Synopsis on

Medicine Recommendation System

Abhinay Hinge

B. K. Birla College, Kalyan

Title of Project:

Performing Sentiment Analysis on Medicine Dataset and Developing a Machine Learning-based Medicine Recommendation System.

Abstract:

The healthcare industry has been rapidly adopting technology to improve patient care and outcomes. In this project, we propose to develop a machine learning-based medicine recommendation system that uses patient data to recommend the most suitable medication. The system will consider a patient's medical history, demographics, and other relevant factors to suggest personalized medicine recommendations. The proposed system has the potential to improve patient outcomes and reduce healthcare costs.

Introduction:

Medication errors are a leading cause of patient harm, and finding the most effective medication for a patient can be a challenging task for physicians. Machine learning algorithms can analyse large datasets of patient data to identify patterns and make personalized medication recommendations. In this project, we propose to develop a medicine recommendation system using machine learning to help physicians make informed decisions and improve patient outcomes.

Scope:

The scope of data includes the collection, analysis, and processing of patient data to develop a medicine recommendation system. The data will include patient medical records, demographics, and other relevant factors such as medication history, allergies, and side effects. The data will be used to train and test machine learning models to develop a personalized medicine

recommendation system. The scope of data does not include the use of data for any other purpose outside the scope of the project.

Objective:

The main objective of this project is to develop a machine learning-based medicine recommendation system that can provide personalized medication recommendations based on a patient's medical history, demographics, and other relevant factors.
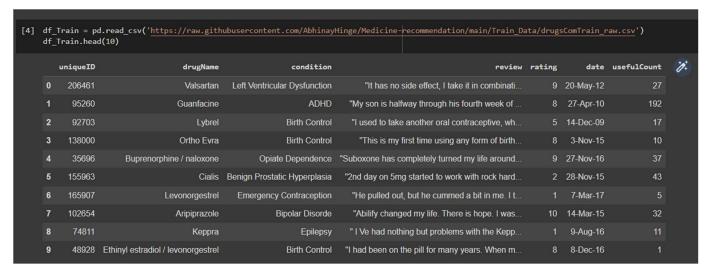
Dataset and Reference:

For this project, we will use publicly available datasets of patient medical records and medication information. We will also reference relevant research articles and studies related to machine learning-based medicine recommendation systems.

Dataset link:

Train Data: https://raw.githubusercontent.com/AbhinayHinge/Medicine-recommendation/main/Train_Data/drugsComTrain_raw.csv

Test Data: https://raw.githubusercontent.com/AbhinayHinge/Medicine-recommendation/main/Test_Data/drugsComTest_raw.csv

```
[4] df_Train = pd.read_csv('https://raw.githubusercontent.com/AbhinayHinge/Medicine-recommendation/main/Train_Data/drugsComTrain_raw.csv')
    df_Train.head(10)
```

	uniqueID	drugName	condition	review	rating	date	usefulCount
0	206461	Valsartan	Left Ventricular Dysfunction	"It has no side effect, I take it in combinati...	9	20-May-12	27
1	95260	Guanfacine	ADHD	"My son is halfway through his fourth week of ...	8	27-Apr-10	192
2	92703	Lybrel	Birth Control	"I used to take another oral contraceptive, wh...	5	14-Dec-09	17
3	138000	Ortho Evra	Birth Control	"This is my first time using any form of birth...	8	3-Nov-15	10
4	35696	Buprenorphine / naloxone	Opiate Dependence	"Suboxone has completely turned my life around...	9	27-Nov-16	37
5	155963	Cialis	Benign Prostatic Hyperplasia	"2nd day on 5mg started to work with rock hard...	2	28-Nov-15	43
6	165907	Levonorgestrel	Emergency Contraception	"He pulled out, but he cummed a bit in me. I t...	1	7-Mar-17	5
7	102654	Aripiprazole	Bipolar Disorde	"Abilify changed my life. There is hope. I was...	10	14-Mar-15	32
8	74811	Keppra	Epilepsy	" I Ve had nothing but problems with the Kepp...	1	9-Aug-16	11
9	48928	Ethinyl estradiol / levonorgestrel	Birth Control	"I had been on the pill for many years. When m...	8	8-Dec-16	1

```
[11] df.info()

     <class 'pandas.core.frame.DataFrame'>
     Int64Index: 215063 entries, 0 to 53765
     Data columns (total 7 columns):
      #    Column        Non-Null Count    Dtype
     ---   ------        --------------    -----
      0    uniqueID      215063 non-null   int64
      1    drugName      215063 non-null   object
      2    condition     213869 non-null   object
      3    review        215063 non-null   object
      4    rating        215063 non-null   int64
      5    date          215063 non-null   object
      6    usefulCount   215063 non-null   int64
     dtypes: int64(3), object(4)
     memory usage: 13.1+ MB
```

```
[12] df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
uniqueID	215063.0	116039.364814	67007.913366	0.0	58115.5	115867.0	173963.5	232291.0
rating	215063.0	6.990008	3.275554	1.0	5.0	8.0	10.0	10.0
usefulCount	215063.0	28.001004	36.346069	0.0	6.0	16.0	36.0	1291.0

```
[13] # summarization of Categorical data
     df[['drugName','condition','review']].describe(include = 'object')
```

	drugName	condition	review
count	215063	213869	215063
unique	3671	916	128478
top	Levonorgestrel	Birth Control	"Good"
freq	4930	38436	39

```
[14] # lets Explore Some of the Important Columns in the dataset

     print("Number of Unique Medicines present in the Dataset :", df['drugName'].nunique())
     print("Number of Unique Medical Conditions present in the Dataset :", df['condition'].nunique())

     print("\nThe Time Period of Collecting the Data")
     print("Starting Date :", df['date'].min())
     print("Ending Date :", df['date'].max())
```

```
     Number of Unique Medicines present in the Dataset : 3671
     Number of Unique Medical Conditions present in the Dataset : 916

     The Time Period of Collecting the Data
     Starting Date : 1-Apr-08
     Ending Date : 9-Sep-17
```

There are 3671 unique medicines with 916 unique conditions

Literature Review:

The literature on this topic is rapidly growing, and this review provides an overview of some of the recent studies on ML-based medicine recommendation systems.

- **One study by Singh et al. (2021)** developed **an ML-based system that recommends personalized treatment plans for patients with hypertension**. The system used a combination of supervised and unsupervised learning algorithms to identify the most effective drug for each patient based on their medical history, demographics, and lifestyle factors. The system was evaluated using clinical data from over 10,000 patients, and the results showed that it outperformed traditional methods of drug selection.

- **Another study by Kwon et al. (2020)** developed an **ML-based system that recommends personalized dosages of immunosuppressive drugs for kidney transplant recipients.** The system used a deep learning algorithm to predict the optimal dosage based on the patient's clinical and genetic data. The system was evaluated using data from over 300 patients and demonstrated improved accuracy compared to traditional dosing methods.

- In a **study by Lee et al. (2020)**, an **ML-based system was developed to predict adverse drug reactions (ADRs) in patients.** The system used a gradient boosting algorithm to predict the likelihood of ADRs based on the patient's clinical and genetic data. The system was evaluated using data from over 150,000 patients, and the results showed that it outperformed traditional methods of ADR prediction.

- Another **study by Chen et al. (2021) developed an ML-based system that recommends personalized treatment plans for patients with depression.** The system used a deep learning algorithm to predict the most effective treatment plan based on the patient's clinical and demographic data. The system was evaluated using clinical data from over 1,500 patients and demonstrated improved accuracy compared to traditional treatment selection methods.

- Finally, a study by **Zhang et al. (2020) developed an ML-based system that recommends personalized treatment plans for patients with type 2 diabetes.** The system used a combination of deep learning and

reinforcement learning algorithms to identify the most effective treatment plan based on the patient's clinical data and treatment history. The system was evaluated using clinical data from over 5,000 patients and demonstrated improved accuracy compared to traditional treatment selection methods.

Methodology:

The proposed system will use machine learning algorithms such as decision trees, random forests, and neural networks to analyse patient data and recommend the most suitable medication. The system will be developed using Python and relevant machine learning libraries such as Scikit-learn and TensorFlow.

Pros and Cons:

The proposed medicine recommendation system has the potential to improve patient outcomes and reduce healthcare costs by reducing medication errors and ensuring patients receive the most effective treatment. However, the system may face challenges related to data privacy and security, and there may be concerns regarding the accuracy and reliability of the recommendations.

Conclusion:

The development of a machine learning-based medicine recommendation system has the potential to improve patient outcomes and reduce healthcare costs by providing personalized medication recommendations based on patient data. However, further research is needed to evaluate the accuracy and reliability of the system and address concerns related to data privacy and security.