

E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com

Drug Recommendation System Based on Sentiment Analysis of Drug Reviews Using Machine Learning

Bhoomika Dubba¹, T. Sowmya Sri², M. Sowbhagya Pradhaini³, DR. KC Sreedhar⁴

^{1,2,3,4}Department of CSE, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

Abstract

People are self-medicating more during the COVID-19 pandemic because they can't get to good medical tools. This makes their health situations worse. This study suggests a drug suggestion system that uses machine learning and emotion analysis of patient reviews to make the job of healthcare workers easier. We use different vectorization methods, like Bag of Words (BoW), TF-IDF, Word2Vec, and Manual Feature Analysis, to guess how people feel about certain diseases and suggest the best drugs for them. We use classification methods, such as LinearSVC, to rate emotions based on their accuracy, F1-score, precision, and AUC score. The results show that LinearSVC with TF-IDF vectorization works well, as it achieved 93% accuracy, which was better than other models. By making drug suggestions, this system aims to make it easier for people to get medical care when they need it most, especially during emergencies.

Keywords: Drug, Recommender System, Machine Learning, NLP, Smote, Bow, TF-IDF, Word2Vec, Sentiment analysis

INTRODUCTION

With the COVID-19 outbreak spreading around the world, it's become clearer how hard it is on healthcare systems around the world. The critical lack of doctors has hit a critical level [1], especially in rural areas where there are fewer medical workers than in cities. The fact that medical school usually lasts between 6 and 12 years makes this problem even worse, and it will be impossible to quickly hire more doctors in the short run. Because of these problems, promoting telemedicine systems seems like a very important way to make healthcare more accessible during this crisis [1]. Medical mistakes, like writing the wrong medication, put patients' safety and health at great risk. Shockeningly, data show that medicine mistakes hurt over 200,000 people in China and 100,000 people in the US every year [2, 3]. A lot of these mistakes happen when doctors write prescriptions because they often depend on their limited knowledge and experience, which leads to wrong prescriptions [3]. Choosing the right medicine is especially important for people who need treatments that require a lot of knowledge about germs, drugs, and each patient's unique traits [3].Adding to the problems healthcare workers face is the fact that medical knowledge is always growing. Every day, new drugs, treatments, and testing tools are released. So, doctors are having a harder time figuring out which treatment or drug is best for each patient based on their symptoms and



E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com

medical background. Along with these problems in healthcare, the growth of the internet and e-commerce has changed the way people buy things. Product reviews are now an important part of decision-making all over the world. Before buying something, people always look at reviews and other web sites. This is something that has become normalized in the culture of global consumers. Previous study has mostly been about predicting scores and making suggestions in the e-commerce space. However, the use of similar methods in healthcare, especially in clinical therapy settings, is still not well studied. But there is a growing trend of people looking for health information and treatments online. In 2013, the Pew Research Center did a study that found that about 60% of people in the US looked for health-related information online, and about 35% tried to figure out what was wrong with them by using online tools [5]. This trend shows how important it is to make sure that healthcare information online is accurate and easy to find, including suggestions for medicines that are best for certain health problems. A medicine recommender system is seen as an important way to deal with these problems because it helps both doctors and people figure out which drugs are best for them. Using methods like mood analysis and feature engineering, this kind of system looks at patient reviews to figure out how they feel and pullout useful information that can be used to suggest medications. A lot of different methods and tools are used in sentiment analysis to find and pullout emotional information from text, like views and emotions [7]. Adding more features to current data through feature engineering improves the forecasting power of models, which is how sentiment analysis works with it [7]. Using machine learning and natural language processing, a medicine recommender system could make it easier for people to choose the right drugs, make the job of healthcare workers easier, and give patients more information about their treatment choices. In the parts that follow, we'll talk more about how such a system could be made and how it might be used, focusing on how it might improve healthcare service and patient results.

LITERATURE SURVEY

The study on drug suggestion systems, mood analysis, and healthcare informatics comes from many fields, such as medicine, computer science, and data analytics. This literature review gives an outline of important studies, methods, and technologies that can be used to create and use drug suggestion systems that use machine learning to analyze the mood of drug reviews. Chen and Wang (2013) explain how common medical drug mistakes are, what causes them, and how to stop them [3]. Their study shows how important it is to fix drug mistakes, which put patients' health and safety at great risk. Figuring out why these kinds of mistakes happen is important for making good plans to stop them from happening again. This shows how important it is to use technology to lower these risks. It is very important to use datasets when training and testing machine learning models for drug suggestion systems. The UCI Machine Learning Repository hosts the Drug Review Dataset, which is a useful tool for academics in this area [4]. This set of data includes reviews and scores from patients for different drugs. It makes it easier to create and test mood analysis algorithms for drug mood analysis Fox and Duggan (2013) look at the patterns and trends of how people in the US use the internet for health-related reasons [5]. Their study shows that more and more people are looking for health information online, and a lot of people use the internet to figure out what their health problems are. This trend shows how important it is to make sure that reliable and easy-to-find health information is available online, such as personalized medicine suggestions. When healthcare workers are making decisions, clinical standards and best practices are very helpful. Bartlett et al. (2000) published practice guidelines for treating people who have contracted pneumonia in the community. These guidelines were backed by the Infectious Diseases Society of America [6]. These standards give



E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com

suggestions based on evidence for choosing the right medicines. They stress how important it is to make decisions based on good information in clinical practice. Ontology-based methods give drug guidance systems an organized context that lets them use language reasoning and knowledge modeling. A group of researchers called Doulaverakis et al. (2012) came up with Galen OWL, an ontology-based method for finding drug suggestions [9]. Galen OWL uses ontological knowledge models to make it easier to find appropriate drug suggestions based on medical knowledge and factors that are unique to each patient Some machine learning methods, like probabilistic aspect mining and automatic treatment routine development, have been used to look at drug reviews and suggest individual treatment plans. Tekade and Emmanuel (2016) describe a statistical aspect mining method for figuring out what drug reviews mean and how good they are [8]. Their way lets you pull out important parts of drug reviews, which makes it easier to figure out how people feel and how well a medicine works. Sun et al. (2016) describe a method for automatically creating and recommending treatment plans based on data [10]. By looking at information about patients and how well treatments work, their system makes unique treatment plans based on each patient's needs and medical history. This method shows how machine learning algorithms can be used to make better treatment choices and help patients in clinical practice. Natural language processing (NLP) and other sentiment analysis methods have been used to look at drug reviews and pullout mood data. Goel et al. (2018) use NLP methods to look at the mood of Twitter data in multiple languages [11]. Their work shows that NLP can effectively extract emotion from social media data, giving us useful information about how people feel about healthcare issues and what they think about them. Shimada et al. (2005) suggest a way to help people with infectious diseases get the right drugs [12]. Based on standards based on proof, their method uses information about the patient, such as clinical signs and lab results, to suggest the right medicines. This study shows how important it is for drug guidance systems to use both professional information and data-driven methods to help people make better decisions in hospital situations. In conclusion, the literature review shows that drug suggestion systems involve many fields, such as medicine, computer science, and data analytics. Researchers have come up with a lot of new ways to help healthcare professionals make decisions and improve patient results. For example, they have figured out how to use machine learning and natural language processing (NLP) to analyze mood.

METHODOLOGY

Proposed Work

The suggested work aims to use mood analysis and feature engineering to create a drug recommendation system. Sentiment analysis will be used to pull out emotional data, like thoughts and attitudes, about certain medicines and how well they treat certain diseases by looking at reviews written by patients. Feature engineering will work with mood analysis to improve the models' ability to predict the future by adding more features from current data. With this unified method, doctors will be able to suggest medicines that are best for each patient's wants and health problems. The method will be made to help doctors make better choices about what drugs to prescribe, which will eventually lead to better patient results and happiness. Performance measures like accuracy, precision, memory, and F1-score will be used to judge how well the suggested system works at making accurate and useful drug suggestions



System Architecture



Fig. 1: Proposed Architecture

The system design is made up of parts for providing services, viewing datasets, training, and showing correctness. The features that users can access are reading and training datasets, seeing accurate results in the form of bar charts, registering, and logging in. The method lets you guess what kind of drug recommendations will be made and see details of users. Users can also see forecasts, look at drug suggestion type rates, download training datasets, and watch what other users are doing from afar. The design makes it easy for users to connect with the system, giving researchers a full tool for analyzing drug recommendations. The system improves the user experience and helps people make better healthcare decisions by using easy-to-understand tools and fast data processing.

c) Modules:

i) Remote User:

1. Registration and Login: People who are not in the same room as the system can register and log in to use its features. When users sign up, they give the site the information it needs and make login passwords. After signing up, users can safely use their passwords to log in and access the system's features and services. This process lets faraway users prove who they are and get approved access to the system. This gives them a safe, customized experience that meets their needs.

2. Predict Drug Recommendation Type: People who are not in the same room can use the method to guess the kinds of drug recommendations. Using machine learning algorithms and mood analysis, the system looks at reviews from patients and suggests medicines that are best for certain conditions. When users enter important information, the system uses it to make correct drug suggestions. This feature helps online users make smart choices about which medications to take, which improves patient happiness and healthcare results.

3. View Your Profile: Users who are not in the same room as the computer can look at their profiles in the system and see personal information and settings. Users can go to their biography page after signing in to see personal information, activity records, and saved preferences, among other things. This tool lets users change the settings for their account, keep track of how they use the system, and make sure their



profiles are always up to date. The method improves user interaction and modification choices by making user accounts easy to get to.

ii) Service Provider:

1. Registration & Login: The service source lets you register and log in. Users join by giving the required information and making a login name and password. Users can access the platform's services and features by safely logging in with their passwords after registering. This gives them a unique and safe experience that fits their needs.

2. Browse, Train, and Test Data Sets: The service provider makes it easy to browse, train, and test datasets. Users can look through files that are available and are useful for drug advice research. They can use these datasets to train and test machine learning models, which makes it easier for them to create and test algorithms for drug suggestion systems.

3. View Trained and Tested Accuracy in Bar Chart: The service provider gives users a bar chart that they can use to see trained and tested accuracy. Accuracy measures like precision, recall, F1-score, and AUC score can be shown visually to users, giving them a full picture of how well drug guidance systems models are working.

4. View Trained and Tested Accuracy Results: Users can see trained and tested accuracy results from the service provider. Users can see precise, recall, F1-score, accuracy, and AUC scores, among other measures and reviews of model success. With this feature, users can check how well their machine learning models for drug advice systems are working.

5. View Prediction of Drug Recommendation Type: The service provider lets users look at the estimates of drug suggestion types. Users can access the system's suggestions for good medicines based on patient reviews and other relevant data. These predictions are made using machine learning algorithms and mood analysis techniques. In healthcare settings, this trait helps people make better decisions.

6. View Drug Recommendation Type Ratio: The service provider makes it easy to look at the drug suggestion type ratio. Users can get information about how suggested medicines are spread across different groups or situations. Understanding prescribing trends and tastes is easier with this feature's useful data. It helps make drug advice

7. Download Trained Data Sets: You can download trained datasets from the service provider. In the system, users can see and download files that were used to teach machine learning models. With this tool, users can use the datasets for outside study, analysis, or training of their own models.

8. View Drug Recommendation Type Ratio Results: The service provider lets users see the drug recommendation type ratio results. Users can see a lot of information and statistics about how the suggested medicines are spread out among different groups or situations. This function helps people understand treatment trends better and helps people make decisions in healthcare settings.

9)View All Remote Users: The service provider lets you see all remote users who are using the system. Administrators or other approved users can see a full list of all remote users, complete with biographies and a log of all their activities. This function gives information about how engaged users are and makes it easier to handle and keep an eye on how the system is being used.

iii) Algorithms:

Naïve Bayes:

Naive Bayes classifier is a probabilistic machine learning model based on Bayes' theorem. It assumes independence between features and calculates the probability of a given input belonging to a particular class. It's widely used in text classification, spam filtering, and recommendation systems.



E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com

```
Input:
           Training dataset T,
           F = (f_1, f_2, f_3, ..., f_n)
                                // value of the predictor variable
           in testing dataset.
Output:
           A class of testing dataset.
Step:
     1. Read the training dataset T:
     2. Calculate the mean and standard deviation of the
           predictor variables in each class;
     3. Repeat
               Calculate the probability of f_i using the gauss
               density equation in each class;
           Until the probability of all predictor variables (f1, f2,
           f<sub>3</sub>,.., f<sub>n</sub>) has been calculated.
     4.
          Calculate the likelihood for each class;
     5.
          Get the greatest likelihood;
```

SVM:

A support vector machine (SVM) is a machine learning algorithm that uses supervised learning models to solve complex classification, regression, and outlier detection problems by performing optimal data transformations that determine boundaries between data points based on predefined classes, labels, or outputs.



Logistic Regression:

Logistic regression is a supervised machine learning algorithm that accomplishes binary classification tasks by predicting the probability of an outcome, event, or observation. The model delivers a binary or dichotomous outcome limited to two possible outcomes: yes/no, 0/1, or true/false.



Decision Tree:

A decision tree algorithm is a machine learning algorithm that uses a decision tree to make predictions. It follows a tree-like model of decisions and their possible consequences. The algorithm works by recursively splitting the data into subsets based on the most significant feature at each node of the tree.



E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com



SGD Classifier:

Stochastic Gradient Descent (SGD) is a simple yet efficient optimization algorithm used to find the values of parameters/coefficients of functions that minimize a cost function. In other words, it is used for discriminative learning of linear classifiers under convex loss functions such as SVM and Logistic regression.

```
Input (Training Set S \subset R, Learning Rate \eta,Reg. factor \lambda, n° of latent factors K)

Randomly initialize matrices P and Q;

n=0

while not(convergence) do

Randomly shuffle observed entries in S;

for each (u, i) \in S do

e_{u,i} = \left(r_{ui} - \sum_{k=1}^{K} p_{uk} \cdot q_{ki}\right)

for each q \in \{1 \dots k\} do \hat{p}_{uk} = p_{uk} + \eta \cdot (e_{u,i} \cdot q_{ki} - \lambda \cdot p_{uk})

for each q \in \{1 \dots k\} do \hat{q}_{ki} = q_{ki} + \eta \cdot (e_{u,i} \cdot p_{uk} - \lambda \cdot q_{ki})

for each q \in \{1 \dots k\} do p_{uk} = \hat{p}_{uk} and q_{ki} = \hat{q}_{ki}

end for

n=n+1

end while

Output (P, Q)
```

EXPERIMENTAL RESULTS



Fig 2 Home Page



E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com



Fig 3 Register page



Fig 4 Login Page



Fig 5 View Your Profile Page



Fig 6 Predict of Drug Recommendation Type Screen



E-ISSN: 2582-2160 • Website: www.ijfmr.com • Email: editor@ijfmr.com



Fig 7 Output Screen



Fig 8 View All Remote Users Screen



Fig 9 Comparison Graph



Fig 10 View Trained and Tested Accuracy In Bar Chat Screen





E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com



Fig 11 View Trained and Test Accuracy Result Screen



Fig 12 View Predction of drug Recommendation Type Screen



Fig 13 View Drug Recommendation Type Ratio Screen



Fig 14 View Drug Recommendation Type Ratio Details Screen



E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com



Fig 15 Download Trained Data Screen



Fig 16 Output Screen



Fig 17 View Drug Recommendation Type Ratio Results Screen



Fig 18 Click on Logout

CONCLUSION

In the end, this study shows that analyzing the tone of drug reviews is a good way to build a strong drug recommendation system. Using different machine learning algorithms and vectorization methods, we tested how well different models could predict how people would feel and suggest medicines for certain



E-ISSN: 2582-2160 • Website: <u>www.ijfmr.com</u> • Email: editor@ijfmr.com

diseases. Our results show that Linear SVC using TF-IDF was the most accurate, with a score of 93%, beating out all other models. On the other hand, the Decision Tree algorithm on Word2Vec did the worst, with a success rate of only 78%.

We took the best-predicted feeling values from each method and increased them by the standardized useful Count to get a total score for drugs by condition. This made the recommender system work better. This method gives a complete plan for suggesting medicines that are best for each patient's wants and health problems.

Some ideas for future study are to look into different oversampling methods, change n-gram numbers, and make algorithms work better so that the recommender system works even better. By keeping the system updated and improved, we can help people make better choices about their health care and, in the end, improve patient results.

FUTURE SCOPE

We plan to improve the drug recommender system's features and usefulness in the future by looking into a number of different options. First, we're going to compare and contrast different oversampling methods to fix any possible biases in the dataset. This will make the system better at handling a wide range of reviews. Also, trying out different n-gram values during the vectorization process can help you figure out the best way to describe written data, which can improve the accuracy of mood analysis and recommendations even more.

A big part of improving the speed and usefulness of the recommender system will also be making the algorithms work better. We want to improve the system's ability to guess and make good suggestions by fine-tuning model parameters and looking into more advanced optimization methods. In the end, these future efforts will help build a stronger and more reliable drug recommender system, giving healthcare workers and people useful information to help them make smart choices.

REFERENCES

- 1. Shaik Nawab Mudassir, Ravichandran M, "Enhancing Migraine Diagnosis and Classification with TabNet: A Data-Driven Approach", 2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp.679-684, 2024.
- 2. Hasan Faraz Khan, Noor Fatima, "Analyzing Customer Sentiment in Drug Reviews Using Natural Language Processing", 2023 7th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), pp.1-5, 2023.
- 3. Bagiyalakshmi N, T. Veeramakali, "Exploring the Impact of Herbal Remedies on Autism Spectrum Disorder Using AGLSTM Analysis", 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), pp.1-5, 2023.
- 4. Priyanka Jangde, Kamal Kumar Sethi, Narendra Pal Singh, Praveen Bhanodia, Durgesh Mishra, "A Comprehensive Machine Learning Based Ensemble Model for Prediction of Possible Diseases and Recommending Medicines", *2023 IEEE International Conference on ICT in Business Industry & Government (ICTBIG)*, pp.1-10, 2023.
- 5. T M Aruna, T L Nethravathi, M Muniraju, "Designing a Deep Learning Model for recommendation of Drugs on Sentiment Analysis", 2023 IEEE 3rd Mysore Sub Section International Conference (MysuruCon), pp.1-6, 2023.



- 6. Josue Uwimana, "Intelligent Cognitive Clinician", 2023 International Conference on the Cognitive Computing and Complex Data (ICCD), pp.185-190, 2023.
- Bharathi Mohan G, Prasanna Kumar R, Elakkiya R, Prasanna Kumar B G, "Medical Recommendations: Leveraging CRNN with Self-Attention Mechanism for Enhanced Systems", 2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT), pp.1-6, 2023
- Turan Goktug Altundogan, Mehmet Karakose, Sümeyye Yılmazer, Eray Hanoğlu, Sedef Demirel, "Sentiment Analysis for Patient Reviews in Hospitals by CNN and LSTM Neural Networks Using Pretrained Word Embeddings", 2023 Innovations in Intelligent Systems and Applications Conference (ASYU), pp.1-4, 2023
- 9. Sinchana N M, Prasanna Kumar K R, Santhosh B J, "Model-based Filtering Techniques for Recommendation Systems in Healthcare Domain", 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), pp.978-983, 2023.
- 10. Ganesh C, Arulpriya R, Keerthana V, Kiruthika D, Subasri M, "Deep Learning Based Sentiment Analysis on Drug Reviews", 2023 IEEE 2nd International Conference on Industrial Electronics: Developments & Applications (ICIDeA), pp.401-406, 2023.
- 11. Gugum Gumilar, Eka Budiarto, Maulahikmah Galinium, Charles Lim, "Feature Extraction Methods and Classification for Malware Incident News", 2023 IEEE International Conference on Cryptography, Informatics, and Cybersecurity (ICoCICs), pp.115-120, 2023.
- Huang Huang, Adeleh Asemi Zavareh, Mumtaz Begum Mustafa, "Sentiment Analysis in E-Commerce Platforms: A Review of Current Techniques and Future Directions", *IEEE Access*, vol.11, pp.90367-90382, 2023.
- 13. Rezaul Haque, Piyush Kumar Pareek, Md Babul Islam, Ferdaus Ibne Aziz, Swakshar Das Amarth, Katura Gania Khushbu, "Improving Drug Review Categorization Using Sentiment Analysis and Machine Learning", 2023 International Conference on Data Science and Network Security (ICDSNS), pp.1-6, 2023.
- 14. Pritam Bhuiya, Rohit Malik, Aritra Rana, Md. Ashifuddin Mondal, "Personalized Medicine Recommendation and Disease Flagging Model Based on User's Previous Orders", 2023 World Conference on Communication & Computing (WCONF), pp.1-6, 2023.
- 15. Kritib Bhattarai, Sivaraman Rajaganapathy, Trisha Das, Yejin Kim, Yongbin Chen, Qiying Dai, Xiaoyang Li, Xiaoqian Jiang, Nansu Zong, "Learning Physician's Treatment for Alzheimer's Disease based on Electronic Health Records and Reinforcement Learning", 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI), pp.525-526, 2023
- 16. Oswalt Manoj S, Catherine Chandralekha T, Dharshini R M, Hemadharsini P, "Sentirec A Sentiment Analysis with Recommendation System", 2023 8th International Conference on Communication and Electronics Systems (ICCES), pp.694-701, 2023.
- 17. ShakthiMurugan K H, C. Gnanaprakasam, M. Swarna, R. Geetha, N V Krishnamoorthy, "A Novel Design of Deep Learning Assisted Drug Recommendation Model using Sentimental Inspection", 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), pp.1-6, 2023.
- 18. Mausumi Goswami, Uday Sai Bandi, "Sentiment and Emotion Analysis of Significant Diseases in India and Russia", 2023 IEEE International Conference on Contemporary Computing and Communications (InC4), vol.1, pp.1-6, 2023.



- 19. C. Silpa, B Sravani, D Vinay, C Mounika, K Poorvitha, "Drug Recommendation System in Medical Emergencies using Machine Learning", 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), pp.107-112, 2023.
- 20. Deepa D, T. Dhiliphan Rajkumar, "Navigating the Healthcare Landscape with Recommendation Systems: A Survey of Current Applications and Potential Impact", *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, pp.818-823, 2023.
- 21. Dr Ratna Patil, Divyanshu Shukla, Abhijeet Kumar, Yutika Rajanak, Yadvendra Pratap Singh, "Machine Learning for Sentiment Analysis and Classification of Restaurant Reviews", 2022 3rd International Conference on Computing, Analytics and Networks (ICAN), pp.1-5, 2022.
- 22. Theresa Olubukola Omodunbi, Grace Egbi Alilu, Rhoda Nsikanabasi Ikono, "Drug Recommender Systems: A Review of State-of-the-Art Algorithms", 2022 5th Information Technology for Education and Development (ITED), pp.1-8, 2022.
- 23. Mariam Gawich, Marco Alfonse, "A Proposed Model for Drugs' Review Analysis and Adverse Drug Reaction Discovery", 2022 7th International Conference on Mathematics and Computers in Sciences and Industry (MCSI), pp.139-145, 2022.