

Leveraging Machine Learning and Patient Reviews for Developing a Drug Recommendation System to Reduce Medical Errors



K. P. Swain¹, S.K. Mohapatra¹, Vinayakumar Ravi^{5,*}, Soumya Ranjan Nayak², Tahani Jaser Alahmadi^{6,*}, Prabhishkek Singh³ and Manoj Diwakar⁴

¹Department of ETC, Trident Academy of Technology, Bhubaneswar, Odisha, India

²School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar 751024, Odisha, India

³School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India

⁴Department of CSE, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

⁵Center for Artificial Intelligence, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia

⁶Department of Information Systems, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

Abstract:

Background: In the rapidly evolving pharmaceutical industry, drug efficacy and safety stand as critical concerns. The vast accumulation of data, including customer feedback, drug popularity, and usage details, offers a rich resource for improving healthcare outcomes.

Aims: The primary aim of this study is to harness machine learning and Natural Language Processing (NLP) techniques to sift through extensive pharmaceutical data, identifying the most effective drugs for various conditions and uncovering patterns that could guide better decision-making in drug efficacy and safety.

Objective: This research seeks to construct a sophisticated model capable of analyzing diverse data points to pinpoint the most efficacious drugs for specific health conditions, thereby providing pharmaceutical companies with data-driven insights to optimize drug safety and effectiveness.

Methods: Employing a blend of Natural Language Processing (NLP) and machine learning strategies, the study analyzes a comprehensive dataset featuring customer reviews, drug popularity metrics, usage information, and other relevant data collected over an extended period. This methodological approach aims to reveal latent trends and patterns that are crucial for assessing drug performance.

Results: The developed model adeptly identifies leading medications for various conditions, elucidating efficacy and safety profiles derived from patient reviews and drug utilization trends. These findings furnish pharmaceutical companies with actionable intelligence for enhancing drug development and patient care strategies.

Conclusion: The integration of machine learning and NLP for the analysis of vast drug-related datasets presents a powerful method for advancing drug efficacy and safety. The insights yielded by the proposed model significantly empower the decision-making processes of the pharmaceutical industry, ultimately fostering improved health outcomes for patients.

Keywords: Drug prediction, Machine learning, NLP, Customer feedback, Customer rating, Human and disease, Health system.

© 2024 The Author(s). Published by Bentham Open.

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

*Address correspondence to these authors at the Department of Information Systems, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; E-mail: tjahmadi@pnu.edu.sa

Cite as: Swain K, Mohapatra S, Ravi V, Nayak S, Alahmadi T, Singh P, Diwakar M. Leveraging Machine Learning and Patient Reviews for Developing a Drug Recommendation System to Reduce Medical Errors. Open Bioinform J, 2024; 17: e18750362291402. <http://dx.doi.org/10.2174/0118750362291402240621044046>



Received: December 31, 2023
Revised: April 17, 2024
Accepted: June 10, 2024
Published: June 25, 2024



Send Orders for Reprints to
reprints@benthamscience.net

1. INTRODUCTION

Healthcare is one of the most important sectors that affect the quality of human life. One of the major concerns in healthcare is ensuring the right treatment for patients, including the prescription of the appropriate medication. Prescription errors, including prescribing the wrong medication or dosage, are a serious problem that can lead to adverse health outcomes, even death. According to a study published by the National Academies of Sciences, Engineering, and Medicine, medication errors affect an estimated 1.5 million people in the United States every year. To address this issue, various approaches have been proposed, including the use of machine learning techniques to build drug recommendation systems [1, 2]. These systems can assist physicians in making informed decisions by analyzing patient data and providing suggestions for the best medication. In recent years, machine learning has become an important tool in healthcare, with a growing number of studies demonstrating its potential in various applications. Machine learning techniques can be applied to patient data, such as medical history, laboratory results, and patient feedback, to generate insights and predictions [3-5]. In this paper, we present a drug recommendation system that uses machine learning techniques to predict the best medication for a given disease. The term “best” is contextually bound to the data source—patient reviews, and our system identifies medications perceived most positively by users in terms of satisfaction, relief, and side effects, as reported in user-generated content. This definition is consistent with our data and analytical approach, employing advanced machine learning models validated through rigorous cross-validation to ensure reliable sentiment prediction. The system is designed as a supplementary tool meant to enhance, not replace, clinical trials, addressing a significant gap in understanding patient perspectives at scale and providing a novel avenue for gathering patient-reported outcomes.

The remainder of this paper is organized as follows. Section 2 presents a literature review of drug recommendation systems using machine learning techniques. Sections 3 and 4 describe the data analysis and data preprocessing used in this study, respectively. Section 5 presents the model building, section 6 recommends useful drugs, and finally, section 7 concludes the paper.

2. LITERATURE REVIEW

Healthcare is one of the most critical industries, and any error could lead to life-threatening consequences. One of the significant challenges in healthcare is medical errors caused by doctors prescribing the wrong medication based on their limited experiences. Medical errors, encompassing a wide range of healthcare failures, from surgical mishaps to the misprescription of medications, are a significant concern, resulting in approximately 250,000 deaths annually in the United States (National Academies of Sciences, Engineering, and Medicine). Notably, Anderson and Abrahamson (2017) discussed the profound impact of medical errors on patient safety, emphasizing the critical need for systems and interventions to mitigate these risks. Among these errors,

incorrect medication prescription stands out as a significant challenge, underscoring the urgency of developing tools like our drug recommendation system to aid healthcare professionals in making more informed decisions [6]. The good news is that with the emergence of machine learning, deep learning, and data mining technologies, these technologies can help explore medical history and reduce medical errors by being doctor-friendly.

In a previous study [7], the authors presented a review of machine learning algorithms for predicting the severity of adverse drug events. They compared different machine learning models and highlighted the need for additional data sources to improve prediction accuracy, whereas another study [8] highlighted various machine learning techniques and their applications in healthcare, including diagnosis, drug discovery, and patient monitoring. In another study [9], the uses of supervised machine learning algorithms were explored to classify adverse drug reactions. One study [10] proposed a hybrid unsupervised and supervised machine learning approach to improve adverse drug event detection. The authors also compared their approach to other methods and showed that their approach outperforms other techniques in detecting adverse drug events.

In a study, the use of machine learning and network analysis was explored to predict drug interactions [11]. Meanwhile, another study [12] provided a review of drug recommender systems, highlighting various approaches to drug recommendation, including collaborative filtering, content-based filtering, and hybrid approaches. A drug recommendation system was also proposed [13] based on disease name recognition and symptom extraction. Here, a real-world dataset was approached, which showed that their system could accurately recommend drugs. One of the studies [14] provided a comprehensive survey of machine learning approaches for predicting adverse drug reactions where different machine learning models were compared and highlighted the need for additional data sources to improve prediction accuracy. A review of machine learning approaches for drug-target interaction prediction was described in another study [15]. In a study [15], a supervised learning approach was employed to predict medication adherence. Here, the authors compared their approach to other methods and showed that their approach outperforms other techniques in predicting medication adherence.

3. METHODS

3.1. Analysis of Data Set

This research uses the Drug Review Dataset from the UCI ML (University of California Irvine machine learning) repository, consisting of 161297 rows and 7 columns with attributes, such as drug name, customer reviews, popularity, and use cases. The dataset includes 3436 unique drugs and 884 unique medical conditions, collected between April 1, 2008, and September 9, 2017. The rating column of the dataset has a positive skew with a median of 7 and a standard deviation that suggests wide variability. The statistical summary of the dataset is given in Fig. (1).

The term “useful count” in Fig. (1) in our study refers to the number of times a particular review was marked as useful by other users. It ranges from 0 to 1,291, with an average of 28, indicating significant variability and providing insight into the perceived helpfulness of the review content to the broader user community. This metric serves as an important parameter in our analysis, helping to quantify the impact and relevance of each review in guiding medication recommendations.

4. DATA PREPROCESSING

After analyzing the dataset, it was found that there are over 6,318 drugs with zero useful count, and the average rating for these drugs is 5.80, indicating that a rating above six is considered good, as shown in Fig. (2a). However, in Fig. (2b), only four drugs have more than 1000 useful counts, all with a rating of 10, two belonging to birth control and two to depression. Levonorgestrel is the most common drug, with 3,657 occurrences. Birth control is the most common condition, with over 28,000 occurrences, and “good” is the most common review with 33 occurrences.

The dataset was checked for missing values using the essential function, and only the condition column was found to have missing values. As the condition is essential for recommending drugs to patients, records with missing values were removed using the drop function, resulting in no missing values in the dataset.

5. MODEL BUILDING

5.1. Unveiling the Hidden Patterns from the Dataset

Using the Seaborn Library's display function, the distribution of rating and useful count balance is checked, which revealed the high distribution for ratings of 8, 9, and 10 and a highly skewed distribution for the useful count, with most drugs having less than 200 useful counts which are shown in Fig. (3). The vertical axis represents the frequency of occurrences for ratings and useful counts across the dataset. The values above 1.0 on the vertical axis indicate relative frequencies of occurrence, not percentages. Specifically, the bars in Fig. (3) do not sum up to 1.0 (or 100%) because they represent the distribution of values as counts normalized by the total number of observations, not as a direct percentage of the total. This method allows us to compare the distribution across different categories more effectively by showing the prevalence of each rating and useful count relative to the total dataset size rather than providing a direct percentage. This approach is particularly useful in datasets with skewed distributions, as it emphasizes the relative commonality or rarity of certain data points without the constraints of a percentage total, which would mask these differences in a larger dataset. Here, a positive linear relationship between the rating and useful count columns was observed, with the number of average uses increasing as the ratings increased.

	rating	usefulCount
count	161297.000000	161297.000000
mean	6.994377	28.004755
std	3.272329	36.403742
min	1.000000	0.000000
25%	5.000000	6.000000
50%	8.000000	16.000000
75%	10.000000	36.000000
max	10.000000	1291.000000

Fig. (1). Statistical summaries for the rating and useful count columns.

```

Analysis on Useless Drugs
-----
The Number of Drugs with No Useful Count : 6318
Number of Good Drugs with Lesser Useful Count : 0
Average Rating of Drugs with No Useful Count : 5.80

Analysis on Useful Drugs
-----
The Number of Drugs with Greater than 1000 Useful Counts : 4
Average Rating of Drugs with 1000+ Useful Counts : 10.0

Name and Condition of these Drugs:

      drugName      condition
0      Mirena      Birth Control
1    Sertraline      Depression
2 Levonorgestrel      Birth Control
3      Zoloft      Depression
    
```

Fig. (2a). Analysis of usefulness of drugs.

	drugName	condition	review
count	161297	160398	161297
unique	3436	884	112329
top	Levonorgestrel	Birth Control	"Good"
freq	3657	28788	33

Fig. (2b). Analysis result for the useful count for drugs.

After that, the relationship between ratings and the usefulness of the drugs listed in the data set is checked. For that, the bar plot with a hot palette is realized, and the output is shown in Fig. (4). This figure shows a bar graph illustrating the correlation between drug ratings (ranging from 1 to 10) and their average useful count. Higher

ratings are associated with greater usefulness, indicating that more users found these reviews helpful in making informed decisions about medications. It is clear from the given chart that there is a positive linear relationship between the rating and useful counter columns. The average useful counts increase as the rating increases.

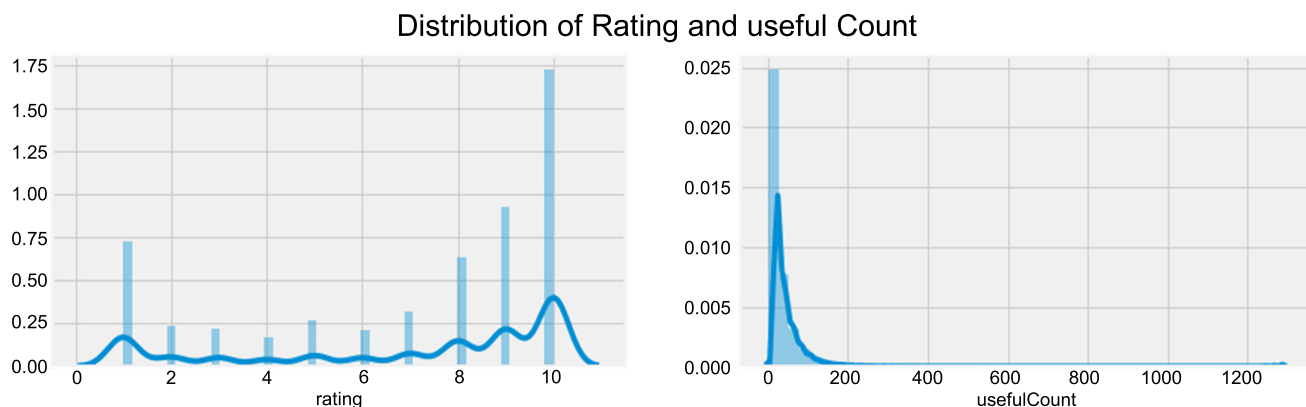


Fig. (3). Distribution of rating and useful-count columns.

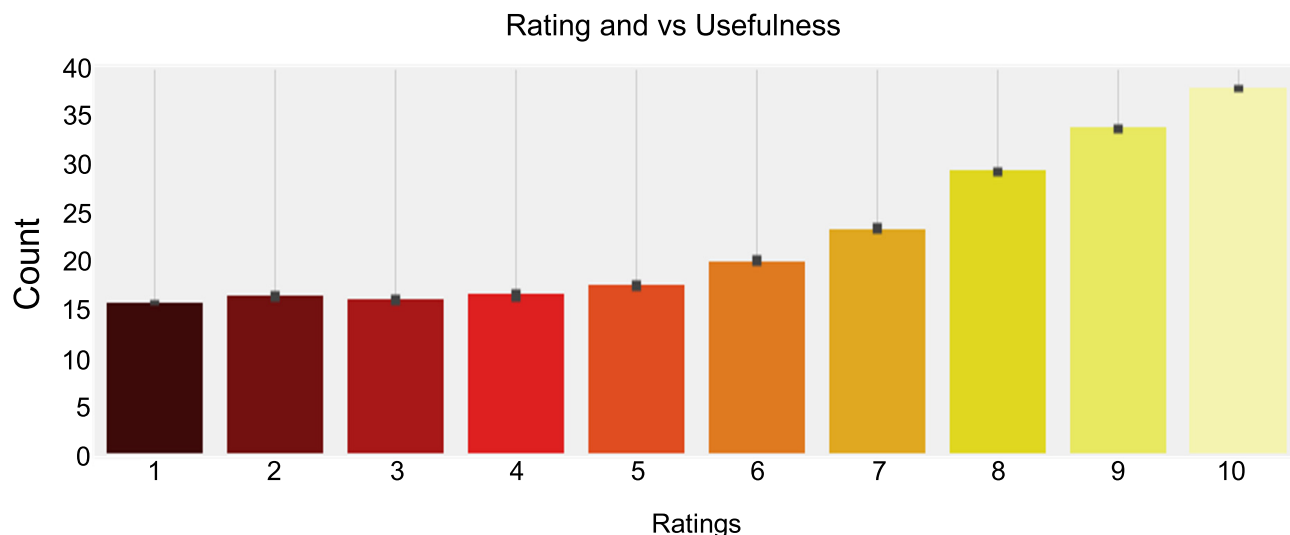


Fig. (4). Relationship between drug ratings and their perceived usefulness.

5.2. Cleaning the Reviews

Cleaning the reviews involves removing impurities and unnecessary elements from the dataset for easy textual analysis. Punctuations, stop words, and numbers have no sentimental meaning, and stop words, such as “he” and “she” are removed. A function is created to remove punctuation and stop words from the review column, using the string library and NLTK (Natural Language Toolkit) library. The resulting clean consumer reviews can then be analyzed.

5.3. Calculating Sentiment from the Reviews

After removing unnecessary punctuations and stop words in the data set, the sentiment of each of the reviews present in the data set is calculated. The sentiment scores

are analyzed by using the VADER (Valence Aware Dictionary and sEntiment Reasoner) lexicon class of the NLTK library. The choice of VADER for sentiment analysis in our study stems from its proficiency in handling informal language, which is common in social media posts and shares similarities with the user-generated content found in online pharmaceutical drug reviews. VADER excels at understanding the nuances of sentiment expressed in short texts, including slang, emoticons, and abbreviations, making it well-suited for analyzing patient feedback where such expressions are prevalent. It was found that sentiment scores were not useful for this project as they were all negative. Therefore, the sentiment columns, along with other unnecessary columns like date, uniqueID, sentiment, review, and len, were dropped from the dataset. The remaining index includes 'drug Name',

'condition', 'rating', and 'useful Count'. Finally, it was found that the sentiment score for the review was completely useless, and so was removed from the data set.

5.4. Calculating the Effectiveness and Usefulness of Drugs

A function is created to calculate the effectiveness and usefulness scores for the drugs to recommend the best and worst drugs for each medical condition. To calculate the effective score, the minimum and maximum ratings are calculated, and then the difference between the minimum rating and all the ratings is divided by the maximum rating minus one. Then, the resulting score is multiplied by five, and if the score is zero, one, or two, then the effective score is zero; otherwise, it is one. This function is applied to create a new column called Effective Score. Then, to calculate the usefulness score, the rating, useful count, and effective score columns are multiplied. After that, the top ten most useful drugs are listed for each condition, with most belonging to the depression, birth control, and weight loss categories, as shown in

Fig. (5). The usefulness of these drugs is cross-checked by reviewing Google reviews and consulting doctors.

5.4.1. Algorithm for the Function to find the Effective Score

```

Find min_rating
Find max_rating
rating = rating - min_rating
rating = rating / (max_rating - 1)
rating = rating * 5 (here 5 is taken as threshold)
if (rating == 0 or rating == 1 or rating == 2)
    eff_score = 0
else
    eff_score = 1
    
```

5.4.2. Algorithm for Usefulness Score

```

Create new column of name eff_score by applying
above function
Usefulness = rating * useful count * eff_score
    
```

	drugName	condition	usefulness
0	Sertraline	Depression	12910
1	Zoloft	Depression	12910
2	Levonorgestrel	Birth Control	12470
3	Mirena	Birth Control	12470
4	Zoloft	Depression	8541
5	Phentermine	Weight Loss	7960
6	Adipex-P	Weight Loss	7960
7	Implanon	Birth Control	7300
8	Viibryd	Depression	6930
9	Vilazodone	Depression	6930

Fig. (5). Top ten most useful drugs.

```
@interact
def check(condition = list(data['condition'].value_counts().index)):
    return data[data['condition'] == condition]['eff_score'].value_counts()
```

condition Birth Control

1 16537
0 12251
Name: eff_score, dtype: int64

condition Depression

1 6643
0 2426
Name: eff_score, dtype: int64

Fig. (6). Useful and useless drugs count for birth control and depression.

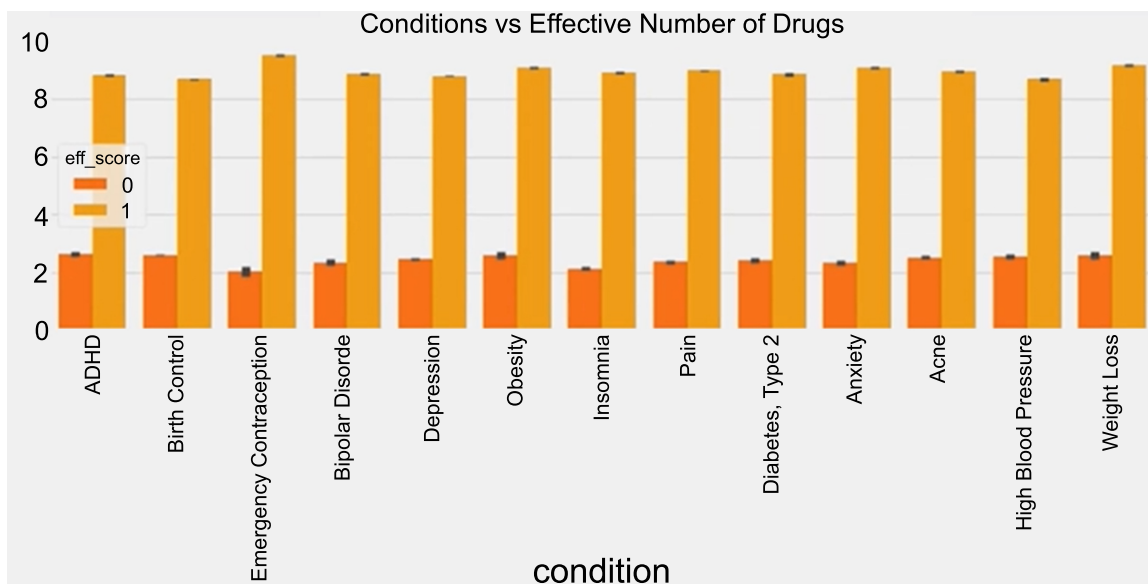


Fig. (7). Comparison of useful and useless drugs across common medical conditions.

5.5. Analyzing the Medical Condition

An interactive function using the IP Widgets Library is created to analyze medical conditions and find the number of useful and useless drugs for each condition. Here, an effective score, which was calculated earlier, is implemented. The analysis revealed that there were around 16537 useful drugs and 12251 useless drugs for birth

control. Also, 6643 drugs were useful, and 2426 drugs were useless for depression. Fig. (6) shows the useful drugs as 1 and the useless drugs as 0.

A bar plot was used to visualize the ratio of useful and useless drugs for popular medical conditions like depression, anxiety, acne, bipolar disorder, insomnia, obesity, ADHD (attention deficit hyperactivity disorder),

emergency contraception, high blood pressure, etc. It was found that the ratio of useful and useless drugs for all of these medical conditions is very similar as shown in Fig. (7). This bar plot contrasts the number of drugs deemed useful (effective score = 1) versus useless (effective score = 0) for various medical conditions such as depression, anxiety, acne, and others. It highlights the variability in drug efficacy perceptions across different health issues. Also, from the plot, it is revealed that around 30 percent of the drugs in each of the medical conditions will be useless.

5.6. List of Most Common Medical Conditions

In the end, to check the list of the most common medical conditions present in the dataset, the value counts function, along with the head function, is realized. Here, it can be marked that birth control, depression, anxiety, acne, etc., are among the top medical conditions. And finally, to check the list of top 10 drugs that were helpful to the highest number of people, the useful counter column is implemented. The outcomes of both cases are shown in Fig. (8a and b), respectively. Fig. (8a) displays the frequency of the top medical conditions mentioned in the drug reviews, indicating the prevalence of conditions like birth control, depression, and anxiety within the dataset, whereas Fig. (8b) lists the drugs that were most frequently marked as useful by the review platform users, showcasing the medications that have been most beneficial according to patient feedback.

6. RESULTS AND DISCUSSION

The outcome of this study, which prescribes the most useful and useless drugs for each of the medical conditions, is shown in Fig. (9). The API (Application programming interface) widgets are used for an interactive function where the input is the list of all the medical conditions listed in the dataset. Inside the function, the names of the top five most useful drugs are printed, along with the top five most useless drugs for a particular medical condition with their usefulness rating. Considering the most common problems these days, like birth control, depression, anxiety, pain, and acne, the result of these top five medical conditions is discussed in Fig. (9).

The primary goal of using NLP to analyze patient reviews is to detect and analyze patterns that could indicate potential issues with medication use, such as unexpected side effects, efficacy concerns, or misunderstandings in usage that could lead to adverse events. These insights are crucial for identifying discrepancies between expected medication outcomes and actual patient experiences. For example, while a medication might be correctly prescribed, patient reviews frequently mentioning adverse reactions or lack of expected relief suggest a need for re-evaluating its use for specific conditions or demographic groups. Our aim is not to base clinical decisions solely on patient satisfaction but to use aggregated patient feedback as a supplementary data source providing real-world evidence of drug performance across diverse populations.

(a)

Number of Unique Conditions : 884

Birth Control	28788
Depression	9069
Pain	6145
Anxiety	5904
Acne	5588
Bipolar Disorder	4224
Insomnia	3673
Weight Loss	3609
Obesity	3568
ADHD	3383

Name: condition, dtype: int64

(b)

	drugName	usefulCount
0	Zoloft	1291
1	Sertraline	1291
2	Levonorgestrel	1247
3	Mirena	1247
4	Zoloft	949
5	Adipex-P	796
6	Phentermine	796
7	Celexa	771
8	Citalopram	771
9	Implanon	730

Fig. (8). (a) The most common medical conditions present in the dataset. (b) List of top 10 drugs which were helpful to the highest number of people.

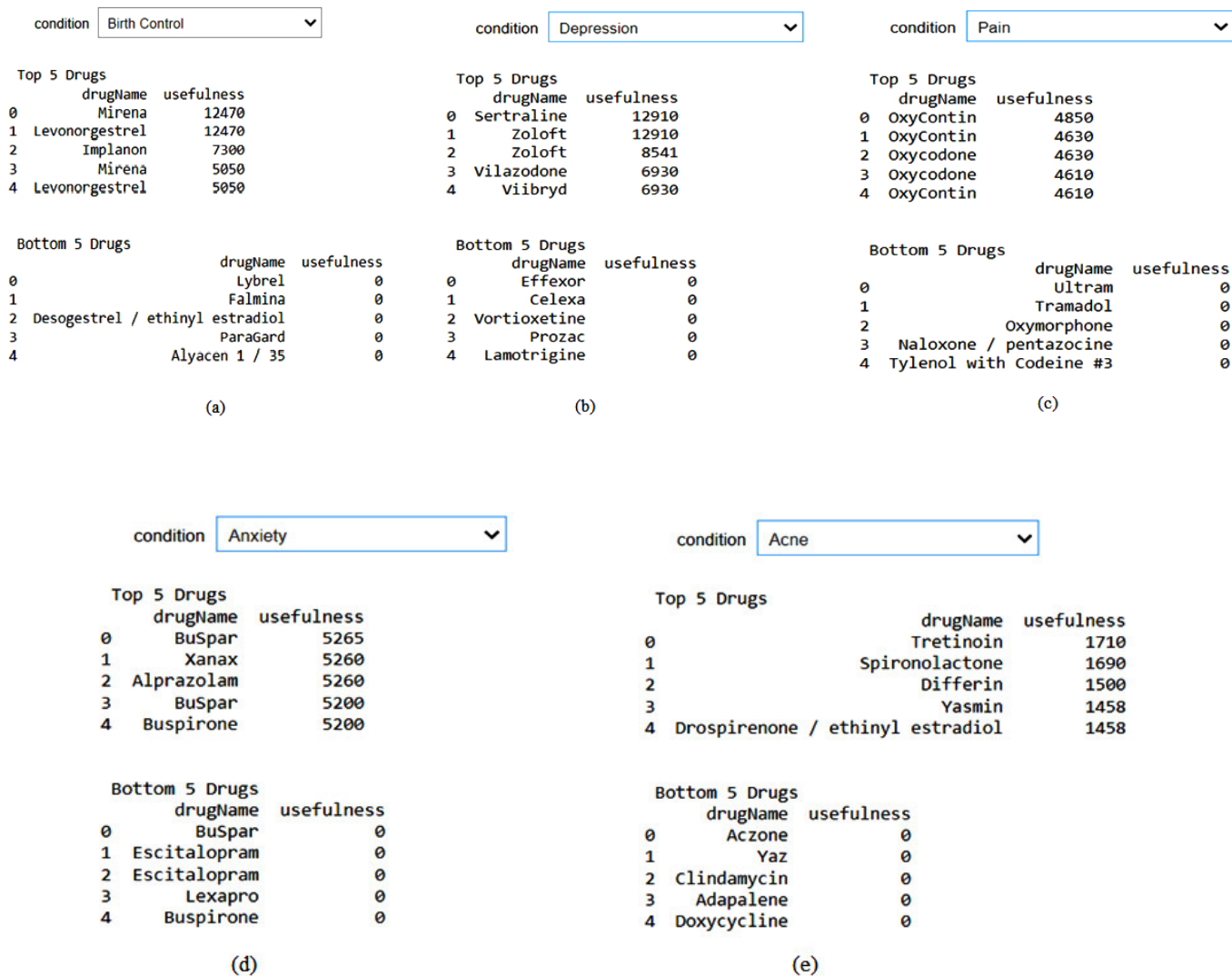


Fig. (9). The result for drug recommendation (a) Birth control (b) Depression (c) Pain (d) Anxiety (e) Acne.

Additionally, this method aids healthcare providers in identifying patterns that may warrant further clinical investigation or a review of prescribing practices, thereby indirectly contributing to reducing medical errors. For instance, if reviews consistently point out misunderstandings about medication dosage, this could highlight a need for enhanced communication or education from healthcare providers. By analyzing patient feedback collectively through NLP, we can gain valuable insights into potential systemic issues, medication efficacy, and safety profiles, assisting healthcare providers in making more informed decisions and ultimately improving patient care and safety.

CONCLUSION

In conclusion, the research presented in this paper

aimed to develop a prescription system for drugs based on sentiment analysis of drug reviews. The system was developed using machine learning classifiers and involved four main steps: analyzing the review dataset, data pre-processing, model building, and recommending the appropriate drug or medicine for a particular disease. The proposed system serves as a tool to support doctors in their disease diagnosis and can be utilized by physicians, psychiatrists, physiologists, and other medical professionals. However, it is important to note that this system is not intended for use by individuals, such as data scientists or data analysts. Overall, the research presented a novel approach to drug prescription and holds potential for future advancements in medical inquiry.

AUTHORS' CONTRIBUTION

It is hereby acknowledged that all authors have accepted responsibility for the manuscript's content and consented to its submission. They have meticulously reviewed all results and unanimously approved the final version of the manuscript.

ABBREVIATION

NLP = Natural Language Processing

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

Not applicable.

CONSENT FOR PUBLICATION

Not applicable

AVAILABILITY OF DATA AND MATERIALS

The data and supportive information are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

Vinayakumar Ravi is the Associate Editorial Board Member of The Open Bioinformatics Journal.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] Gräßer F, Tesch F, Schmitt J, Abraham S, Malberg H, Zaunseder S. A pharmaceutical therapy recommender system enabling shared decision-making. *User Model User-adapt Interact* 2022; 32(5): 1019-62. <http://dx.doi.org/10.1007/s11257-021-09298-4>
- [2] Mohapatra Mahima, Nayak Mamata, Mahapatra Saswati. A machine learning based drug recommendation system for health care. *Graduate Res Eng Technol* 2022; 1(6)
- [3] Javaid M, Haleem A, Pratap Singh R, Suman R, Rab S. Significance of machine learning in healthcare: Features, pillars

- and applications. *Int J Intell Networks* 2022; 3: 58-73. <http://dx.doi.org/10.1016/j.ijin.2022.05.002>
- [4] Deloar Hossain Md, Shafiul Azam Md. Drugs rating generation and recommendation from sentiment analysis of drug reviews using machine learning. 2020 *Emerging Technology in Computing, Communication and Electronics (ETCCE)*, . Bangladesh, 21-22 December 2020, pp. 1-6.
 - [5] Stephenson Natalie, Shane Emily, Chase Jessica, *et al.* Survey of machine learning techniques in drug discovery. *Curr Drug Metab* 2019; 20(3): 185-93.
 - [6] Anderson JG, Abrahamson K. Your health care may kill you: medical errors. *Stud Health Technol Inform* 2017; 234: 13-7. PMID: 28186008
 - [7] Sharda Shreya. Machine learning based recommendation system: A review. *Int J Next-Gen Comput - Spec Issue* 2021; 12(2): 134-44.
 - [8] Bissoyi S, Mishra BK, Patra MR. Recommender systems in a patient centric social network — A survey. *International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, . Paralakhemundi, India, 03-05 October 2016, pp. 386-389. <http://dx.doi.org/10.1109/SCOPES.2016.7955858>
 - [9] Das P, Mazumder DH. An extensive survey on the use of supervised machine learning techniques in the past two decades for prediction of drug side effects. *Artif Intell Rev* 2023; 56(9): 9809-36. <http://dx.doi.org/10.1007/s10462-023-10413-7> PMID: 36819660
 - [10] Zhang M, Geng G. Adverse drug event detection using a weakly supervised convolutional neural network and recurrent neural network model. *Information* 2019; 10(9): 276. <http://dx.doi.org/10.3390/info10090276>
 - [11] You J, McLeod RD, Hu P. Predicting drug-target interaction network using deep learning model. *Comput Biol Chem* 2019; 80(June): 90-101. <http://dx.doi.org/10.1016/j.compbiolchem.2019.03.016> PMID: 30939415
 - [12] Tran TNT, Felfernig A, Trattner C, Holzinger A. Recommender systems in the healthcare domain: state-of-the-art and research issues. *J Intell Inf Syst* 2021; 57(1): 171-201. <http://dx.doi.org/10.1007/s10844-020-00633-6>
 - [13] Chen J, Li K, Rong H, Bilal K, Yang N, Li K. A disease diagnosis and treatment recommendation system based on big data mining and cloud computing. *Inf Sci* 2018; 435: 124-49. <http://dx.doi.org/10.1016/j.ins.2018.01.001>
 - [14] Dey S, Luo H, Fokoue A, Hu J, Zhang P. Predicting adverse drug reactions through interpretable deep learning framework. *BMC Bioinformatics* 2018; 19(S21) (Suppl. 21): 476. <http://dx.doi.org/10.1186/s12859-018-2544-0> PMID: 30591036
 - [15] Antoniadi AM, Du Y, Guendouz Y, *et al.* Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: A systematic review. *Appl Sci* 2021; 11(11): 5088. <http://dx.doi.org/10.3390/app11115088>