Mined
Hackathon

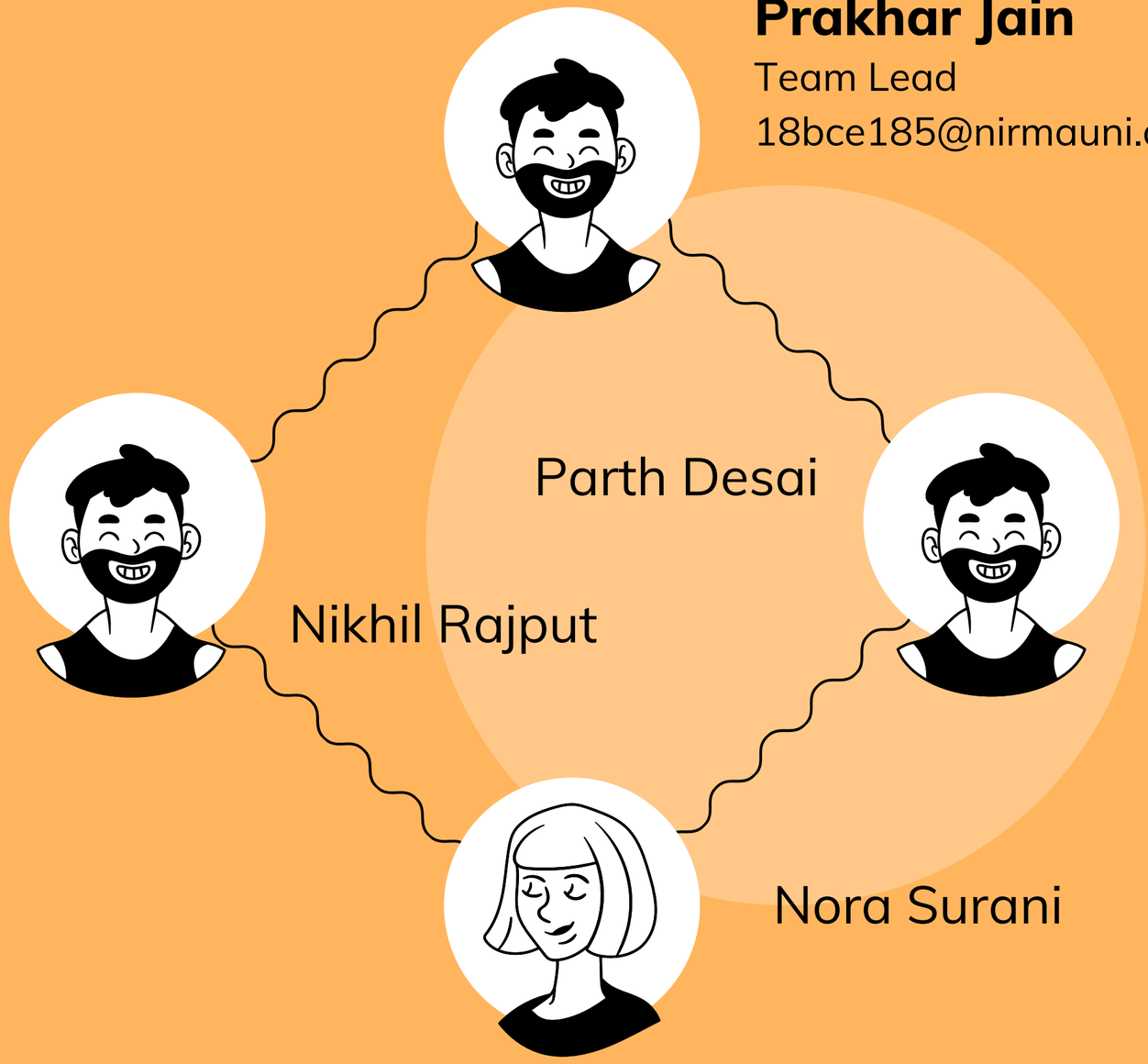
MEDICAL TEXT ANALYSIS

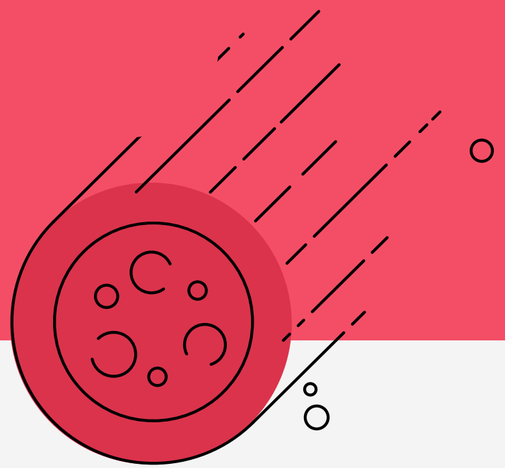
PITCH DECK v01

Team DeepBlue

Nirma University

Prakhar Jain
Team Lead
18bce185@nirmauni.ac.in





Problem

Develop an NLP module to identify the keywords related to a patient's investigation, medication and chief complaint from a free text in the text box. Highlight the extracted content and feed them as input in EMR’s Chief complaint, Investigation and Medication module.

PITCH DECK V1.0

Complaints

1 "I have high fever with chills, since 3 days. The fever increases at night. I also have generalized body ache and diziness"

2 "I have itching in the front lower abdominal reigon since 1 month, on ithching the area becomes red and rashes develops.Pain increases when the skin is dry and decreases when i apply moisturising cream."

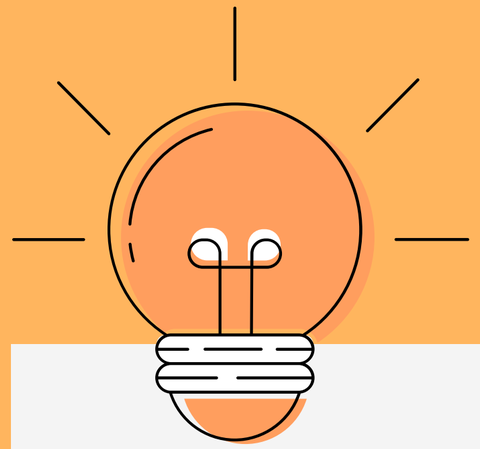
3 I am having continous pain in the stomach since 2 days. I am also experiencing constipation and delay emptying of stomach. The pain increases after eating food and decreases after lying down.

DATA

46
Rows

7
Categories

314
Data Points



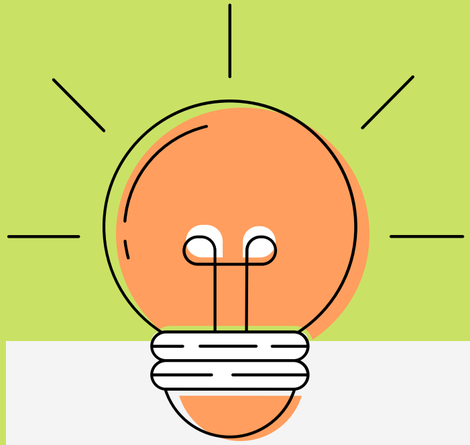
Data Processing

1 The dataset was cleaned by removing stop words, excess whitespace, punctuation.

2 Word Tokenization was implemented for better classification of symptoms

3 Mined and modelled a bag of words for symptoms and checked similarity with the tokenized words from the story while checking for symptoms.

PITCH DECK V1.0



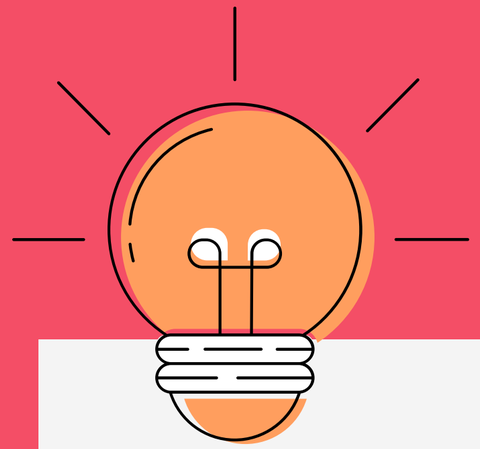
Data Classification

1. Successfully able to extract information like duration, severity and symptoms from the dataset.

2. The classification was done majorly through various Regex Techniques and the essential information is now readily available in a summarized format for future processing.

3. Furthermore the entire project has been deployed on a web app and as an android application

PITCH DECK V1.0



Approach

Our NLP approach uses both text and sentence tokenization.

No third party NLP libraries were used for the project and all pattern matching was done through Regex

This helps us serve results instantly as opposed to using popular libraries like nltk or spacy, giving us the name InstantMD. We preserving sentence structure while also searching the tokenized words. We can then extract symptoms using a list of symptom keywords and the chief complaint by matching with the part of the anatomy that the patient is experiencing symptoms in. This is done through the anatomy keywords. Other factors are similarly extracted using keywords and rules designed by us and implemented through Regex.

Future Expansion

1. Using already available dataset and research papers, we will be able to diagnose the disease as well.

2. The model can be trained on FastAI and using Multi-Label Text Classification so that it keeps learning by itself and increasing the accuracy of it's diagnosis.

3. Using PyTorch, FastAI and MelSpectograms we will be able to successfully extract audio from call recordings or voice texts for better reach

Functioning Chatbots

OneRemission

SafeDrug Bot

Florence

SymptomMate

WebMD

Family Doctor

Symptom Diagnosing Services

DEMO WEBSITE

Welcome to Instant MD

Instant MD is a Investigation, Medication and Chief complaint recognition application using NLP

Problem Statement

Abstract

Being able to have machines understand unstructured textual content already plays a big part nowadays in our life. NLP can contribute largely to the advancement of medical science. NLP is used to extract information from free text narratives written by a variety of healthcare providers. Here we approach natural language processing algorithm where we

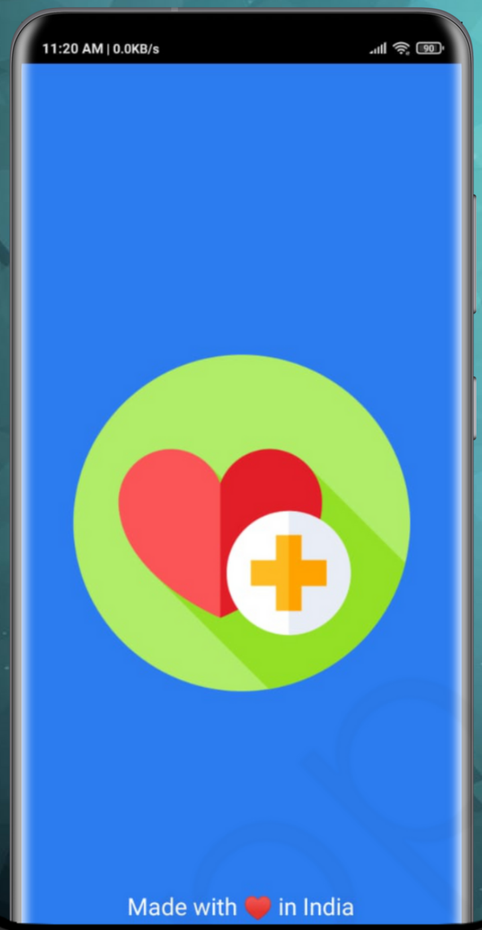
Name	Gender	Story
Prakhar Jain	M	i have pain in shoulder last 1 month. the area also have mild swelling. the pain increases on lifting weight and decreases on applying hot pack or pressure.

Factor	Status	Description
Chief Complaint	✔ Found	pain, area: shoulder
Symptoms	✔ Found	['pain', 'swelling']
Duration	✔ Found	1 month
Severity	✔ Found	mild
Associated Symptom	✔ Found	have mild swelling
Aggravating Factors	✔ Found	increases on lifting weight
Relieving Factors	✔ Found	decreases on applying hot pack or pressure

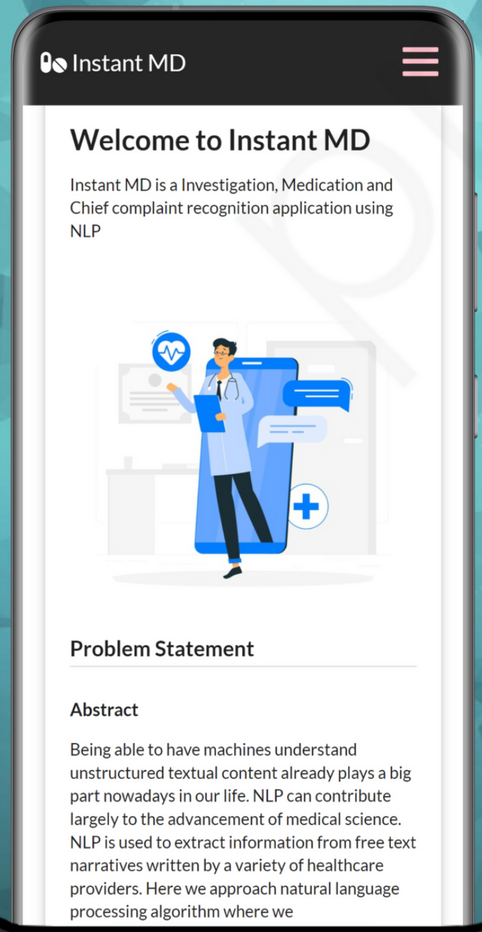
Store in Database

DEMO APP

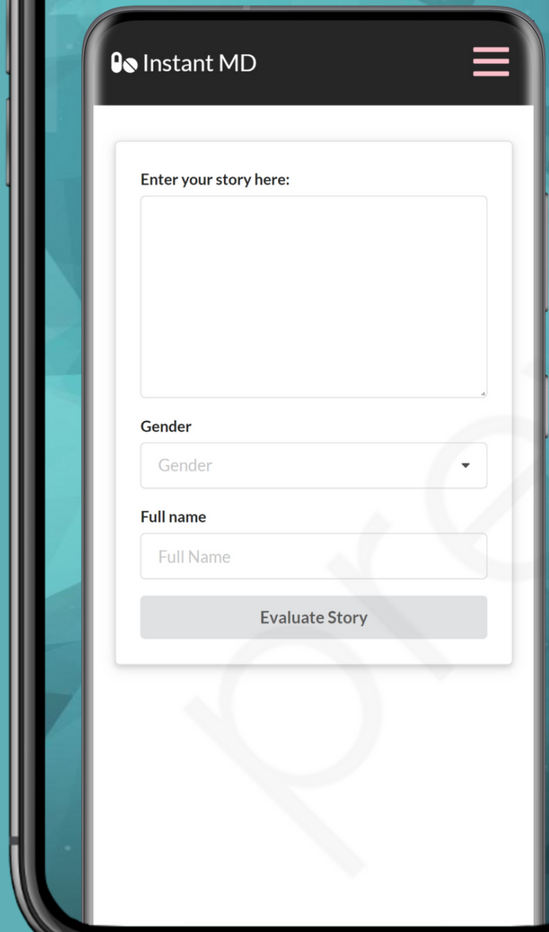
INSTANT MD MOBILE APP



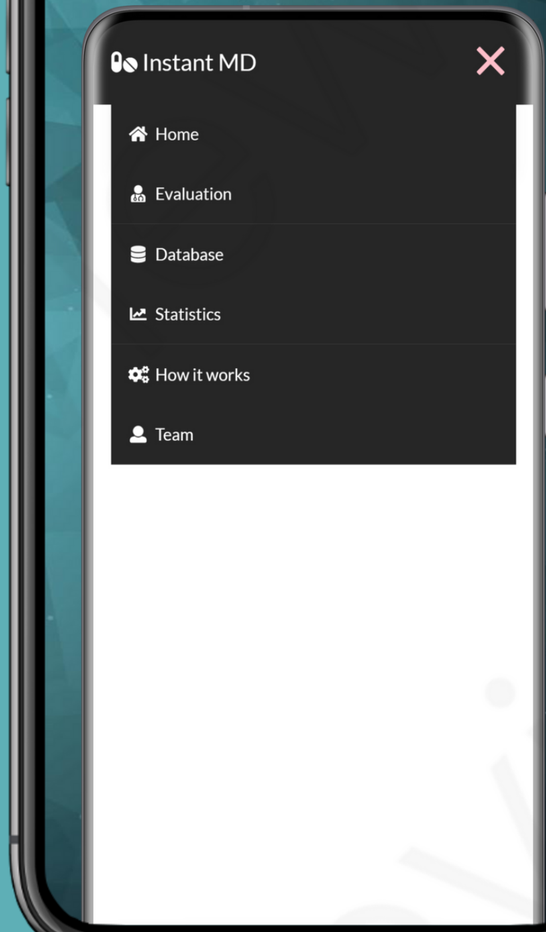
HOME PAGE



EVALUATE YOUR STORIES



GET REPORT INSTANTLY



REFERENCES

DATASET

https://github.com/sekharvth/symptom-disease
https://github.com/infermedica/js-symptom-checker-example

PREPROCESSING

Medical Symptoms Text and Audio Classification | Kaggle

Text Data Cleaning - tweets analysis | Kaggle