e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss.1, 2022
 UGC CARE Listed (Group -I) Journal

Medication Suggestion System using Machine Learning techniques over Pharmacy Reviews

Mohammed Arshed Ali Khan [1] Mohammed Waheeduddin Hussain [2] Md Ateeq Ur Rahman [3]

[1] Research Scholar, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.

[2] Professor, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.

[3] Professor, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.

Abstract - Since the coronavirus emerged, it has been increasingly difficult to get legitimate therapeutic resources, such as the scarcity of specialists and healthcare professionals, appropriate equipment and medications, etc. There are many deaths as a result of the medical profession as a whole being in turmoil. Due to a lack of availability, people began taking medication on their own without the proper consultation, which made their health situation worse than usual. Recently, machine learning has proven useful in a variety of applications, and creative work automation is on the rise. This research article aims to propose a system for prescribing medications that can significantly reduce the workload of specialists. In this research project, we develop a drug recommendation system that makes use of patient reviews to forecast sentiment using a variety of vectorization techniques, including Bow, TF-IDF, Word2Vec, and manual feature analysis, which can support the recommendation of the best medication for a given disease by various classification algorithms. Precision, recall, f1score, accuracy, and AUC score were used to assess the anticipated sentiments. The Sequential Model and XGBoost classifier surpass all other models with roughly 95% accuracy, according to the data. We implemented this model in a real-world setting in which users can log in, submit their symptoms, and receive a list of medicines that are recommended for their condition.

Index Terms — Drug, Recommender System, Machine Learning, NLP, Sentiment analysis.

I. INTRODUCTION

On the Web, one of the most hotly debated subjects is health information. 59% of Americans have gone online for health-related information, according to a Pew Internet and American Life Project survey, with 35% of respondents concentrating on online medical problem 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, many hospitals lack either doctors or medical experts, and (ii) expert diagnosis mostly depends on the expert's expertise, particularly for those naive novices who find it difficult to avoid errors. While this is going on, the majority of diagnosis case data in hospitals is still being maintained and has not been 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 and recommender systems.

Numerous recommender system ideas have been put out since the mid-1990s, and numerous types of recommender system software have lately been created for a wide range of applications. Traditional recommender systems include hybrid recommendation systems, collaborative filtering (CF), content-based (CB), and knowledge-based (KB) strategies, all of which have some drawbacks. The recommendations in CB are overly specific, and the issues in CF are cold-start, sparseness, and scalability. We create our own architecture for a medical recommender system and use data mining and machine learning techniques to address these issues.

II. SYSTEM ANALYSIS

Problem Statement:

The world is experiencing a doctor shortage due to the exponential increase in corona virus cases, particularly in rural areas where there are fewer experts than in urban areas. A doctor must complete their education between six and twelve years. Therefore, it is impossible to increase the number of doctors in a short period. Clinical errors occur often today. Every year, medication errors have an impact on over 200 000 people in China and 100,000 people in the USA. Since specialists only have a limited amount of information, they frequently prescribe the wrong medication (more than 40% of the time). Choosing the best prescription is important for people who require medical professionals with extensive knowledge of microscopic conditions. 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.

Objective:

This research project aims to propose a system for prescribing medications that can significantly lessen the workload of specialists. In this study, we develop a drug recommendation system that makes use of patient reviews to forecast sentiment using a variety of vectorization techniques, including Bow, TF-IDF, Word2Vec, and manual feature analysis, which can support the recommendation of the best medication for a given disease by various classification algorithms.

Aim of the Project:

A medication recommender system is absolutely necessary in order to aid doctors and patients in expanding their knowledge about medications for particular medical conditions. A recommender

framework is a common system that makes an item recommendation to the user based on their benefit and needs. These frameworks use consumer surveys to analyze the responses and offer recommendations based on the respondents' precise needs. The drug recommender system uses sentiment analysis and feature engineering to conditionally provide medications based on patient reviews.

Scope of the Project:

Sentiment analysis is a collection of methods, procedures, and tools for locating and disentangling emotional content in languages, such as attitudes and opinions. In contrast, the process of "feathering engineering" involves adding new features to the ones that already exist in order to enhance model performance.

Proposed System:

A medication recommender system is absolutely necessary in order to aid doctors and patients in expanding their knowledge about medications for particular medical conditions. A recommender framework is a common system that makes an item recommendation to the user based on their benefit and needs. These frameworks use consumer surveys to analyze the responses and offer recommendations based on the respondents' precise needs. The drug recommender system uses sentiment analysis and feature engineering to conditionally provide medications based on patient reviews. Sentiment analysis is 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 Sequential Model and XGBoost classifier surpass all other models with roughly 95% accuracy, according to the data. We implemented this model in a real-world setting in which users can log in, submit their symptoms, and receive a list of medicines that are recommended for their condition.

Advantages:

- Able to predict drug recommender systems uses sentiment analysis and feature engineering most accurately.
- By receiving a list of medicines that are recommended for their condition Minimize loss and maximize profits.

III. PROPOSED MODULAR IMPLEMENTATION

The Algorithm/ Technique used:

The dataset is fit a model after performing Data Pre-processing and Feature Transformation. The training set is fed into the algorithm to learn how to predict values. Testing data is given as input after Model Building a target variable to predict. The models are built using a technical approach implementation for a Drug recommendation system using the Drug review dataset.

- Data cleaning and visualization
- Feature Extraction
- Data Preparation using BOW, TF-IDF, and Word2Vec
- Splitting the data set into test data and training data
- Data modeling using Sequential and XGBoost classifier
- Data Evaluation and prediction

- Building the Drug recommendation system

Below is the proposed modular implementation of the project. It consists of two modules:

1. Admin
2. User

Admin Module:

The admin of the system is responsible for the activities like:

1. Get stock data from yahoo finance
2. Data Analysis of the dataset
3. Splitting the dataset for training and testing
4. Training the model for multiple algorithms
5. Review the performance of the algorithms on the given dataset
6. Create the model using LSTM algorithm.

User Module:

The user of the system can utilize the machine learning services that are offered like:

1. Logging into the system
2. Enter stock details to predict future trends
3. Receive prediction for future trend

IV. PROJECT EXECUTION

Upload Dataset:

On this page, the administrator of the system can upload datasets that are used for training the machine learning models. The admin has to select the file by clicking on the Choose file button and click on the upload button to upload the file to the server. Once the upload is complete, a success message would be displayed that the file is successfully uploaded. For this project we are using Train_3.csv as a dataset.

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss.1, 2022

Exploratory Data Analysis:

Exploratory Data Analysis is performed on the dataset in order to clean the dataset for any missing data, identify patterns, identify the relationships of various parameters of the outputs with the help of graphs, statistics etc. so that Common Conditions Analysis, Common Drugs Analysis, Common Reviews Analysis, Available Drugs per Condition Analysis, Drugs Used Over For Conditions Analysis, Condition Count Analysis can be performed.

Common Conditions Analysis:

The below graph shows the Common Conditions graph most common type of crimes from Training dataset drugsComTrain_raw.csv File.

Common Drugs Analysis:

The below graph shows the Common Drugs Analysis graph for past one year Stock Data of Microsoft company from Training dataset drugsComTrain_raw.csv File.

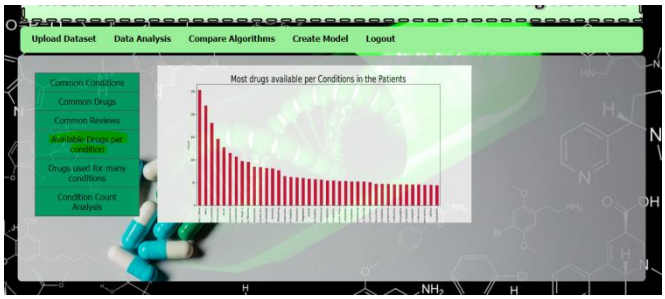
e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss.1, 2022

Common Reviews Analysis:

The below graph shows the Common Reviews Analysis graph for past one year Stock Data of Microsoft company from Training dataset drugsComTrain_raw.csv File.

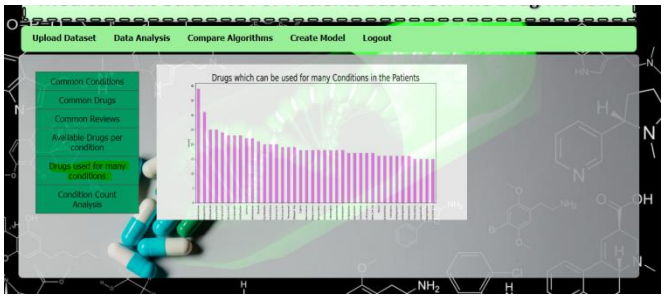
Available Drugs per Condition Analysis:

The below graph shows the Available Drugs per Condition Analysis graph that shows different crimes recorded over the years from Training dataset drugsComTrain_raw.csv File.



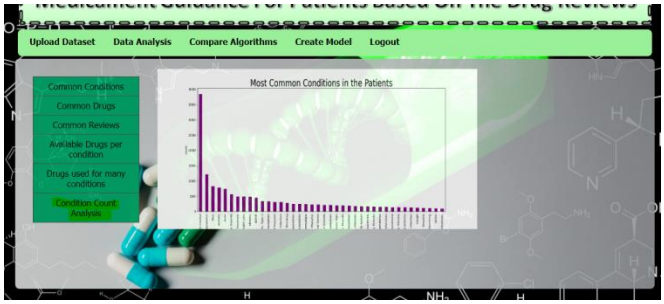
Drugs Used Over For Conditions Analysis:

The below graph shows the Drugs Used For Many Conditions Analysis graph for Proportions of crime during regular days Vs holidays from Training dataset drugsComTrain_raw.csv File.

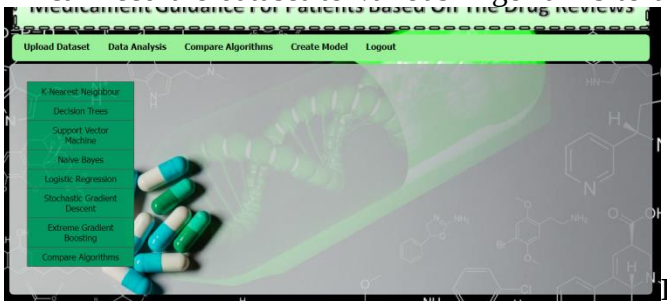


Condition Count Analysis:

The below graph shows the Condition Count Analysis graph for Proportions of crime during regular days Vs holidays from Training dataset drugsComTrain_raw.csv File.



Compare Algorithms:

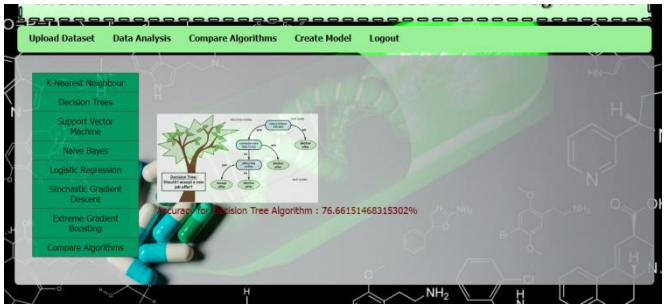
On this page, the admin can feed the dataset to various Algorithms to train them and get the test accuracy for each algorithm. 

K-Nearest Neighbour:

When the dataset is feed to K-Nearest Neighbour algorithm we observe that the test accuracy is 72.90%.

Decision Trees:

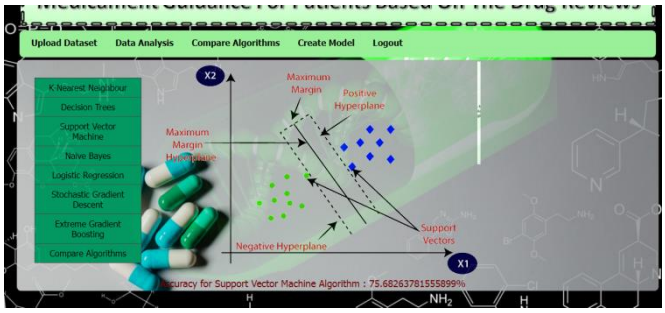
When the dataset is feed to Decision Trees algorithm we observe that the test accuracy is 76.66%.



e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss.1, 2022
  UGC CARE Listed (Group -I) Journal

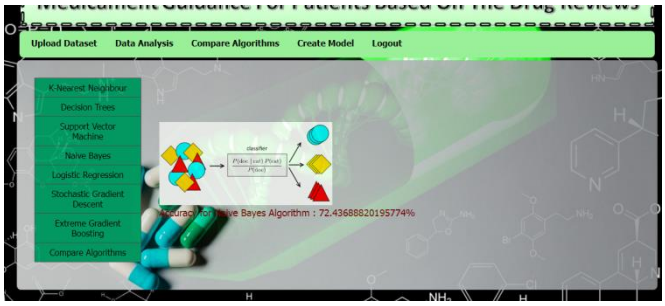
Support Vector Machine:

When the dataset is feed to Random Forest algorithm we observe that the test accuracy is 75.68%.



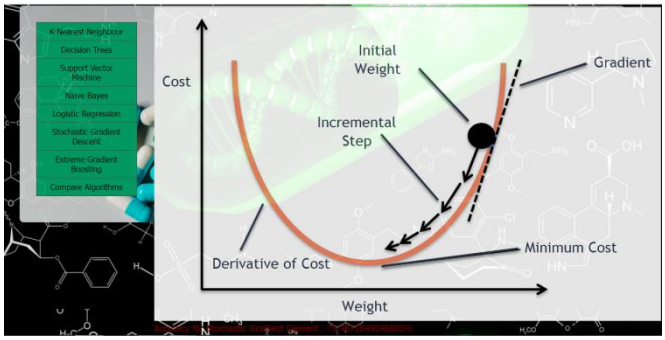
Naive Bayes:

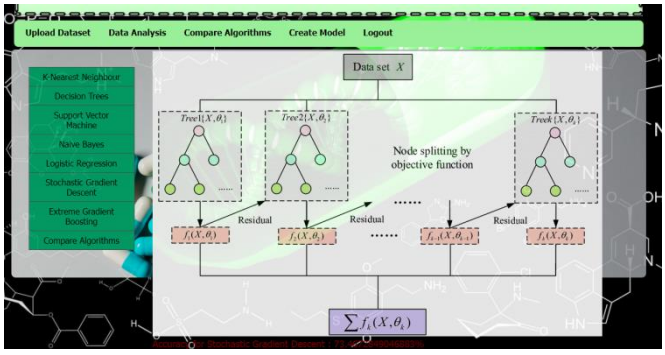
When the dataset is feed to Naive Bayes algorithm we observe that the test accuracy is 72.43%.



Stochastic Gradient Descent:

When the dataset is feed to Stochastic Gradient Descent algorithm

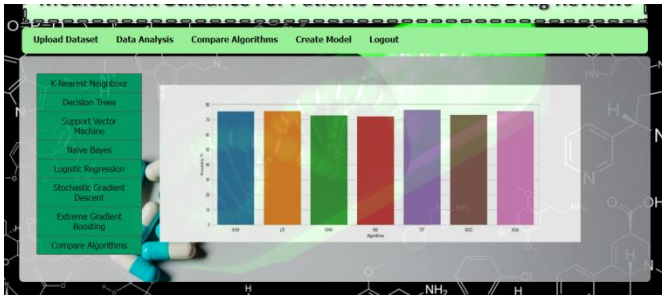




Compare Algorithms:

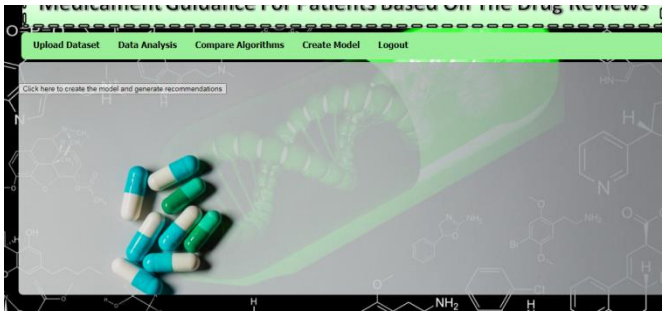
This screen shows the comparison of various test accuracies of the Algorithms.

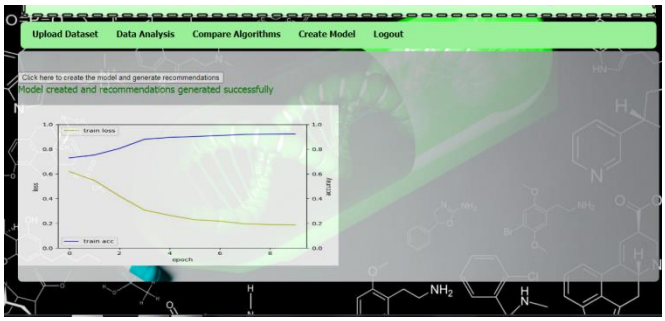
e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss.1, 2022
  UGC CARE Listed (Group -I) Journal



Create Model:

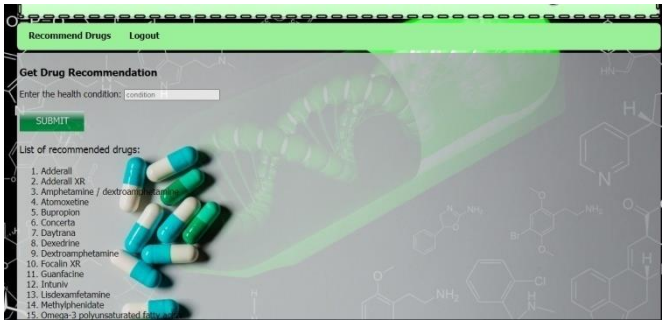
This screen shows the creation of Model for better optimized of system.





Prediction

This is in User Home Page for the user module. The user need to login into the system with his credentials in order to facilitate Drug prediction over Drug dataSet.



V. CONCLUSION

In this project we successfully predicted the drugs using Medicament Guidance For Patients Based On The Drug Reviews system built. To achieve, we have taken research article aims to propose a system for prescribing medications that can significantly reduce the workload of specialists. In this research project, we develop a drug recommendation system that makes use of patient reviews to forecast sentiment using a variety of vectorization techniques, including Bow, TF-IDF, Word2Vec, and manual feature analysis, which can support the recommendation of the best medication for a given disease by various classification algorithms. Precision, recall, f1score, accuracy, and AUC score were used to assess the anticipated sentiments. The Sequential Model and XGBoost classifier surpass all other models with roughly 95% accuracy, according to the data. We implemented this model in a real-world setting in which users can log in, submit their symptoms, and receive a list of medicines that are recommended for their condition.

REFERENCES

[1]Telemedicine, https://www.mohfw.gov.in/pdf/Telemedicine.pdf
[2] Wittich CM, Burkle CM, Lanier WL. Medication errors: an overview for clinicians. Mayo Clin Proc. 2014 Aug;89(8):1116-25.
[3] CHEN, M. R., & WANG, H. F. (2013). The reason and prevention of hospital medication errors. Practical Journal of Clinical Medicine, 4.
[4] Drug Review Dataset,
https://archive.ics.uci.edu/ml/datasets/Drug% 2BReview%2BDataset%2B%2528Drugs.com%2529#
[5] Fox, Susannah, and Maeve Duggan. ”Health online 2013. 2013.”
URL:http://pewinternet.org/Reports/2013/Healthonline.aspx
[6] Bartlett JG, Dowell SF, Mandell LA, File TM Jr, Musher DM, Fine MJ. Practice guidelines for the management of community-acquired pneumonia in adults. Infectious Diseases Society of America. Clin Infect Dis. 2000 Aug;31(2):347-82. doi: 10.1086/313954. Epub 2000 Sep 7. PMID: 10987697; PMCID: PMC7109923.
[7] Fox, Susannah & Duggan, Maeve. (2012). Health Online 2013. Pew Research Internet Project Report.
[8] T. N. Tekade and M. Emmanuel, ”Probabilistic aspect mining approach for interpretation and evaluation of drug reviews,” 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, 2016, pp. 1471-1476, doi: 10.1109/SCOPES.2016.7955684.

[9] Doulaverakis, C., Nikolaidis, G., Kleontas, A. et al. GalenOWL: Ontology-based drug recommendations discovery. J Biomed Semant 3, 14 (2012). https://doi.org/10.1186/2041-1480-3-14
[10] Leilei Sun, Chuanren Liu, Chonghui Guo, Hui Xiong, and Yanming Xie. 2016. Data-driven Automatic Treatment Regimen Development and Recommendation. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
(KDD '16). Association for Computing Machinery, New York, NY, USA, 1865–1874. DOI:https://doi.org/10.1145/2939672.2939866