Chronic Illness: Symptoms, Treatments & Triggers

Sometimes you will be in control of your illness & other times you will sink into despair, & that's OK! Freak out, Forgive yourself, & try again tomorrow

Anvay Govind Pandit, Shibin Tazhe Veettil, Seyedhooman Sajjadi

INTRODUCTION

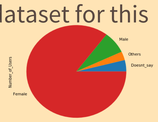
Many people experience suffering chronic diseases and symptoms throughout their life. Chronic symptoms usually need persistent tracking & management due to their nature. Patients usually go through a long period of treatment with different medications and remedies while tracking the effect.

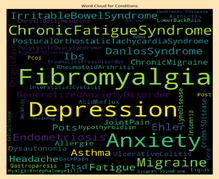
Analyzing & **interpreting** the relationships between **symptoms**, **treatments**, & the **environmental factors** to find meaningful patterns can help doctors and patients both.

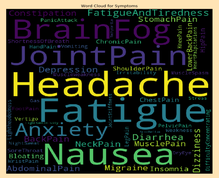
DATASET

We utilized the Flaredown Chronic Illness dataset for this project. Some key statistics:

- *Number of unique Users*: 22071
- *Total Number of data points:* **2,483,304**



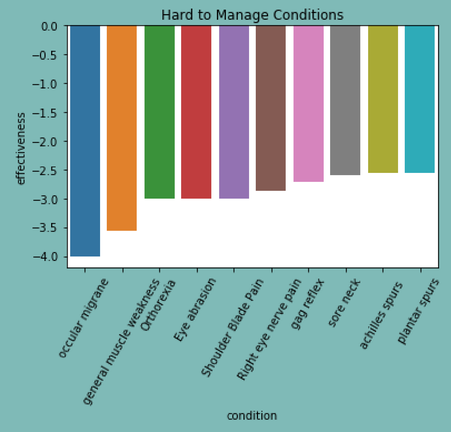


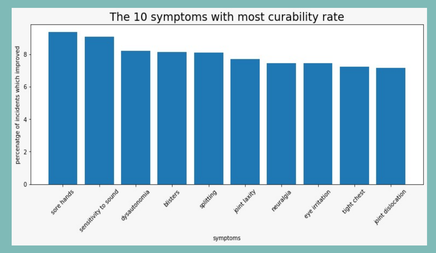


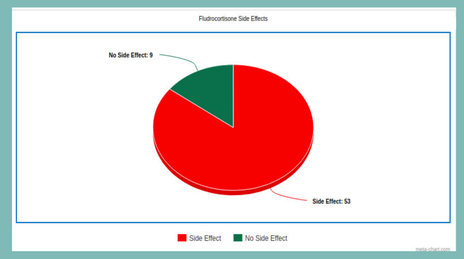
- *Common Symptoms :* ***Fatigue, Headache, Nausea***
- *Common Conditions****: Fibromyalgia, Depression, Anxiety***
- *Common Treatments:* ***Ibuprofen, Vitamin D, Tramadol***

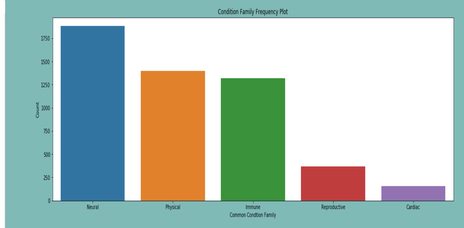
METHODOLOGY

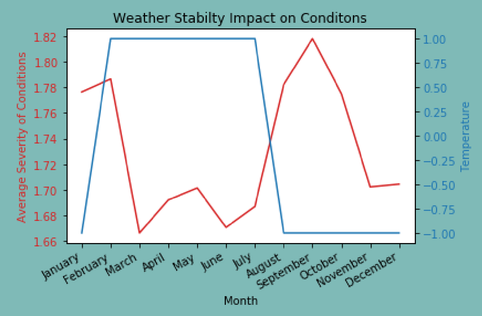
- Data Exploration – PCA & Clustering of Conditions/Symptoms
  - To figure out which conditions/symptoms live near each other
- Association Rule Mining (Apriori) with Items as **Symptoms+Treatments** & Transaction as **Users**
- Decision Tree Classifiers to predict conditions based on symptoms
- Treatment Recommendation for conditions with Effectiveness Score (SVD++)
  - The Recommendation System Predicts the effectiveness of new treatments for existing conditions
- Result Analysis
  - To Analyze the results and collect unique insights related to the above methods

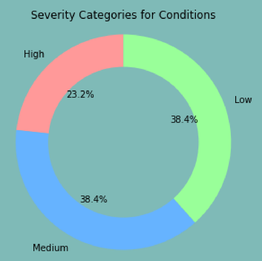












TREATMENT RECOMMENDATION & CONDITION PREDICTIONS

Matrix Factorization based (SVD++) Recommendation Symptom to recommend effective treatments based on conditions

For instance Circadin is an effective treatment for Fatigue and Dizziness
RMSE – 0.97 based on effectiveness score

Decision Tree Based (XgBoost) to predict conditions based on symptoms
F1 Score – 74 %

SIDE EFFECTS & OTHER INSIGHTS

Treatment	Side Effects
(loratadine)	(headache)
(fludrocortisone)	(nausea,headache)
(loratadine)	(fatigue, headache)
(lansoprazole)	(abdominal pain)
(melatonin)	(fatigue, headache)

- Above are few examples side effects we observed in rule mining , parameters are S=0.001, C = 0.4 and Lift = 1
- Another interesting insight we observed is that the severity of any condition is low when there is a stable weather

DATA PREPROCESSING

- Replaced null age values with the median age
- We used the NLTK synonym database to remove :
  - Symptom synonymity (like 'fatigue' = 'tiredness', 'drowsiness' = 'sleepiness' etc)
  - Condition Synonymity(like 'anxiety' = 'anxiousness', 'drowsy' = 'yawning' etc)
  - Treatment Synonymity (like 'advil' = 'ibuprofen' = 'motrin', 'sertraline', 'zoloft' etc)
- Cleaned the text tags using text cleaning techniques

CHALLENGES

- The Scale of the data was very large with around 2 million rows, increasing the processing and training time.
- Data pre-processing was a challenge because large number of treatments were identical like (advil, ibuprofen,motrin)
- The fine tuning of the support, confidence and lift parameters was an arduous process.

WHAT'S NEXT?

- Combine Information from Tags to get more insightful rules.
- Analyze Food data to extend features for the decision tree classifier.
- Implicit recommendation Systems for effective treatments