PHARMA FEEDBACK RECOMMENDER

Mohammed Shekeeb.k and Dr.shidha M V

***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.**

I. 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, past clinical history. By combining clinical knowledge with the collective wisdom of patients as expressed in user reviews, this project aims to contribute to the evolution of drug recommendation system, fortering a more patient-centric approach to health decision making.

II. Literature survey

The work in this article focuses on pharmaceutical errors, which are reviewed for the general practitioner, with an emphasis on terminology, definitions, incidence, risk factors, disclosure and legal implications. A number of variables can contribute to medication errors, including those related to the drug, the patient, and the health care provider. One or more of the outcomes that doctors may face after making medication errors include losing the trust of their patients, civil lawsuits, criminal charges, and medical board discipline. Various approaches have been tried in the prevention of pharmaceutical errors with varying degrees of success. Physicians' ability to provide safe care to their patients can be improved by learning more about medication errors. In the more than 10 years since the last Community-Acquired Pneumonia (CAP) proposal from the American Thoracic Society (ATS) / Infectious Diseases Society of America, the guideline development process has changed and new clinical data (IDSA) have been created. Given the proliferation of information on diagnostic, treatment, and management decisions for the care of patients with CAP, we intentionally limited the scope of this framework to cover judgments ranging from the medical diagnosis of pneumonia to the discontinuation of antibiotic therapy and the wearing of chest imaging. This article offers a brief summary of facet mining methods as they are used in the search for new drugs. It is essential for the pharmaceutical industry to carry out research aimed at detecting adverse drug reactions as quickly as possible. It is a difficult task to identify important themes from short and noisy reviews. As a solution to this problem, a Probabilistic Aspect Mining Model (PAMM) is proposed to find aspects and objects related to class labels. Due to the special characteristic of PAMM, it focuses on discovering features specific to one class rather than simultaneously discovering features for all categories during each operation. Drug-drug and drug-disease interactions can be difficult to identify and finding the necessary information can be challenging due to the vast number of drugs already on the market and ongoing pharmaceutical research. Although international standards have been created to facilitate efficient information exchange, such as ICD-10 classification and UNII registration, healthcare personnel still need to be regularly informed to effectively identify drug interactions prior to prescribing. In previous publications, the use of Semantic Web technology was proposed as a solution to this problem. The work in this paper focuses on drug recommendation using a Graph Convolution Network, which mainly uses the mechanism of information propagation and embedding propagation layers to model high-order connectivity and

elaborate representation learning. The proposed system includes three key components, namely, an embedding layer, an information propagation layer, and a prediction layer. The work focuses on accuracy rather than evaluation of the recommendation system. The proposed system in this paper focused on recommending a parameter that is effective when using a curing parameter recommendation system. The proposed system includes a voting method that is developed by seven machine learning algorithms. These ML models are trained as classifiers primarily to recommend a candidate representative medical data set. The file with the highest frequency is selected as the recommended representative data set. Long short-term memory networks are used for heating curve prediction presets. In this research, a deep learning approach is proposed for health-based healthcare datasets. This approach automatically identifies what food should be served to which person based on condition and other parameters such as age, race, body weight, calories, fat, sodium, protein, fiber and cholesterol. The main focus of this study framework is the integration of deep learning and machine learning methods such as regression analysis, naive baye, recurrent neural networks, multi-level perceptrons, gated recurrent units, and long-short-term memory (LSTM). The properties of these IoMT samples were evaluated and further processed before applying machine learning, deep learning and other learning-based methods.

III. METHODOLOGIES

The dataset used in this research is Drug Review Dataset (Drugs.com) taken from the UCI ML repository [4]. This dataset contains six attributes, name of drug used (text), review (text) of a patient, condition (text) of a patient, useful count (numerical) which suggest the number of individuals who found the review helpful, date (date) of review entry, and a 10-star patient rating (numerical) determining overall patient contentment. It contains a total of 215063 instances. The proposed system in this paper focused on recommending a parameter that is effective when using a curing parameter recommendation system. The proposed system includes a voting method that is developed by seven machine learning algorithms. These ML models are trained as classifiers primarily to recommend a candidate representative medical data set. The file with the highest frequency is selected as the recommended representative data set. Long short-term memory networks are used for heating curve prediction presets. Susannah et al. [7] In this research, a deep learning approach is proposed for health-based healthcare datasets. This approach automatically identifies what food should be served to which person based on condition and other parameters such as age, race, body weight, calories, fat, sodium, protein, fiber and cholesterol. The main focus of this study framework is the integration of deep learning and machine learning methods such as regression analysis, naive baye, recurrent neural networks, multi-level perceptrons, gated recurrent units, and long-short-term memory (LSTM). The properties of these IoMT samples were evaluated and further processed before applying machine learning, deep learning and other learning-based methods.

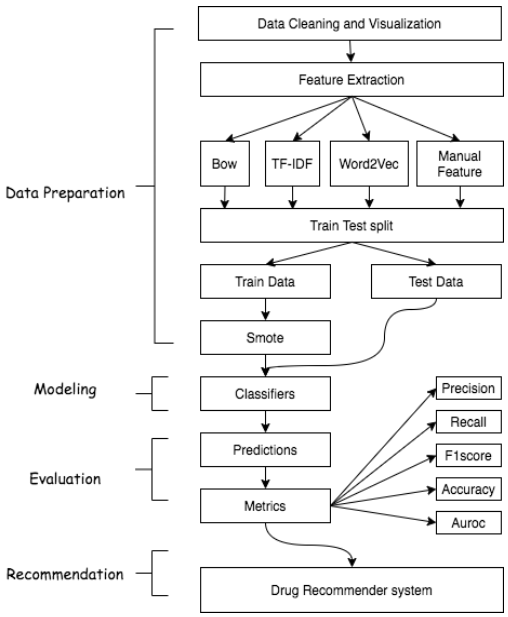

Fig1.Flowchart of the proposed model

It shows the proposed model used to build a medicine recommender system. It contains four stages, specifically, Data preparation, classification, evaluation, and Recommendation.

A. Data Cleaning and Visualisation

Applied standard Data preparation techniques like checking null values, duplicate rows, removing unnecessary values, and text from rows in this research. Subsequently, removed all 1200 null values rows in the conditions column. We make sure that a unique id should be unique to remove duplicacy.

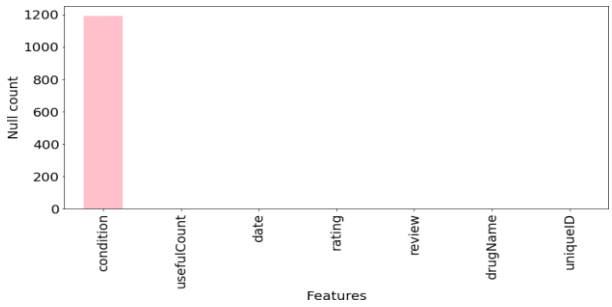

Fig. 2. Bar plot of the number of null values versus attributes

Fig. 3 shows the top 20 conditions that have a maximum number of drugs available. One thing to notice in this figure is that there are two green-colored columns, which shows the conditions that have no meaning. The removal of all these sorts of conditions from final dataset makes the total row count

equals to 212141.

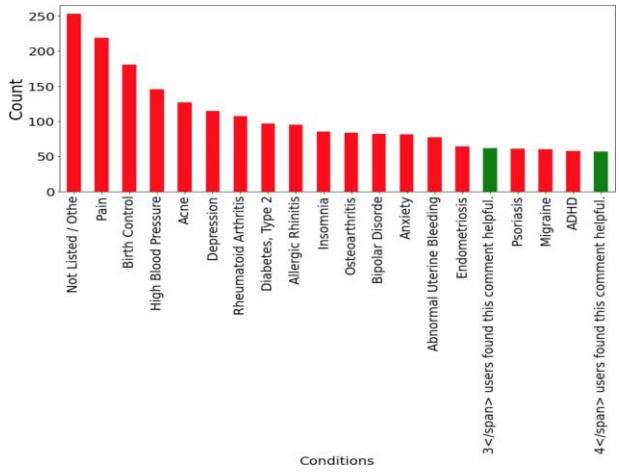

Fig. 3. Bar plot of Top 20 conditions that has a maximum number of drugs available

Fig. 4 shows the visualization of value counts of the 10-star rating system. The rating beneath or equivalent to five featured with cyan tone otherwise blue tone. The vast majority pick four qualities; 10, 9, 1, 8, and 10 are more than twice the same number. It shows that the positive level is higher than the negative, and people's responses are polar.

The condition and drug column were joined with review text because the condition and medication words also have predic- tive power. Before proceeding to the feature extraction part, it is critical to clean up the review text before vectorization. This process is also known as text preprocessing. We first cleaned the reviews after removing HTML tags, punctuations, quotes, URLs, etc. The cleaned reviews were lowercased to avoid duplication, and tokenization was performed for converting the texts into small pieces called tokens. Additionally, stopwords,

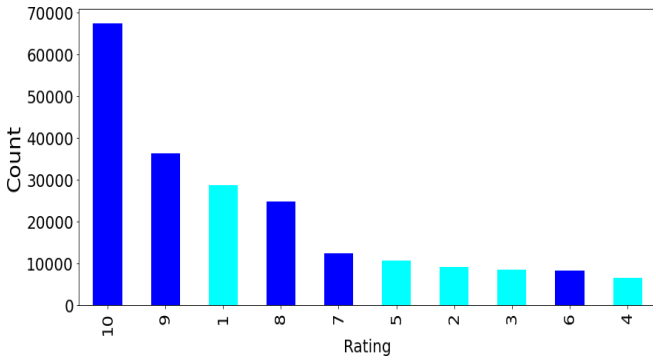

Fig. 4. Bar plot of count of rating values versus 10 rating number

for example, "a, to, all, we, with, etc.," were removed from the corpus. The tokens were gotten back to their foundations by performing lemmatization on all tokens. For sentiment analysis, labeled every single review as positive and negative based on its user rating. If the user rating range between 6 to 10, then the review is positive else negative.

B. Feature Extraction

After text preprocessing, a proper set up of the data required to build classifiers for sentiment analysis. Machine learning algorithms can't work with text straightforwardly; it should be changed over into numerical format. In particular, vectors of numbers. A well known and straightforward strategy for feature extraction with text information used in this research is the bag of words (Bow) , TF-IDF, Word2Vec. Also used some feature engineering techniques to extract features manually from the review column to create another model called manual feature aside from Bow, TF-IDF, and Word2Vec.

C. Train Test Split

We created four datasets using Bow, TF-IDF, Word2Vec, and manual features. These four datasets were split into 75% of training and 25% of testing. While splitting the data, we set an equal random state to ensure the same set of random numbers generated for the train test split of all four generated datasets. prevent a synthetic minority over-sampling technique (Smote) to prevent the class imbalance problem. Smote is an over- sampling technique that synthesized new data from existing data. Smote generates the new minority class data by linear interpolation of randomly selected minority instance 'a' in combination with its k nearest neighbor instance 'b' in the feature space. Table II shows the total distribution of data on final dataset i.e. after data cleaning. Fig. 6 shows the projection of non-smote and smote using t-distributed stochastic neighbor embedding (t-SNE) [21] of 1000 rows on manual features data. It displays that there are more orange points in the non-smote t-SNE projection, which represents the majority class domi- nance. It also shows that there has been an increment in blue points after using smote that brings out the balance between a majority and minority class that curbs the predominance of the majority class.

E. Classifiers

Distinct machine-learning classification algorithms were used to build a classifier to predict the sentiment. Logistic Regression, Multinomial Naive Bayes, Stochastic gradient descent, Linear support vector classifier, Perceptron, and Ridge classifier experimented with the Bow, TF-IDF model since they are very sparse matrix and applying tree-based classifiers would be very time-consuming. Applied Decision tree, Ran- domForest, LGBM, and CatBoost classifier on Word2Vec and manual features model. A significant problem with this dataset is around 210K reviews, which takes substantial computational power. We selected those machine learning classification al- gorithms only that reduces the training time and give faster predictions.

F. Metrics

The predicted sentiment were measured using five metrics, namely, precision (Prec), recall (Rec), f1score (F1), accuracy (Acc.) and AUC score [23]. Let the letter be: T_P = True positive or

occurrences where model predicted the positive sentiment truly, T_n = True negative or occurrences where model predicted the negative class truly, F_p = False positive or occurrences where model predicted the positive class falsely, F_n = False negative or occurrences where model predicted the negative class falsely, Precision, recall, accuracy, and f1score.

G. Drug Recommender system

After assessing the metrics, all four best-predicted results were picked and joined together to produce the combined prediction. The merged results were then multiplied with normalized useful count to generate an overall score of drug of a particular condition. The higher the score, the better is the drug. The motivation behind the standardization of the useful count was looking at the distribution of useful count in Fig. 7; one may analyze that the contrast among the least and most extreme is around 1300, considerable. Moreover, the deviation is enormous, which is 36. The purpose behind is that the more medications individuals search for, the more individuals read the survey regardless of their review is positive or negative, which makes the useful count high. So while building the recommender system, we normalized useful count by conditions

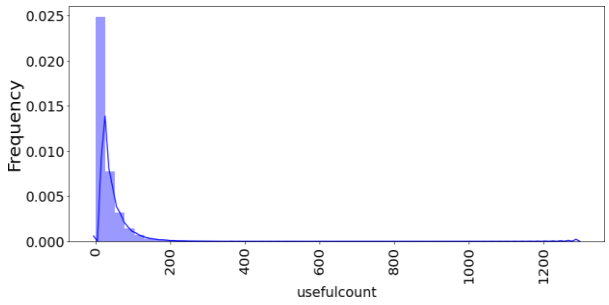

Fig. 5. Distribution of Useful Count

IV.RESULTS

In this work, each review was classified as positive or negative, depending on the user's star rating. Ratings above five are classified as positive, while negative ratings are from one to five-star ratings. Initially, the number of positive ratings and negative ratings in training data were 111583 and 47522, respectively. After applying smote, we increased the minority class to have 70 percent of the majority class examples to curb the imbalances. The updated training data contains 111583 positive classes and 78108 negative classes. Four different text representation methods, namely Bow, TF-IDF, Word2Vec, Manual feature and ten different ML algorithms were applied for binary classification.

The drug review sample used in this study was obtained from the UCI ML resource. This data consists of six components: the name of the drug used, the patient's rating, the patient's condition, a valuable number that indicates the number of people who experienced the benefit of the rating, the date the review was written, and a 10-star patient rating that indicates how the patient is doing overall satisfactory. According to the user's star rating, each review in this work has been categorized as positive or negative. Positive reviews are reviews with five or more stars, while negative reviews vary from one to five stars. In Figure 2, we can see the top medical conditions with the largest number of treatment options. One factor to observe in this figure is the fact that there are two green bars that indicate criteria that are of little importance.

condition	drugName	Score
Acne	Retin-A	0.069334
Acne	Atralin	0.088545
Acne	Magnesium hydroxide	0.088545
Acne	Retin A Micro	0.097399
Birth Control	Mono-Linyah	0.005448
Birth Control	Gildess Fe 1.5 / 30	0.005987
Birth Control	Ortho Micronor	0.006149
Birth Control	Lybrel	0.027766
High Blood Pressure	Adalat CC	0.303191
High Blood Pressure	Zestril	0.305851
High Blood Pressure	Toprol-XL	0.362589
High Blood Pressure	Labetalol	0.367021
Pain	Neurontin	0.158466
Pain	Nortriptyline	0.171771
Pain	Pamelor	0.231829
Pain	Elavil	0.304513
Depression	Remeron	0.124601
Depression	Sinequan	0.146486
Depression	Provigil	0.240185
Depression	Methylin ER	0.328604

IV. Discussion

The results procured from each of the four methods are good, yet that doesn't show that the recommender framework is ready for real-life applications. It still need improvements. Predicted results show that the difference between the positive and negative class metrics indicates that the training data should be appropriately balanced using algorithms like Smote, Adasyn [24], SmoteTomek [25], etc. Proper hyperparameter optimization is also required for classification algorithms to improve the accuracy of the model. In the recommendation framework, we simply just added the best-predicted result of each method. For better results and understanding, require a proper ensembling of different predicted results. This paper intends to show only the methodology that one can use to extract sentiment from the data and perform classification to build a recommender system.

V. Conclusion

Reviews are becoming an integral part of our daily lives; whether go for shopping, purchase something online or go to some 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 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.

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] 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.

[3] 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.

[4] Doulaverakis, C., Nikolaidis, G., Kleontas, A. et al. GalenOWL: Ontology-based drug recommendations discovery. J Biomed Semant 3, 14 (2012).

[5] Gao, Xiaoyan, Fuli Feng, Heyan Huang, Xian-Ling Mao, Tian Lan, and Zewen Chi. “Food recommendation with graph convolutional network.” Information Sciences 584 (2022): 170-183.

[6] Chen, Yu-Xiu, Li-Chih Wang, and Pei-Chun Chu. “A medical dataset parameter recommendation system for an autoclave process and an empirical study.” Procedia Manufacturing 51 (2020): 1046-1053.

[7] Fox, Susannah and Duggan, Maeve. (2012). Implementing a Machine Learning Model to Realize an Effective IOMTAssisted Client Nutrition Recommender System. Pew Research Internet Project Report.

[8] J. Ramos et al., “Using tf-idf to determine word relevance in document queries,” in Proceedings of the first instructional conference on machine learning, vol. 242, pp. 133-142, Piscataway, NJ, 2013.

[9] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique, 2011, Journal of Artificial Intelligence Research, Volume 16, 2020.