Drug Recommendation based on Sentiment Analysis using ML

Dr.Krishnareddy K R[1], Dr. Aravinda T V[2], Kumarswamy H[3] , Ramesh B E[4] , Avinash G M[5] , M R Desai[6], Vishwas C N[7]

[1]Professor, Department of Computer Science and Engineering S.J.M.I.T. Chitradurga, Karnataka
[2]Professor, Department of Computer Science and Engineering S.J.M.I.T. Chitradurga, Karnataka
[3]Assistant Professor, Department of Computer Science and Engineering S.J.M.I.T. Chitradurga, Karnataka
[4]Associate Professor, Department of Computer Science and Engineering S.J.M.I.T. Chitradurga, Karnataka
[5]Assistant Professor, Department of Computer Science and Engineering S.J.M.I.T. Chitradurga, Karnataka
[6]Assistant Professor, Department of Computer Science and Engineering S.J.M.I.T. Chitradurga, Karnataka
[7]Student, IV[th] Sem Mtech, Department of Computer Science and Engineering S.J.M.I.T. Chitradurga, Karnataka

---

Abstract - As e-commerce has grown, an increasing number of consumers prefer to buy drugs online for convenience's sake. As a result, the difficulties with health and medical diagnostics are becoming more well known. Because medical professionals write prescriptions based on their very limited experience, they frequently make drug mistakes. Technologies like data mining or recommender systems offer opportunities to investigate potential insight from diagnosis background records and aid physicians in reliably prescribing medications to reduce medication mistake. This effort aims to propose a medication recommendation system that can significantly lessen the workload of professionals. We create a system that, using ML classification algorithms, predicts whether a medicine can be taken or not based on the reviews.

***Key Words*:** Stress, Machine learning and IT professionals.

1. INTRODUCTION

On the Web, one of the most hotly debated subjects is health information. According to a 2017 survey by the Pew Internet or American Life Project, 59% of Americans have searched online for health-related information, with 35% of the respondents concentrating on online medical condition diagnosis. The findings reveal that an increasing number of people are concerned about the issue of medical diagnostics and health. However, drug errors continue to claim the lives of many people. According to the administration's assessment, drug errors cause more than 200,000 deaths annually in China and even 100,000 deaths in the United States Doctors are responsible for more than 42% of drug errors because they write prescriptions based on their limited experience. The following facts could contribute to these problems:

For serious illnesses,

(i) many hospitals lack doctors or medical specialists;

(ii) expert diagnosis mostly depends on the expert's experience, particularly for those naive novices who find it difficult to avoid errors.

While this is going on, the majority of diagnosis cases in hospitals remains in use and yet to be mined, making it impossible to discover the worth of the data. Massive amounts of data are produced by hospital information systems (HIS), making it difficult to extract potentially valuable insights from diagnosis case data. A viable approach to solving these difficult problems is through the use of data mining as well as recommender systems.

A recommender framework was a common system that makes an item recommendation to the user based on their benefit and need. These frameworks use the consumer surveys to analyze the responses and offer recommendations based on the respondents' precise needs. The project at hand Based on patient reviews, medication use is either advised or discouraged utilizing sentiment analysis & feature engineering. Sentiment analysis represents a succession of approaches, techniques, and instruments for identifying and separating emotional information from language, such as opinions and attitudes. In contrast, the process of "feathering engineering" involves adding new features to the ones that already exist in order to enhance model performance.

The work is to propose a medicine recommendation system that uses patient reviews topredict the sentiment using various vectorization processes like Bow, TF-IDF,Word2Vec, and Manual Feature Analysis, which can help recommend the to drug for a given disease by different classification algorithms. The predicted sentiments were evaluated by precision, recall, f1score, accuracy, and AUC score.

2. LITERATURE SURVEY

In 2014, **Zhang, Yin & Zhang, Dafang & Hassan, Mohammad & Alamri, Atif & Peng, Limei** I suggest a unique cloud-assisted drug recommendation system (CADRE) that can match users with top-N relevant medications based on their symptoms. In CADRE, the medications are initially grouped into a number of groups based on the functional description data, after which a simple personalized drug recommendation is created via user collaborative filtering. We then provide a cloud-assisted approach for enhancing enduser Quality in Experience (or QoE of drug recommendation by modeling or representing the relationship between the user, symptom, and medicine by tensor decomposition. This is done in light of the shortcomings in collaborative filtering algorithm, that include computing expensive, cold temperatures start, and data sparsity. Finally, an experimental investigation based on a genuine dataset downloaded from the Internet evaluates the suggested approach.

In 2015, Danushka Bollegala, Takanori Maehara and Kenichi Kawarabayashi I put forth an unsupervised approach for acquiring word representations that are precise in capturing the domain-specific facets of word semantics. They begin by choosing a selection of frequently occurring words that appear in the two domains as "emph" pivots. Then, they optimize a goal function that imposes two constraints: (a) pivots must reliably forecast the co-occurring non-pivots in documents from both the source and the target domains, plus (b) word representations learned for pivots must be similar across the two domains. They also suggest a technique for performing domain adaptation on the learned word representations. They demonstrate the highest sentiment classification accuracy rates across all domain-pairs inde a benchmark dataset using their suggested method, which greatly exceeds benchmark competitors' baselines, including the most advanced domain-insensitive word representations.

In 2016, Y. Bao and X. Jiang Create and develop a framework for a general medicine recommender system that incorporates data mining techniques. The database platform module, data prep module, recommendation model part, model assessment module, and data visualization modules make up the medication recommender system. Based on the medical data, they look at the SVM (Support Vector Classifier), BP neural network, and ID3 decision tree models for recommending medications. To improve performance, experiments are conducted to fine-tune the parameters underlying each algorithm. In order to achieve a suitable trade-off between model accuracy, system efficiency, and model scalability, the SVM recommendation model is chosen for the drug recommendation module in the supplied open dataset. To assure the accuracy of the diagnostic and the caliber of the service, they also suggest an error-checking process. The system can recommend medications with good efficiency, accuracy, and scalability, according to experimental data .

In 2016, J. Li, H. Xu, X. He, J. Deng and X. Sun In order acquire vector-based tweet representations and recommend hashtags, introduce a novel deep neural network model. To put it more specifically, they create distributed word representations using a skip-gram model and then train semantic phrase vectors using a convolutional neural network. The long short-term memory neural network with recurrent connections (LSTMRNN) is then trained using the text vectors. Without any feature engineering, they classify hashtags using the generated tweet vectors as features. The proposed LSTM-RNN model beats state-of-the-art approaches, according to experiments using real-world data from Twitter to suggest hashtag.

The LSTM unit also achieves the best performance when compared to ordinary RNN or gated Recurrent Unit (GRU).

3. ARCHITECTURE

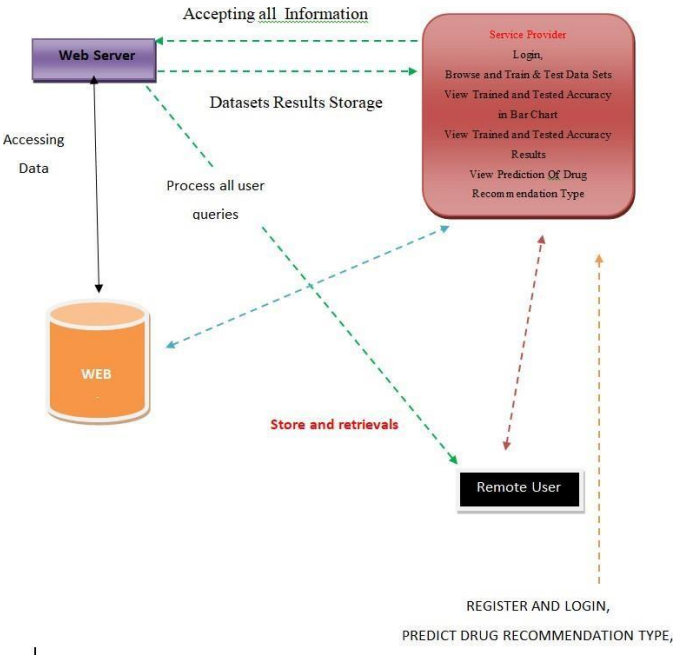

Fig -1: System architecture

The above figure 1 explains the architecture that would be used for developing a software product. The architecture diagram provides an overview of an entire system, identifying the main components that would be developed for the product and their interfaces.

4. METHODOLOGY

A suggested block diagram for a pharmaceutical recommender system is shown in Figure 2. Data preparation, designation, evaluation, and recommendation are the four processes that make up this process. Normal data preparation techniques are used in the project, including checking for null values, duplicating rows, deleting pointless values, and removing text from rows.

It is crucial to clean off the review language before vectorization before moving on to the extraction of characteristics section. Text preparation is another name for this procedure.

Here, cleaning the reviews after eliminating HTML tags, punctuation, quotations, URLs, etc. is the major procedure. To avoid duplication, edited reviews were lowercased, then tokenization was used to cut the sentences into discrete units called tokens.

After text processing, a suitable setup of the data needed to create sentiment analysis classifiers. Text should be converted to numbers so that algorithms based on machine learning can process it. Text cannot be processed directly by these algorithms. specifically, numerical vectors. a well-known and simple technique for feature extraction from text-based data. Then it uses a machine learning classification method, which takes less time to train and produces predictions more quickly.

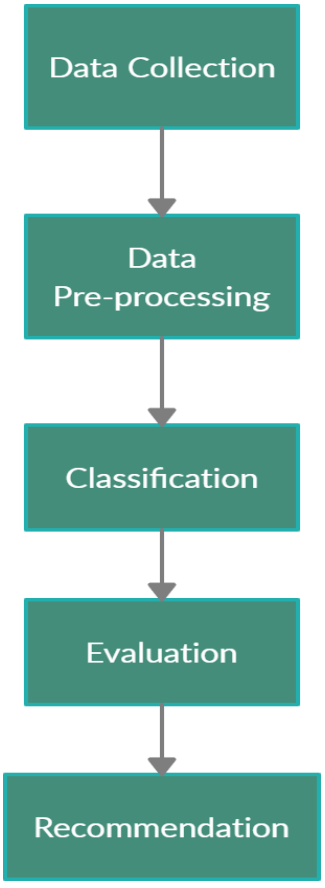

Fig -2: Methodology steps

5. RESULTS

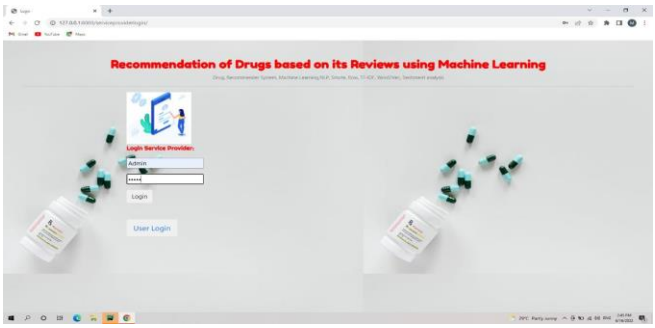

Fig -3 : Service Provider Login Page

Figure 3, shows the home page of the application.it has three menus register,admin and user login for admin and user.

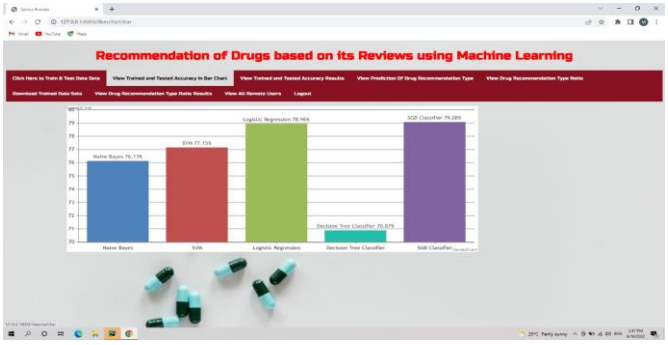

Fig -4: view trained and tested accuracy in bar chart page

Figure 4 shows the accuracy of the trained and tested datasets through the different agorithms.

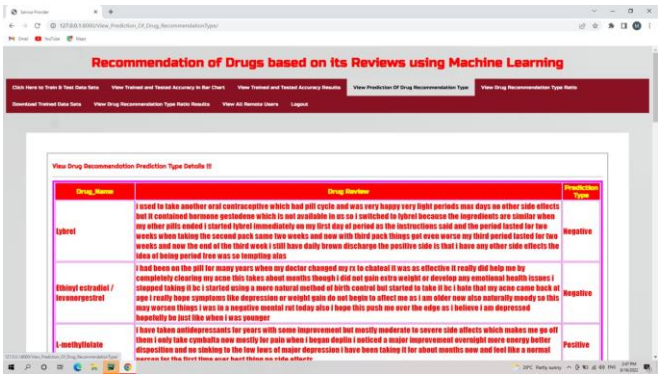

Fig -5 : View Prediction of Drug Recommendation Type Page

Figure 5 represents predicted results of the drug predicted by remote users in admin page.

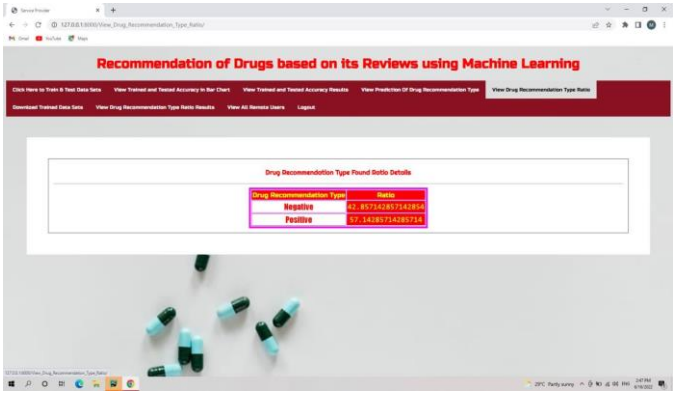

Fig -6 : View Prediction of Drug RecommendationType Ratio Results Page.

Figure 6 shows the over all remote users who are all registered and predicted the drug.

Fig -7 : Prediction of Drug Recommendation Type Page.

Figure 7 shows that prediction of drug recommendation type page ,where in this page user can enter the drug name and drug review to predict the drug.

6. CONCLUSION

Reviews are becoming integral part of our lives. Whenever we purchase something online, we first check reviewsto decide whether to buy that thing or not. In this proposed system, we are building a drug recommendation system in which based on the reviews we can decide whether the drug can be used or not. We first train the model by using five algorithms which are are Naïve Bayes, SVM (Support Vector Machine), Logistic Regression, Decision Tree Classifier, SGD (Stochastic Gradient Descent) Classifier. And based on the Accuracy of algorithms and correct prediction Naïve Bayes and SVM (Support Vector Machine) algorithms are used to

predict whether the drug can be used or not. If the model results review as positive, then the drug is safe to be used and if the model results review as negative then the drug is not safe to be used. Advantages of this system is that this system is more effective since it presents the proposed algorithm used in natural language processing responsible for counting the number of times of all the tokens in review or document and the system has exact sentiment analysis prediction techniques for Data Cleaning and Visualization. Disadvantages of this system is that this system cannot be used to predict drugs for severe diseases as the collection of information of such diseases is not so easy. 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. Zhang, Yin & Zhang, Dafang & Hassan, Mohammad & Alamri, Atif & Peng, Limei. (2014). CADRE: Cloud-Assisted Drug Recommendation Service for Online Pharmacies. Mobile Networks and Applications. 20. 348-355. 10.1007/s11036-014-0537-4.
2. Danushka Bollegala, Takanori Maehara and Kenichi Kawarabayashi. Unsupervised Cross-Domain Word Representation Learning, 2015; arXiv:1505.07184.
3. 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.
4. 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: 10.11IJCNN.2016.7727385.