A project report on

DRUG RECOMMENDATION SYSTEM USING LSTM MACHINE LEARNING ALGORITHM

Submitted in partial fulfillment for the award of the degree of

Master of Computer Applications (MCA)

by

Md Nouman.S (21MCA0278)

SCHOOL OF INFORMATION TECHNOLOGY & ENGINEERING

April, 2023

DECLARATION

I hereby declare that the thesis “DRUG RECOMMENDATION SYSTEM USING LSTM MACHINE LEARNING ALGORITHM” submitted by me, for the awards of the degree of Master of Computer Applications VIT is a record of bonafide work carried out by me under the supervision of Dr Vanitha M.

I further declare that work reported in this thesis has not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

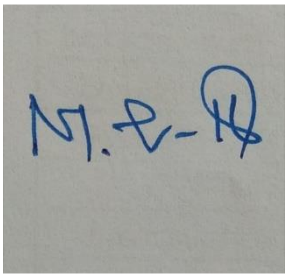
Place: Vellore.

DATE :07/04/2023 Signature of the candidate.

CERTIFICATE

This is to certify that the thesis entitled "DRUG RECOMMENDATION SYSTEM USING LSTM MACHINE LEARNING ALGORITHM" submitted by MD NOUMAN.S (21MCA0278) School of Information Technology and Engineering VIT, for the award of the degree Master of Computer Applications is a record of bonafide work carried out by him under my supervision.

The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The Project report fulfils the requirements and regulations of VIT and in my opinion meets the necessary standards for submission.



Signature of the Guide. **Signature of the Hod.**

Internal Examiner. **External Examiner.**

ABSTRACT

An application that uses electronic health records (EHR) to forecast and propose medicines that are right for patients is known as a drug recommendation system utilising the LSTM algorithm. The system uses a recurrent neural network (RNN) called a Long Short-Term Memory (LSTM) method, which can represent sequential data and capture long-term dependency issues with the data input.The LSTM-based drug recommendation system can provide clinicians with tailored prescription recommendations by examining patient demographics, medical history, symptoms, and lab test findings. The approach can help doctors make well-informed choices on the prescription of pharmaceuticals that are catered to the patient's particular needs and medical background. The drug recommendation system can enhance patient outcomes and lower healthcare expenditures by using this tailored approach. The drug recommendation system can help doctors make decisions about prescription prescribing that are supported by the best available data, improving patient outcomes and minimising drug side effects. The method can also lower healthcare expenditures by decreasing the possibility of needless prescriptions and lowering the expenses related to managing adverse drug responses.

ACKNOWLEDGEMENT

It is my pleasure to express with a deep sense of gratitude to **Dr Vanitha M**, Associate Professor Sr, School of Information Technology and Engineering (SITE), Vellore Institute of Technology, for his/her constant guidance, continual encouragement, understanding; more than all, he taught me patience in my endeavour. My association with him/her is not confined to academics only, but it is a great opportunity on my part to work with an intellectual and expert in the field of Computer Application.

I would like to express my gratitude to Dr. G. VISWANATHAN, Chancellor, Shri. SANKAR VISWANATHAN, Vice President, Dr. RAMBABU KODALI, Vice Chancellor, Prof. NARAYANAN, PRO-Vice Chancellor and Dr. SUMATHY S, Dean, School of Information Technology and Engineering, for providing me with an environment to work in and for his inspiration during the tenure of the course.

In a jubilant mood, I express ingeniously my whole-hearted thanks to Dr. VANITHA M, HOD, all teaching staff and members working as limbs of our university for their not-self-centred enthusiasm coupled with timely encouragements showered on me with zeal, which prompted the acquirement of the requisite knowledge to finalize my course study successfully. I would like to thank my parents for their support. It is indeed a pleasure to thank my friends who persuaded and encouraged me to take up and complete this task. Last, but not least, I express my gratitude and appreciation to all those who have helped me directly or indirectly toward the successful completion of this project.

Place: Vellore

Date: 12/03/2023

MD NOUMAN.S

Name of the student

TABLE OF CONTENTS

CHAPTER 5

IMPLEMENTATION & TESTING

Chapter-6

Results

Chapter-7

Conclusion & Future Work

Chapter-8

Reference

LIST OF FIGURES

LIST OF TABLE

Chapter - 1

Introduction

1.1 INTRODUCTION

Especially for doctors and other healthcare professionals who are in charge of providing medications to patients, recommending pharmaceuticals is an important duty in the field of medicine. Recently algorithms like Long-Short—Term-Memory (LSTM) have demonstrate promising outcomes in identifying and prescribing appropriate medications for patients. Recurrent- neural -networks (RNNs) with the ability to recognise long--term dependencies and sequentially patterns are ideally suited for modelling time-series data, such as that seen in electronic health records (EHRs). An LSTM-based drug recommendation system can give clinicians individualised prescription recommendations that are suited to the patient's particular needs and medical history by utilising EHR data, which includes patient demographics, medical history, symptoms, and lab test results. This is how LSTM-based drug recommendation systems ensuring that patients receive the most efficient and suitable medications for their diseases has the potential to enhance patient outcomes and lower healthcare costs.

1.2 PROJECT STATEMENT

• To guarantee that patients receive high-quality, cost-effective treatment, doctors must overcome a variety of hurdles.

• Since medical errors can have devastating effects on patients, clinicians are under pressure to lower their frequency. Errors can be challenging to prevent due to the healthcare industry's growing complexity.

• Medical professionals around the world frequently lack staff, supplies, and prescriptions. It may be challenging to give patients the best care possible as a result.

1.3 OBJECTIVES

- Effectiveness: Ensuring that the medication is successful in treating the ailment for which it is being prescribed.
- Safety: Ensuring that the patient can take the prescribed medication without experiencing any serious side effects.
- Considering patient-specific elements including age, weight, medical history, and any additional drugs they could be taking.
- Cost-effectiveness: Taking into account the drug's price and the patient's capacity to afford it.
- Accessibility: Taking into account the drug's accessibility to the patient and its proximity to the patient.
- Adherence: Taking into account the patient's propensity for following the recommended course of treatment and suggesting a medication that is convenient and simple for the patient to take..

1.4 SCOPE OF THE PROJECT

- Creating User Interface Development and user interface for the drug recommendation system is created. The user interface should enable clinicians to enter patient information and view drug recommendations.
- test using actual data in order to assess its precision, effectiveness, and usability. In order to evaluate the system's performance, it may also be contrasted with other clinical guidelines or drug recommendation systems.
- Improve accuracy

Chapter-2

Literature Survey

2.1 SUMMARY OF THE EXISTING SYSTEM

The author described in this paper[1] uses patient reviews to predict sentiment using a variety of vectorization techniques, including Bow, TF-IDF, Word2Vec, and manual feature analysis. These techniques can assist in recommending the best medication for a given disease by using various classification algorithms. The findings demonstrate that the performance of the classifier LinearSVC employing TF-IDF vectorization and algortim can be enhanced.

The suggested model in this paper[2] is made up of a number of machine learning algorithms, including Logical Regression, XgBoost, and Naive Bayes approaches, all of which were trained using a set of medical records. The result is totally dependent on the data collection and implementation strategy used, and it could change if more processes are added to improve any of the models mentioned above. The unimplemented future project was a web application.

The author of [3] paper has created a method that will help drug users and medical professionals select and locate the right medications. The datasets were trained using the semi-supervised clustering and classification technique. The deep learning method was used to more accurately categorise the medications with side effects. Moreover, give reviews on medications that come with the highest praise.

The author of [4] paper proposes the Drug Recommendation Model based on Message Propagation Neural Network (abbreviated as DRMP) in this study. The drug-drug interaction (DDI) knowledge is incorporated into the proposed model in order to reduce the rate of DDIs in drugs that are advised. The presented model is enlarged to a Bayesian Neural Network in order to implement several recommendations (BNN).

IN [5]the authors has present Medical sentiment analysis is a difficult and complex task, as the patients who submit the reviews are typically non-professional users and prefer to use informal language, according to the author of the [6] research. It uses a bidirectional LSTM algorithm. The complexity of medical text information is more than that of product-related language because of the greater lexical diversity and indirect representation of sentiment.

In [6] paper, the author proposed a method To increase the equity and safety of treating infectious diseases, has provide systems that propose drugs using layered artificial neural networks This kind of technology might be helpful in recommending secure medications to patients, particularly in times of medical emergency. By altering the threshold value, the design leverages statistical analysis to enhance accuracy and balances fairness.

In [7] study, the author uses DNN to build a model and determine the high-order relationships between diseases. We contend that low-order illness relations cannot be neglected, drawing inspiration from DeepFM [18]. The model generates a score that represents the likelihood that the user would contract the disease by incorporating relations that are both high-order and low-order. As the possible disease's prognosis is solely based on statistics and the characteristics of disease arre inexplicable, we must combine some expert knowledge with certain techniques to increase the likelihood that the prognosis will come true. To improve the precision of the forecast, one method is to use the symptoms of the patients and additional information. the other is taking into account the explicit characteristics and additional expert knowledge of diseases to increase predictability with particular tactics.

In [8]The author has discusses how Diabetes Mellitus is one of the most common illnesses, affecting a large number of people. The current protocol at crisis centres is to gather the necessary data for diabetes discovery through a variety of testing, and appropriate treatment is given ward on discovery. Important article on how crucial learning programmes are used to treat common HRV and undiagnosed diabetes, according to our most recent statistics. This is possible if specialists and medical facilities provide enough comprehensively measured information to investigate the community. Anticipated information can serve as a patient's notification signal if the specialist exercises sufficient restraint and caution.

In [9]th paper developer has provided a lot of beneficial services and applications that have greatly enhanced the healthcare sector and made people's life easier and better.
In the suggested design, six IoMT sensors are connected to a network of four hosts, and their data is provided to the MQTT broker via a network of switches. The six IoMT sensors include blood pressure, pulse rate, ECG, SpO2, body temperature, GSR, and health data.To forecast the host's state of health, five machine learning (ML) algorithms—including KNN, decision trees, k-means clustering, linear regression, and LSTM—are trained on the preprocessed data. In terms of accuracy, recall, and F1-Score, LSTM leads the pack with a score of 78%, or 71.Moreover, five crucial vital indicators out of the four To ascertain whether a certain patient is at risk, Five vital indicators from the four hosts—BP systolic, BP diastolic, pulse rate, body temperature, and SpO2—are compared with a covid usage case to assess the risk of developing covid.Five vital signs from the four hosts—BP systolic, BP diastolic, pulse rate, body temperature, and SpO2—are matched with a covid use case to evaluate whether a particular patient is at risk of developing covid.

The [10] paper study's author explains experimental results that show learning--based and lexicon--based strategies complement one another better than they do individually.
TextBlob has additionally generated excellent outcomes, obtaining accuracy rates of 96%% with MLP when used with TF-IDF and with LR when used with TF U TF-IDF.

It combines lexicon-based and learning-based approaches for labelling and categorization to accomplish this.As a data annotation technique for medication reviews, TextBlob, VADER, and AFFIN are each evaluated for effectiveness.Two traditional approaches—TF and TF-IDF—as well as one modified approach—the merger of TF and TF-IDF—are used to extract features from the annotated data.

In order to determine the most effective ways to represent persons, illnesses, and treatments in a shared low-dimension space, the author of [11] developed graph-based embedding models.The proposed framework differs from the majority of the previous studies in that it can effectively recommend recently produced medicines to patients.Advising new patients on safe drugs

The developer of the [12] paper discusses lstm, which stands for Neurocomputing and machine learning have both been significantly changed by long short-term memory. Some internet sources claim that this methodology has significantly enhanced Google Translate's machine translations, Alexa's responses, and speech recognition. Facebook also uses this neural network, and as of 2017, it was performing more than 4 billion LSTM-based translations every day. Convolution and pooling layers were utilised in such hybrid models to drastically eliminate representational redundancy while reducing the problem's dimensionality. Nevertheless, because there are so many networks that can be integrated

The author of the [13] study tackled the two primary bottlenecks of deep learning algorithms by choosing the optimum architecture with a regularisation method based on dropouts and the best parameter values through cross validation.Experimental results are used to produce modern performance, and they also offer some incredibly insightful information.the majority of the time in data with complex structures, such image and text, as they may benefit from the structure already there in the data.Unfortunately, a sizable amount of the structure information is also lost as a result of data conversions.

The writer suggested in [14].Sentiment analysis is an automated process that uses computation to separate positive and negative points of view in text.Sentiment analysis is widely used to obtain meaningful data for other types of analysis.DL methods like LSTM perform better for sentiment classification with 85% accuracy when there are more training data.In the future, we want to broaden our research to consider different embedding models across a variety of datasets.

In [15], the author describes a system for modelling contextually linked words that expands on the weighted word representation technique and incorporates linguistic constraints.Also, we train five well-known classifiers for the same task: -SVM, Decision Tree, -Random -Forests, -Naive Bayes, and -K-Nearest -Neighbor.There isn't currently a sentiment lexicon in use..

In [16] paper An important study direction that intends to develop medication recommendations in accordance with patients' electronic health records (EHRs) is suggested by the author in [16]. The majority of current approaches either simply prescribe changes to EHRs for the most recent admission, neglecting the patient's prior records, or they don't completely take into account the relationships between the clinical events from each admission. Clinical events in EHRs include intricate structural relationships and temporal interdependence., these techniques have demonstrated their limits, which leads in poor recommendation quality and a lack of temporal prediction capability

2.2 CHALLENGES PRESENT IN EXISTING SYSTEM

- Limited availability of data: Drug recommendation systems require a largee amount of dataa to trains effectively. However, data may be limited due to issues such as patient privacy and data protection laws. This makes it challenging to train accurate models.

- Complexity of the data: Drug recommendation systems require a comprehensive understanding of various factors such as patient history, medical conditions, allergies, and drug interactions. The complexity of this data makes it difficult to accurately represent and model in a system.
- User acceptance: Healthcare providers may be hesitant to adopt drug recommendation systems if they do not trust the recommendations or if they perceive them as being too complex or difficult to use.

- Website :they never created a user friendly website to intract to the patient

2.3 DISADVANTAGES OF THE EXISTING SYSTEM

- Accuracy level is not as good to use in real word
- No user friendly website
- No credential for patients to inquriry about the drug which doctor gave to them

Chapter-3

Requirements

3.1 HARDWARE REQUIREMENTS

- Processor: i5
- Ram:8GB
- Hard Disk: 1 TB

3.2 SOFTWARE REQUIREMENTS

I) Software used:

- Collab
- Vs code

II) Language used:

- Python
- Flask
- React j

Chapter-4

Analysis And Design

4.1 PROPOSED METHODOLOGY

- The goal of the proposed system is to create a drug recommendation system that is more accurate than the current system.
- To efficiently extract the features, BOW, TF-IDF, and Word2vec feature extraction algorithms are applied.
- LSTM and other algorithms will be altered to increase accuracy in comparison to the current system.
- The user can enter certain features to determine whether the drug is reliable and safe to use. This is part of the drug quality prediction process.
- This prediction aids in preventing the use of incorrect medications that were recommended due to a shortage of doctors or low servicing experience.

4.1.2 ADVANTAGES OF PROPOSED SYSTEM

- An affordable and practical drug suggestion solution
- Better precision.
- Helps save lives

4.2 SYSTEM ARCHITECTURE

The below Fig. 1.1 is a ARCHITECTURE of my project its shows the working process and flow of the project

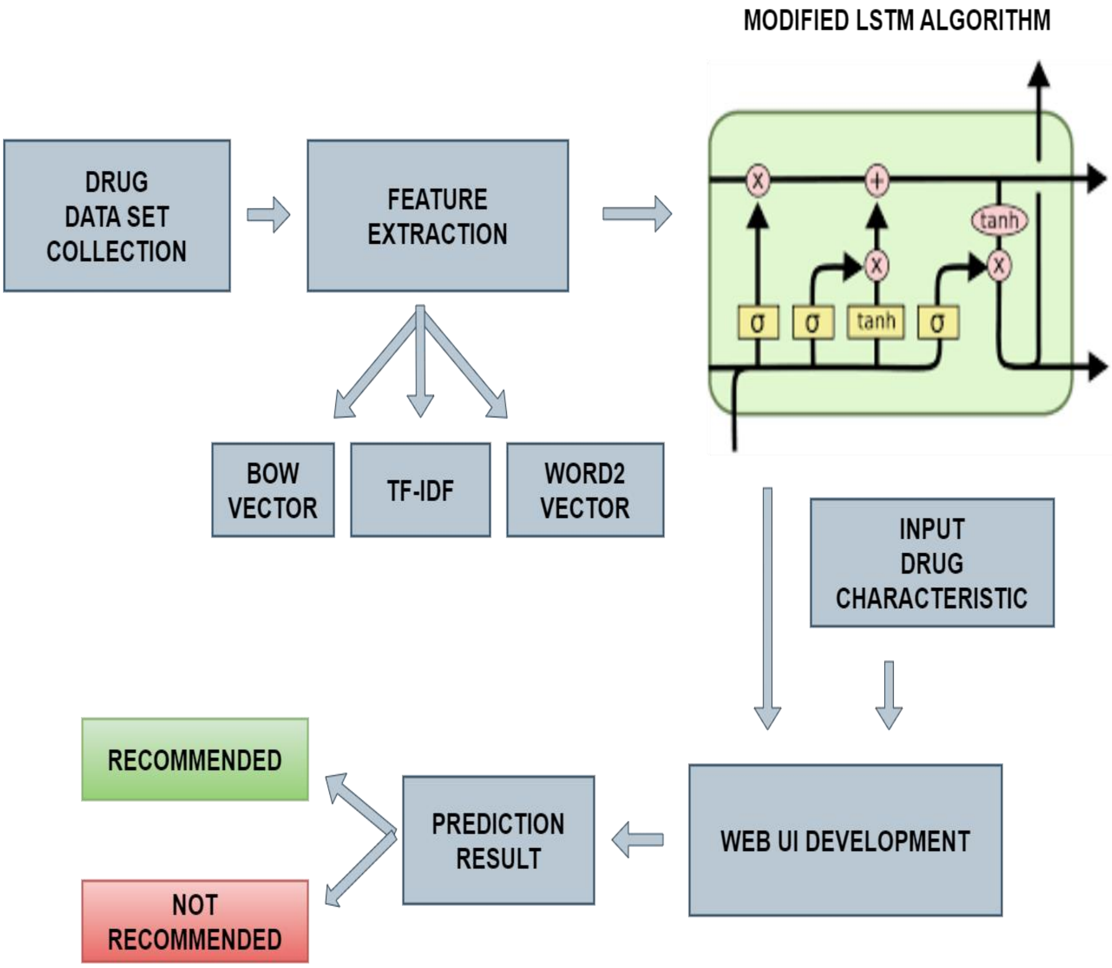


Fig. 1.1 Architecture Diagram Of The Project

4.3 MODULE DESCRIPTION:

DRUG DATASET COLLECTION

This dataset was used for the Winter 2018 Kaggle University Club Hackathon and is now publicly available. See Acknowledgments section for citation and licensing. Note: The types of data and recommendation based solutions provided by the contestants are purely for NLP learning purposes.

DATASET COLLECTION DFD

The fiq 2.1 given below is the screenshot of the dataset

1	uniqueID	drugName	condition	review	rating	date	usefulCount
2	163740	Mirtazapine	Depression	"I've tried a few antidepressants over the years (citalopram, fluoxetine, amitriptyline), but none of those helped with my depressic	10	28-Feb-12	22
3	206473	Mesalamine	Crohn's Disease,	"My son has Crohn's disease and has done very well on the Asacol. He has no complaints and shows no side effects. He has taken a	8	17-May-09	17
4	159672	Bactrim	Urinary Tract Inf	"Quick reduction of symptoms"	9	29-Sep-17	3
5	39293	Contrave	Weight Loss	"Contrave combines drugs that were used for alcohol, smoking, and opioid cessation. People lose weight on it because it also helps contr	9	05-Mar-17	35
6	97768	Cyclafem 1 /	Birth Control	"I have been on this birth control for one cycle. After reading some of the reviews on this type and similar birth controls I was a bit appreh	9	22-Oct-15	4
7	208087	Zyclara	Keratosis	"4 days in on first 2 weeks. Using on arms and face. Put vaseline on lips, under eyes and in nostrils to protect from cream. So far no reacti	4	03-Jul-14	13
8	215892	Copper	Birth Control	"I've had the copper coil for about 3 months now. I was really excited at the thought of not taking hormones. I'm good with pa	6	06-Jun-16	1
9	169852	Amitriptyline	Migraine Preven	"This has been great for me. I've been on it for 2 weeks and in the last week I only had 3 headaches which went away with 2 Tylenol	9	21-Apr-09	32
10	23295	Methadone	Opiate Withdrav	"Ive been on Methadone for over ten years and currently,I am trying to get off of this drug. Ive been decreasing my does 2 mgs per month	7	18-Oct-16	21
11	71428	Levora	Birth Control	"I was on this pill for almost two years. It does work as far as not getting pregnant however my experience at first was it didn't make	2	16-Apr-11	3
12	196802	Paroxetine	Hot Flashes	"Holy Hell is exactly how I feel. I had been taking Brisdelle for 1.5 years. The hot flashes did indeed subside - however, the side affects of	1	22-Feb-17	17
13	31947	Miconazole	Vaginal Yeast Inf	"Honestly its day one on the 3 day treatment. Yes it burns a bit and it does leak out if you dont lay down after insertion. But im faithful it w	6	07-May-15	7
14	4907	Belviq	Weight Loss	"This is a waste of money. Did not curb my appetite nor did it make me feel full."	1	23-Sep-14	57
15	66736	Seroquel	Schizoaffective I	"No problems, watch what you eat."	10	08-Oct-14	19
16	97013	Ambien	Insomnia	"Ditto on rebound sleepless when discontinued. I have done very strange things with no memory including taking additional Ambien. It h	2	13-Jan-15	44
17	213376	Nuvigil	Narcolepsy	"A doctor in the ER prescribed me 200 mg of Provigil when I was first diagnosed with Narcolepsy. It didn't seem to have any effect o	9	30-Jun-10	14
18	151674	Chantix	Smoking Cessati	"I smoked for 50+ years. Took it for one week and that was it. I didn't think it was possible for me to quit. It has been 6 years now.	10	14-Feb-15	26
19	33173	Microgestin F	Acne	"So I was on Ginanvi for about 3 months before I switched over to this pill due to the high cost of Ginavi (I don't have insurance). Gir	3	22-Jun-17	1
20	30401	Klonopin	Bipolar Disorde	"This medication helped me sleep, but eventually it became ineffective as a sleep aid. It also helps me calm down when in severe stress,	6	14-Jul-09	24
21	152490	Ciprofloxacin	Urinary Tract Inf	"After just 1 dose of this ciprofloxacn, I felt 99% better."	10	09-Jun-10	9
22	231397	Trazodone	Insomnia	"If I could give it a 0, I would absolutely do so. Started at 50mg, and felt WIRED. Wanted to get up and clean the house! Bumped it to 100r	1	18-Oct-16	15
23	38116	EnteraGam	Irritable Bowel S	"I am so happy with the samples provided by my Endocrinologist. The only thing I am so sad about is that I cannot afford the prohibitive cc	9	01-Jan-15	43

Fig. 2.1 Drug Dataset Collection

DATASET PRE-PROCESSING

- Data preparation, a crucial stage in deep learning, is converting raw data into a format appropriate for a neural network's analysis.

  Data preprocessing aims to enhance the effectiveness and quality of the input data so that the network can learn more effectively and generate more accurate results.
- Preparing the raw data to be used with a machine learning model is known as data pre-processing

FEATURES EXTRACTION

1. BOW VECTOR

The most straightforward method of translating text into numbers is the Bag of Words (BoW) paradigm. We can visualise a sentence as a bag of words vector, much like the term itself (a string of numbers).

Table I
Bow vector example

	1 this	2 moview	3 is	4 very	5 scary	6 and	7 Long	8 not	9 slow	10 spooky	11 Good	Length of the review (in word)
Review 1	1	1	1	1	1	1	1	0	0	0	0	7
Review 2	1	1	2	0	0	1	1	0	1	0	0	8
Review 3	1	1	1	0	0	0	1	0	0	1	1	6

2. TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) is a popular technique used in natural language processing and deep learning to represent the importance of a word in a document or a corpus of documents. TF-IDF is based on two factors:

Term Frequency (TF): The frequency of a word in a document is the number of times that word appears in the document. The term frequency of a word is the ratio of the number of times the word appears in a document to the total number of words in the document.

Inverse Document Frequency (IDF): The inverse document frequency of a word is a measure of how important that word is in the corpus of documents. It is calculated as the logarithm of the ratio of the total number of documents to the number of documents containing the word.

Table II

Tf-Idf Example

Word	TF		IDF	TF*IDF	
	A	B		A	B
The	1/7	1/7	log(2/2)=0	0	0
Car	1/7	0	log(2/1)=0.3	0.043	0
Truck	0	1/7	log(2/1)=0.3	0	0.043
Is	1/7	1/7	log(2/2)=0	0	0
Driven	1/7	1/7	log(2/2)=0	0	0
On	1/7	1/7	log(2/2)=0	0	0
The	1/7	1/7	log(2/2)=0	0	0
Road	1/7	0	log(2/1)=0.3	0.043	0
Highway	0	1/7	log(2/1)=0.3	0	0.043

3. WORD2VECTOR

Word2vec is a model or technique that creates word embedding for better word representation.

A significant number of accurate syntactic and semantic word associations are captured using this method of natural language processing.

It is a two-layer shallow neural network that, after being trained, can identify words that are synonymous and suggest new words for incomplete phrases..

1. I enjoy flying.
2. I like NLP.
3. I like deep learning.

The resulting counts matrix will then be:

$$X = \begin{array}{cc} & \begin{array}{cccccccc} I & like & enjoy & deep & learning & NLP & flying & . \end{array} \\ \begin{array}{c} I \\ like \\ enjoy \\ deep \\ learning \\ NLP \\ flying \\ . \end{array} & \begin{bmatrix} 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Fig. 3.3. Word2vec

LSTM ALGORITHM WORKING DIAGRAM

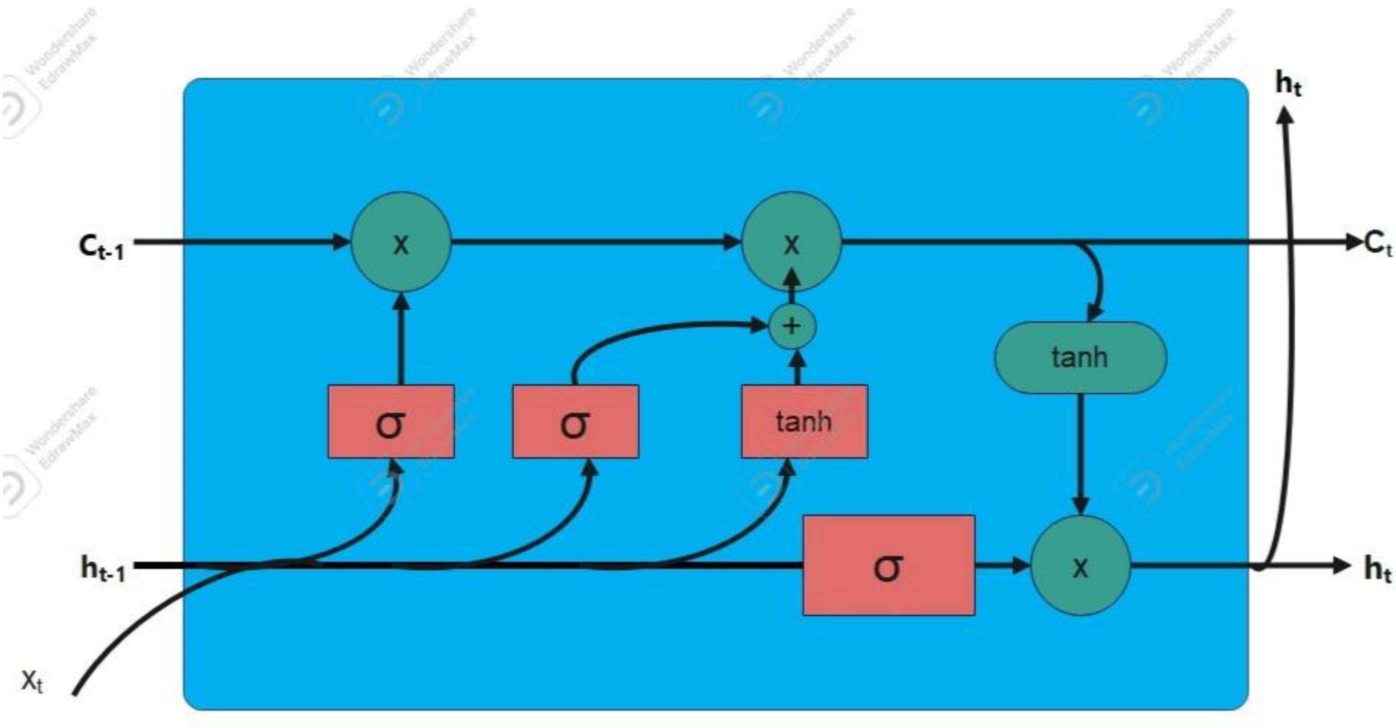


Fig. 4.1. Lstm diagram working process

TRAINING WITH ALGORITHMS

- The steps involved in raining with the LSTM algorithm are as follows:

- Data Preparation: Setting up the data for training is the initial step. The data must first be cleaned, prepped, and divided into training, validation, and test sets.

- Model Architecture: The LSTM model's architecture must be established as the next stage. The number of LSTM layers, the quantity of hidden units in each layer, and the activation functions to be employed must all be specified.

- Model Compilation: After specifying the model architecture, the model must be assembled. This entails defining the evaluation metric, the optimizer method, and the loss function that will be optimised.

- After the model has been created, the following step is to train it. the model employing training information. The model learns to modify its weights to reduce the loss function during training. Up until the model converges, the training process can be repeated over several epochs.

- Evaluation: The model's performance is assessed after training using validation and test data. This entails calculating numerous evaluation measures, including F1-score, recall, accuracy, and precision.

- A model can be fine-tuned if the performance is not up to par by changing the model architecture or hyperparameters like learning rate, batch size, and number of epochs.

- Deployment: The model can be used to make predictions on new data after it has been trained and assessed. Predictions on fresh input sequences are made using the learned model in this process.

Chapter 5

Implementation & Testing

5.1 DATA SET

1	uniqueID	drugName	condition	review	rating	date	usefulCount
2	163740	Mirtazapine	Depression	"I've tried a few antidepressants over the years (citalopram, fluoxetine, amitriptyline), but none of those helped with my depressic	10	28-Feb-12	22
3	206473	Mesalamine	Crohn's Disease,	"My son has Crohn's disease and has done very well on the Asacol. He has no complaints and shows no side effects. He has taken a	8	17-May-09	17
4	159672	Bactrim	Urinary Tract Inf	"Quick reduction of symptoms"	9	29-Sep-17	3
5	39293	Contrave	Weight Loss	"Contrave combines drugs that were used for alcohol, smoking, and opioid cessation. People lose weight on it because it also helps contr	9	05-Mar-17	35
6	97768	Cyclafem 1 /	Birth Control	"I have been on this birth control for one cycle. After reading some of the reviews on this type and similar birth controls I was a bit appreh	9	22-Oct-15	4
7	208087	Zyclara	Keratosis	"4 days in on first 2 weeks. Using on arms and face. Put vaseline on lips, under eyes and in nostrils to protect from cream. So far no reacti	4	03-Jul-14	13
8	215892	Copper	Birth Control	"I've had the copper coil for about 3 months now. I was really excited at the thought of not taking hormones. I'm good with pa	6	06-Jun-16	1
9	169852	Amitriptyline	Migraine Preven	"This has been great for me. I've been on it for 2 weeks and in the last week I only had 3 headaches which went away with 2 Tylenol	9	21-Apr-09	32
10	23295	Methadone	Opiate Withdrav	"Ive been on Methadone for over ten years and currently,I am trying to get off of this drug. Ive been decreasing my does 2 mgs per month	7	18-Oct-16	21
11	71428	Levora	Birth Control	"I was on this pill for almost two years. It does work as far as not getting pregnant however my experience at first was it didn't make	2	16-Apr-11	3
12	196802	Paroxetine	Hot Flashes	"Holy Hell is exactly how I feel. I had been taking Brisdelle for 1.5 years. The hot flashes did indeed subside - however, the side affects of	1	22-Feb-17	17
13	31947	Miconazole	Vaginal Yeast Inf	"Honestly its day one on the 3 day treatment. Yes it burns a bit and it does leak out if you dont lay down after insertion. But im faithful it w	6	07-May-15	7
14	4907	Belviq	Weight Loss	"This is a waste of money. Did not curb my appetite nor did it make me feel full."	1	23-Sep-14	57
15	66736	Seroquel	Schizoaffective I	"No problems, watch what you eat."	10	08-Oct-14	19
16	97013	Ambien	Insomnia	"Ditto on rebound sleepless when discontinued. I have done very strange things with no memory including taking additional Ambien. It h	2	13-Jan-15	44
17	213376	Nuvigil	Narcolepsy	"A doctor in the ER prescribed me 200 mg of Provigil when I was first diagnosed with Narcolepsy. It didn't seem to have any effect o	9	30-Jun-10	14
18	151674	Chantix	Smoking Cessati	"I smoked for 50+ years. Took it for one week and that was it. I didn't think it was possible for me to quit. It has been 6 years now.	10	14-Feb-15	26
19	33173	Microgestin F	Acne	"So I was on Ginanvi for about 3 months before I switched over to this pill due to the high cost of Ginavi (I don't have insurance). Gi	3	22-Jun-17	1
20	30401	Klonopin	Bipolar Disorde	"This medication helped me sleep, but eventually it became ineffective as a sleep aid. It also helps me calm down when in severe stress,	6	14-Jul-09	24
21	152490	Ciprofloxacin	Urinary Tract Inf	"After just 1 dose of this ciprofloxacn, I felt 99% better."	10	09-Jun-10	9
22	231397	Trazodone	Insomnia	"If I could give it a 0, I would absolutely do so. Started at 50mg, and felt WIRED. Wanted to get up and clean the house! Bumped it to 100r	1	18-Oct-16	15
23	38116	EnteraGam	Irritable Bowel S	"I am so happy with the samples provided by my Endocrinologist. The only thing I am so sad about is that I cannot afford the prohibitive cc	9	01-Jan-15	43

Fig. 5.1. Dataset Screenshot

5.2 SAMPLE CODE

```
from google.colab import drive
drive.mount('/content/drive')

!unzip /content/drive/MyDrive/Files.zip -d .
# Drug Sentiment Analysis

import all the necessary packages
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns
```

```
%matplotlib inline
from matplotlib import style
style.use('ggplot')
Load Data
read the train and test data
test = pd.read_csv(r"./dataset/drugsComTest_raw.csv") #test data
train = pd.read_csv(r"./dataset/drugsComTrain_raw.csv") #train data
#check the head of train data

train.head()

#check the head of test data

test.head()

#check the shape of the given dataset

print(f'train has {train.shape[0]} number of rows and {train.shape[1]} number of columns')
print(f'train has {test.shape[0]} number of rows and {test.shape[1]} number of columns')
merge = [train,test]
merged_data = pd.concat(merge,ignore_index=True)
merged_data.shape   #check the shape of merged_data
merged_data.info()

### Check number of unique values in drugName and condition

#check number of unique values in drugName

print('Total no. of Drugs:     ', merged_data['drugName'].nunique())

#check number of unique values in condition

print('Total no. of Condition: ', merged_data['condition'].nunique())

### Check the top 20 conditions

#plot a bargraph to check top 20 conditions

plt.figure(figsize=(12,6))
conditions = merged_data['condition'].value_counts(ascending = False).head(20)
plt.bar(conditions.index,conditions.values)
plt.title('Top-20 Conditions',fontsize = 20)
plt.xticks(rotation=90)
plt.ylabel('count')
plt.show()
```

Plot the bottom 20 conditions

#plot a bargraph to check bottom 20 conditions

```
plt.figure(figsize=(12,6))
conditions_bottom = merged_data['condition'].value_counts(ascending = False).tail(20)
plt.bar(conditions_bottom.index,conditions_bottom.values)
plt.title('Bottom-20 Conditions',fontsize = 20)
plt.xticks(rotation=90)
plt.ylabel('count')
plt.show()
```

Check top 20 drugName

#plot a bargraph to check top 20 drugName

```
plt.figure(figsize=(12,6))
drugName_top = merged_data['drugName'].value_counts(ascending = False).head(20)

plt.bar(drugName_top.index,drugName_top.values,color='blue')
plt.title('drugName Top-20',fontsize = 20)
plt.xticks(rotation=90)
plt.ylabel('count')
plt.show()
```

Check bottom 20 drugName

#plot a bargraph to check top 20 drugName

```
plt.figure(figsize=(12,6))
drugName_bottom = merged_data['drugName'].value_counts(ascending = False).tail(20)
plt.bar(drugName_bottom.index,drugName_bottom.values,color='blue')
plt.title('drugName Bottom-20',fontsize = 20)
plt.xticks(rotation=90)
plt.ylabel('count')
plt.show()
```

Checking Ratings Distribution

```
ratings_ = merged_data['rating'].value_counts().sort_values(ascending=False).reset_index().\
            rename(columns = {'index' :'rating', 'rating' : 'counts'})
ratings_['percent'] = 100 * (ratings_['counts']/merged_data.shape[0])
print(ratings_)
```

 # Setting the Parameter

```
sns.set(font_scale = 1.2, style = 'darkgrid')
plt.rcParams['figure.figsize'] = [12, 6]
```

```
#let's plot and check
sns.barplot(x = ratings_['rating'], y = ratings_['percent'],order = ratings_['rating'])
plt.title('Ratings Percent',fontsize=20)
plt.show()
```

Check number of Drugs per condition

```
#lets check the number of drugs/condition
merged_data.groupby('condition')['drugName'].nunique().sort_values(ascending=False
).head(20)

span_data =
merged_data[merged_data['condition'].str.contains('</span>',case=False,regex=True)
== True]
print('Number of rows with </span> values : ', len(span_data))
noisy_data_ = 100 * (len(span_data)/merged_data.shape[0])
print('Total percent of noisy data {} %  '.format(noisy_data_))
```

#drop the nosie

```
merged_data.drop(span_data.index, axis = 0, inplace=True)
```

Now let's look at the not listed/other

#check the percentage of 'not listed / othe' conditions

```
not_listed = merged_data[merged_data['condition'] == 'not listed / othe']
print('Number of not_listed values : ', len(not_listed))
percent_not_listed = 100 * len(not_listed)/merged_data.shape[0]
print('Total percent of noisy data {} %  '.format(percent_not_listed))

# drop noisy data
merged_data.drop(not_listed.index, axis = 0, inplace=True)

# after removing the noise, let's check the shape
merged_data.shape

### Now Check number of drugs present per condition after removing noise

#lets check the number of drugs present in our dataset condition wise
conditions_gp =
merged_data.groupby('condition')['drugName'].nunique().sort_values(ascending=False
)

#plot the top 20
# Setting the Parameter
condition_gp_top_20 = conditions_gp.head(20)
sns.set(font_scale = 1.2, style = 'darkgrid')
plt.rcParams['figure.figsize'] = [12, 6]
sns.barplot(x = condition_gp_top_20.index, y = condition_gp_top_20.values)
```

```
plt.title('Top-20 Number of drugs per condition',fontsize=20)
plt.xticks(rotation=90)
plt.ylabel('count',fontsize=10)
plt.show()
```

Check bottom 20 drugs per conditions

```
#bottom-20
condition_gp_bottom_20 = conditions_gp.tail(20)
#plot the top 20

sns.barplot(x = condition_gp_bottom_20.index, y = condition_gp_bottom_20.values,color='blue')
plt.title('Bottom-20 Number of drugs per condition',fontsize=20)
plt.xticks(rotation=90)
plt.ylabel('count',fontsize=10)
plt.show()
```

Now let's check if a single drug can be used for Multiple conditions

#let's check if a single drug is used for multiple conditions

```
drug_multiple_cond = merged_data.groupby('drugName')['condition'].nunique().sort_values(ascending=False)
print(drug_multiple_cond.head(10))
```

Check the number of drugs with rating 10

#Let's check the Number of drugs with rating 10.

```
merged_data[merged_data['rating'] == 10]['drugName'].nunique()
```

Plot top-20 drugs with rating 10

#Check top 20 drugs with rating=10/10

```
top_20_ratings = merged_data[merged_data['rating'] == 10]['drugName'].value_counts().head(20)
sns.barplot(x = top_20_ratings.index, y = top_20_ratings.values )
plt.xticks(rotation=90)
plt.title('Top-20 Drugs with Rating - 10/10', fontsize=20)
plt.ylabel('count')
plt.show()
```

Check for what condition Levonorgestrel is used for

```
merged_data[merged_data['drugName'] == 'Levonorgestrel']['condition'].unique()
```

Top 10 drugs with 1/10 Rating

#check top 20 drugs with 1/10 rating

```
top_20_ratings_1 = merged_data[merged_data['rating'] == 1]['drugName'].value_counts().head(20)
sns.barplot(x = top_20_ratings_1.index, y = top_20_ratings_1.values )
plt.xticks(rotation=90)
plt.title('Top-20 Drugs with Rating - 1/10', fontsize=20)
plt.ylabel('count')
plt.show()
```

Now we will look at the Date column

 # convert date to datetime and create year and month features

```
merged_data['date'] = pd.to_datetime(merged_data['date'])
merged_data['year'] = merged_data['date'].dt.year  #create year
merged_data['month'] = merged_data['date'].dt.month #create month
```

Check Number of reviews per year

#plot number of reviews year wise

```
count_reviews = merged_data['year'].value_counts().sort_index()
sns.barplot(count_reviews.index,count_reviews.values,color='blue')
plt.title('Number of reviews Year wise')
plt.show()
```

Check average rating per year

```
#check average rating per year
yearly_mean_rating = merged_data.groupby('year')['rating'].mean()
sns.barplot(yearly_mean_rating,yearly_mean_rating.values,color='green')
plt.title('Mean Rating Yearly')
plt.show()
```

Per year drug count and Condition count

#check year wise drug counts and year wise conditions counts

```
year_wise_condition = merged_data.groupby('year')['condition'].nunique()
sns.barplot(year_wise_condition.index,year_wise_condition.values,color='green')
plt.title('Conditions Year wise',fontsize=20)
plt.show()
```

****We expect that as the the conditions has increased. Drugs should have also increased. Let's check that out.****

#check drugs year wise

```
year_wise_drug = merged_data.groupby('year')['drugName'].nunique()
sns.barplot(year_wise_drug.index,year_wise_drug.values,color='green')
plt.title('Drugs Year Wise',fontsize=20)
plt.show()
```

Data Pre-Processing

 # check the null values

```
merged_data.isnull().sum()
```

 # drop the null values

```
merged_data.dropna(inplace=True, axis=0)
```

Pre-Processing Reviews

****Check the first few reviews****

#check first three reviews

```
for i in merged_data['review'][0:3]:
    print(i,'\n')

pip install nltk

import nltk
nltk.download('stopwords')

#import the libraries for pre-processing
from bs4 import BeautifulSoup
import nltk
import re
from nltk.corpus import stopwords
from nltk.stem.snowball import SnowballStemmer

stops = set(stopwords.words('english')) #english stopwords

stemmer = SnowballStemmer('english') #SnowballStemmer

def review_to_words(raw_review):
    # 1. Delete HTML
    review_text = BeautifulSoup(raw_review, 'html.parser').get_text()
    # 2. Make a space
    letters_only = re.sub('[^a-zA-Z]', ' ', review_text)
    # 3. lower letters
    words = letters_only.lower().split()
    # 5. Stopwords
    meaningful_words = [w for w in words if not w in stops]
    # 6. Stemming
```

```
    stemming_words = [stemmer.stem(w) for w in meaningful_words]
    # 7. space join words
    return( ' '.join(stemming_words))
```

#apply review_to_words function on reviews

```
merged_data['review'] = merged_data['review'].apply(review_to_words)
print(merged_data['review'] )
```

Now we will create our target variable "Sentiment" from rating

#create sentiment feature from ratings

#if rating > 5 sentiment = 1 (positive)

#if rating < 5 sentiment = 0 (negative)

```
merged_data['sentiment'] = merged_data["rating"].apply(lambda x: 1 if x > 5 else 0)
```

Label Encoding

```
from tensorflow.keras.utils import to_categorical
sentiment_label = to_categorical(merged_data.sentiment)
```

TF-IDF

#import all the necessary packages

```
from sklearn.feature_extraction.text import TfidfVectorizer
import tensorflow as tf
vectorizer = TfidfVectorizer(max_features=500)
reviews_corpus = vectorizer.fit_transform(merged_data.review).toarray()
features = vectorizer.get_feature_names_out()
reviews_corpus.shape
```

Split train and test data

```
from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test = train_test_split(reviews_corpus, sentiment_label, test_size=0.30, random_state=42)
```

LSTM

```
from keras.models import Sequential
from keras.layers import Activation,BatchNormalization, Bidirectional, Dense, Dropout, Embedding, Flatten, Conv1D, MaxPooling1D, GlobalMaxPool1D, LSTM
from tensorflow.keras.optimizers import Adam, SGD, RMSprop
from tensorflow.keras import utils as np_utils

model = Sequential()
```

```
model.add(LSTM(100,input_shape=(reviews_corpus.shape[1],1)))
model.add(Dense(2, activation='softmax'))
#model.compile(loss='categorical_crossentropy',
 #              optimizer=Adam(0.001),
  #             metrics=['accuracy'])
model.compile(optimizer= 'adam' , loss= 'categorical_crossentropy',
metrics=['accuracy'])
model.summary()
```

%%time

 # Checkpointer to save best model during training

```
import tensorflow as tf
from keras.callbacks import ModelCheckpoint, ReduceLROnPlateau
checkpoint_filepath=(r"/content/drive/MyDrive/drug_weight.h5")
checkpointer = ModelCheckpoint(checkpoint_filepath, verbose=1,
save_best_only=True, monitor='val_accuracy')
# reduce_lr = ReduceLROnPlateau(monitor='val_accuracy', factor=0.2,
#                         patience=5, min_lr=0.001)

#model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
# model.compile(loss='categorical_crossentropy',
#               optimizer=Adam(0.001),
#               metrics=['accuracy'])
history = model.fit(X_train,Y_train,
              batch_size=200,
              epochs=1,
              validation_data=(X_test,Y_test),
              verbose=1,
              callbacks=[checkpointer])


from sklearn.metrics import classification_report
preds = model.predict(X_test)
labels = ['Negative', 'Positive']
print(classification_report(Y_test.argmax(axis=1), preds.argmax(axis=1),
target_names=labels))

from sklearn.metrics import confusion_matrix
import seaborn as sns
CM = confusion_matrix(preds.argmax(axis=1), Y_test.argmax(axis=1))
labels=['Positive','Negative']
# drawing confusion matrix
sns.heatmap(CM, center = True , annot=True, fmt="d" ,cmap="Blues",
xticklabels=labels, yticklabels=labels)
plt.show()

!pip install pyngrok
!pip install flask-ngrok
```

```
!pip install flask-cors==3.0.7
!ngrok authtoken 2GiH8URoZbuvn0eIqW4C8br8caN_73m7xUQYdwzsqxo9Jo6EX
```

Flask code to connect backend to frontend

```
from flask_ngrok import run_with_ngrok
from flask import Flask
from flask import Flask, app, request
import json
from keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from flask_cors import CORS, cross_origin
from keras.models import Model, load_model
import warnings
warnings.filterwarnings('ignore')

app = Flask(__name__)
cors = CORS(app)
run_with_ngrok(app)   #starts ngrok when the app is run
@app.route('/drug', methods=['GET','POST'])
# @cross_origin()
def login():
   result = input(request.json['uri'])
   return result
def input(uri):

   perprocess_data = review_to_words(uri)

   model = load_model('/content/drive/MyDrive/drug/word_weight.h5')

   tokenizer = Tokenizer()
   tokenizer.fit_on_texts(perprocess_data)
   text_to_val = pad_sequences(tokenizer.texts_to_sequences(perprocess_data), 
maxlen=100)
   prediction = model.predict(text_to_val)[0]
   labels = ['Not Recommend', 'Recommend']
   predicted_label = labels[int(prediction.argmax())]

   return ({"data":predicted_label})

if __name__ == '__main__':
   app.run()
```

5.3 SAMPLE OUTPUT

PYTHON DATA VISUALIZATION

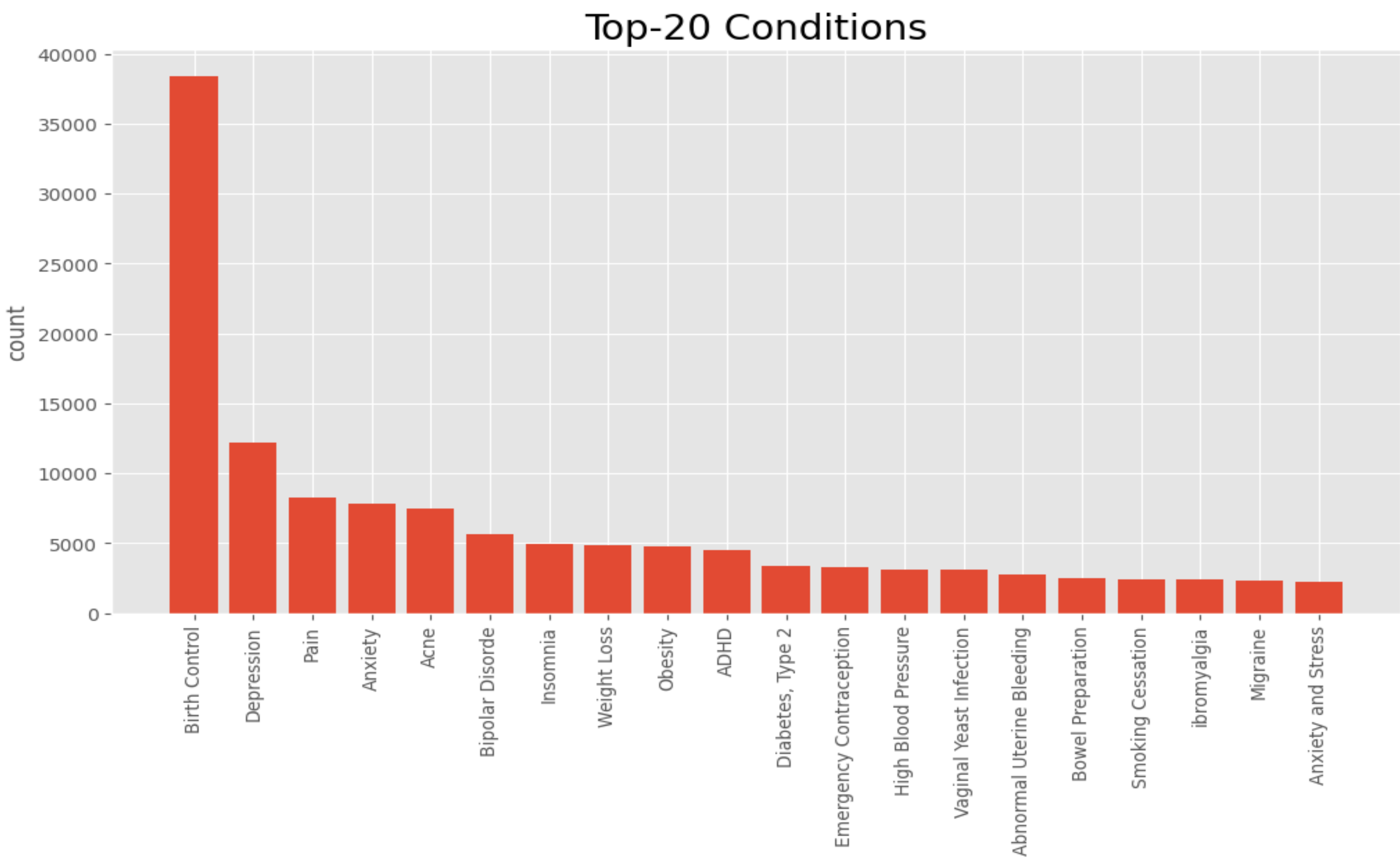


Fig. 6.1. Plot A Bar Graph To Check Top 20 Conditions

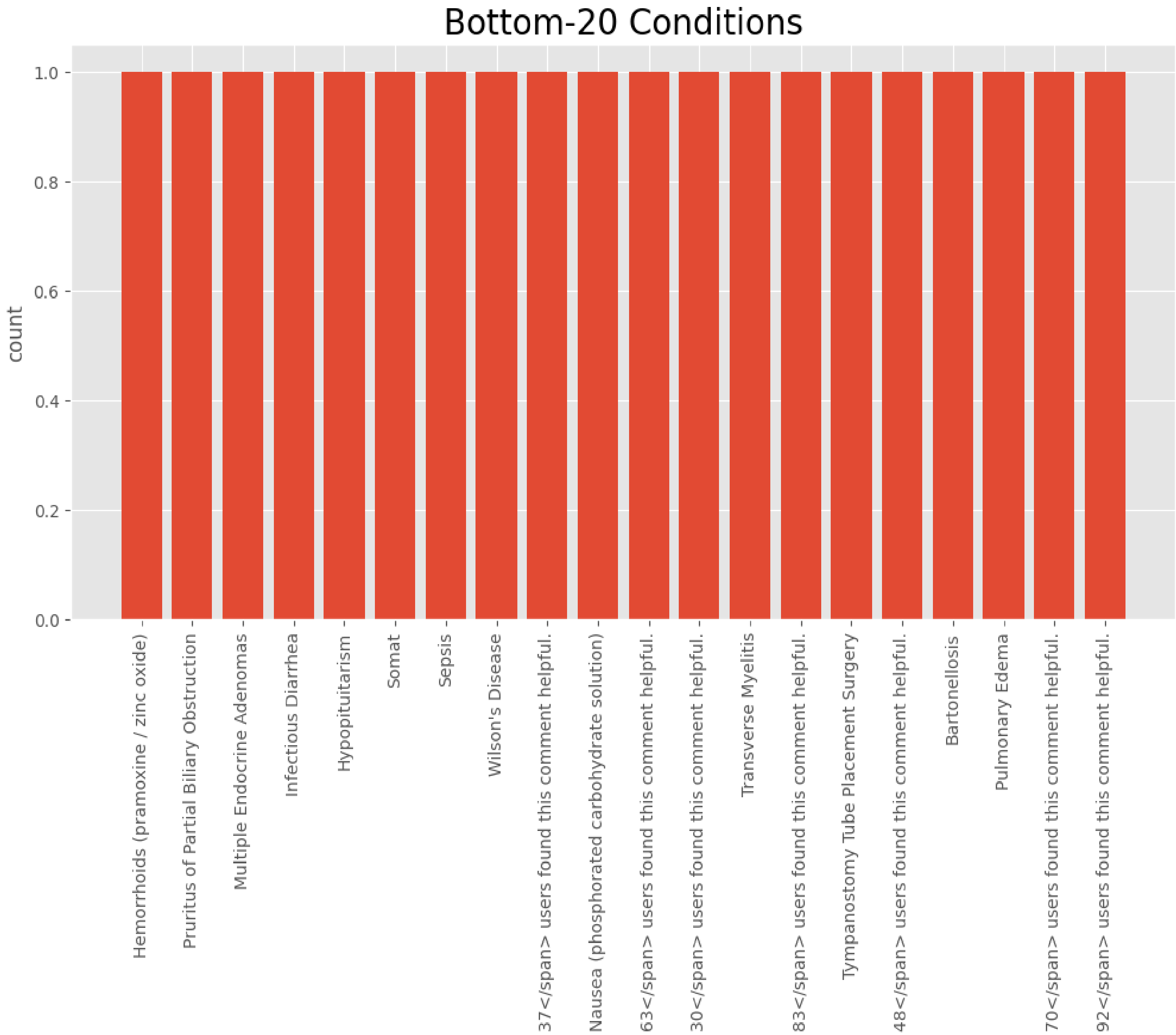


Fig. 6.2. Plot A Bar Graph To Check Bottom 20 Conditions

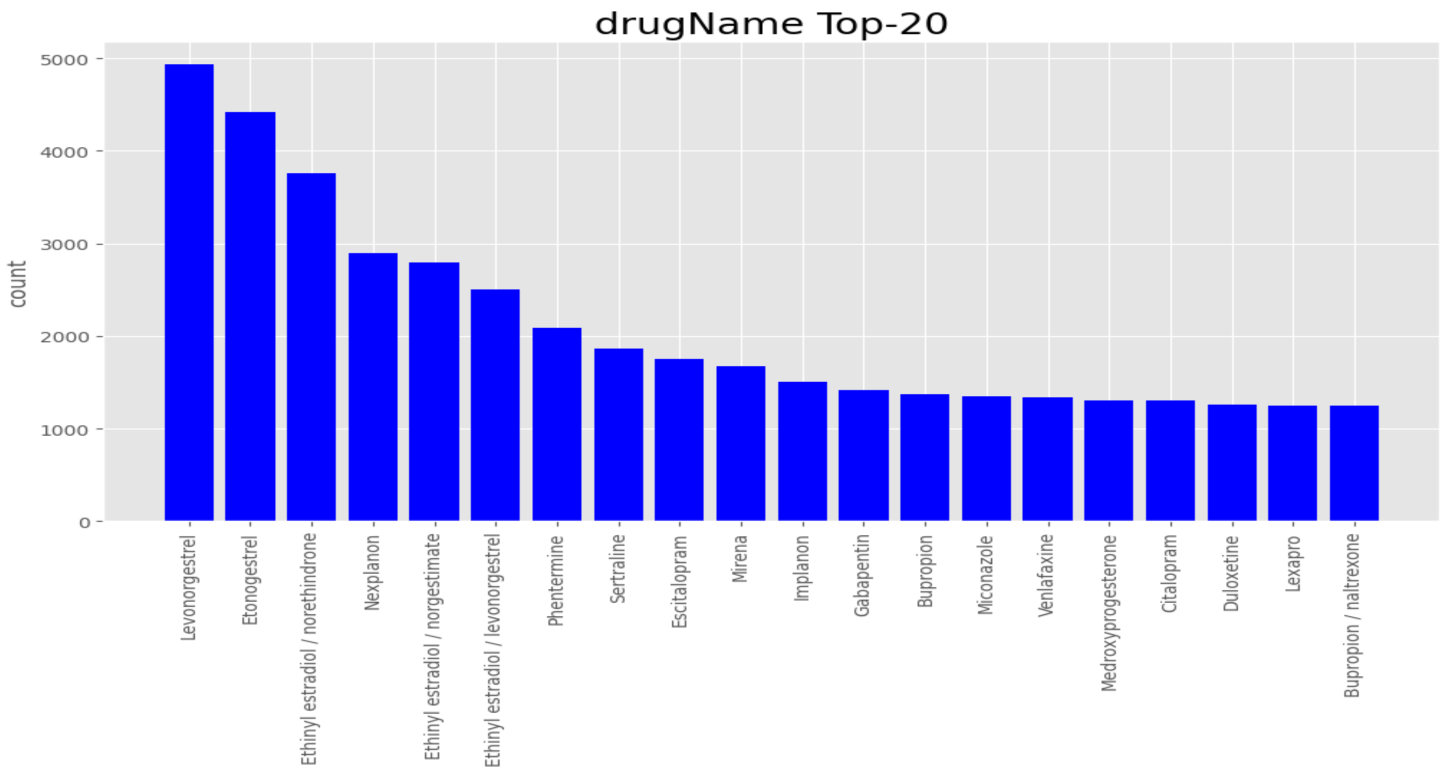


Fig. 6.3. Plot A Bar Graph To Check Top 20 Drugname

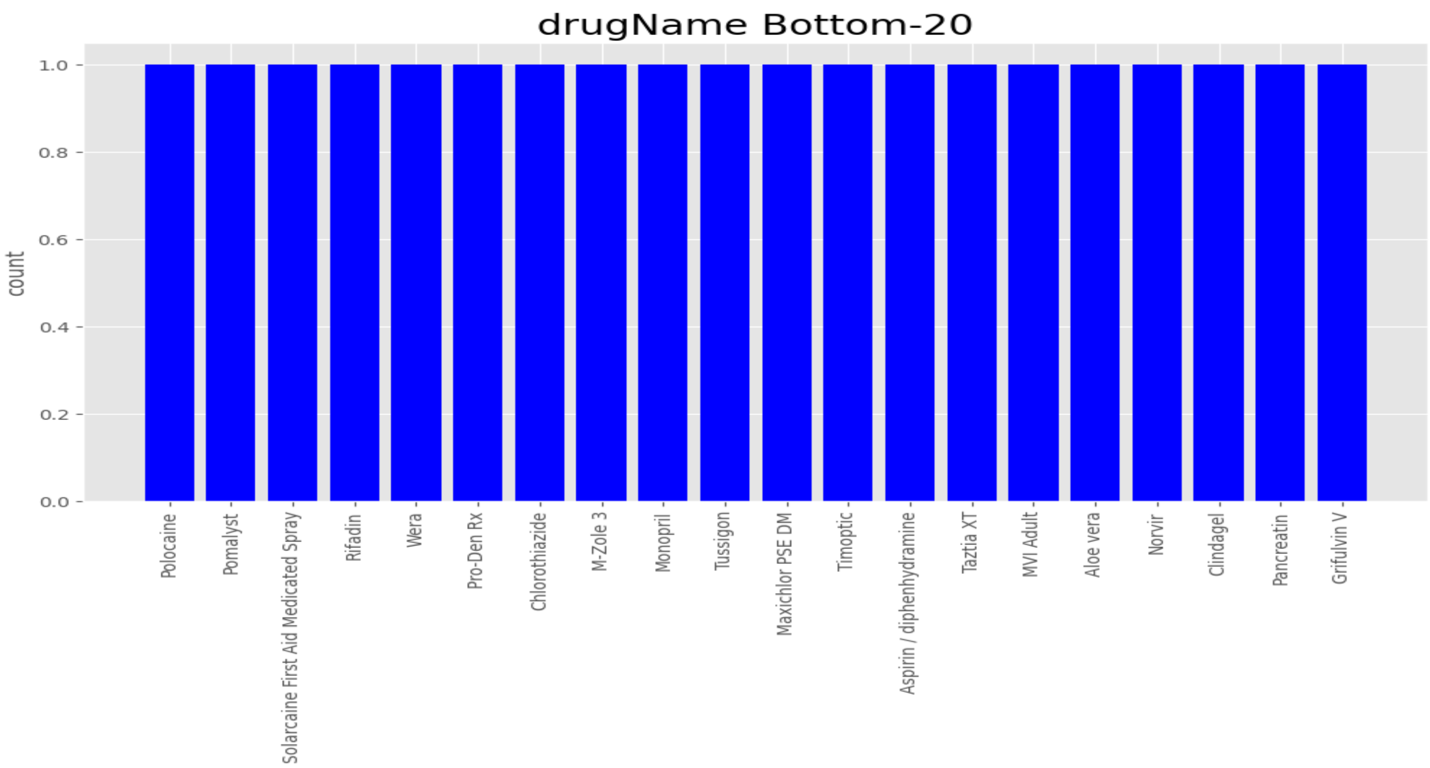


Fig. 6.4. Plot A Bar Graph To Check Botom 20 Drugname

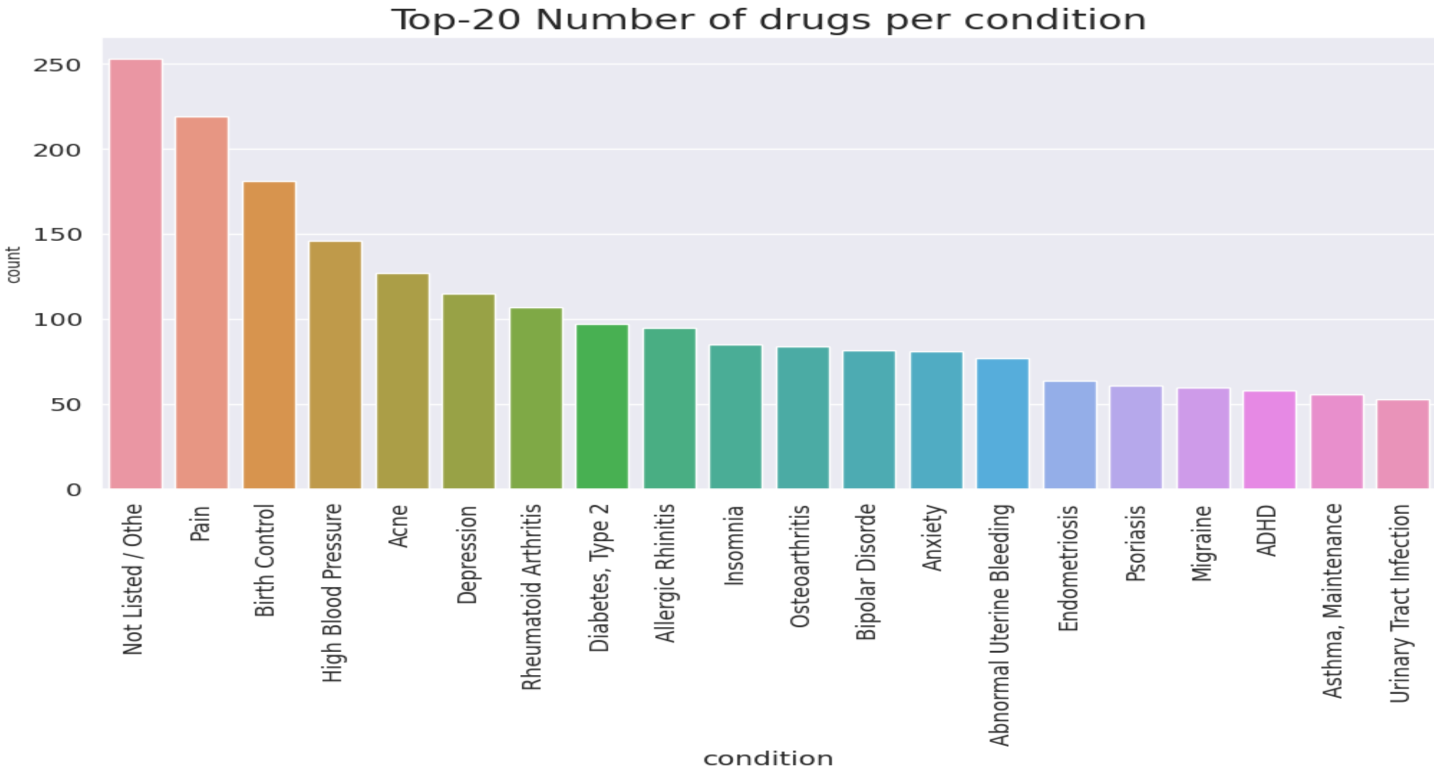


Fig. 6.5.Check The Number Of Drugs Present In Our Dataset Top Condition Wise

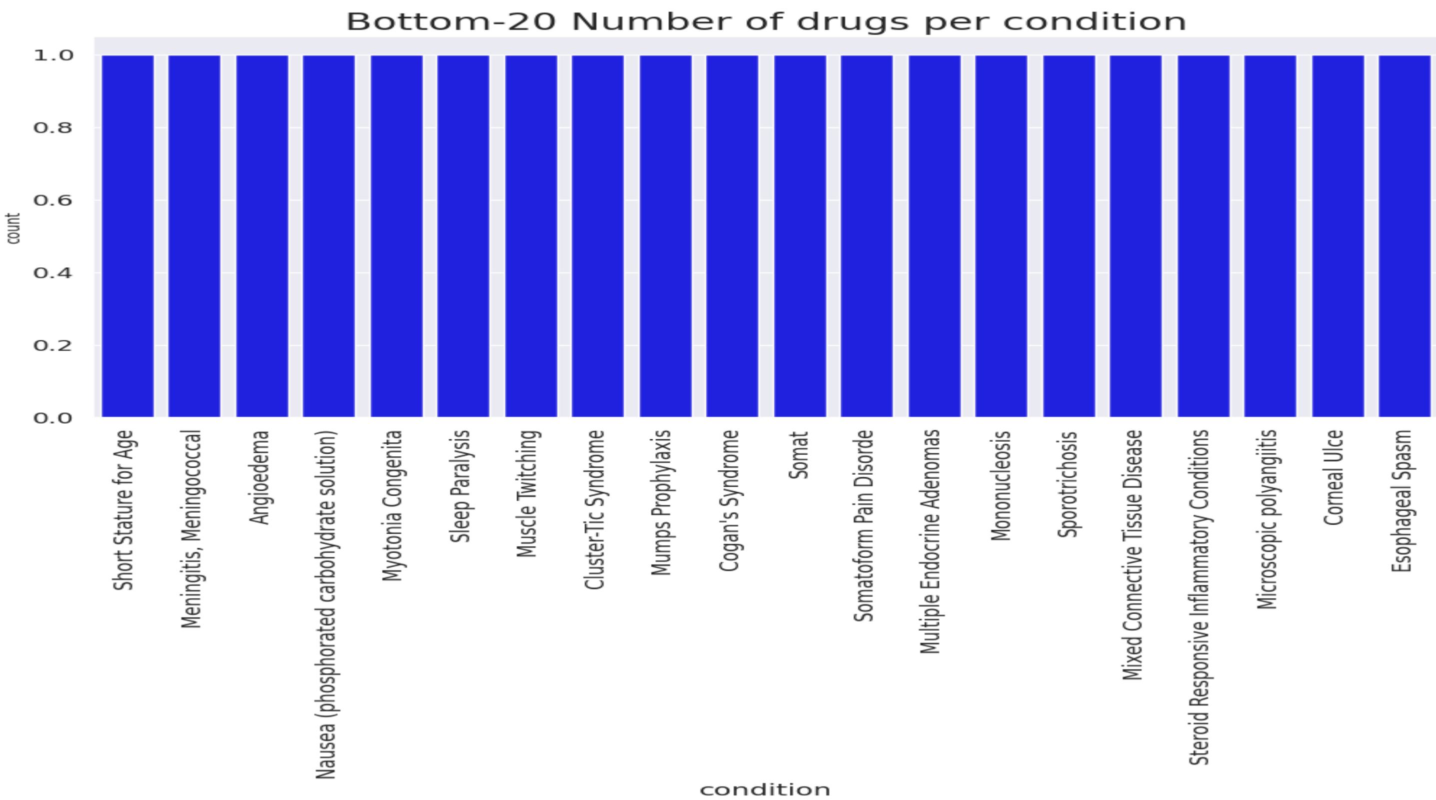


Fig. 6.6. The Number Of Drugs Present In Our Dataset Bottom Condition Wise

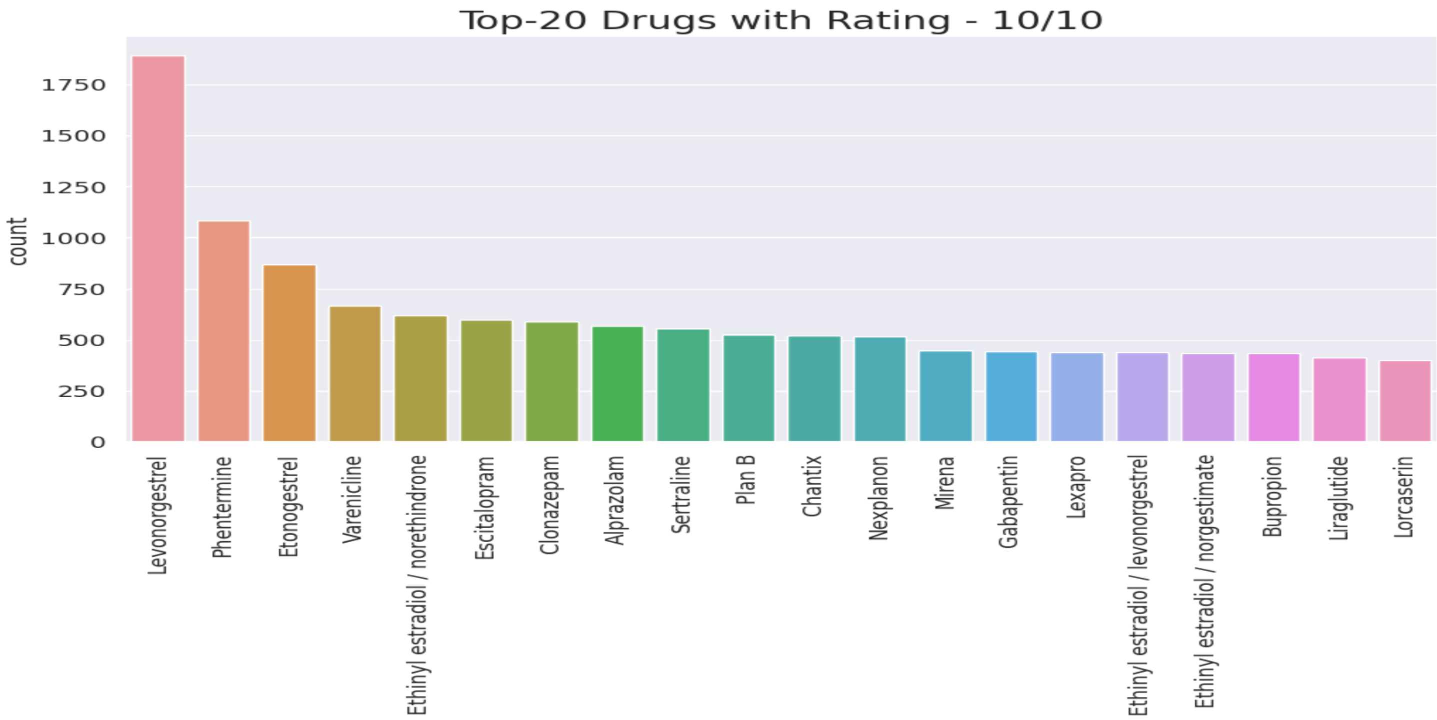


Fig. 6.7. Top 20 Drugs With Rating=10/10

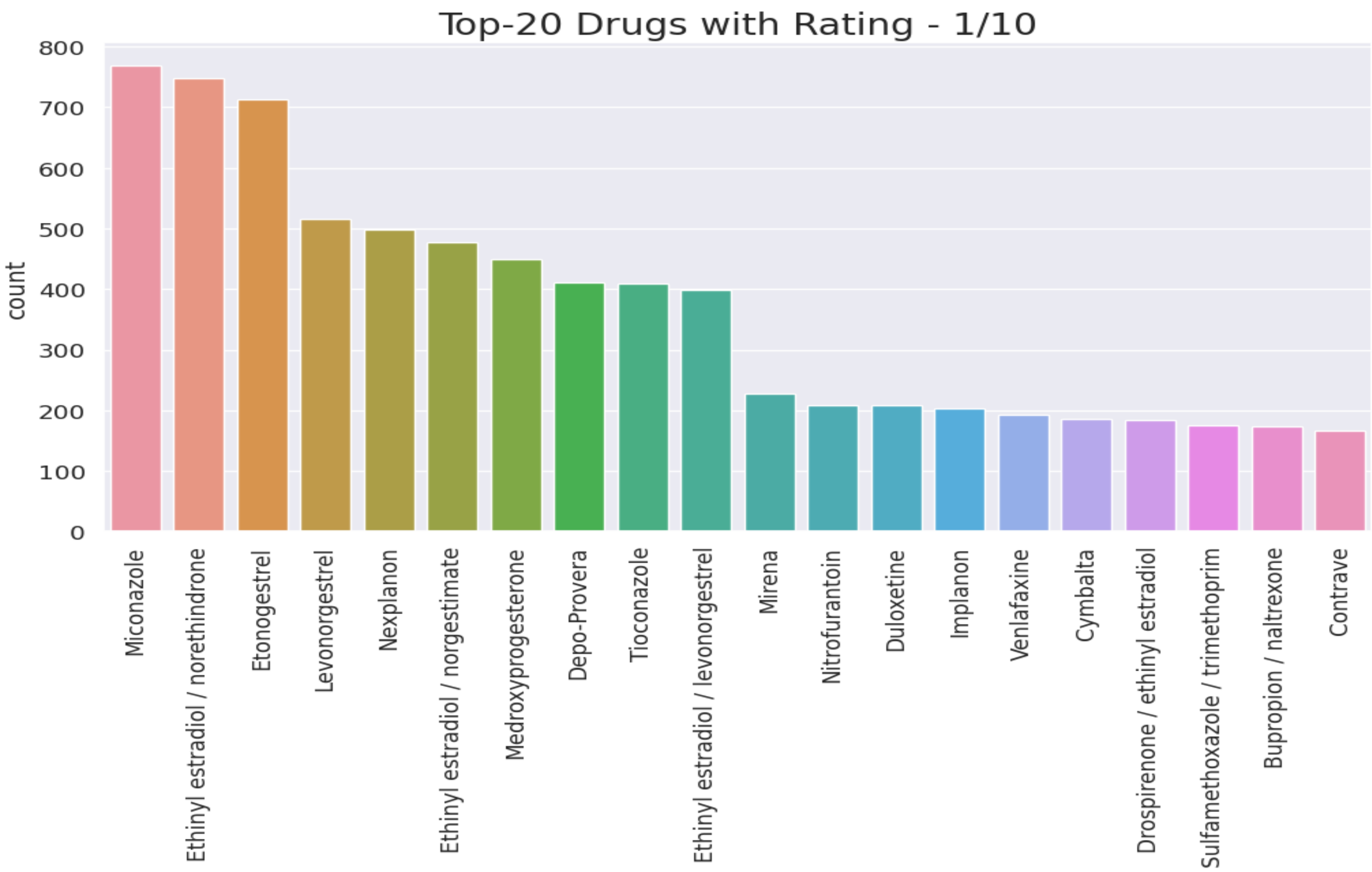


Fig. 6.8. Top 20 Drugs With 1/10 Rating

WEB APPLICATION LOGIN PAGE

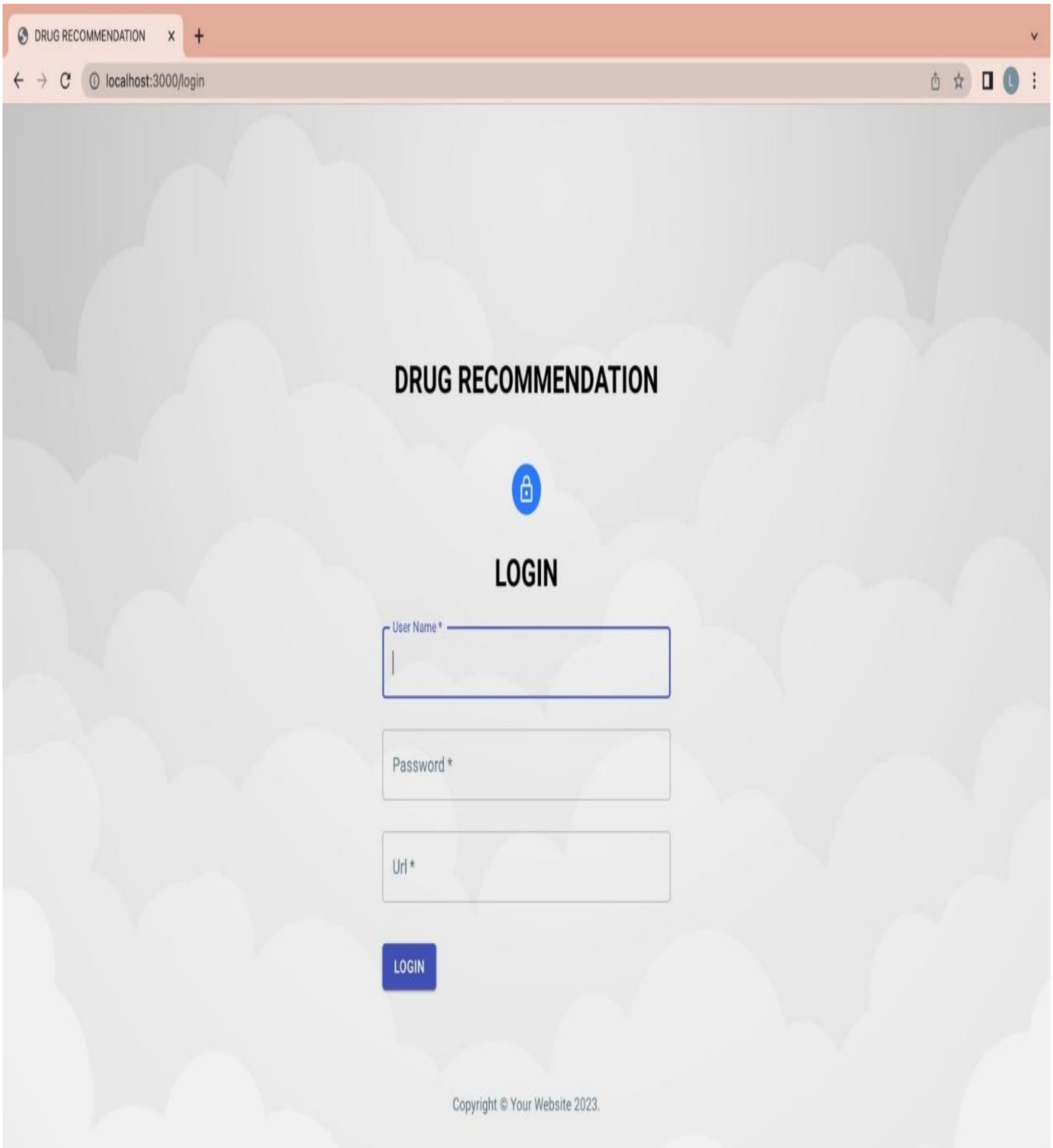


Fig. 6.9. Overview Of Website

PREDICTION OF DRUG RECOMMENDATION OR NOT

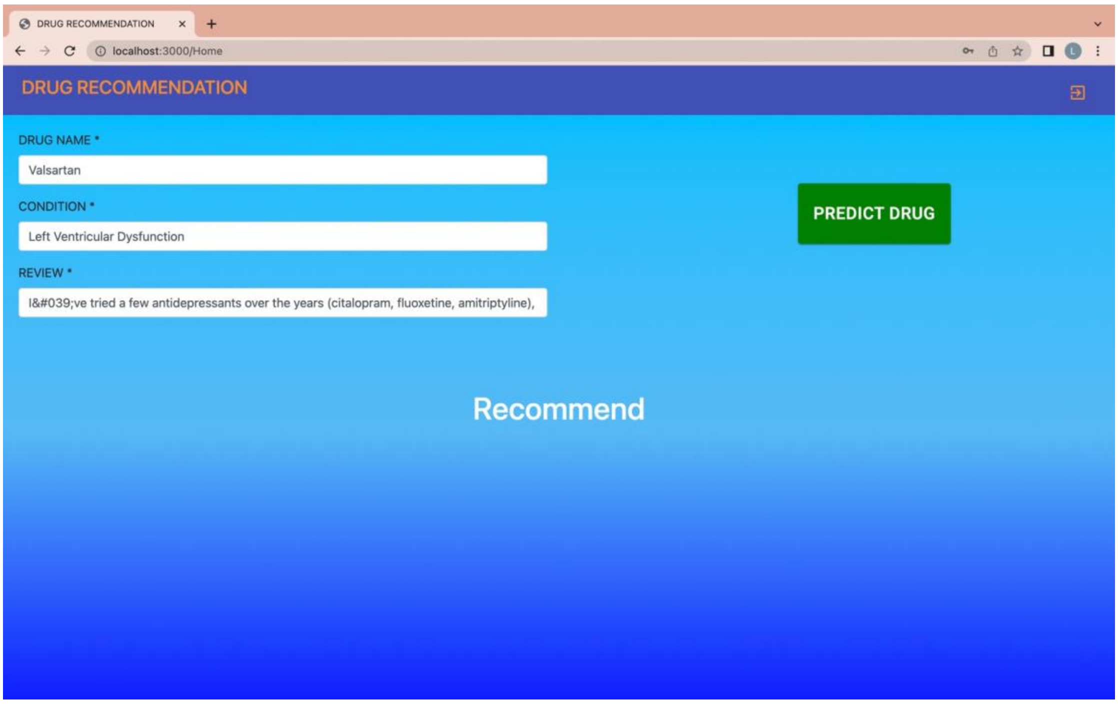


Fig. 6.10. Final Output After The Prediction-1

ABOUT THE WEBSITE

Fig. 6.9 and Fig. 6.10. are the screenshot of the final design of the website

5.4 TEST PLAN & DATA VERIFICATION

Software testing is the process of examining a system to look for flaws, gaps, or requirements that aren't there compared to what is actually needed. With this project, we tested each and every module using every input scenario imaginable. Due to the restricted amount of modules in this project, we utilised the manual testing concept..

TEST CASES (Unit Testing)

Table III
Test cases

Test case id	**Input 1 Drug name**	**Input 2 Condition**	**Input 3 Review**	**Recommended / Not Recommended**
1.	Mirtazapine	Depression	"I've tried a few antidepressants over the years (citalopram, fluoxetine, amitriptyline), but none of those helped with my depression, insomnia & anxiety. My doctor suggested and changed me onto 45mg mirtazapine and this medicine has saved my life. Thankfully I have had no side effects especially the most common - weight gain, I've actually lost alot of weight. I still have suicidal thoughts but mirtazapine has saved me."	Recommended
2.	Mesalamine	Crohn's Disease, Maintenance	"My son has Crohn's disease and has done very well on the Asacol. He has no complaints and shows no side effects. He has taken as many as nine tablets per day at one time. I've been very happy with the results, reducing his bouts of diarrhea drastically."	Recommended
3.	Bactrim	Urinary Tract Infection	"Quick reduction of symptoms"	Recommended
4.	Contrave	Weight Loss	"Contrave combines drugs that were used for alcohol, smoking, and opioid cessation. People lose weight on it because it also helps control over-eating. I have no doubt	Not-Recommended

			that most obesity is caused from sugar/carb addiction, which is just as powerful as any drug. I have been taking it for five days, and the good news is, it seems to go to work immediately. I feel hungry before I want food now. I really don't care to eat; it's just to fill my stomach. Since I have only been on it a few days, I don't know if I've lost weight (I don't have a scale), but my clothes do feel a little looser, so maybe a pound or two. I'm hoping that after a few months on this medication, I will develop healthier habits that I can continue without the aid of Contrave."	
5.	Cyclafem 1 / 35	Birth Control	"I have been on this birth control for one cycle. After reading some of the reviews on this type and similar birth controls I was a bit apprehensive to start. Im giving this birth control a 9 out of 10 as I have not been on it long enough for a 10. So far I love this birth control! My side effects have been so minimal its like Im not even on birth control! I have experienced mild headaches here and there and some nausea but other than that ive been feeling great! I got my period on cue on the third day of the inactive pills and I had no idea it was coming because I had zero pms! My period was very light and I barely had any cramping! I had unprotected sex the first month and obviously didn't get pregnant so I'm very pleased! Highly recommend"	Not-Recommended
6.	Zyclara	Keratosis	"4 days in on first 2 weeks. Using on arms and face. Put vaseline on lips, under eyes and in nostrils to protect from	Not-Recommended

			cream. So far no reaction at all. I know I have many pre cancer and thought I would light up like a Christmas tree but so far so good. Maybe it's coming but time will tell."	
7.	Copper	Birth Control	"I've had the copper coil for about 3 months now. I was really excited at the thought of not taking hormones. I'm good with pain however I nearly fainted with insertion, couldn't belive how painful it was; the doctor did say it is very painful for some. Well 3 months in, my periods last 11 days and I'm in pain for about 15 days with random twangs especially in the left side and I'm considering whether I want to put up with the intense pain and heavy periods. I'd recommend this 100% to somebody who doesn't already have heavy painful periods but right now it just isn't for me"	Not-Recommended
8.	Amitriptyline	Migraine Prevention	"This has been great for me. I've been on it for 2 weeks and in the last week I only had 3 headaches which went away with 2 Tylenol. I was having chronic daily headaches that wouldn't go away no matter what I took. I'm still a little sleepy during the day, but I know that will get better. I take 10mg at night."	Recommended
9.	Methadone	Opiate Withdrawal	"Ive been on Methadone for over ten years and currently,I am trying to get off of this drug. Ive been decreasing my does 2 mgs per month for over a year. I am at 3 mgs and really starting to feel the withdraw.I don't plan to get my next 30 doses.because its almost rediculous how little it does for me. I have 3 does doses of 3 mg and Im terrified.	Recommended

			Can anyone give me some truthful encouragement?....."	
10.	Levora	Birth Control	"I was on this pill for almost two years. It does work as far as not getting pregnant however my experience at first was it didn't make a huge difference then 6 or 7 months into it my sex drive went down, along with being very very dry, my moodiness increased drastically. I would cry one second and then get angry with my husband over anything and everything. My skin has gotten a lot worse, I broke out in places I never had in the last week. So now I am on Yaz."	Recommended
11.	Paroxetine	Hot Flashes	"Holy Hell is exactly how I feel. I had been taking Brisdelle for 1.5 years. The hot flashes did indeed subside - however, the side affects of this medicine coupled with the fact Noven was acquired by YET another pharmaceutical company - YOU CAN'T PLACE A REP IN THE AREA, DISTRIBUTE YOUR DRUGS, AND THEN FIRE HER-AND NOT REPLACE THEREFORE there is NO medicine or support here. You dumped this drug in the Dr's hands and walked away. After calling Sebula - you act like you don't even care. You have made it impossible to obtain this. I happen to think this is illegal. I just decided to wean myself off this and Premarin. It has been nothing short of a nightmare. If you don't need this drug- DON'T START. Seriously."	Recommended
12.	Miconazole	Vaginal Yeast Infection	"Honestly its day one on the 3 day treatment. Yes it burns a bit and it does leak out if you dont lay down after insertion. But im faithful it will work."	Not-Recommended

13.	Belviq	Weight Loss	"This is a waste of money. Did not curb my appetite nor did it make me feel full."	Recommended
14.	Seroquel	Schizoaffective Disorde	"No problems, watch what you eat."	Recommended
15.	Levonorgestrel	Birth Control	"I went in to have my Skyla placed yesterday morning. After reading all of these reviews I was hyperventilating and crying on my way there, bc I did not want to experience "the worst pain of my life". However, it was a complete waste of tears and energy. Before the procedure, I started doing heavy breathing, similar to pregnancy breaths, bc I was so nervous and I can tell you I felt NOTHING. There was a little bit of pressure, but nothing compared to what I've read on here. Probably a 2/10. It was over in 30 seconds, and as soon as she was done, the pressure vanished. I was shocked. I went to work from 5:45-11:30, & I managed with mild cramping.I cant speak for long term yet, but don't let other women scare you out of this, its very worth it!"	Not-Recommended
16.	Clonidine	ADHD	"My 5 year old son was diagnosed with ADHD just yesterday, the Behavior Specialist said his was one of the worst cases that she had seen in a while, she had suggested putting him on a stimulant medication, I told her i would like to a non-stimulant medication first and she prescribed him Kapvay. My son took it for the first time last night before bed, he went right to sleep and when he woke up this morning he was the calmest most pleasant, helpful and nicest he had ever been in his life. I could not believe the overnight change. I'm so glad it worked so fast, he has not gotten in	Recommended

			trouble once today which is a new record! His teachers are going to be thrilled on Monday! Thank you to the makers of Clonidine!"	
17.	Ethinyl estradiol / norgestimate	Birth Control	"I'd never been on birth control up until a few months ago, and I was given Ortho Tri-Cyclen Lo as a starter, because I did not want it to have much hormones. I went through two whole packs and decided to switch because of extremely low sex drive, extreme mood swings and increased appetite (I would eat a full meal and two hours later my stomach would be growling again). My relationship began to suffer due to these side effects so I stopped taking it and am back to normal. However, this pill DID have pros: I didn't get pregnant, it regulated my period and caused no breakouts or acne. But be wary of weight gain, decreased libido and mood swings."	Recommended
18.	Nicoderm CQ	Smoking Cessation	"I will say this about the patch. It work for me. We're there side effects? yes. A small bit of bruising and a small rash.I was willing to put up with those side effects because I did not have one craving for a cigarette. I was very happy with the program and have remained smoke free for three years boxing day passed."	Not-Recommended
19.	Levonorgestrel	Emergency Contraception	"on March 21-25 I had my period. On March 26 I had an accident. My cycle ranges from 28-33 days. I took plan B within 1 hour that same day March 26. That day I got cramps and got really tired. The next day I experienced diahrrea. On March 31st I started bleeding again which scared me and it lasted 4 days. The bleeding was really light.	Recommended

			I spent all April worrying. I took over 10 home pregnancy test and they were all negative. Well today I started bleeding again which makes me 3 days late from my actual period. Yet on time from the plan B period. I never want to go through this ever again."	
20.	Celecoxib	Osteoarthritis	"Celebrex did nothing for my pain."	Recommended
21.	Fluoxetine	Major Depressive Disorde	"I have Major Depressive Disorder, Bipolar Disorder, and anxiety, I have never felt more calm and in-control of my emotions. Usually every night I'd fall into a fit of thinking of self-harm, I have thought of them no more, I still get sad, but it's sadness, not full on depression. My anxiety has decreased so much, I still get nervous, but it is not even close to how bad it was. An added bonus, is for me, it makes me drowsy, so I get more sleep. I have never been so happy in my life. Prozac is a life saver."	Recommended
22.	Depakote	Bipolar Disorde	"General tiredness with the medication but no manic episodes. I guess the trade-off is fair."	Recommended
23.	Riboflavin	Migraine Prevention	"I take 400 mg a day and it helps."	Recommended

Chapter-6

Results

6.1 RESEARCH FINDINGS

Table IV

Tf-Idf

	Precision	Recall	F1-Score	Support
Not Recommend	0.65	0.01	0.03	15884
Recommend	0.70	1.00	0.82	37291
Accuracy			0.70	53175
Macro Avg	0.68	0.51	0.43	53175
Weighted Avg	0.69	0.70	0.59	5317

Table V
Bow Vector

	Precision	Recall	F1-Score	Support
Not Recommend	0.52	0.12	0.20	15884
Recommend	0.72	0.95	0.82	37291
Accuracy			0.70	53175
Macro Avg	0.62	0.54	0.51	53175
Weighted Avg	0.66	0.70	0.63	53175

Table VI
Word2Vec

	Precision	Recall	F1-Score	Support
Not Recommend	0.80	0.69	0.74	19154
Recommend	0.87	0.93	0.90	44656
Accuracy			0.85	63810
Macro Avg	0.84	0.81	0.82	63810
Weighted Avg	0.85	0.85	0.85	63810

ABOUT THE TABLE MENTION ABOVE:

After Training The Data Set With Different Feature Extraction The Final Report Are Given In Table IV, Table V, Table VI We Can See The Accuracy Level We Got Among The Three Feature .Word2vec Has The Highest Accuracy Outcome Comparing Other Extraction Thus We Will Use Word2vec In The Backend Work To Predict The Outcome Of The Project

6.2 CONFUSION_MATRIX:

The Given below fiq7.1, fiq7.2, fiq7.3 are the confusion matrix of the three extraction features we got after training is demonstrated in the below fiqures

TF –IDF

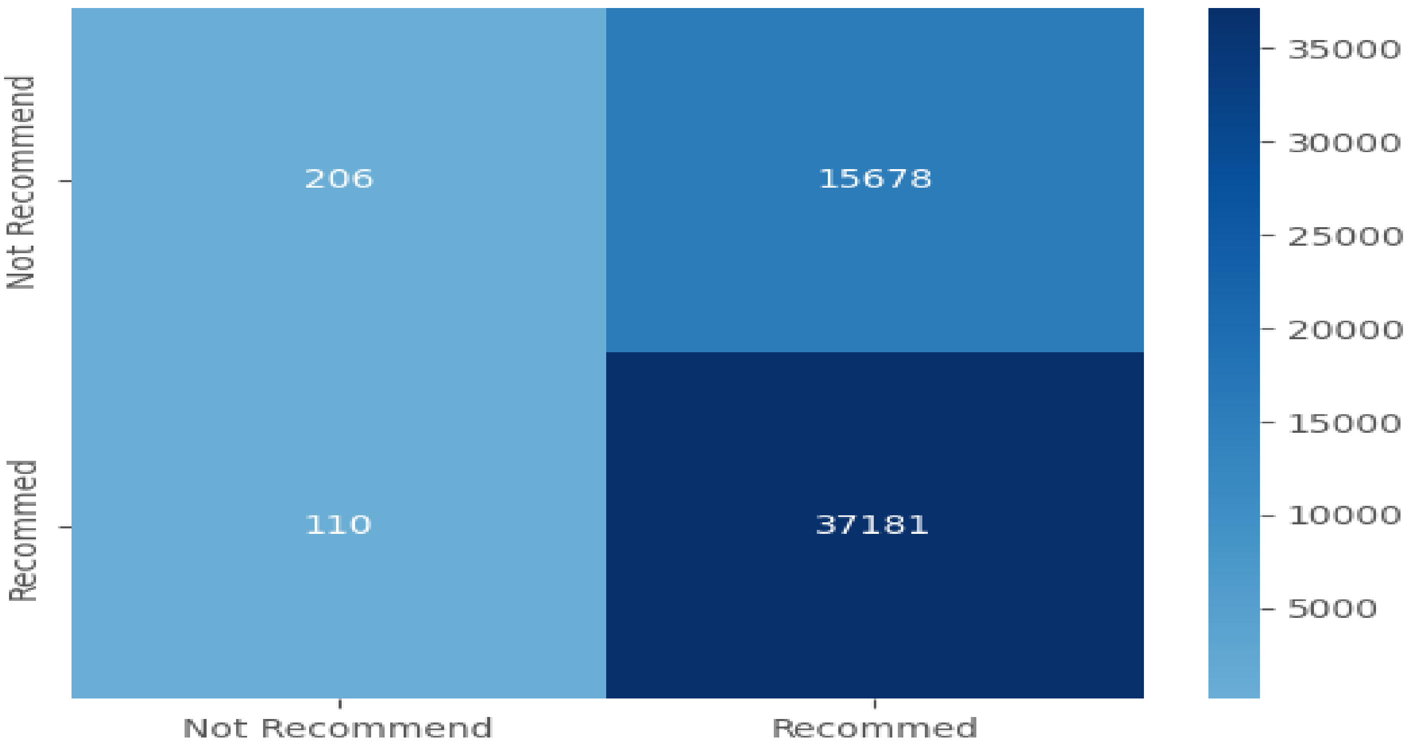


Fig. 7.1. Confusion Matrix Of Tf-Idf

BOW VECTOR:

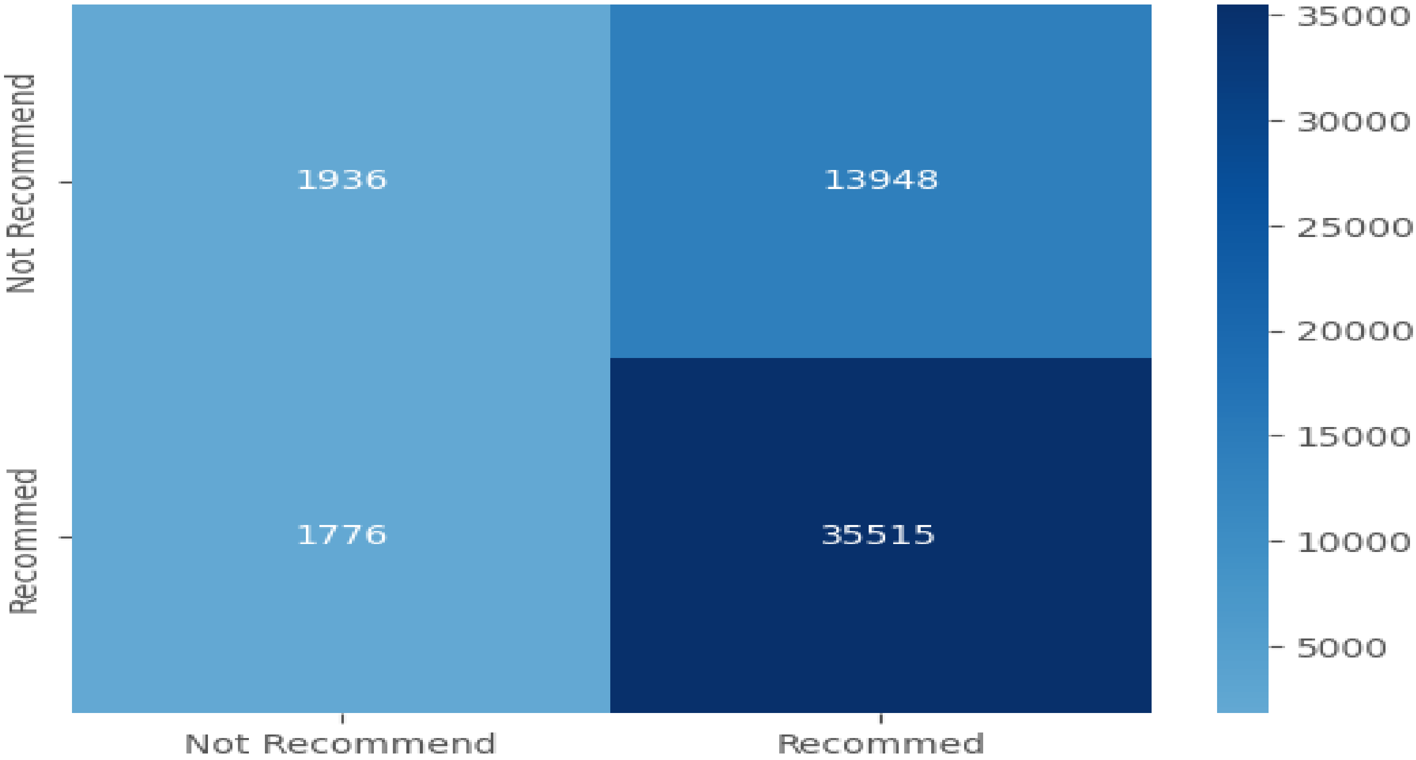


Fig. 7.2. Confusion Matrix Of Bow Vector

WORD2VEC:

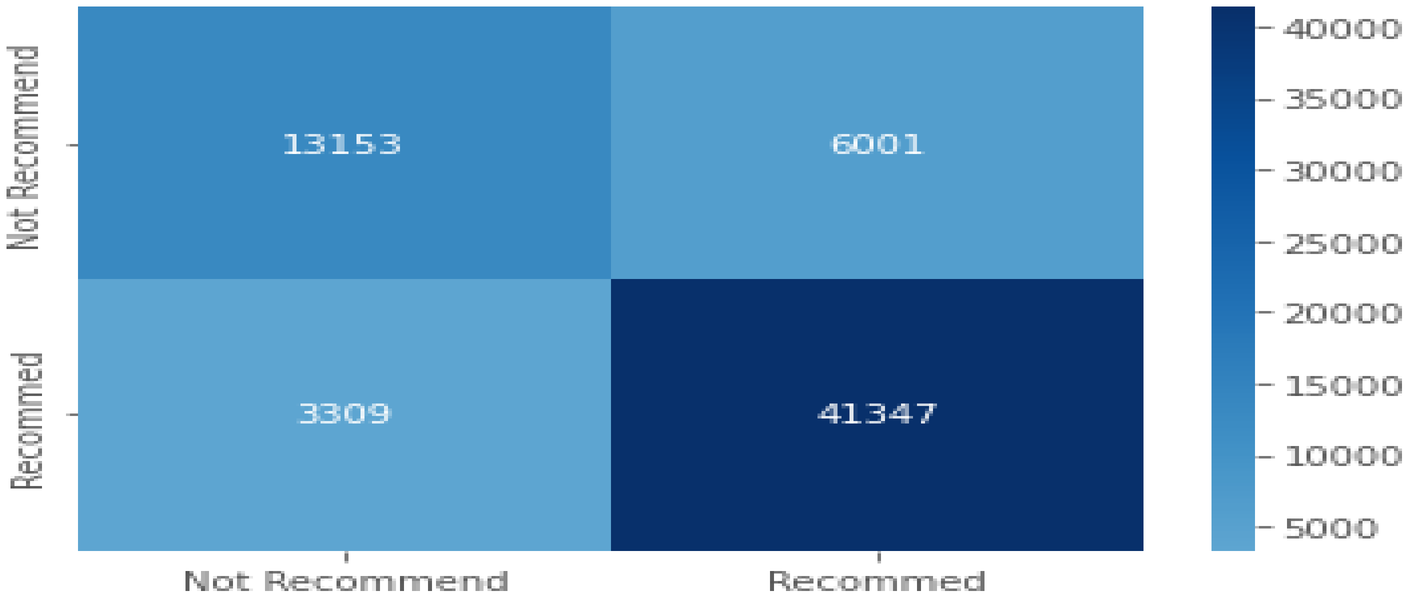


Fig. 7.3. Confusion Matrix Of Word2Vector

Chapter-7

Conclusion & Future Work

CONCLUSION:

The LSTM algorithm-based drug recommendation system has demonstrated considerable promise in reliably identifying the best medications for individuals based on their medical histories and symptoms. In comparison to conventional machine learning algorithms, the use of LSTM enables the system to efficiently capture the temporal correlations and patterns in the input data, improving prediction accuracy. Yet, both the quality and accessibility of the data, as well as the model's features, have a significant impact on the system's effectiveness. For the system to be safe and effective in real-world situations, more testing and validation through clinical studies are also necessary.Overall, the LSTM algorithm-based drug recommendation system shows potential in enhancing the effectiveness and precision of drug prescription and customised medicine, ultimately leading to better patient outcome with user friendly website , and the future work we will added in the websites to show symtoms if the drug is not recommended

Chapter-8

Reference

8.1 REFERENCES

[1] Garg, S. (2021, January). Drug recommendation system based on sentiment analysis of drug reviews using machine learning. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 175-181). IEEE..

[2] Marappan, R., Bhaskaran, S., Ashwadh, S., & Raj, H. A. (2022). Extraction of Drug Review Polarity Using Sentimental Analysis. Journal of Applied Mathematics and Computation, 6(2), 167-177.

[3] Ramya, S. P., Sumitha, B., Ranjani, R., & Ahamed, M. A. (2022, August). A Comparative Study on Aspects Level Drug Reviews using Back Propagation Neural Networks. In 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 1341-1344). IEEE.

[4] Ren, Y., Shi, Y., Zhang, K., Wang, X., Chen, Z., & Li, H. (2022). A Drug Recommendation Model Based on Message Propagation and DDI Gating Mechanism. IEEE Journal of Biomedical and Health Informatics.

[5] Basiri, M. E., Chegeni, R. S., Karimvand, A. N., & Nemati, S. (2020, April). Bidirectional lstm deep model for online doctor reviews polarity detection. In 2020 6th International Conference on Web Research (ICWR) (pp. 100-105). IEEE.

[6] Bhimavarapu, U., Chintalapudi, N., & Battineni, G. (2022). A Fair and Safe Usage Drug Recommendation System in Medical Emergencies by a Stacked ANN. Algorithms, 15(6), 186
.

[7] Hong, W., Xiong, Z., Zheng, N., & Weng, Y. (2019). A medical-history-based potential disease prediction algorithm. IEEE Access, 7, 131094-131101.

[8] Kumar, L., & Johri, P. (2022, April). Designing a health care application for prediction diabetes using LSTM, SVM and deep learning. In 2022 2nd

International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 1319-1323). IEEE.

[9] Neog, H., Duttab, P. E., & Medhic, N. (2022). Health condition prediction and covid risk detection using healthcare 4.0 techniques. Smart Health, 100322.

[10] Saad, E., Din, S., Jamil, R., Rustam, F., Mehmood, A., Ashraf, I., & Choi, G. S. (2021). Determining the efficiency of drugs under special conditions from users' reviews on healthcare web forums. IEEE Access, 9, 85721-85737.

[11] Gong, F., Wang, M., Wang, H., Wang, S., & Liu, M. (2021). SMR: medical knowledge graph embedding for safe medicine recommendation. Big Data Research, 23, 100174
.

[12] Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. Artificial Intelligence Review, 53(8), 5929-5955.

[13] Saha, B. N., & Senapati, A. (2020, July). Long Short Term Memory (LSTM) based Deep Learning for Sentiment Analysis of English and Spanish Data. In 2020 International Conference on Computational Performance Evaluation (ComPE) (pp. 442-446). IEEE.

[14] Murthy, G. S. N., Allu, S. R., Andhavarapu, B., Bagadi, M., & Belusonti, M. (2020). Text based sentiment analysis using LSTM. Int. J. Eng. Res. Tech. Res, 9(05)..

[15] Yadav, A., & Vishwakarma, D. K. (2020, September). A weighted text representation framework for sentiment analysis of medical drug reviews. In 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM) (pp. 326-332). IEEE.

[16] Su, C., Gao, S., & Li, S. (2020). GATE: graph-attention augmented temporal neural network for medication recommendation. IEEE Access, 8, 125447-125458.