Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023

# SENTIMENT ANALYSIS OF DRUG REVIEWS AND RECOMMENDATION USING MACHINE LEARNING

**Mr. S.Vamsi [1], Ms. Ramya Pedamalla [2]**

**#1 Assistant Professor In The Department Of AI & IT at DVR & Dr HS MIC College of Technology (Autonomous), Kanchikacherla, NTR District,A.P.**

**#2 MCA Student In The Department Of Computer Applications at DVR & Dr HS MIC College of Technology (Autonomous), Kanchikacherla, NTR District, A.P.**

***Abstract*—** Clinical blunders are very regular nowadays. In today's digital era healthcare is one among the major core areas of the medical domain. People trying to find suitable health-related information that they are concerned with. The Internet could be a great resource for this kind of data, however you need to take care to avoid getting harmful information. Nowadays, a colossal quantity of clinical information dispersed totally across different websites on the Internet prevents users from finding useful information for their well-being improvement. Many people in the medical community perish because of the widespread sorrow. As a result of the shortage, people started medicating themselves without first consulting a professional, worsening the health crisis. Machine learning has proven useful in many areas, and new research and development in the field of automation has recently increased in pace and scope. The goal of this research is to introduce a drug recommender system that can significantly lessen specialists' workload. In this study, we developed a medicine recommendation system that predicts sentiment based on patient reviews by employing a number of vectorization processes, including Bow, TF-IDF, Word2Vec, and Manual Feature Analysis, and thus aids in the selection of the best drug for a given disease as determined by a number of different classification algorithms. Precision, recall, f1score, accuracy, and area under the curve (AUC) were used to rate the anticipated emotions. The findings demonstrate that the classifier Linear SVC with TF-IDF vectorization achieves the highest accuracy compared to the other models.

## INTRODUCTION

The study is based on the fact that the recommended drug should depend upon the patient's capacity. Since the emergence of the corona virus, there has been a dramatic increase in the difficulty with which authorised clinical resources, such as doctors, nurses, diagnostic tools, and medications, may be obtained. Especially in rural areas, where there are fewer specialists than in urban areas, countries are experiencing a lack of doctors while the number of corona virus cases increases dramatically. Depending on the medical school you attend, it might take anywhere from six to twelve years to become a fully qualified doctor. Therefore, the number of medical professionals cannot be increased rapidly. In this trying time, the use of a Telemedicine framework should be promoted extensively. Nowadays, medical mistakes occur frequently. Each year, medication errors harm around 200 thousand people in China and 100,000 people in the United States. In over 40 percent of cases, doctors make mistakes when writing prescriptions because they tailor the treatment to the patient based on their own limited understanding. Patients in need of specialists who have extensive knowledge of microorganisms, antibacterial drugs, and patients are in a position to make informed decisions about which medication to use. Every day, more and more research becomes available, and with it, additional treatments and tests that can be used by clinical professionals. Accordingly, it becomes increasingly difficult for clinicians to decide which treatment or medications to prescribe a patient based on indications, past clinical history, and other factors. The proliferation of the internet and e-commerce websites has made product reviews a vital part of the buying process everywhere. People everywhere have gotten in the habit of doing some preliminary research in the form of online reviews and shopper comparison sites before making any major purchases. Most previous research has focused on rating expectation and proposals in the E-Commerce industry, but the realm of medical care or clinical remedies has been rarely attended to. The number of people seeking out a diagnosis for themselves or a loved one online has increased. It was shown in a 2013 poll by the Pew American Research Center that around 60% of adults looked for health-related topics online, and roughly 35% of users sought for diagnosing health disorders. There is a critical need for a medication recommender framework that can aid doctors and patients in expanding their understanding of the effects of medications on individual diseases. A recommender framework is a commonplace application that makes product suggestions based on the user's stated preferences and needs. These models utilise customer surveys to categorise responses and make tailored recommendations. Using sentiment analysis and feature engineering, the drug recommender system determines under what circumstances a given medication should be prescribed. The term "sentiment analysis" refers to a set of techniques for identifying and extracting linguistic expressions of emotion, such as opinion and attitude. Alternatively, Featuring engineering enhances model efficiency by creating new features from preexisting ones. There are five parts to this examination: There will be a place labelled "Introduction" where you can provide a brief overview of why this study Previous studies in this field are briefly discussed in the "Related Works" section, and the research techniques used herein are described in the "Methodology" section. The framework's constraints are presented in the Discussion part, and the results of the applied models are evaluated using the Evaluation section.

## LITERATURE REVIEW

**[1] 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, 80.32%.
**[2] Bartlett JG et al.** In more than 10 years since the last Community-Acquired Pneumonia (CAP) proposal from the American Thoracic Society (ATS) / Infectious Diseases Society of America, the process for creating guidelines has altered, and new clinical data have been created (IDSA). Due to the expansion of information regarding the diagnostic, treatment, and managerial decisions for the patient care with CAP, we purposefully limited the extent of this framework to cover judgments from the point of medical diagnosis of pneumonia to the end of antibiotic treatment and carry chest image processing.
**[3] Jiugang Li et al.** constructed a hashtag recommender framework that utilizes the skipgram 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.
**[4] Susannah et al.** In this research, a deep learning approach for health-based medical datasets is proposed. This approach automatically identifies what meal should be supplied to which person depending on the condition and other parameters like age, race, body weight, calories, fat, sodium, protein, fiber and cholesterol. The integration of deep learning and machine learning methods such as regression analysis, naive bayes, recurrent neural networks, multilevel perceptrons, gated recurrent

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023

units, and long short-term memory (LSTM) is the main goal of this study framework. The characteristics of these IoMT samples were evaluated and further processed prior to using machine learning, deep learning, and other learning-based methods.

## IMPLEMENTATION

1) Upload Drug Review Dataset: using this module we will upload dataset to application
2) Read & Preprocess Dataset: using this module we will read all reviews, drug name and ratings from dataset and form a features array.
3) TF-IDF Features Extraction: features array will be input to TF-IDF algorithm which will find average frequency of each word and then replace that word with frequency value and form a vector. If word not appear in sentence then 0 will be put. All reviews will be consider as input features to machine learning algorithm and RATINGS and Drug Name will be consider as class label.
4) Train Machine Learning Algorithms: using this module we will input TF-IDF features to all machine learning algorithms and then trained a model and this model will be applied on test data to calculate prediction accuracy of the algorithm.
5) Comparison Graph: using this module we will plot accuracy graph of each algorithm
6) Recommend Drug from Test Data: using this module we will upload disease name test data and then ML will predict drug name and ratings.

## PROPOSED WORK

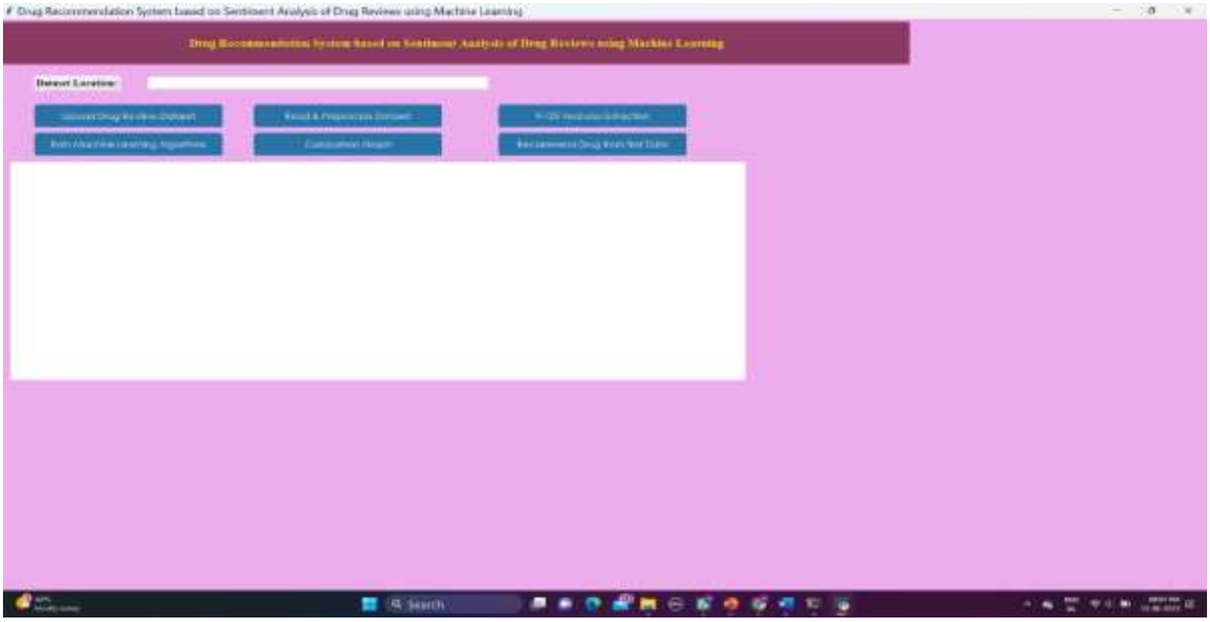
Even with the advent of high-powered computers, medical professionals have continued to have a need for technologies such as surgical representation processes and x-ray photography. Medical records, the environment, blood pressure, and other variables all play a role in this strategy, therefore the doctor's knowledge and experience are still essential. No model has successfully assessed the enormous number of factors that are considered as whole variables necessary to understand the complete functioning process itself. Using a medical decision support system is the only way to overcome this limitation. This system can help doctors make the right choice. The term "medical decision support system" can be used to describe either the effort put forth to ascertain the possibility of a disease or ailment, or the conclusion reached after doing so.

**Why It's Important:**

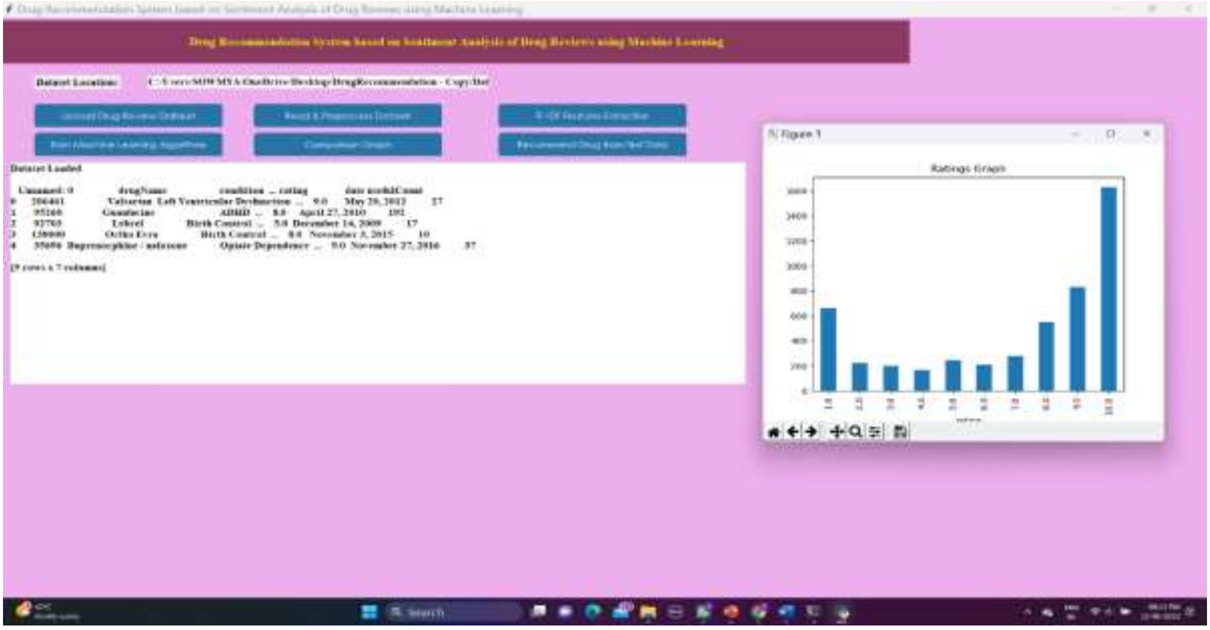
Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023

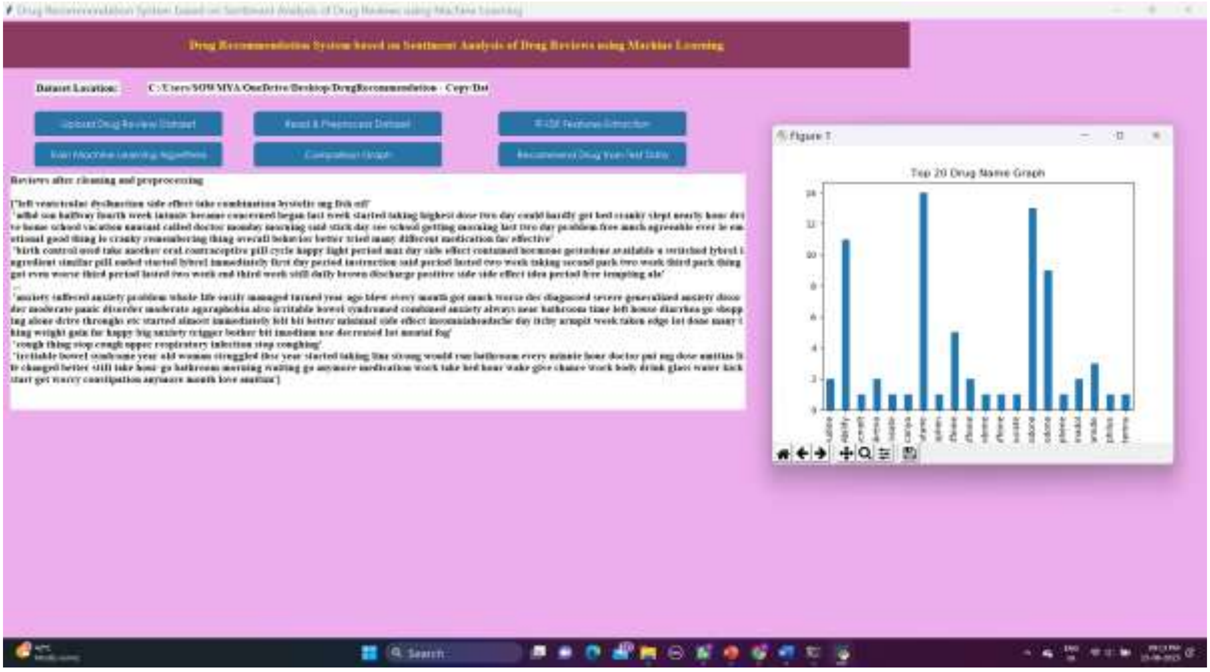
Depending on the circumstances, such as with uncommon diseases, making a medical decision can be a highly specialized and challenging task. Stress, exhaustion, and a lack of sleep are all possible contributors, as can a lack of resources and a lack of knowledge on the part of medical professionals. It's possible for a standard algorithm to examine all of the determinants, such as the patient's current health status, previous medical history, family medical history, and other aspects relevant to the patient's medical file. If there are numerous potential explanations for anything, differential diagnosis can be used to zero in on the most likely one. This strategy calls for an elimination procedure or data collection that reduces the likelihood of potential situations to ze

**SAMPLE SCREENSHOTS**

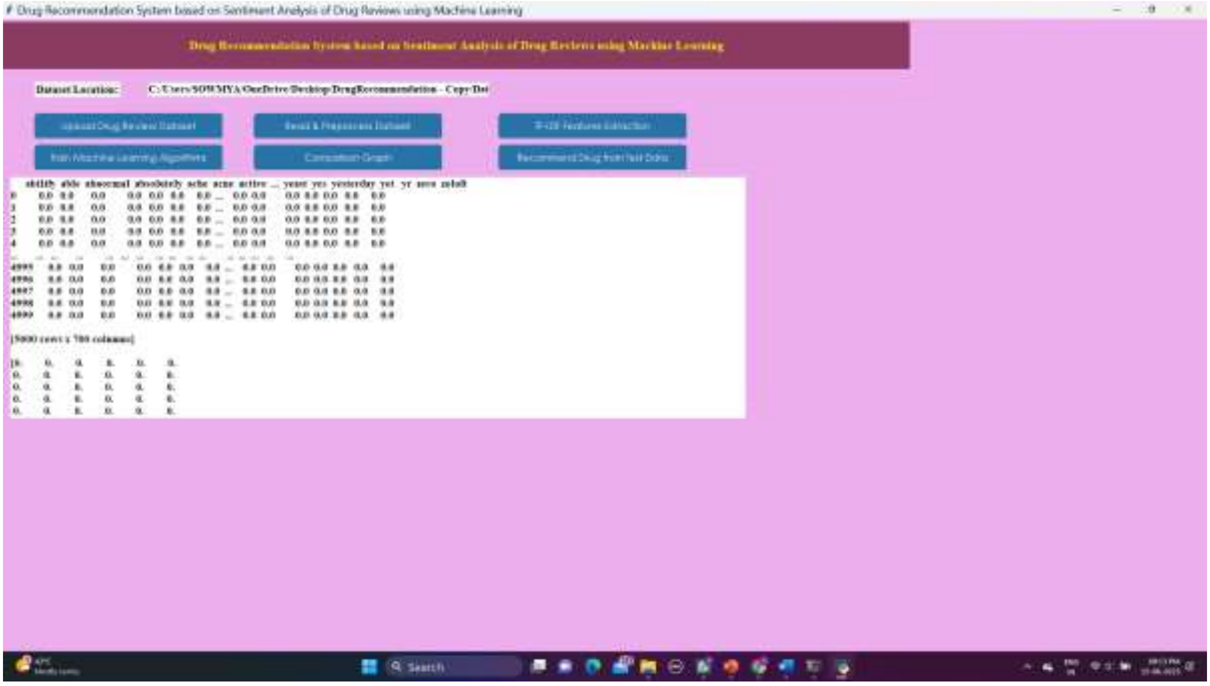


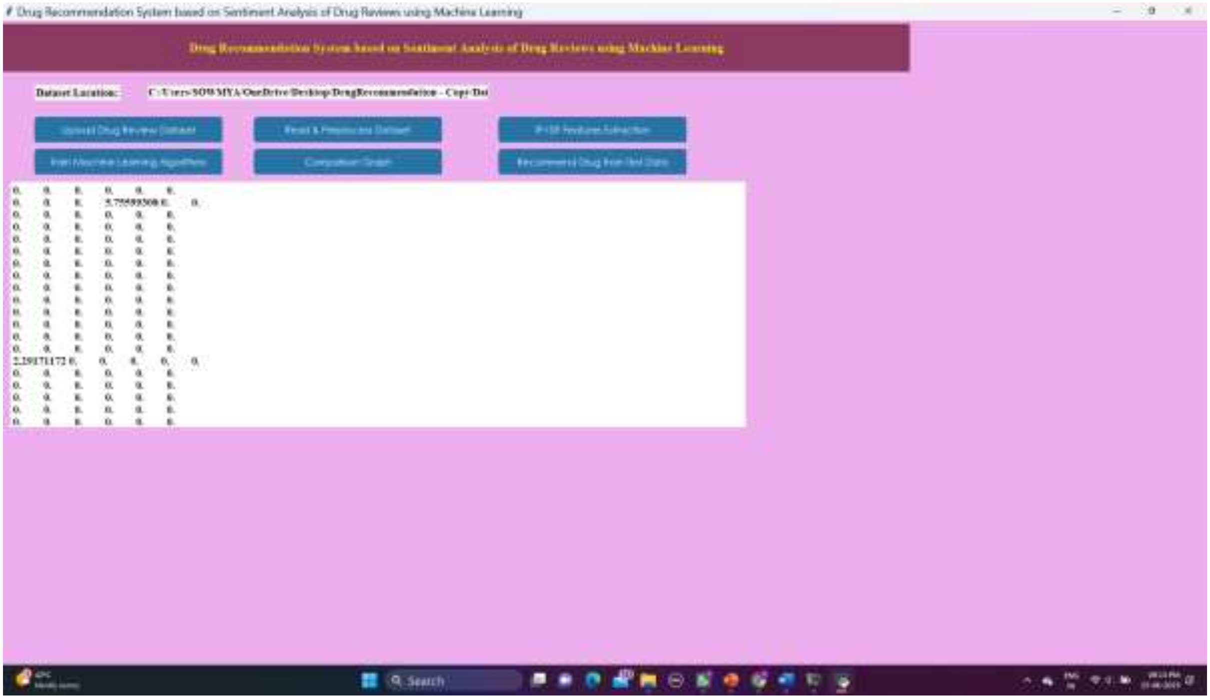
Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023



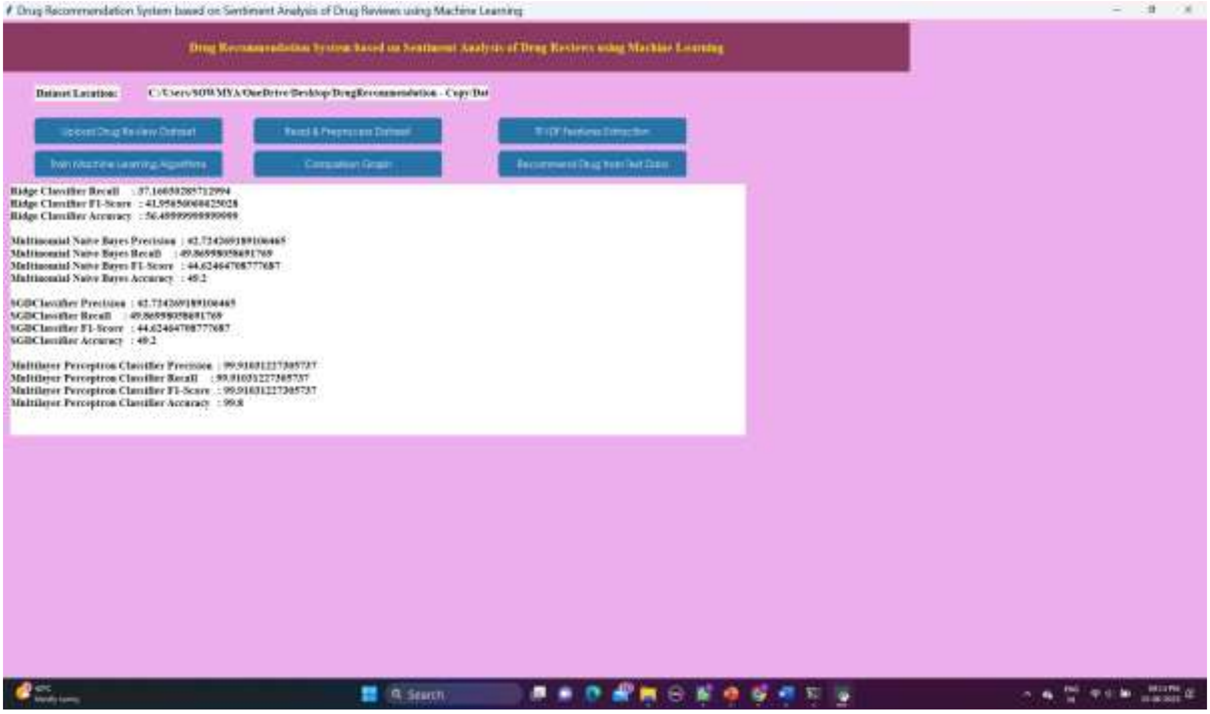



Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023

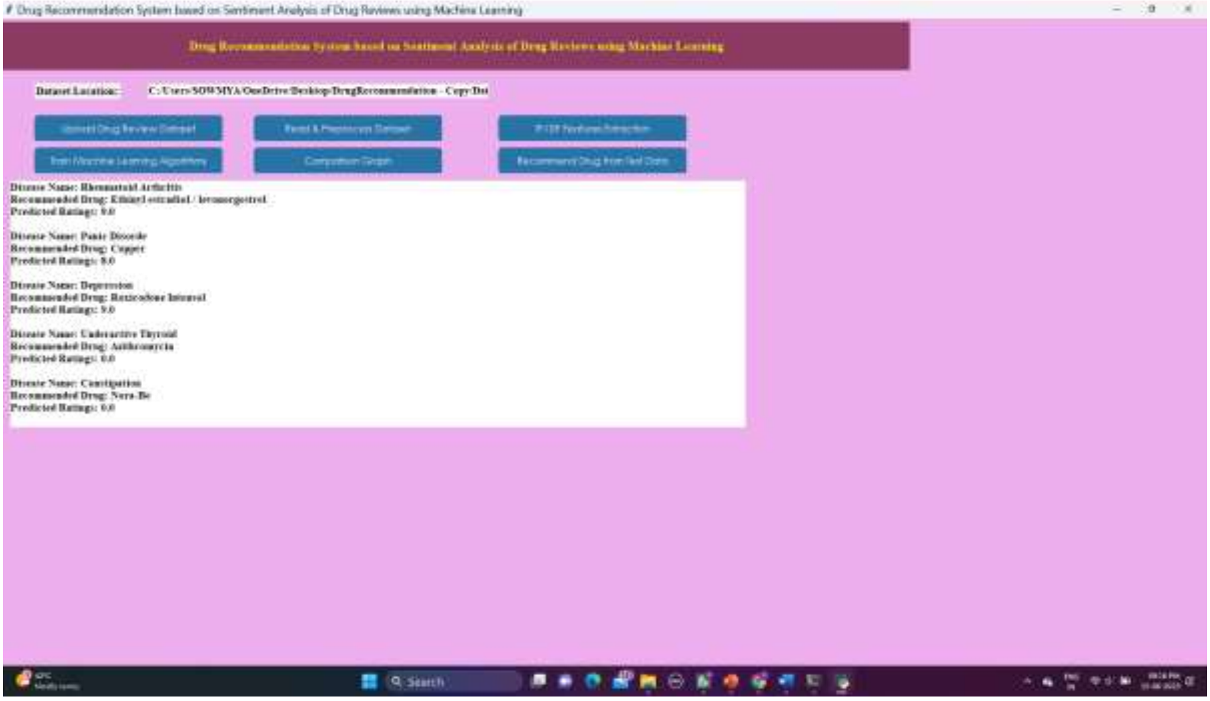


Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023




Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023




## CONCLUSION

In this paper, we propose a 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, Perception, Multinomial Naive Bayes, Ridge classifier, Stochastic gradient descent, LinearSVC, applied on Bow, TF-IDF, and classifiers such as Decision Tree, Random Forest, Lgbm, and Catboost 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, Perception 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] Telemedicine,https://www.mohfw.gov.in/pdf/Telemedicine.pdf

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

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

[4] Drug Review Dataset, https://archive.ics.uci.edu/ml/datasets/Drug%2BReview%2BDataset%2B%2528Drugs.com%2529#

[5] Fox, Susannah, and Maeve Duggan. ”Health online 2013. 2013.” URL: http://pewinternet.org/Reports/2013/Health-online.aspx

[6] 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.Clin Infect Dis. 2000 Aug;31(2):347-82. doi: 10.1086/313954. Epub 2000 Sep 7. PMID: 10987697; PMCID: PMC7109923.

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

[8] T. N. Tekade and M. Emmanuel, ”Probabilistic aspect mining approach for interpretation and evaluation of drug reviews,” 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi,2016, pp. 1471-1476, doi: 10.1109/SCOPES.2016.7955684.

[9] 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

[10] Leilei Sun, Chuanren Liu, ChonghuiGuo, Hui Xiong, and YanmingXie. 2016. Data-driven Automatic Treatment Regimen Development and Recommendation. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’16). Association for Computing Machinery, New York, NY, USA, 1865–1874.DOI:https://doi.org/10.1145/2939672.2939866

[11] V. Goel, A. K. Gupta and N. Kumar, ”Sentiment Analysis of Multilingual Twitter Data using Natural Language Processing,” 2018 8th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2018, pp. 208-212, doi: 10.1109/CSNT.2018.8820254.

[12] Shimada K, Takada H, Mitsuyama S, et al. Drug-recommendation system for patients with infectious diseases. AMIA AnnuSymp Proc. 2005;2005:1112.

[13] Y. Bao and X. Jiang, ”An intelligent medicine recommender system framework,” 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, 2016, pp. 1383-1388, doi: 10.1109/ICIEA.2016.7603801.

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

[15] J. Li, H. Xu, X. He, J. Deng and X. Sun, ”Tweet modeling with LSTM recurrent neural networks for hashtag recommendation,” 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 1570-1577, doi: 10.1109/IJCNN.2016.7727385.