCE

1. https://textblob.readthedocs.io/en/dev/.

2. https://archive.ics.uci.edu/ml/datasets/Drug%2BReview%2BDataset%2%2528Drugs.com%2529#
3. https://www.mohfw.gov.in/pdf/Telemedicine.pdf
4. Wittich CM, Burkle CM, Lanier WL. Medication errors: an overview for clinicians. Mayo Clin Proc. 2014 Aug;89(8):1116-25.
5. Bartlett JG, Dowell SF, Mandell LA, File TM Jr, Musher DM, Fine MJ. Practice guidelines for the management of community acquired pneumonia in adults. Infectious Diseases Society of America. ClinInfect Dis. 2000 Aug;31(2):347-82. doi: 10.1086/313954. Epub 2000 SePMID: 10987697; PMCID: PMC7109923.
6. Fox, Susannah & Duggan, Maeve. (2012). Health Online 2013.Pew Research Internet Project Report.
7. van der Maaten, Laurens & Hinton, Geoffrey. (2008). Viualizing data using t- SNE. Journal of Machine Learning Research. 9. 2579-2605.
8. N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique, 2011, Journal Of Artificial Intelligence Research, Volume 16, pages 321-357,2002; arXiv:1106.1813. DOI: 10.1613/jair.953.
9. Z. Wang, C. Wu, K. Zheng, X. Niu and X. Wang, ”SMOTETomek Based Resampling for Personality Recognition,” in IEEE Access vol.7, pp. 129678- 129689, 2019, doi: 10.1109/ACCESS.2019.2940061.