Disease Treatment Analysis for Knowledge based System by using Natural Language Processing

Divya Kere[1], Sagar Manchadi[2], Prof. Savita Sangam[3]

Shivajirao S. Jondhale College of Engineering, Dombivli, Maharashtra, India[1,2,3]

Abstract: The machine learning becomes a very important field in many research domain and just become a very important tool in medical domain field. with the help of machine learning, computer based system is integrated with medical field in order to get a better and a well organized medical treatment care. It shows a machine learning based methodology for making an application that is capable of identifying and improving healthcare information. It selects sentences from published medical papers that shows disease and treatments and finds semantic relation exist between them.

Keywords: Data mining, Electronic healthcare record system, Machine learning, natural language processing.

I. INTRODUCTION

NOW a day's people are takes more care about their health and want to be more secure about of health and healthcare .tools that helps us to manage and keep track of our health are Google health 1 and Microsoft health vault 2 are reasons to get information about health. but this information is very rich in manner but gives less knowledge. data can be obtained in a large no but its difficult to find out relation between disese and treatment .

Here we basically focused on two tasks:
1. Identifying the sentence published in medical paper.
2. Making semantic relation between disease and treatment. This contains cure, prevention and side effects. Knowledge based system is also used here for solving difficult complex problem.

II. LITERATURE SURVEY

In Jason, "Tackling The POOR Assumption Bayes Text Classifier" but it don't give 100% proper output.
1. In T.Mouratis, S.Kotsiantis, " the accuracy off discremitive bayasias classifier in text classification used correctly audit it and give high output.
2. In b.rasio m.a hearst "semantic relation in bioscience text" here for entity reorganization here we use the hidden models. After that natural language processing is converted into biomedical data is mapped in structural form.
3. In l.hunter biomedical language processing what's beyond pubmed sentence processing natural language processing is used for biomedical sentence processing.

III. PROBLEM DEFINITION

Here our system provides technique that the doctors in deciding better medical decision. We take the information from medical sites in the form of medical paper that gives information about treatment as well. but this information are very large and in today life it is very hard to read and extract main points from all the big article so here we are making an application that gives a semantic relation between disease treatment and cure. Basically its act as a virtual doctor.

IV. PROPOSED METHODOLOGY

A. Functionality Requirement

1. log in: In this module, 4 person like users Doctors, Patients, Hospital Staff's, Medical experts to upload all the details of the diseases details and the treatment details. For that purpose they first registered his/her personal information and login into page to upload these information in medical files.

2. Upload files: it is very useful for the many people. They can easily get the information data's from the upload database. In this module, we upload the medical disease information, with the treatment information, etc., each disease have separate id to generate.

3. Search engine: In this module, the search engine is designed to search and collecting information on the databases like World Wide Web. The search engines results are generally presented in a list of results and are referred as SERP (search engine result pages). The information may consist of pages, images, information, large data videos and other different types of files. Some search engines also mine data present in the databases or open directory .web directories, which are maintained and understand only by human editors, search engines also gives information and maintain that real-time information by running on an algorithm on a crawler.
4.verify files: In this module, finally admin verify and check the upload files details. To check that who

uploaded the files and information. What is the purpose of the files and information. And give full authority to those files. And save the file into the database.

V.OPERATING ENVIRONMENT

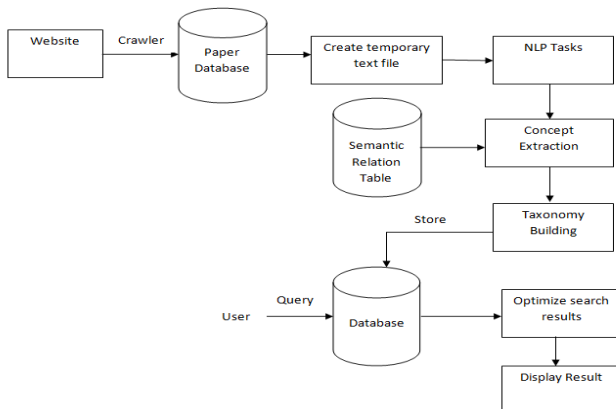|	
System	: Pentium 2.4GHz.
Hard Disk	: 40 GB.
Floppy Drive	: 1.4 Mb.
Monitor	: 16 VGA Colour.
Mouse	: Logitech.
Ram	: 512 Mb.
Operating system	: Windows XP.
Coding Language	: JAVA.
Data Base	: MYSQL.

VI. SYSTEM ARCHITECTURE

Fig.1. Block Diagram of System Architecture

A. Working:
1. Module: pages are downloaded from med sites and converted into zip and then using parser converted into xml and finally into text.
2. Module: after all the conversion irrelevant data get eliminated.
3. Module: we have to collect proper information the data we want from sites and find relation between treatment cure and prevention.
4. Module: after receiving relation we can store data in database.
5. Module: finally we have to analyze results.

VII. CONCLUSION

Here natural language processing is used to get relevant data and to obtain semantic relation we used machine learning algorithms. hence we have seen that all the hectic work is done by machine itself no human intervention is required over here.

REFERENCES

[1]. Russell S, Norvig P, (1994), Artificial Intelligence: A Modern Approach. NewJersey, Prentice-Hall,Inc., ASimon & Schuster Company, Englewood Cliffs, New Jersey 07632, ISBN 0-13-103805-2.Derek Ma, Qiu hau Lin, Tong Zhang, "Mobile Based Text Detection and Translation".

[2]. O. Frunza, D.Inkpen and T. Tran, "Machine learning approach for identifying disease-treatment relations inshore text", IEEE Computer Society, June 2011, pp-no. 1041 -4347.Implementing Optical Character Recognition on the Android Operating System for Business Cards By Sonia Bhaskar, Nicholas Lavassar, Scott Green EE 368 Digital Image Processing.

[3]. L.Hunter,B.Cohen," Perspective Biomedical Language Processing :What's Beyond PubMed? ", Molecular Cell 21, March 2006, pp-no. 589–594.

[4]. H.Yana, Y.Jiangb, J.Zhenge, C.Pengc,Q.Lid ,"A multilayer perceptron-based medical decision support system for heart disease diagnosis", Expert Systems and Applications 30, May 2006,pp-no. 272–281.

[5]. O. Frunza and D. Inkpen, "Textual Information in Predicting Functional Properties of the Genes Proc. Workshop Current Trends in Biomedical Natural Language Processing in conjunction with Assoc. for Computational Linguistics (ACL '08), 2008.

[6]. B. Rosario and M.A. Hearst, "Semantic Relations in Bioscience Text," Proc. 42nd Ann. Meeting on Assoc. for Computational Linguistics, vol. 430, 2004.

[7]. Jeff Pasternack, Don Roth "Extracting Article Text From Web With Maximum Subsequence Segmentation", WWW 2009 MADRID

[8]. Abdur Rehman, Haroon.A.Babri, Mehreen saeed," Feature Extraction Algorithm For Classification Of Text Document", ICCIT 2012

BIOGRAPHIES

Ms. Divya D. Kere is a student of Computer Engineering in Shivajirao S. Jondhale College of Engineering.

Mr. Sagar A.Manchadi is a student of Computer Engineering in Shivajirao S. Jondhale College of Engineering.