



SmartMEDs

Outliers Team

| Artificial Intelligence Course

Together for Tomorrow!
Enabling People

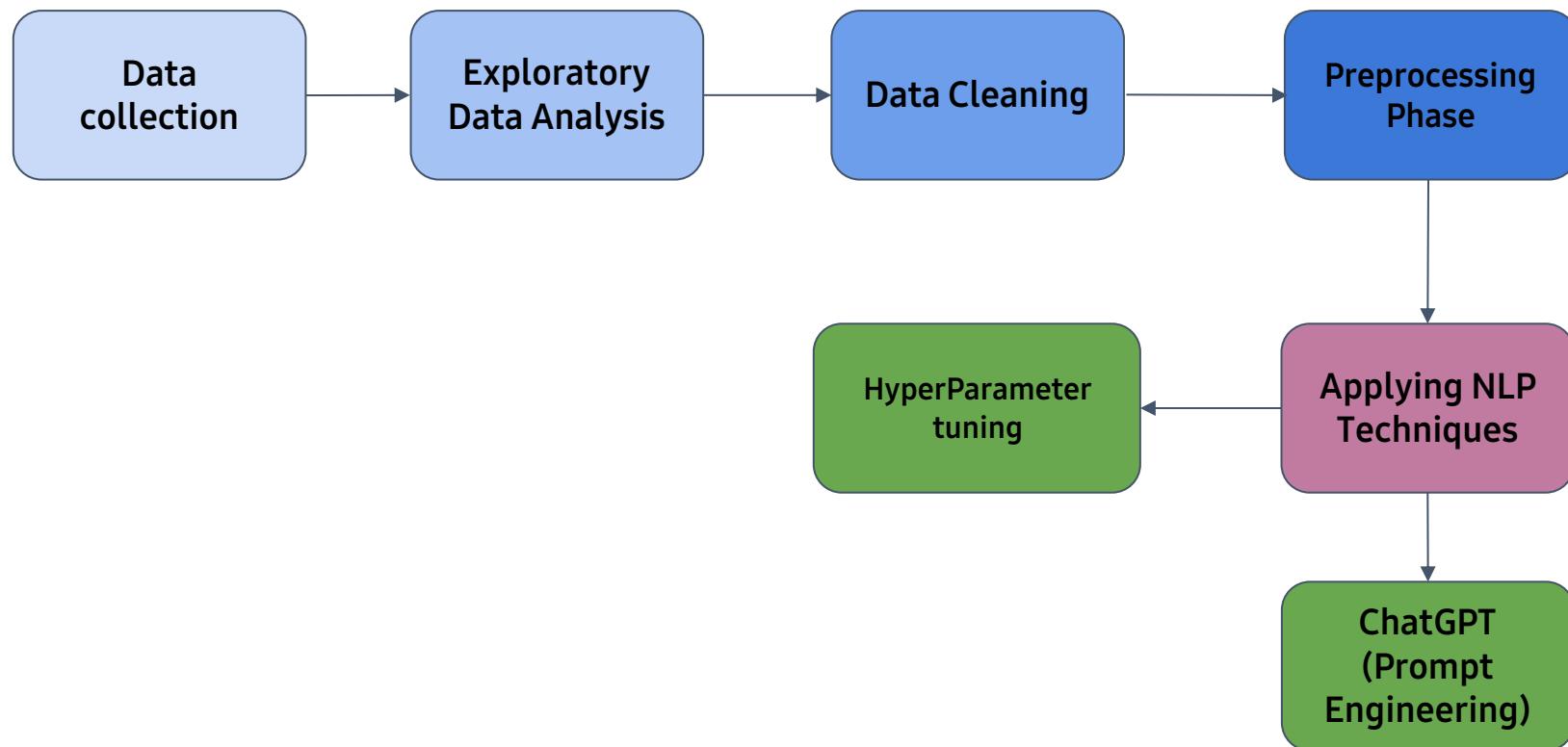
Education for Future Generations

Project Objective

- SmartMeds is a recommendation system that gives patients the *top recommended* medications which are reviewed by the *community*. It aims to reduce expenses (consultation fees) and wait times for patients.
- The model used can recommend you the medication, and also can work vice-versa i.e. if you provide it with a drug name it can return you the possible condition.
- The model is deployed in Custom ChatBOT that can help you get as many information you want using certain indicators in a calm and nice tune conversational way



Project Process



The Data

UCI Machine Learning Repository
and it contains

215.063 instance

7 columns

The Data

ID

Drug name

Condition

Rating

Date

useful-count

	Unnamed: 0	drugName	condition	review	rating	date	usefulCount
0	206461	Valsartan	Left Ventricular Dysfunction	"It has no side effect, I take it in combinati...	9.0	May 20, 2012	27
1	95260	Guanfacine	ADHD	"My son is halfway through his fourth week of ...	8.0	April 27, 2010	192
2	92703	Lybrel	Birth Control	"I used to take another oral contraceptive, wh...	5.0	December 14, 2009	17
3	138000	Ortho Evra	Birth Control	"This is my first time using any form of birth...	8.0	November 3, 2015	10
4	35696	Buprenorphine / naloxone	Opiate Dependence	"Suboxone has completely turned my life around...	9.0	November 27, 2016	37

The Data Cleaning/Preprocessing

The Data contained a lot of invalid and null values, which required comprehensive manual examination and cleaning

```
df['condition'] = df['condition'].replace('Bipolar Disorde', 'Bipolar Disorder')
#Bleeding Disorde
df['condition'] = df['condition'].replace('Bleeding Disorde', 'Bleeding Disorder')
#'Body Dysmorphic Disorde'
df['condition'] = df['condition'].replace('Body Dysmorphic Disorde', 'Body Dysmorphic Disorder')
#'Borderline Personality Disorde'
df['condition'] = df['condition'].replace('Borderline Personality Disorde', 'Borderline Personality Disorder')
#'Brain Tum'
df['condition'] = df['condition'].replace('Brain Tum', 'Brain Tumor')

#Breast Cance
df['condition'] = df['condition'].replace('Breast Cance', 'Breast Cancer')
#Cance
df['condition'] = df['condition'].replace('Cance', 'Cancer')
#'Colorectal Cance'
df['condition'] = df['condition'].replace('Colorectal Cance', 'Colorectal Cancer')
#Corneal Ulce
df['condition'] = df['condition'].replace('Corneal Ulce', 'Corneal Ulcer')
#Dermatitis Herpeti
df['condition'] = df['condition'].replace('Dermatitis Herpeti', 'Dermatitis Herpetiformis')
#'Dissociative Identity Disorde'
df['condition'] = df['condition'].replace('Dissociative Identity Disorde', 'Dissociative Identity Disorder')
#'Duodenal Ulce'
df['condition'] = df['condition'].replace('Duodenal Ulce', 'Duodenal Ulcer')
#'Endometrial Cance'
df['condition'] = df['condition'].replace('Endometrial Cance', 'Endometrial Cancer')
#'Gastric Cance'
df['condition'] = df['condition'].replace('Gastric Cance', 'Gastric Cancer')

#'Gastrointestinal Stromal Tum'
df['condition'] = df['condition'].replace('Gastrointestinal Stromal Tum', 'Gastrointestinal Stromal Tumor')
#'Generalized Anxiety Disorde'
df['condition'] = df['condition'].replace('Generalized Anxiety Disorde', 'Generalized Anxiety Disorder')
```

condition	review	rating	date
2 users found this comment helpful.	"I'm 16 and I have been on Loestrin 24 f...	3.0	2010-11-03
4 users found this comment helpful.	"I'm 24 years old and have always had a p...	1.0	2016-03-27
3 users found this comment helpful.	"I took Yaz for a little over 2 years. From a...	3.0	2010-06-01
4 users found this comment helpful.	"Took this pill for 1.) Acne and 2.) Birth Con...	2.0	2014-06-24
11 users found this comment helpful.	"I have suffered with low back pain - 2 surger...	9.0	2009-03-19
...
3 users found this comment helpful.	"I have been taking this pill for less than a ...	8.0	2010-09-24
5 users found this comment helpful.	"I got the Mirena put in last month. And holy ...	7.0	2013-02-12
21 users found this comment helpful.	"It was a nightmare.I had the worse side effec...	1.0	2015-04-14
28 users found this comment helpful.	"My 10 year old son took Dulera for asthma. I...	1.0	2012-02-29
5 users found this comment helpful.	"Wonderful drug, after being on narcotics for ...	9.0	2009-12-29

The Cleaning & Preprocessing

1. Remove values HTML entities in 'conditions' column (using RegEx)
2. Remove escape characters in the 'review' column to make it suitable for further processing
3. Out of Range (drug names, html entity values etc in 'condition' column)
4. Date Parsing
5. Stop words removal, tokenization and lemmatization on 'review' column
6. Natural Language Processing sentiment intensity analyzer
1. Label Encoding

review

"It has no side effect, I take it in combinati...

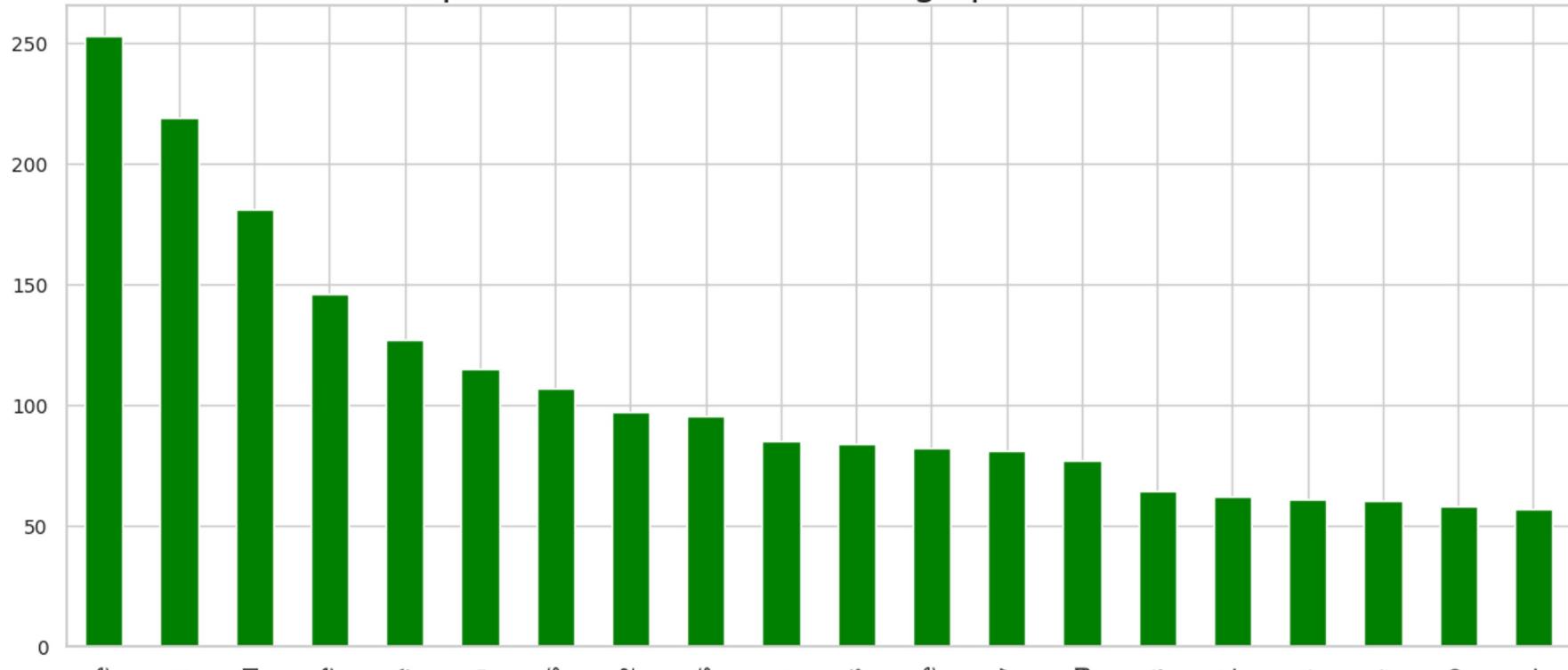
"My son is halfway through his fourth week of ...

"I used to take another oral contraceptive, wh...

"This is my first time using any form of birth...

'Suboxone has completely turned my life around...

Top20 : The number of drugs per condition.



3 users found this comment helpful.

4 users found this comment helpful.

The Cleaning & Preprocessing

Using over
100 lines of
code for
manual
cleaning and
analysis

```
df['condition'] = df['condition'].replace('Idiomyalgia', 'Primary Myalgia')
df['condition'] = df['condition'].replace('Binge Eating Disorde', 'Binge Eating Disorder')
#Benign Essential Trem
df['condition'] = df['condition'].replace('Benign Essential Trem', 'Benign Essential Tremor')
#Bipolar Disorde
df['condition'] = df['condition'].replace('Bipolar Disorde', 'Bipolar Disorder')
#Bleeding Disorde
df['condition'] = df['condition'].replace('Bleeding Disorde', 'Bleeding Disorder')
#Body Dysmorphic Disorde
df['condition'] = df['condition'].replace('Body Dysmorphic Disorde', 'Body Dysmorphic Disorder')
#Borderline Personality Disorde
df['condition'] = df['condition'].replace('Borderline Personality Disorde', 'Borderline Personality Disorder')
#Brain Tum
df['condition'] = df['condition'].replace('Brain Tum', 'Brain Tumor')

#Breast Cance
df['condition'] = df['condition'].replace('Breast Cance', 'Breast Cancer')
#Cance
df['condition'] = df['condition'].replace('Cance', 'Cancer')
#Colorectal Cance
df['condition'] = df['condition'].replace('Colorectal Cance', 'Colorectal Cancer')
#Corneal Ulce
df['condition'] = df['condition'].replace('Corneal Ulce', 'Corneal Ulcer')
#Dermatitis Herpeti
df['condition'] = df['condition'].replace('Dermatitis Herpeti', 'Dermatitis Herpetiformis')
#Dissociative Identity Disorde
df['condition'] = df['condition'].replace('Dissociative Identity Disorde', 'Dissociative Identity Disorder')
#Duodenal Ulce
df['condition'] = df['condition'].replace('Duodenal Ulce', 'Duodenal Ulcer')
#Endometrial Cance
df['condition'] = df['condition'].replace('Endometrial Cance', 'Endometrial Cancer')
#Gastric Cance
df['condition'] = df['condition'].replace('Gastric Cance', 'Gastric Cancer')
```

Encoding All the Values

Users > ahmed > Downloads > drug+review+dataset+drugs+com > drugsComTrain_raw.tsv

```

1 | drugName condition review rating date usefulCount
2 206461 Valsartan Left Ventricular Dysfunction ""It has no side effect, I take it in combination of Byst
3 95260 Guanfacine ADHD ""My son is halfway through his fourth week of Intuniv. We became concerned when
4 We have tried many different medications and so far this is the most effective.""" 8.0 April 27, 2010 19
5 92703 Lybrel Birth Control ""I used to take another oral contraceptive, which had 21 pill cycle, and was
6 The positive side is that I didn't have any other side effects. The idea of being period free was so
7 138000 Ortho Evra Birth Control ""This is my first time using any form of birth control. I'm glad
8 35696 Buprenorphine / naloxone Opiate Dependence ""Suboxone has completely turned my life around. I fee
9 155963 Cialis Benign Prostatic Hyperplasia ""2nd day on 5mg started to work with rock hard erections h
10 165907 Levonorgestrel Emergency Contraception ""He pulled out, but he cummed a bit in me. I took the Pl
11 102654 Aripiprazole Bipolar Disorde ""Abilify changed my life. There is hope. I was on Zoloft and Cloni
12 74811 Keppra Epilepsy ""I've had nothing but problems with the Keppra : constant shaking in my arms
13 48928 Ethinyl estradiol / levonorgestrel Birth Control ""I had been on the pill for many years. When my
14 29607 Topiramate Migraine Prevention ""I have been on this medication almost two weeks, started out on 2
15 75612 L-methylfolate Depression ""I have taken anti-depressants for years, with some improvement but mo
16
17 I only take Cymbalta now mostly for pain.
18
19 When I began Deplin, I noticed a major improvement overnight. More energy, better disposition, and no sink
20 191290 Pentasa Crohn's Disease ""I had Crohn's with a resection 30 years ago and have been mostly i
21 221320 Dextromethorphan Cough ""Have a little bit of a lingering cough from a cold. Not giving me much
22 98494 Nexplanon Birth Control ""Started Nexplanon 2 months ago because I have a minimal amount of contrac
23 I've never had acne problems in my life, and immediately broke out after getting it implanted. Sex dr
24 81890 Liraglutide Obesity ""I have been taking Saxenda since July 2016. I had severe nausea for about a
25 48188 Trimethoprim Urinary Tract Infection ""This drug worked very well for me and cleared up my UTI in
26 219869 Amitriptyline ibromyalgia ""I've been taking amitriptyline since January 2013 after being di
27 212077 Lamotrigine Bipolar Disorde ""I've been on every medicine under the sun (it seems) to manage
28 119705 Nilotinib Chronic Myelogenous Leukemia ""I have been on Tasigna for just over 3 years now (300mg
29 12372 Atripla HIV Infection ""Spring of 2008 I was hospitalized with pneumonia and diagnosed with Lyme di
30 231466 Trazadone Insomnia ""I have insomnia, it's horrible. My story begins with my PCP prescribin
31 227020 Etonogestrel Birth Control ""Nexplanon does its job. I can have worry free sex. The only thing i

```

Opening a folder will close all currently open editors. To keep them open, add a folder instead.

You can clone a repository locally.

Clone Repository

To learn more about how to use Git and source control in VS Code read our docs.

Users > ahmed > Downloads > proccesed_fr.csv >

	drug_name	Unnamed: 0	condition	nltkF
1	,3421,0,500,2,2			
2	1,1536,1,80,3,2			
3	2,1983,2,175,3,1			
4	3,2446,3,175,3,1			
5	4,553,4,613,3,2			
6	5,715,5,170,1,2			
7	6,1874,6,314,2,1			
8	7,284,7,174,1,2			
9	8,1779,8,324,1,1			
10	9,1263,9,175,3,1			
11	10,3272,10,542,3,1			
12	11,1805,11,273,3,2			
13	12,2554,12,251,1,1			
14	13,988,13,249,3,1			
15	14,2302,14,175,3,1			
16	15,1915,15,601,3,1			
17	16,3332,16,867,1,1			
18	17,214,17,346,1,2			
19	18,1826,18,174,2,1			
20	19,2317,19,223,1,1			
21	20,333,20,387,3,1			
22	21,3297,21,474,3,2			
23	22,1272,22,175,1,1			
24	23,1256,23,746,3,1			
25	24,3250,24,874,1,1			
26	25,384,25,216,1,1			
27	26,1145,26,423,1,2			
28	27,922,27,80,1,1			
29	28,327,28,637,3,2			
30	29,1658,29,541,1,1			
31	30,2000,31,272,2,1			

PROBLEMS 211 OUTPUT DEBUG CONSOLE TERMINAL

Svn installation not found.

Models used

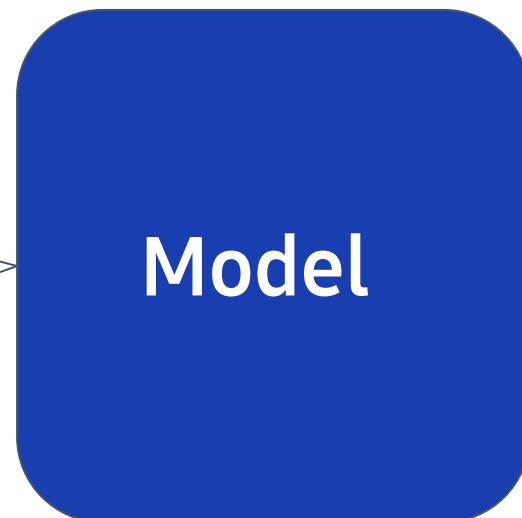
Model name	Sklearn model name	Hyperparameter	Accuracy	Mean Absolute Error (degrees)	F-1 Score	Recall	Precision
Random Forest	RandomForest Classifier	Used with RandomizedSearchCV to find the best Max depth= 18 No. of estimators=255	0.68	85.01	0.68	0.68	0.68
Random Forest Regressor	RandomForest Regressor	Max depth= 18 No. of estimators=255	0.69	82.54	0.69	0.69	0.69
Decision Tree	DecisionTreeClassifier	Criterion = entropy max depth = [10,20, 25,50] used the best parameters using gridsearchCV	0.72	75.94	0.72	0.72	0.72
KNN	KNeighborsClassifier	Using gridsearch with cross validation, the best n_neighbors was found to be n-neighbors=9	0.69	82.54	0.69	0.69	0.69

The Application

API App



ChatGPT



Model

Conversational
Model

Challenges & Future Enhancement

- 1. Imbalanced Data** - Some of the classes in our data was just a noise and removing it wasn't that easy
- 2. API** - We configured the existing API to add symptoms to our data and it needed some parameters and more configuration
- 3. Modeling** - Due to the complexity of the data and its volume, our machines couldn't handle training on more powerful models, as some models are taking over 8 hours to train and also crashes sessions

- 1. Train the data** in a more powerful model to improve the accuracy and relevance of medication suggestions. This includes considering additional factors such as patient demographics or medical history and fine-tuning the weighting of different factors in the recommendation process.
- 2. Apply API** For a larger database and more information retrieval
- 3. Implement functionality** to provide real-time updates and alerts to users, such as notifications about new medications or changes in recommendations based on updated information or user feedback



Together for Tomorrow! Enabling People

Education for Future Generations

©2023 SAMSUNG. All rights reserved.

Samsung Electronics Corporate Citizenship Office holds the copyright of book.

This book is a literary property protected by copyright law so reprint and reproduction without permission are prohibited.

To use this book other than the curriculum of Samsung Innovation Campus or to use the entire or part of this book, you must receive written consent from copyright holder.