

Major Project Report

On

**PREDICTION OF PHYSICIANS FOR PATIENT DIAGNOSIS**

Submitted by

**B Muni Rohith - 1MS17IS037**

**Anmol Kapoor – 1MS17IS026**

**Dhanuja H.N - 1MS18IS404**

Under the Guidance of

**Prof.**

**Department of Information Science and Engineering**

**MSRIT**

**Date of Submission:**

**Abstract**

This work provides a predictive model for selecting the most appropriate health care practitioners nearby who can diagnose a patient. First, identification of the doctors who can diagnose a patient is done. Second, probabilities are used to provide a ranking of each physician. Then the top physicians with higher probability to diagnose the disease are picked. For each physician, the top five specialists in each field in the nearby location from Pacto’s user are identified. These specialists are filtered through web scraping on Practo. These specialists are ranked using sentimental analysis on the reviews from the patients who previously visited them. In order to evaluate our results, random forest and logistic regression models are used. Then the construction of a basic user interface to suggest select Practo specialists based on the results is done. In conclusion, it is asserted that all selected specialists are able to diagnose the patient to an extent and that some specialists have a greater ability to diagnose the disease than others.

**Keywords- Open data, Logistic Regression, Random Forest, Practo, Sentiment Analysis, Flask**

**Literature Survey**

In this paper **[1]**, they classify the data according to the requirements and then by applying association rules on it, they predict the diseases. With given symptoms of the patients, predicting disease and recommendation of the prescription of the obtained diseases done. But this prediction is done only for the prescription and not the actual recommendation of the doctors, and also, the prescription is done for very few diseases. The other paper is again a recommendation system **[2]**. The goal is to develop a recommendation system for identifying KOLs for any specific disease with unsupervised learning models. Now, this system can actually make recommendations to pharmaceutical companies and patients but isn’t directly connected to doctors or pharmaceuticals. The last paper **[3]** , gives a proposed Intelligent HRS using Convolutional Neural Network (CNN) deep learning methods. The system finds recommended hospitals by calculating the similarity of patients choices. But here, the execution time is very high. If the similarity between the patients’ choices is very less, the data would lose the vital information, and furthermore, the output doesn’t really deal with actual doctors’ recommendation, but instead, gives out the best possible hospitals according to the patient’s choice history.

**Conclusion of Literature Survey**

As seen in the last section, it has been noticed that a lot of research work and recommendation systems have been made but they’ve been done either with disease prediction or with only hospitals in mind. And the few papers which have actually considered a direct relation between a patient and a doctor have considered only a single doctor to be recommended, instead of a list of ranked doctor fields or doctors. Now, when considering the availability factor of a physician, the recommendation system would have to consider that aspect of a doctor as well. Another thing that was observed was that most of these papers collaborated with various hospitals and used their data for prediction. Now, this really puts a break in bench-marking this project’s results, and hence, brings us to the part where the decision to use multiple models to compare the results is made. Part of the goal is to consider this lack of unavailability of a doctor recommended by a recommendation system and provide trusted alternatives who can diagnose the patient properly.

* **Problem Statement**

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

* **Objectives**
* Determining the type of doctor required to diagnose the patient based on the patient details.
* 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 ranking based on online reviews.
* Mapping the Practo profiles with the predicted type of doctors.