

Prediction of Physicians For Patient Diagnosis

Under the guidance of -

Prof. Ananthanarayana V. S.

Submitted by -

Bharath Simha

- 16IT146

Bharat Sharma

- 16IT210

D. Praneetha

- 16IT215

Literature Survey

Paper	Year	Authors	Methodology	Results	Limitations
Disease Prediction and Doctor Recommendation System [1]	March 2018	Tejaswita P. Vaidya Dhanashri Gujar Rashmi Biyani	Predicts disease using Naïve Bayes. Suggesting the details of the disease specialists based on Success rate.	Accuracy above 80 percent. Doctors can be chosen on success rate or fare.	Very few testing and training symptoms and limited filters.
A Hybrid Recommender System for Patient-Doctor Matchmaking in Primary Care [2]	March 2019	Qiwei Han Manas Gaur Leid Zejnilovic	Find doctors visited by similar patients who visits the same doctor.(Hybrid)	Generate a list of doctors for each patient ranked by the predicted trust.	Personalized doctor recommendations.Not for the public(Individual centralized).
Prediction System for Diseases and Suggestion of Appropriate Medicines[3]	December 2017	Disha Mahajan Mrudula Phalak Saniya Pathan	The Classification of data according to requirements and then by applying Association on it for the prediction of diseases	With given symptoms of the patients, predicting disease and recommendation of the prescription of the obtained disease.	Not available with doctors (MEDICINES ONLY) and medicine is prescribed for very few diseases.



Paper	Year	Authors	Methodology	Results	Limitations
A Recommender system for identifying the right doctor [4]	December 2016	Li Guo Bo jin Haoyu yang	To develop a recommender system for identifying KOLs for any specific disease with unsupervised learning	To identify KOL system can make recommendation to pharmaceutical companies and patients.	Not directly connected to doctors or pharmaceuticals.
Deep Learning Based Health, recommender system using Collaborative Filtering.[5]	May 2019	Abhaya Sahoo Chittaranjan Pradhan , Rabindra Kumar	This paper gives a proposed Intelligent HRS using Convolutional Neural Network (CNN)	Systems find recommended hospitals by calculating the similarity of patients' choices.	Recommendation of hospitals not doctors as such.



Outcome of Literature Survey

- The works have been done either with disease prediction or only hospitals recommendation in mind.
- Closely related papers predict a single doctor.
- Most of these papers collaborated with various hospitals and used their data for prediction.
- Hence, the decision was made to take unavailability of doctors into account And thus, the result was to output a list of doctors.



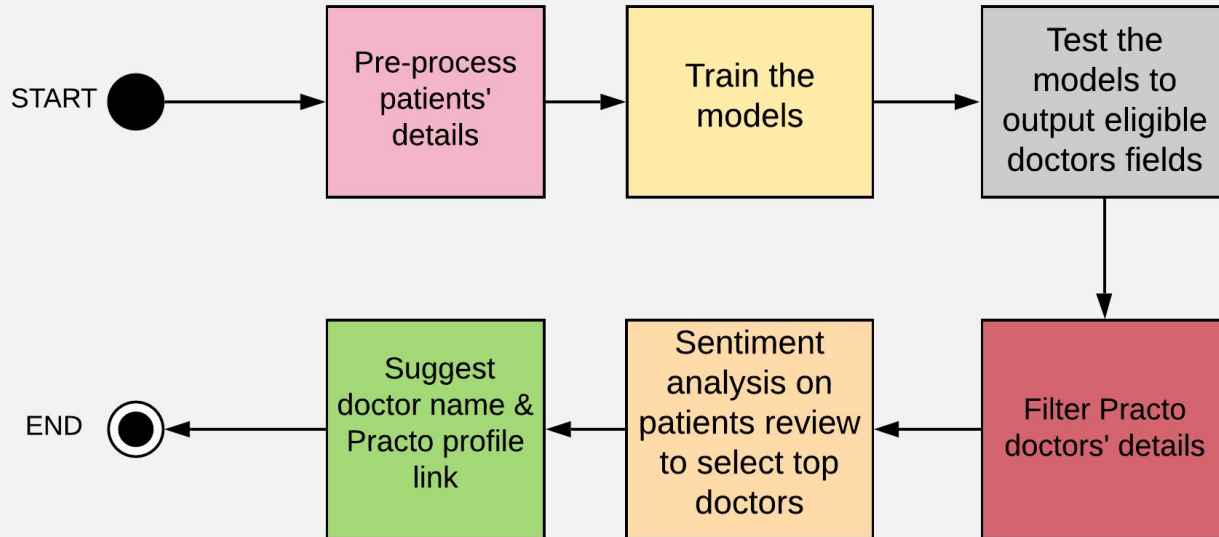
Problem Statement

An efficient predictive model for selecting an appropriate group of healthcare practitioners based on patient details.

Objectives:

- Determining the type of doctor required to diagnose the patient based on patient details i.e. joining two datasets.
- Construct an eligible group of type of doctors to which a patient can refer to, in order.
- Collecting doctors details from open-source data and using a model to construct their rating based on online reviews.
- Mapping the Practo profiles with the predicted type of doctors.

Methodology

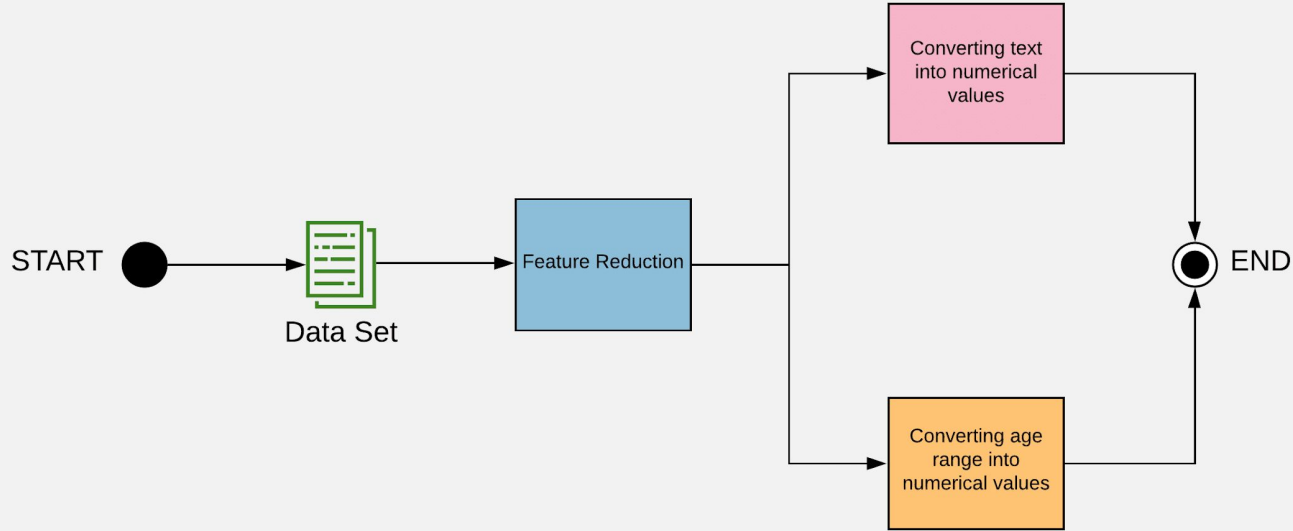




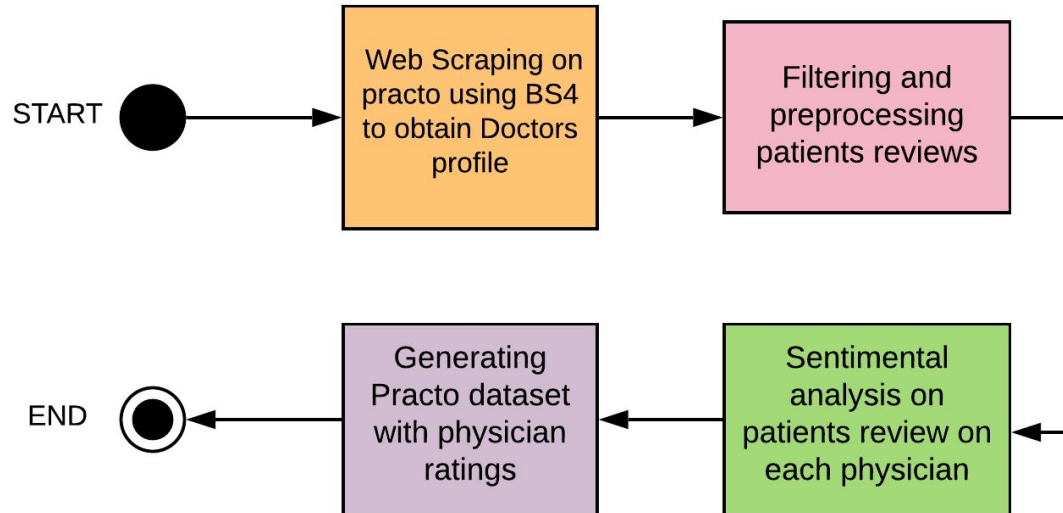
DATASET

- **The Health Statewide Planning and Research Cooperative System (SPARCS):**
It contains Medical information of patient characteristics, diagnoses, treatments and services of patients whose lesional and/or functional status is considered to be stable.
- It has more than 1 lakh observations and 39 variables in its raw state which includes many missing, redundant and irrelevant values. All variables not correlated with the objectives of the study were removed. To solve the problem of missing values, we have used the simple and direct approach that consists of the complete elimination of entries that have missing values.

Patients' Details Pre-processing



Doctors' Reviews Pre-processing





WORK DONE

- Patients' details pre-processing.
- Predicted physicians class using classification models.
- Successfully extracted beta-values.
- Collected doctors profile and ran sentiment analysis on them.
- Implemented a user interface.



Results & Discussion

- Collection of two different datasets was done:
 - Dataset1 : After analysing SPARCs dataset we reduced the features.
 - Dataset2 : We created this dataset on our own. It contains 4 symptoms, the disease and the type of doctor required to diagnose the patient.

gastroesophageal reflux disease	pain	pain chest	burning sensation	hyponatremia	gastroenterologist
dehydration	fever	diarrhea	vomiting	hypotension	general physician
cardiomyopathy	shortness of breath	orthopnea	hypokinesia	jugular venous distention	cardiologist
chronic kidney failure	vomiting	orthopnea	hyperkalemia	oliguria	nephrologist

- Dataset 1 and 2 were combined using a description column , *Diagnosis Description*, in the SPARCS dataset.
- All the textual data was changed to numerical values.
- 11 types of doctors.

18 to 29	F	124	Minor	cough	fever	decreased translucency	shortness of breath	pulmonologist
70 or Older	F	123	Minor	cough	wheezing	shortness of breath	chest tightness	Allergist
70 or Older	M	122	Major	wheezing	cough	shortness of breath	chest tightness	Allergist

	Age	Gender	S1	S2	S3	S4	Doctor
0	1	1	35	41	11	1	9
1	4	0	7	19	9	41	0
2	3	0	33	35	4	24	0
3	3	1	29	17	48	20	4
4	0	1	19	10	50	25	10



Logistic Regression

- Probability of a certain class or event existing.
- Python's *Sklearn* library.
- Precision and Accuracy for all the 11 classes for logistic regression model using One vs Rest Classifier.
- Store in a Pickle file.

	Precision	Accuracy
0	0.82	0.86
1	0.82	0.86
2	0.82	0.86
3	0.82	0.86
4	0.83	0.87
5	0.82	0.86
6	0.82	0.86
7	0.87	0.90
8	0.82	0.86
9	0.83	0.87
10	0.83	0.86



- Probabilities of all 11 classes to be able to diagnose the patients for some 3 test cases for logistic regression model using One vs Rest classifier.
- Function called *model.predict_proba()*.

	Allergist	Endocrinologist	General Physician	Cardiologist	Gastroenterologist	Nephrologist	Neurologist	Pediatrician	Psychiatrist	Pulmonologist	Rheumatologist
0	0.905516	0.897717	0.921901	0.913620	0.913422	0.920392	0.897668	0.999573	0.919471	0.912018	0.912724
1	0.912265	0.935740	0.920992	0.939244	0.927671	0.903604	0.899341	0.745244	0.905352	0.933433	0.931883
2	0.881476	0.906254	0.919435	0.912892	0.907052	0.891713	0.883865	0.999990	0.905735	0.908178	0.920634



Random Forest -

- Precision and Accuracy for all the 11 classes for random forest model using One vs Rest Classifier.
- Python's *Sklearn* library.
- Store in a Pickle file.

	Precision	Accuracy
0	0.84	0.86
1	0.82	0.86
2	0.82	0.86
3	0.82	0.86
4	0.82	0.86
5	0.82	0.86
6	0.82	0.86
7	0.89	0.90
8	0.82	0.86
9	0.82	0.86
10	0.82	0.86



- Probabilities of all 11 classes to be able to diagnose the patients for some 3 test cases for random forest model using One vs Rest classifier.

	Allergist	Endocrinologist	General Physician	Cardiologist	Gastroenterologist	Nephrologist	Neurologist	Pediatrician	Psychiatrist	Pulmonologist	Rheumatologist
0	0.931355	0.938899	0.933221	0.814891	0.944916	0.934478	0.883007	1.000000	0.932431	0.731708	0.946216
1	0.969033	0.904769	0.898911	0.966588	0.854188	0.905392	0.972800	0.638555	0.937987	1.000000	0.939697
2	0.979752	0.854084	0.855233	0.887698	0.864322	0.849877	0.882552	1.000000	0.958358	0.979421	0.905205

- Scraping of doctors' profile details from Practo using BeautifulSoup4.

	Name	Specialization	Location	Link
1	Dr. Sheela Chakravarthy	Internal Medicine	bangalore	https://www.practo.com/bangalore/doctor/sheela-chakravarthy-internal-medicine?specialization=Internal Med
2	Dr. Sheetal Kamat	Internal Medicine	bangalore	https://www.practo.com/bangalore/doctor/sheetal-kamat-internal-medicine?specialization=Internal Medicine&
3	Dr. B Rajashekar	General Physician	bangalore	https://www.practo.com/bangalore/doctor/dr-rajashekhar-general-physician?specialization=General Physician&
4	Dr. Raja Selvarajan	Diabetologist	bangalore	https://www.practo.com/bangalore/doctor/dr-raja-selvarajan-diabetologist?specialization=Diabetologist&pract
5	Dr. Sharat Honnatti	general physician	bangalore	https://www.practo.com/bangalore/doctor/dr-sharat-honnatti-general-physician?specialization=&practice_id=
6	Dr. Ashok M N	General Physician	bangalore	https://www.practo.com/bangalore/doctor/dr-ashok-m-n-cardiologist?specialization=General Physician&practi
7	Dr. Tharanath S	General Physician	bangalore	https://www.practo.com/bangalore/doctor/dr-tharanath-s-general-physician?specialization=General Physician&
8	Dr. Pankaj Singhai	Internal Medicine	bangalore	https://www.practo.com/bangalore/doctor/dr-pankaj-singhai-1-internal-medicine?specialization=Internal Medi
9	Dr. Shalini Joshi	Internal Medicine	bangalore	https://www.practo.com/bangalore/doctor/shalini-joshi-internal-medicine?specialization=Internal Medicine&pr
10	Dr. Mohan Badagandi	General Physician	bangalore	https://www.practo.com/bangalore/doctor/dr-mohan-badagandi-diabetologist?specialization=General Physicia
11				



Rating Doctors

- The sentiment of each each review is calculated using **Textblob**.
- The Average sentiment of all the reviews converted to percentage is used as doctors rating.

Sentiment value results for some of the reviews:

↳ Met for fever He is really great During this viral season I recommend him : 0.8

Doctor left in spite of me having appointment when called by the hospital staff he refused to acknowledge deplorable : -0.3

Physician Prediction

Age: Below 17 Years ▾

Gender: ☒ Male ☐ Female

Symptom 1: dizziness ▾

Symptom 2: vomiting ▾

Symptom 3: cough ▾

Symptom 4: fever ▾

Choose a Location: Bangalore ▾

Submit

Age : Below 17 Years , Gender : Male , S1 : dizziness , S2 : vomiting , S3 : cough , S4 : fever , Location : bangalore

Predicted Doctors:

Dr. Baijini C C



Symptom 1: dizziness

Symptom 2: vomiting

Symptom 3: cough

Symptom 4: fever

Choose a Location: Bangalore

Submit

Age : Below 17 Years , Gender : Male , S1 : dizziness , S2 : vomiting , S3 : cough , S4 : fever , Location : bangalore

Predicted Doctors:

Dr. Rajini C C

Dr. Arjun SK

Dr. Lokanadha Reddy

Dr. Bharath S

Dr. Sreenath S Manikanti

Activities Firefox Web Browser Wed 15:29

Dr. Rajini C C - Pediatrician - Book Appointment Online, View Fees, Feedbacks | Practo - Mozilla Firefox


Dr. Rajini C C - Pediatrician X +

https://www.practo.com/bangalore/doctor/dr-rajini-c-c-pediatrician-2/recommend

practo Doctors Book an appointment Chat Chat with top doctors Pharmacy Medicines & health products Diagnostics Book tests & checkups For Providers Security & help Login / Signup

Bangalore Search doctors, clinics, hospitals, etc. Fed up of endless wait? Look for clinic with Prime

Home > Bangalore > Pediatrician > KR Puram > Dr. Rajini C C



Dr. Rajini C C

MBBS, Diploma in Child Health (DCH), DNB - Paediatrics
Pediatrician, 9 Years Experience Overall (6 years as specialist)

Medical Registration Verified

100% (22 votes)

Dr. Rajini C C MBBS, DCH, DNB Paediatrics (Manipal) consultant paediatrician & neonatologist specialized in paediatric - newborn care, vaccination/immunization, [more...](#)

[Share your story](#)

VIDEO CONSULTATION ₹ 400

Today 17 Slots Available Tomorrow 29 Slots Available

Evening (9 slots)

6:30 PM	6:40 PM	6:50 PM	7:00 PM
7:10 PM	7:20 PM	7:30 PM	7:40 PM
7:50 PM			

PICK A TIME SLOT

Tirumala Newborn Kids And Ortho Clinic

5 ★ ₹400 Max 15 mins wait

KR Puram

Info **Stories(11)** Consult Q&A Healthfeed

Patient Stories for Dr. Rajini C C

These are patient's opinions and do not necessarily reflect the doctor's medical capabilities. [Read more](#)

Offer for health condition/treatment



Individual Contribution

- Bharat Sharma - Practo Data Scraping + Data collection + Models + User Interface
- Bharath Simha - Pre-processing + Data collection + RF + Textblob
- D. Praneetha - Data Mapping + Data collection + LR + Beta Values + Pickle file storage

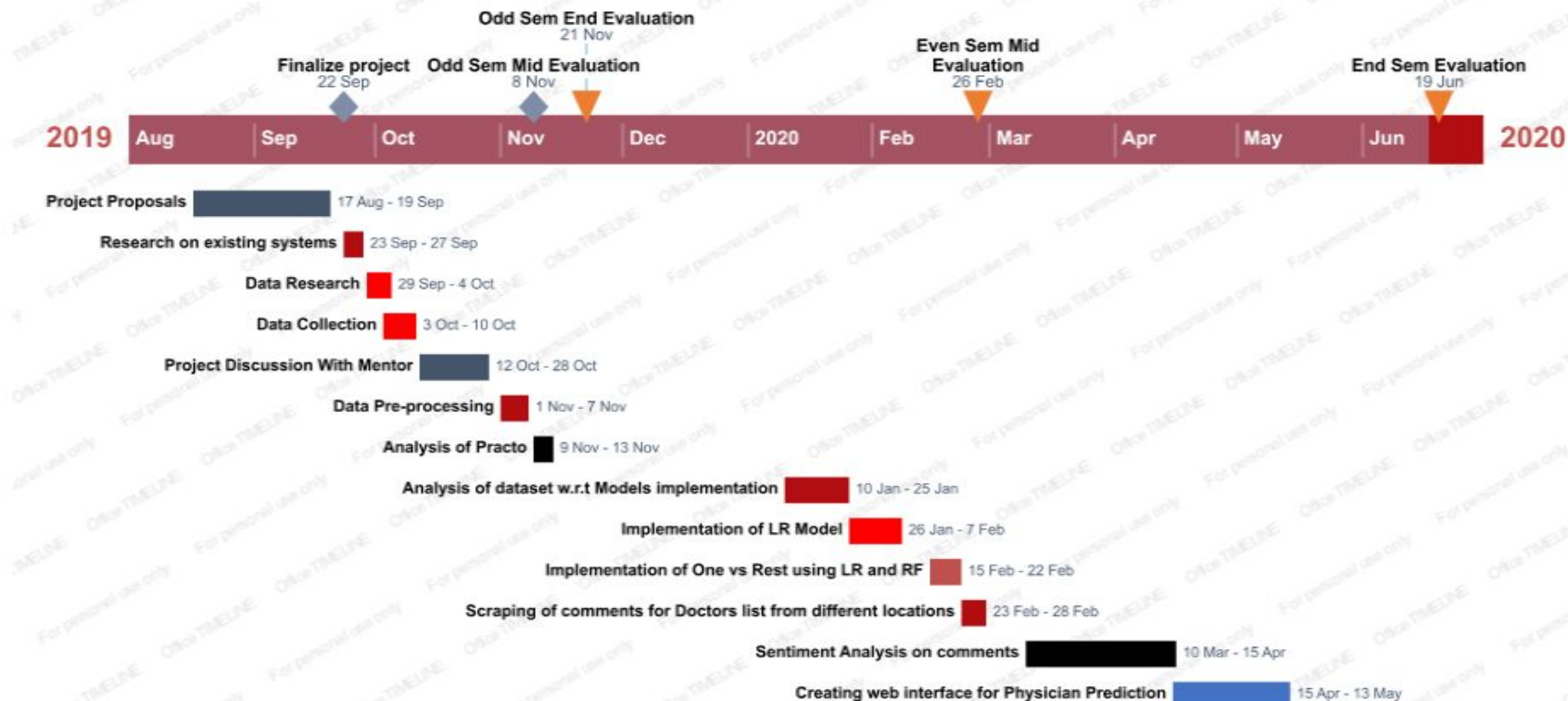


Conclusion & Future Work

Manual collection of the symptoms, diseases and type of doctors' was done. The two data sets were mapped together and were then pre-processed. Scraping is used for collecting the required Practo data of doctors. The objective to predict the doctor's field is successfully achieved. Sentiment analysis is successfully done on the Practo comments for the respective doctors. And finally, a basic user interface is constructed to bring together both the parts of this project i.e. predicting doctor's field and performing sentiment analysis on Practo data.

The future work is to improve the user interface as well as convert this into an android application. Data from other websites along with Practo can be considered as well. More symptoms and doctor fields can be added.

Timeline





References

- [1] Tejaswita P. Vaidya ,Dhanashri Gujar and Rashmi Biyani, Disease Prediction and Doctor Recommendation System.”, International Research Journal of Engineering and Technology, Volume 05, Issue 03, March 2018
- [2] Ramakrishnan, Naren Keller, Benjamin Mirza, Batul Grama, Ananth Karypis, George. (2001). When being Weak is Brave: Privacy in Recommender Systems. Computing Research Repository - CORR.
- [3] Disha Mahajan , Mrudula Phalak , Saniya Pathan,” Prediction System for Diseases and Suggestion of Appropriate Medicines”,International Journal of Innovations Advancement in Computer Science, Volume 6, Issue 11, December 2017
- [4] Guo,Li Jin,Bo Yao,Cuili Yang,Haoyu Huang,Degen Wang,Fei.(2016).Which Doctor to Trust: A Recommender System for Identifying the Right Doctors. Journal of Medical Internet Research. 18. e186. 10.2196/jmir.6015.
- [5] Sahoo,A.K.;Pradhan,C.;Barik,R.K.;Dubey,H.DeepReco: Deep Learning Based Health Recommender System Using Collaborative Filtering. Computation 2019.
- [6] Federico Cesconi, BERT - Next Generation topic detection and sentiment analysis explained to business people, Available: <https://www.linkedin.com/pulse/bert-nextgeneration-topic-detection-sentiment-analysis-cesconi>, Visited: 28th October 2019



**Thank
You!!!**