Analyzing Patient Reviews for Recommending Treatment Using NLP and Deep Learning-Based Approaches

Tauheed Shahid, Suraj Singh, Shatmanyu Gupta, and Shanu Sharma(✉)

Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad, India
{tauheed.18bcs1108,suraj.18bcs1028, shatmanyu.18bcs1030}@abes.ac.in, shanu.sharma16@gmail.com

Abstract. Nowadays technological advancement can be seen in the medical field starting from medical devices, data collection, analysis to diagnosis and treatment recommendations for diseases. Drug and treatment recommendation is one of the most popular applications, which is now observed and used by everyone in this digital era. These types of recommendation systems usually require a huge set of data from the patients and an efficient Machine learning-based model to conclude significant insights that can help in the prediction of the best possible medications for a particular disease. The health-related recommendation systems can be proved as a significant tool in the healthcare sector for speeding up various decision-making processes such as health insurance, clinical pathway-based treatment methods, and assisting doctors by recommending drugs using a patient's health profile. People frequently utilize social media to explore their health issues, therefore there is now a wealth of information on social networks that may be leveraged to create various health-related recommendation systems. In this research, a deep learning-based based drug recommendation system using N-Gram is provided, which uses patient review data as input and sentiment analysis to choose the appropriate treatment for the ailment. The accuracy achieved, shows the efficacy of the proposed approach for the real time applicability of the model.

Keywords: Recommender system · Drugs recommendation · Deep learning · N-grams · Reviews

1 Introduction

In the past few years, health was the most discussed topic on the internet due to the Global Covid outbreak. Due to the lockdown and availability and popularity of the virtual space, almost all people were looking for health-related solutions on the internet [1]. Thus, the majority of people log on for health-associated troubles to educate themselves. On the other side, the young, as well as the old generation, is also searching out fitness-related facts on the web. At some point, part of the population is also involved in searching for similar stories to resolve or to assist their mental health-related issues [2]. Thus, it can

V. Sugumaran et al. (Eds.): AIR 2022, CCIS 1738, pp. 179–190, 2022.
https://doi.org/10.1007/978-3-031-23724-9_17

be seen that nowadays, people are actively participating in sharing and discussing their health-related issues on the internet through various platforms, and thus a wide range of related data is available on the internet. This type of patient-related data available on social media can be proved as a significant tool in the healthcare sector such as suggesting health insurance, clinical pathway-based treatment methods, and drug recommendations based on the patient's health profile or assisting doctors [3].

A drug recommendation system is a framework for the healthcare sector that recommends the best drugs for a particular disease by analyzing data related to patients such as their background, reviews, other diseases, etc. [4, 5]. Any healthcare system requires studying huge data of the patients to conclude significant insights and help in the prediction of the best possible medications for the disease. With the advancements in machine learning (ML) and deep learning (DL) technologies, this huge amount of patient data can be used to extract meaningful insights for the betterment of the healthcare sector [6, 7]. Over the past decade, researchers have been analyzing the emotional impact of user experience and the severity of adverse drug reactions by extracting sentiment and semantic information from patient data [8, 9].

In this paper, a drug recommendation framework is proposed, and its operation is illustrated. The framework makes use of modern technologies, such as machine learning, natural language processing, sentiment analysis, etc., to uncover the interesting records that are hidden in the data and minimize medical errors made by doctors when prescribing medications. A database module, data preparation, data visualization, recommendation, and a section for model evaluation make up the proposed framework's many modules. The publicly accessible dataset on Kaggle [10] is used to create the suggested recommender architecture using machine learning N-Gram and Lightgbm algorithms. The main objective of the proposed work is to provide an optimized model for the medication suggestion framework to achieve the measurements like great exactness, adaptability, and proficiency.

With the goal of offering a better platform for the automation of Drug Recommendation, the work on the proposed system is presented as Sect. 2 represents the background and related work present on the Drug Recommendation System. Section 3 represents the adopted methodologies for the proposed system. Section 4 represents the outcome of our proposed system and the last section i.e.; Sect. 5 represents the conclusion and future scope of the presented work.

2 Background and Related Work

In today's era, the most significant and researched topic on the internet is health care or health-related information. In this digital space, everyone is looking for quick solutions, so the majority of people go online for health-related issues to educate themselves [1]. Nearly 60% of adults, according to a study [2], look online for adequate health information, with 35% of respondents focusing solely on online illness diagnosis. Previous studies have also shown that users are often looking for stories from "patients like them" on the Internet, which is hard to find among their friends and family [3, 4]. This kind of affected person-related facts available on social media may be proved as a substantial tool in the healthcare area from a specific point of perspective including health insurance, clinical pathway-based remedy strategies, and other drugs primarily based on the affected person's fitness profile or assisting doctors.

To draw important conclusions and aid in the prediction of the best treatments for a condition, a healthcare system needs to analyze vast amounts of patient data. With the development of machine learning (ML) and deep learning (DL) technologies, it is now possible to use this enormous amount of patient data to derive insightful information that will improve the healthcare industry [6, 7]. By collecting sentiment and semantic data from patient data, researchers have spent the last ten years examining the emotional influence of user experience and the severity of adverse drug reactions [8, 9].

For as long as decade, analysts have been examining the enthusiastic effect of client experience and the seriousness of antagonistic medication responses by extricating feeling and semantic data [11]. Past examinations have shown that wellbeing-related client-produced content is helpful according to various perspectives. One of the benchmark papers in this space was composed by Jane Sarasohn-Kahn [2]. It expresses that clients are regularly searching for stories from "patients like them" on the Internet, which is elusive among their loved ones. For as long as decade, analysts have been examining the enthusiastic effect of client experience and the seriousness of antagonistic medication responses by extricating feeling and semantic data [12, 15]. Leilei Sun [1] inspected enormous scope treatment records to find the best treatment solution for patients. The thought was to utilize a productive semantic bunching calculation assessing the similitudes between treatment records. In like manner, the creator made a system to evaluate the sufficiency of the proposed treatment. This construction can endorse the best treatment regimens to new patients according to their segment areas what's more, unexpected problems. An Electronic Medical Record (EMR) of patients assembled from various centers for testing. The outcome shows that this system further develops the fix rate. Xiaohong Jiang et al. [1] analyzed three unmistakable calculations, choice tree calculation, support vector machine (SVM), and backpropagation brain network on treatment information. SVM was picked for the drug proposition module as it performed really well in every one of the three novel limits - model precision, model capability, model adaptability. Furthermore, proposed the slip-up actually take a look at framework to guarantee investigation, accuracy and organization quality.

Because of an absence of trust and nature of client communicated clinical language, broad examination in the clinical and wellbeing area has not been finished [13, 14]. Along these lines, we expect to construct a stage where patients and clinicians can look by side effects and get drug proposals, symptoms of medications and acquire bits of knowledge into patients' portfolio.

3 Proposed Framework

The pipeline used to develop the proposed drug recommendation framework is presented in Fig. 1.

Fig. 1. Process for the development of proposed framework

3.1 Dataset Description

The dataset used in the experiments is a patient medication survey dataset containing ascribes like unique Id, drug name, condition (disease of the patient), date, helpful count, audits, and appraisals given by the patient as shown in Table 1.

Table 1. Data heads and their attributes

Attributes	Description of attribute	Type of attributes
DrugName	It is a categorical attribute which specifies the name of the drug prescribed by the practitioner	Categorical
Condition	It is a categorical attribute which states that in which condition or disease the drug is prescribed by the practitioner	Categorical
Review	It is the text review given by patient about the drug	Text
Rating	It is the overall patient satisfaction score	Numerical
Date	It is a date attribute which tells the date of entry of review	Date
Useful count	It is the number of patients that find the review useful	Numerical

3.2 Data Understanding and Pre-processing

It comprises of dataset collection, reviewing, cleaning and dataset preprocessing. The real-world information is raw information which can be splitted and unorganized and is unable to be used for training of the model. So, information cleaning is used to

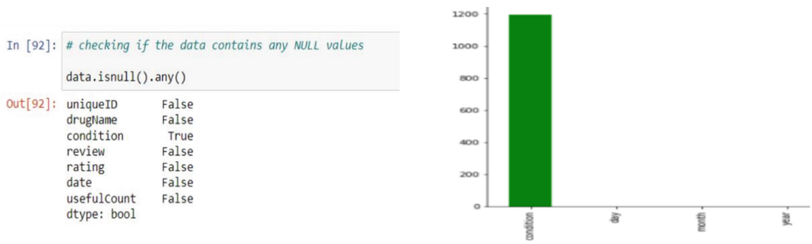

Fig. 2. Checking and counting null values

clean information. It consists of null/missing values processing, correlation analysis and removing duplicate data. The result is presented in Fig. 2.

Data Exploration: Following steps are followed during Data Exploration

(i) Analyzing patient ids to check if a patient has written more than one review
(ii) Find count of drugs for each condition by analyzing condition and quantity of drugs.

Data Cleaning; Following steps are followed during Data Cleaning

(i) Find the count of missing or empty fields for all the dimensions.
(ii) These none values can be removed, ignored or filled so the rows with missing values can be deleted.
(iii) Remove the redundancy in the dataset to normalize the data
(iv) Delete the rows with only drug as only one drug is unable to recommend the best one.
(v) Words like don't need, never, etc. should be removed from stop words as they don't possess any specific results about the attitude of review. At last removal of stop words is done to clean the reviews.

3.3 Data Visualization

Data Visualization is the process to visualize the data and relationships among different attributes and how one attribute depends on the other Some snapshots of relationships of attributes are shown below:

(i) Standard Data Cleaning techniques are applied to look at invalid qualities, copy columns, eliminating anomalies, and text from lines in this examination. Hence, eliminated each of the 1200 invalid qualities lines in the conditions segment.
(ii) Figure 3 shows top conditions with maximum number of drug available. After removing conditions that have no meaning, the dataset reduces to 212141 rows.

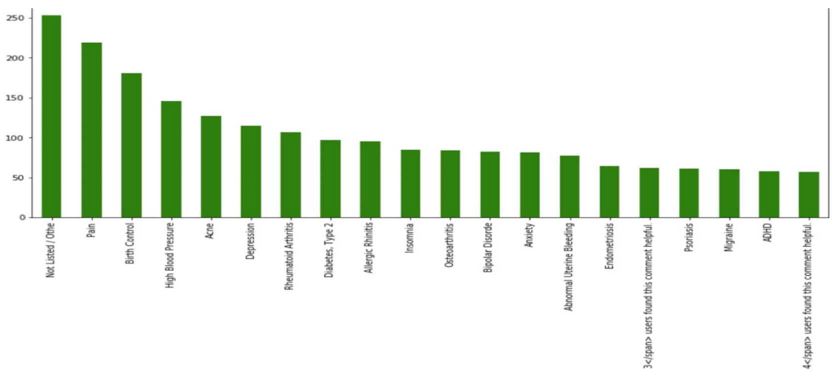

Fig. 3. Drugs per conditions

3.4 Approaches

Sentiment Analysis-Sentiment analysis is a technique for NLP that determines the emotional undertone of a body of text. This is a common method used by companies to gather and classify opinions. To mine text for sentiment and subjective information, it uses data mining, machine learning, and artificial intelligence. In Sentiment Analysis, predefined labels positive or negative is given to text document [16].

N-gram model-N-grams are a chunk of subsequent words, by studying these sequences we can efficiently understand the context in which a particular word is used. Example – the word 'book' can be used in different contexts like – to 'book' tickets, read the 'book'. Here the word book is used as the verb in the first phrase while as a noun in the latter. So, in order to efficiently understand the context of the word, N-grams look at the after the word and before the word and then determine if the word is used as a noun or verb in the sentence or in other context. N in N-grams denotes the number of words machine will look at before and after the target word. Ex- This 'book', A 'book', Your 'book' are all examples of bi-grams where before word 'book' is a noun. Bi-grams are the two pairs of words to look at before and after the target word while sliding over the words. The context can be extended by going to tri-grams which means looking at three pairs of words before and after the target word [17, 18].

Feeling investigation helps assessing the exhibition of items or administrations from client created substance. Vocabulary based opinion investigation approaches are liked 28 over learning-based ones when preparing information isn't satisfactory. Existing vocabularies contain just unigrams alongside their feeling scores. It is seen that opinion n-grams shaped by joining unigrams with intensifiers or invalidations show further developed outcomes. Such opinion n-gram vocabularies are not openly accessible. This paper presents a procedure to make such a vocabulary called Senti-N-Gram. Proposed rule-based methodology removes the n-grams opinion scores from an irregular corpus containing item surveys and relating numeric rating in 10-point scale. The scores from this computerized system are contrasted and that of the human annotators utilizing ttest and viewed as genuinely same N-grams can also be used to capture words in positive or negative context or viceversa. Example – 'the staff were not friendly, terrible really'. In this sentence 'Not friendly' and 'friendly terrible' is enough context to elucidate that the word 'friendly' is used in a negative context. In isolation the word 'friendly' is positive in when we are looking forward 'terrible' and backward 'not' which cancels out the positive meaning of the word [19, 20].

4 Implementation and Results

In this section, various results obtained during the development of model are presented. Figure 4, presented the view of the dataset used in the experiment after pre-processing of the data. It contains the patient unique Id and the six attributes as discussed in previous section.

```
# checing the sample of new dataset
data.sample(5)
```

	uniqueID	drugName	condition	review	rating	date	usefulCount
112607	170298	Quetiapine	Generalized Anxiety Disorde	"\r\n\r\n please tell the ones who is sufferin...	10	25-Jul-16	45
141096	136119	Acamprosate	Alcohol Dependence	"I was a two bottle plus red wine drinker ever...	9	28-Jan-14	97
35338	170518	Quetiapine	Bipolar Disorde	"I have been taking seroquel for a little over...	10	8-Oct-15	16
26858	209685	Lupron Depot	Endometriosis	"I started this about 6 -7 weeks ago after lap...	8	11-Oct-16	10
121226	122848	Linaclotide	Constipation, Chronic	"I've had chronic constipation for as lon...	9	23-Mar-14	113

Fig. 4. Dataset description

Some of the results obtained during visualization of the dataset are presented in Fig. 5 and 6, which shows the popular drugs, and most common conditions among patients respectively.

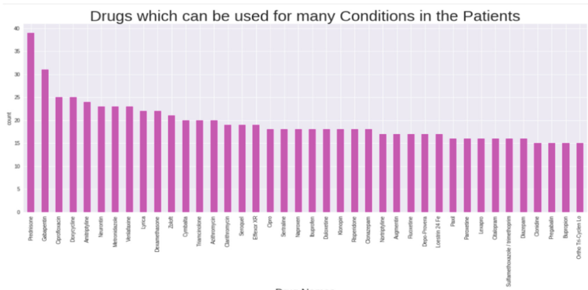

Fig. 5. Popular drugs

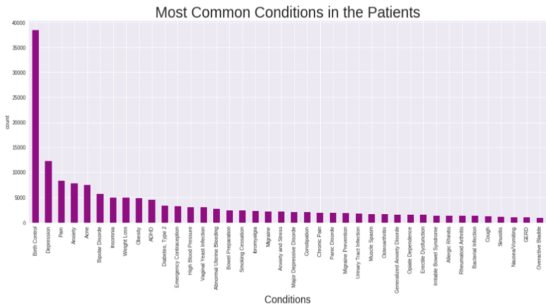

Fig. 6. Most common conditions in patients

Furthermore, during sentiment analysis of reviews, the world cloud of all reviews is presented in Fig. 7, followed by the world cloud of positive and negative reviews in Fig. 9 and 10 respectively. The weightage of positive and negative reviews is presented in Fig. 8.

After the classification of positive and negative sentiments, the analysis through 1–4 g is performed to check that which corpus best classifies the emotions. After analysis of

Fig. 7. World cloud of reviews

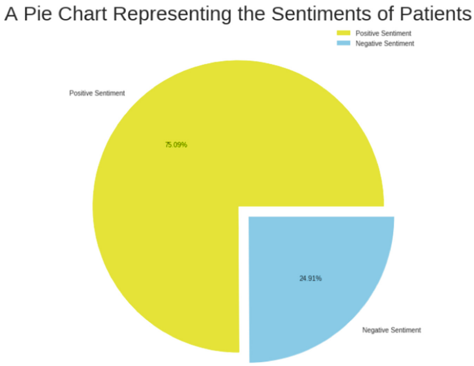

Fig. 8. Sentiment visualization of patients

Fig. 9. World cloud of positive reviews

1–4 g, it has been observed that 4-g classifies emotions much better than other grams, thus 4 g is used for further processing.

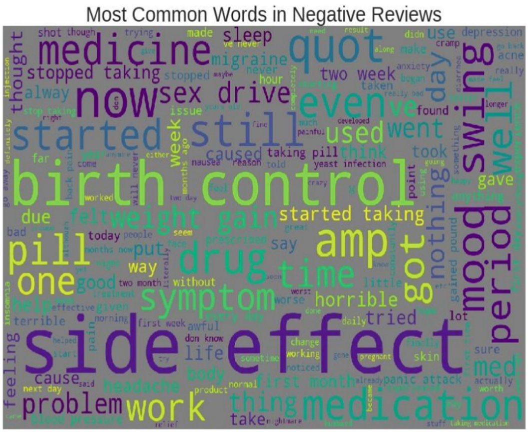

Fig. 10. World cloud of negative reviews

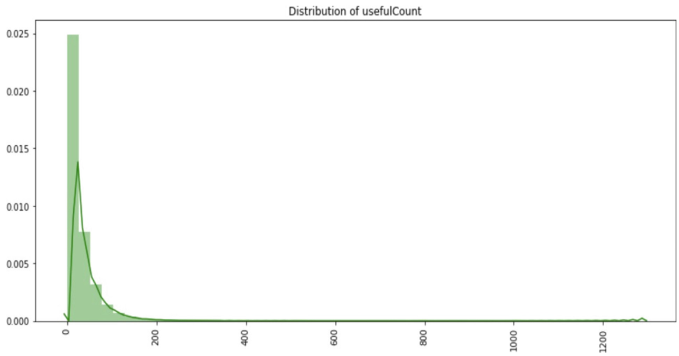

Fig. 11. Distribution of useful counts

From the distribution of useful Count shown in Fig. 11, it can be observed that the contrast between least and greatest is 1291, which is high. Likewise, the deviation is gigantic, which is 36. The justification behind this is that the more medications individuals search for, the more individuals read the survey regardless of whether their items are fortunate or unfortunate, which makes the most of the help exceptionally high. Thus, when designing the model, standardization is done on conditions.

Furthermore, for the prediction of reviews through sentiment analysis, various ML models have been used and tested such as Random Forest, Naïve Bayes, and Linear Regression. For this purpose, the raw text of the review is converted into a numeric data representation through a vectorization approach so that it can be used by a numerical classifier. Vectorization is done via the TF-IDF method which involves creating tokens (i.e. individual words or groups of words extracted from the text). Once the list of tokens is created, they are assigned an index integer identifier which allows them to be listed. It can help in providing the count of number of words in the document and normalize them in such a way that de-emphasizes words that appear frequently (like "a", "the", etc.). This creates what is known as a bag (multiset) of words. Such a representation associates a real-valued vector to each review representing the importance of the tokens (words) in the review. This represents the entire corpus of reviews as a large matrix where each row of the matrix represents one of the reviews and each column represents a token occurrence. Term-Frequency Inverse Document-Frequency (TF-IDF) is a way of handling the excessive noise due to words such as "a", "the", "he", "she", etc. Clearly such common words will appear in many reviews, but do not provide much insight into the sentiment of the text and their high frequency tends to obfuscate words that provide significant insight into sentiment.

For emotion analysis using a word dictionary, deep learning with an n-gram approach is used, where Harvard emotional dictionary is adopted. For feature analysis idea of attempting to derive importance out of the vectorized features by k-clustering by similarity. To compensate the limitation of natural language processing, Lightgbm machine learning model can be used, and reliability can be further secured through useful count The accuracy of the various tested models is presented in Table 2, and the output of the prediction task for a particular condition is presented in Fig. 12.

		total_pred
		mean
condition	drugName	
ADHD	Adderall	0.070960
	Adderall XR	0.042328
	Adzenys XR-ODT	0.010250
	Amantadine	0.011098
	Amphetamine	0.013925
	Amphetamine / dextroamphetamine	0.046908
	Aptensio XR	0.005885
	Armodafinil	0.028856
	Atomoxetine	0.047597
	Bupropion	0.083736
	Catapres	0.044449
	Clonidine	0.059211
	Concerta	0.059579
	Cylert	0.014713
	Daytrana	0.031764
	Desoxyn	0.133611
	Desvenlafaxine	0.006131
	Dexedrine	0.064658
	Dexmethylphenidate	0.041450
	Dextroamphetamine	0.052630
	Dextrostat	0.045610
	Dyanavel XR	0.016457
	Evekeo	0.008692
	Focalin	0.046282
	Focalin XR	0.044685
	Guanfacine	0.070408
	Intuniv	0.078066
	Kapvay	0.127024
	Lisdexamfetamine	0.045679
	Metadate CD	0.037710

Fig. 12. Prediction results

Table 2. Accuracy of tested models

Vectorizer	Model	Accuracy
TF-IDF	Linear regression	78.2%
	Naive Bayes	75.2%
	Random forest	83.1%
	N-Gram based model	89.4

5 Conclusion and Future Work

Reviews have now become part of our daily life, such as E-commerce reviews, Restaurant Reviews, fashion reviews, etc. Inspired by this, Sentiment Analysis on drug reviews is presented in this paper, where various ML-based approaches have been used to analyze the sentiments behind patient reviews to recommend drugs for a particular condition using Sentiment Analysis on Reviews. The result shows the Deep Learning frame work using N gram achieved an accuracy of 89.4%. As future work productivity of proposal framework can be expanded by including age of the individual, segment data during the preparation stage. Additionally, the brand and the substance contents accessible in the medication can work on the suggested prescriptions.

References

1. Sun, L., Liu, C., Guo, C., Xiong, H., Xie, Y.: Data-driven automatic treatment regimen development and recommendation. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016). https://doi.org/10.1145/2939672.2939866
2. Sarasohn-Kahn, J.: The Wisdom of Patients: Health Care Meets Online Social Media (2008)
3. Guidelines for Telemedicine. https://www.mohfw.gov.in/pdf/Telemedicine.pdf. Accessed 20 Dec 2021
4. Shani, G., Gunawardana, A.: Evaluating recommendation systems. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P. (eds.) Recommender Systems Handbook, pp. 257–297. Springer, Boston (2011). https://doi.org/10.1007/978-0-387-85820-3_8
5. Bao, Y., Jiang, X.: An intelligent medicine recommender system framework. In: IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), June 2016. https://doi.org/10.1109/iciea.2016.7603801
6. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A literature review and classification of recommender systems research. Expert Syst. Appl. **39**(11), 10059–10072 (2012). https://doi.org/10.1016/j.eswa.2012.02.038
7. Mu, R.: A survey of recommender systems based on deep learning. IEEE Access **6**, 69009–69022 (2018). https://doi.org/10.1109/ACCESS.2018.2880197
8. Calero Valdez, A., Ziefle, M., Verbert, K., Felfernig, A., Holzinger, A.: Recommender systems for health informatics: state-of-the-art and future perspectives. In: Holzinger, A. (ed.) Machine Learning for Health Informatics. LNCS (LNAI), vol. 9605, pp. 391–414. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-50478-0_20

9. Fernandez-luque, L., Karlsen, R., Vognild, L.K.: Challenges and opportunities of using recommender systems for personalized health education. Stud. Health Technol. Inform. **150**(903), 903–907 (2009)
10. Drug Review Data Set. https://www.kaggle.com/datasets/jessicali9530/kuc-hackathon-winter-2018. Accessed 12 Jan 2022
11. Goel, V., Gupta, A.K., Kumar, N.: Sentiment analysis of multilingual Twitter data using natural language processing. In: 8th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, pp. 208–212 (2018). https://doi.org/10.1109/CSNT.2018.8820254
12. Shimada, K., et al.: Drug-recommendation system for patients with infectious diseases. In: AMIA Annual Symposium Proceedings, p. 1112 (2005). PMID: 16779399; PMCID: PMC1560833
13. Pandey, S.C.: Data mining techniques for medical data: a review. In: 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), pp. 972–982 (2016). https://doi.org/10.1109/SCOPES.2016.7955586
14. Tekade, T.N., Emmanuel, M.: Probabilistic aspect mining approach for interpretation and evaluation of drug reviews. In: International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, pp. 1471–1476 (2016). https://doi.org/10.1109/SCOPES.2016.7955684
15. Doulaverakis, C., Nikolaidis, G., Kleontas, A., Kompatsiaris, I.: GalenOWL: ontology-based drug recommendations discovery. J. Biomed. Semant. **3**(1), 14 (2012). https://doi.org/10.1186/2041-1480-3-14
16. Popescu, A.-M., Etzioni, O.: Extracting product features and opinions from reviews. In: Kao, A., Poteet, S.R. (eds.) Natural Language Processing and Text Mining, pp. 9–28. Springer, London (2007). https://doi.org/10.1007/978-1-84628-754-1_2
17. Gopalakrishnan, V., Ramaswamy, C.: Patient opinion mining to analyze drugs satisfaction using supervised learning. J. Appl. Res. Technol. **15**(4), 311–319 (2017). https://doi.org/10.1016/j.jart.2017.02.005
18. Gräßer, F., Kallumadi, S., Malberg, H., Zaunseder, S.: Aspect-based sentiment analysis of drug reviews applying cross-domain and cross-data learning. In: Proceedings of the 2018 International Conference on Digital Health, April 2018. https://doi.org/10.1145/3194658.3194677
19. Ozsoy, M.G., Özyer, T., Polat, F., Alhajj, R.: Realizing drug repositioning by adapting a recommendation system to handle the process. BMC Bioinform. **19**(1) (2018). https://doi.org/10.1186/s12859-018-2142-1
20. Leveraging N-grams to Extract Context From Text. https://towardsdatascience.com/leveraging-n-grams-to-extract-context-from-text-bdc576b47049. Accessed 10 Mar 2022