

Contents

Overview

Project Objective

Workflow

Data Pre-processing

NLP Processing

Topic Modelling

Model Fitting

Logistic Regression

Naive Bayes

Conclusion

Future Work

Overview

Pharmaceuticals provide humongous variety of drugs.

- Multiple drugs available for each condition.
- Patient have little information about drugs.
- It makes difficult for them to land on the suitable drug.

Project objective

Our objective is to analyze drugs, based on reviews.

To Help patients decision based on their concern.

Workflow





UCI machine learning repository

Preprocessing



Clean data
Remove stop words
Lemmatizing
CountVectorizer

Topic Modeling



Latent Dirichlet Allocation

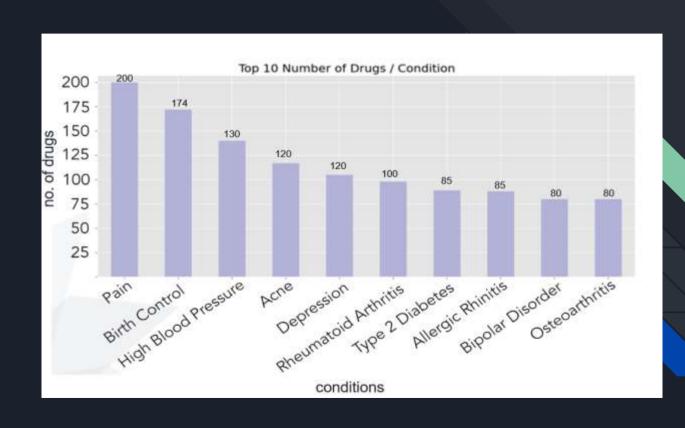
Data Pre-Processing

- Data Cleaning
- NLP Natural Language Preprocessing

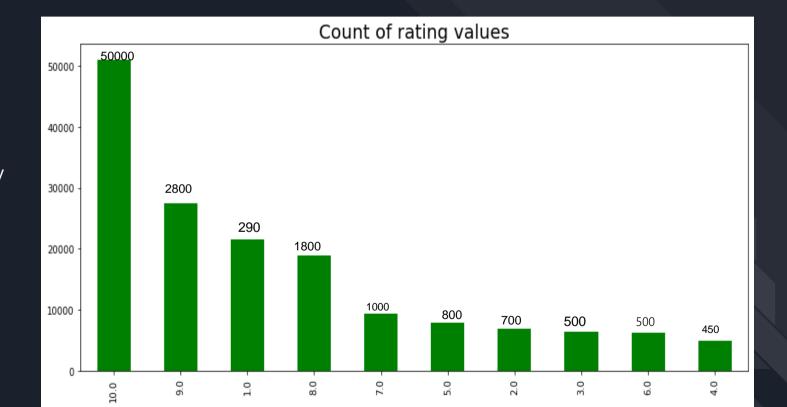
NLP Processing

- Removal of StopWords
- Used Stemming
- Used Lemmatizing

Top 10 Number of Drugs/Conditions



Rating Count Distribution



Rating

Frequency

Topic Modelling

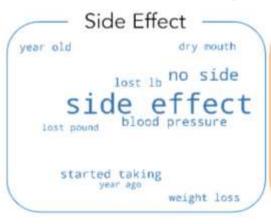
- LDA - Latent Dirichlet Analysis

It can help with the following:

- discovering the hidden themes in the collection.
- classifying the documents into the discovered themes.
- using the classification to
 organize/summarize/search the document

Results of Topic Modelling

Topic Modelling





started taking

panic attack

no longer

much better

depression anxietycould not

year old year ago

back pain

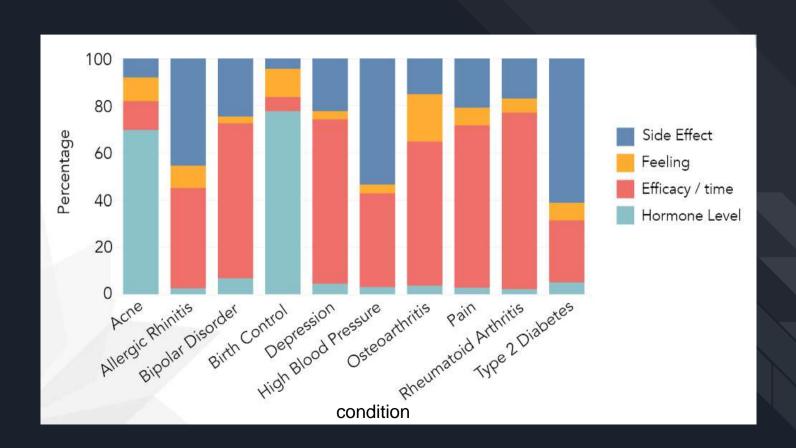
feel like

Birth Control

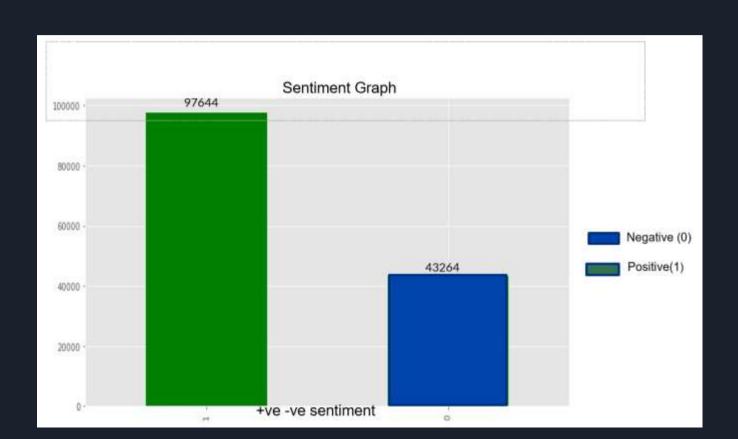
gained weight
no weight weight gain
taking pill

birth control
first month
mood swing
sex drive
side effect
gained pound

Topic Distribution of Top 10 Conditions



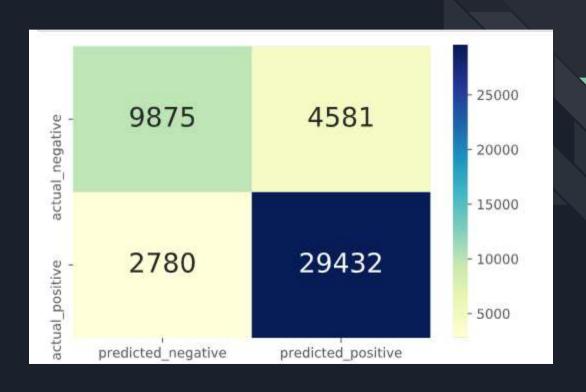
Sentiment Graph



Model Fitting

- Logistic Regression
- Naive Bayes

Logistic Regression(Unigram): confussion Matrix



accuracy 0.84

Logistic Regression for(bigram)- Best Fitted



Conclusion Predicted Values

	LogReg1	LogReg2	NB1	NB2
Accuracy	0.842	0.919	0.797	0.846
Precision	0.864	0.932	0.857	0.895
Recall	0.914	0.952	0.848	0.880
F1 Score	0.888	0.942	0.852	0.887

Business Value and Future Work

- Build a Recommender System
 - Help patients Decision Making
 - Provide benchmark to drug provider
- Topic Modelling for reviews across different conditions

Thank You!!