Identification of Disease, Treatment Relationship, Using Machine Learning and Natural Language Processing

Mr. Mohammed Faiz A. K. Patel[1], Mr. Sourav Sajeevan[2], Mr. Akash S. Pansambal[3]

Guide Name: Professor Shital Jadhav[4]

[1]*Student, Department of Computer Science Engineering, G.H Raisoni Institute of Engineering and Management, M.S, India*

[2]*Student, Department of Computer Science Engineering, G.H Raisoni Institute of Engineering and Management, M.S, India*

[3]*Student, Department of Computer Science Engineering, G.H Raisoni Institute of Engineering and Management, M.S, India*

[4]*Assistant Professor, Department of Computer Science Engineering, G.H Raisoni Institute of Business Management, M.S, India*

ABSTRACT

Machine Learning has gained momentum in almost any domain of research and just recently has become a reliable tool in the medical domain. It can be briefly defined as enabling computers to make successful predictions using past experiences. That's why we are implementing software that will take symptoms of a patient as an input in the form of natural language or text. And from machine learning the software will predict his/her disease and help him/her with the proper method of treatment. As now-a-days new diseases are evolving our software will reduce their daily stress. Also, in small villages where doctors are not available 24/7 our software will prove helpful for them and will advise them to take essential steps for medications.

Keyword: - *NLP, Image Processing, Naïve Bayes, Computer science, Artificial Intelligence, Machine Learning, Natural Language processing, disease, disease identification, cure, preventions*

1. INTRODUCTION

Now a day's people are more aware about their health and healthcare. In spite of their busy schedules, they want information regarding to their health for each and everything in a suitable way. People want Fast access to reliable information and in a manner that is suitable to their habits and work-flow. Medical field has grown in a wider to such an extent that information about latest discoveries is published day by day. The proposed system gives more reliable information and classification performances regarding Medline database. Our proposed technique provides the doctors in making better medical decisions. The tool that is built with the techniques such as Natural Language Processing (NLP) and Machine Learning (ML) has capability to find all relevant short text information regarding diseases and treatments. This work presents various Machine Learning (ML) and information for classifying short texts and relation between diseases and treatments.

2. LITERATURE SURVEY

Disease Treatment using machine learning and natural language processing are now-a-days becoming important for every doctors. Its demand in the market is also increasing as it relives the doctors from almost 40-50 % of their stressful work. Also many deadly viruses and diseases are evolving which makes it even more difficult for doctors to treat the patients. For our project we explored few projects that were related to our project they are as following.

"Identify Disease Treatment Relationship in Short Text Using Machine Learning Approach" by Keerrthega. M. C Ms. D. Thenmozhi. It is very useful for everyone as it gives information only of the area of interest. Since, machine learning approaches are widely used in many medical applications like protein-protein interaction, extraction of medical knowledge and health care, they propose a machine learning approach which is supervised that uses several features from medical papers and medical abstracts by doing natural language processing technique for identifying disease treatment relations namely cure, prevent and side effect.

"Identifying Disease-Treatment Relation Using ML and NLP Approach" Machine learning, data science and Medline, Bio text. It provides reliable & efficient medical information in short-text. Now a day's people are more aware about their health and healthcare. In spite of their busy schedules they want information regarding to their health for each and everything in a suitable way. People want fast access to reliable information and in a manner that is suitable to their habits and work-flow. This paper was by Dhamne Kalpesh Mistari Sachin Dahite Sushil Dalvi Suraj R. S. Shirsath.

"ML-Based Approach for Identifying Disease-Treatment Relations in Short Texts" by P. Shirisha K. Shirisha. Machine language and Natural Language Processing. The interests are in line with the tendency of having a personalized medicine that has one in which each patient has its medical care tailored to its needs. This Work provides the foundation for development of technology framework that makes easy to find all the relevant information regarding treatment and diseases.

3. SYSTEM ARCHITECTURE

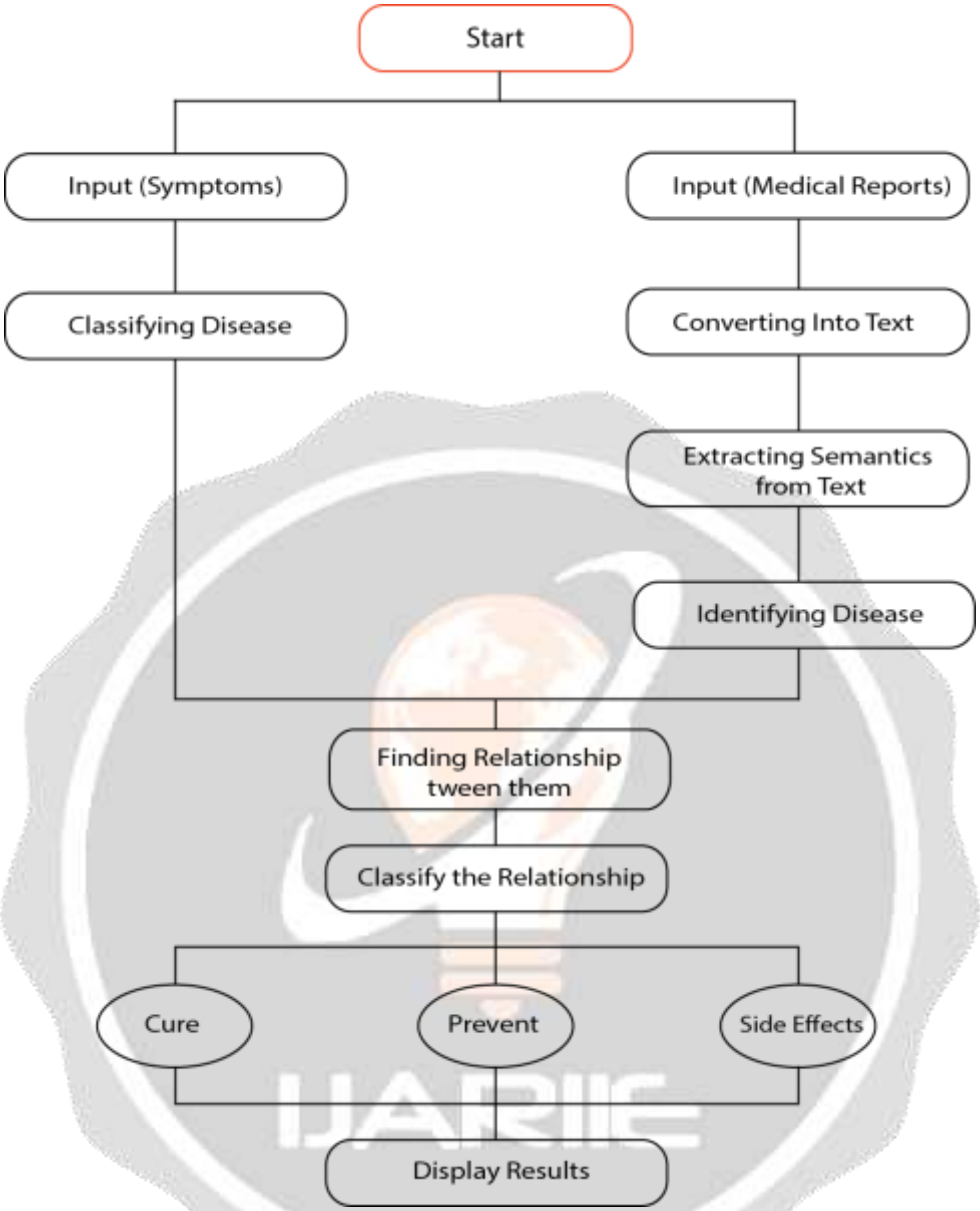
The proposed system consists of taking of an input and showing results regarding to it.

One way is taking inputs in the form of symptoms

- Initially the user will enter symptoms
- On that basis it will classify the disease and will find cure and preventions
- In the end it will show the results.

The other way of input is uploading a picture or document of any type of medical reports.

- It will scan it and will convert it into the text
- After that it will find important semantic keywords and their values
- On that basis it will show results first will identify disease and after that cure and preventions
- Finally it will give user and output.



3.1 DESCRIPTION

Firstly, the software will start. Our software will be using Medline database for the processing. The software will be taking article (report) as an input, then the same report will then be converted into text file which will be used for extracting semantics and then it will be divided into two sections Disease and Treatment. Then software will relate them and give output in three sections which are: - CURE, PREVENT and SIDE EFFECT and will provide a final report.

4. CONCLUSION

This project will help many people to identify disease and treatment in case of an emergency, when they are not able to contact any physician and will also help to find treatment and cure as well. This will be a machine learning model it will get trained by many algorithms and techniques but it will not be perfect as it will have some errors. But surely

it will be more effective for predicting disease. We are working hard to make it as good as possible in an efficient way in order to keep errors minimum. It will be a web application so that it can work on any OS on browsers and also will be responsive.

5. REFERENCES

[1] B. Rosario and M.A. Hearst, "Semantic Relations in Bioscience Text," Proc. 42nd Ann. Meeting on Assoc. for Computational Linguistics, vol. 430.

[2] R. Bunescu and R. Mooney, "A Shortest Path Dependency Kernel for Relation Extraction," Proc. Conf. Human Language Technology and Empirical Methods in Natural Language Processing (HLT/EMNLP), pp. 724-731, 2005.

[3] R. Bunescu, R. Mooney, Y. Weiss, B. Scho¨ lkopf, and J. Platt,"Subsequence Kernels for Relation Extraction," Advances in Neural Information Processing Systems, vol. 18, pp. 171-178, 2006.

[4] A.M. Cohen and W.R. Hersh, and R.T. Bhupatiraju, "FeatureGeneration, Feature Selection, Classifiers, and Conceptual Drift for Biomedical Document Triage," Proc. 13th Text Retrieval Conf.(TREC), 2004.

[5] M. Craven, "Learning to Extract Relations from Medline," Proc.Assoc. for the Advancement of Artificial Intelligence, 1999.