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

**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.

**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 toplevel 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, pastclinical 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. Individualsworldwide become adjusted to analyze 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 center 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.

**EXISTING SYSTEM**

The existing system aims to address the challenges arising from the scarcity of medical resources during the COVID-19 pandemic by proposing a drug recommendation system based on sentiment analysis of drug reviews using machine learning techniques.

It leverages patient reviews to predict sentiment through various vectorization methods like Bag of Words (BoW), TF-IDF, Word2Vec, and Manual Feature Analysis. Classification algorithms are then employed to recommend the top drug for a given disease.

The evaluation metrics used include precision, recall, F1 score, accuracy, and AUC score. The study concludes that the LinearSVC classifier using TF-IDF vectorization achieves the highest accuracy of 93%.

This system aims to alleviate the burden on healthcare professionals by providing automated drug recommendations based on patient feedback.

The incorporation of various vectorization processes and classification algorithms demonstrates a thorough approach to the problem.

The reported performance metrics, particularly the 93% accuracy achieved by the LinearSVC classifier using TF-IDF vectorization, indicate promising results.

These systems often use static databases with predefined rules and may include some level of artificial intelligence to match symptoms with treatments.

**PROPOSED SYSTEM**

The proposed system utilizes machine learning techniques to analyze large volumes of patient reviews, which are collected from various medical forums and websites.

These reviews are processed using several vectorization techniques, namely Bag of Words (BoW), TF-IDF, Word2Vec, and Manual Feature Analysis, to convert the unstructured text data into structured form suitable for machine learning algorithms.

This transformation allows the system to capture and quantify the sentiments expressed by patients regarding their experiences with different medications.

For predicting sentiments, the system employs a range of classification algorithms, including LinearSVC, which has demonstrated superior performance in preliminary tests. Specifically, the LinearSVC model using TF-IDF vectorization achieved an impressive 93% accuracy, along with high scores in precision, recall, F1 score, and AUC score.

This high level of accuracy indicates that the model is very effective at interpreting the sentiments of drug reviews and can reliably identify patient sentiment towards various medications.

The proposed drug recommendation system based on sentiment analysis of drug reviews using machine learning sounds promising, especially in light of the challenges faced by the medical community during the COVID-19 pandemic.

By leveraging patient reviews and sentiment analysis, the system aims to provide valuable insights for medication recommendations, potentially easing the burden on healthcare professionals.

The use of various vectorization techniques and classification algorithms demonstrates a thorough approach to model development and evaluation. Achieving a 93% accuracy with the LinearSVC classifier using TF-IDF vectorization is impressive and indicates the system's effectiveness in predicting sentiments accuracy.

**LITERATURE REVIEW**

**[1]**Despite being a significant contributor to patient morbidity and death, medication mistake is sometimes an obscure and underestimated subject. [1]This article offers an overview of pharmaceutical errors with a focus on terminology and definitions, incidence, risk factors, avoidance techniques, disclosure, and legal ramifications for physicians in practice.[1] Any mistake that happens during the pharmaceutical use procedure is considered a medication error. The Institute of Medicine estimates that 1 in 131 outpatient and 1 in 854 inpatient deaths are related to medication mistakes.[1] Medication errors can be caused by a variety of factors, including medication-related ones (such as similar-sounding names and low therapeutic index), patient-related ones (such as poor renal or hepatic function, impaired cognition, and polypharmacy), and health care professional-related ones (such as cognitive biases and the use of abbreviations in prescriptions and other communications).[1] Medical board sanctions, legal lawsuits, criminal charges, and a loss of patient trust are some of the repercussions that doctors may experience following pharmaceutical errors. [1]There has been mixed effectiveness in preventing pharmaceutical errors through the use of information technology, improved drug labeling, and medication reconciliation, among other strategies. [1]Patients want quick, in-person admission of errors along with an apology and an explanation of measures being made to avoid them in the future. [1]Acquiring knowledge regarding drug errors could potentially improve the ability of medical practitioners to treat patients safely.

**[2]**These days, clinical errors are extremely common. Healthcare is one of the main focal points of the medical domain in the current digital era. People looking for pertinent information on health issues that they are worried about. [2] For this kind of knowledge, the Internet may be a terrific resource, but you must exercise caution to prevent yourself from finding false information. [2] These days, individuals cannot locate helpful information to improve their well-being on the Internet due to the massive amount of clinical data that is completely scattered throughout several websites. [2] The widespread grief causes many deaths in the medical sector. People began self-medicating due to the lack rather of first seeing a doctor, which made the health situation worse.[2] Numerous fields have found application for machine learning, and recent studiesand the rate and extent of automation development have recently grown. The introduction of a medication recommendation system that can greatly reduce the strain of specialists is the aim of this project. [2] In this work, we created a medication recommendation system that uses several vectorization processes, such as Bow, TF-IDF, Word2Vec, and Manual Feature Analysis, to predict sentiment based on patient reviews. This helps patients choose the optimal medication for their condition, as determined by various classification algorithms.[2] The predicted emotions were rated using precision, recall, f1score, accuracy, and area under the curve (AUC). The results show that, in comparison to the other models, the classifier Linear SVC with TF-IDF vectorization obtains the maximum accuracy.

**[3**] 35 percent of American adults claim to have at some point looked up specific medical conditions online in an attempt to determine whether they or someone else might have them. [3] These results are based on a nationwide poll conducted by the Internet & American Life Project at the Pew Research Center. We refer to people who looked up solutions online throughout this report as "online diagnosers." 46% of online diagnosers respond in the affirmative when asked if the information they obtained online made them believe they needed medical assistance. [3] Thirty-eight percent of online diagnosers said it was something they could handle at home, while eleven percent claim it was somewhere in the middle.[3] 35 percent of American people claim to have said that at some point.They intentionally went online to attempt to determine what kind of medical ailment they or someone else might have. These results are based on a nationwide poll conducted by the Internet & American Life Project at the Pew Research Center. We refer to people who looked up solutions online throughout this report as "online diagnosers." 46% of online diagnosers respond in the affirmative when asked if the information they obtained online made them believe they needed medical assistance. [3]Thirty-eight percent of online diagnosers said it was something they could handle at home, while eleven percent claim it was somewhere in the middle.

**[4]**Drug-drug and drug-disease interactions can be a challenging issue to manage because of the growing number of pharmaceuticals on the market and the continuous research being done in the pharmaceutical industry, which makes it challenging to find pertinent information. Despite the development of international standards like the UNII registration and the ICD-10 classification to facilitate effective knowledge sharing, medical personnel must stay up to speed in order to identify drug interactions before prescribing. [4] To address this issue, the application of Semantic Web technologies has been suggested in previous publications. Outcomes GalenOWL, a semantic-enabled online service, is presented in this work and can provide real-time drug-drug and drug-disease interaction finding. [4] To make this type of service possible, health data andt was necessary to convert language into ontological words and combine it with relevant medical knowledge. International standards, like the previously mentioned UNII and ICD-10, serve as the foundation for the common representation of medical data, and a rule base that utilizes these standards is used to describe medical knowledge regarding medication interactions. Along with presenting specifics of the system architecture, an overview of the challenges faced is provided. [4] A side-by-side analysis of the developed ontology-based system and a corresponding system built with a conventional business logic rule engine provides information about the benefits and limitations of each implementation. Final Thoughts Research has shown that utilizing Semantic Web technology is a suitable fit for creatingmedication recommendation frameworks.[4] Medical knowledge may be efficiently captured by ontologies, and knowledge about medication interactions can be captured and encoded via rule-based reasoning.

**[5**] As e-commerce has grown, more and more people are choosing to buy medications online due to convenience.[5] Nevertheless, buying medicines carelessly without the required prescription guidance is a big problem. [5] Our research presents a novel cloud-assisted drug recommendation system called CADRecommended, which is able to provide users with top-N relevant drugs based on their symptoms. [5] Using user collaborative filtering, we first cluster the medications into multiple groups in CADRE based on the functional description information. Then, we construct a basic personalized drug suggestion.[5] Next, taking into account the collaborative filtering algorithm's drawbacks, We suggest a cloud-assisted method for improving end-user Quality of Experience (QoE) of drug recommendation by modeling and describing the link of the user, symptom, and medicine using tensor decomposition. [5]This technique addresses issues like computing expense, cold start, and data sparsity. Lastly, an experimental research based on an actual dataset that was scraped from the Internet is used to assess the suggested methodology.

**CONCLUSION**

Reviews are becoming an integral part of our daily lives; whether go for shopping, purchase something online or go tosome 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, LinearSVC, applied onBow, TF-IDF, and classifiers such as Decision Tree, Random Forest, Lgbm, and Catboost were applied onWord2Vec 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%), LinearSVC on TF-IDF (93%), LGBM on Word2Vec (91%), Random Forest on manual features (88%), and multiply them by the normalized usefulCount 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.

**REFERENCES**

**[1]**Wittich CM, Burkle CM, Lanier WL. Medication errors: an overview for clinicians. Mayo Clin Proc. 2014 Aug;89(8):1116-25.

**[2]**CHEN, M. R., & WANG, H. F. (2013). The reason and prevention of hospital medication errors. Practical Journal of Clinical Medicine, 4.

**[3**] Fox, Susannah & Duggan, Maeve. (2012). Health Online 2013. Pew Research Internet Project Report.

**[4]**Doulaverakis, C., Nikolaidis, G., Kleontas, A. et al. GalenOWL: Ontology-based drug recommendations discovery. J Biomed Semant 3, 14 (2012). <https://doi.org/10.1186/2041-1480-3-14>.

**[5]**Zhang, Yin & Zhang, Dafang & Hassan, Mohammad & Alamri, Atif & Peng, Limei. (2014). CADRE: Cloud-Assisted Drug REcommendation Service for Online Pharmacies. Mobile Networks and Applications. 20. 348-355. 10.1007/s11036-014-0537-4.