

ABSTRACT

To prevent losses in the yield one of the best ways is identification of plant diseases. In general, the studies of plant diseases observe visually the pattern on the plant. To get the accurate treatment for appropriate disease it is very hard to identify the patterns manually. This is a very time-consuming process and should require an expert with lot of experience. In this proposed application both the disease and treatment is being found by inputting various patterns which is physically observed on any kind of plant. For any plant there are 5 levels of disease occurrences: Stem Level, Leaves Level, Seed Level, Lessions Level, Plant Level. All the physical inputs are collected and gathered according to each level of plants and are substituted in the medical directory which has been created. Analysis results of the proposed method are written using Machine Learning approach.

TABLE OF CONTENTS		
CHAPTER No.	TITLE	PAGE No
	ABSTRACT	i
	LIST OF FIGURES	ii
	LIST OF TABLES	iii
1	INTRODUCTION	1
	1.1 AIM OF THE PROJECT	2
	1.2 Methodology	4
	1.3 Significance of Work	4
	1.4 Report Organization	4
2	LITERATURE SURVEY	5
	2.1 Introduction	6
	2.2 Related Work	6
3	DESIGN ASPECTS	8
	3.1 Feasibility Study	9
	3.2 Existing System	10
	3.3 Proposed System	10
	3.4 System Design	11
	3.5 Data flow Diagram	13
	3.6 UML Diagram	15
	3.7 Database Design	22
4	IMPLEMENTATION ASPECTS	27
	4.1 Software Requirement Specification	28
	4.2 Functional Requirement	29
	4.3 Non Functional requirement	30
	4.4 Model Used	21
	4.5 Hardware Specification	34
	4.6 Software Specification	34
	4.7 Theoretical Background	34

	4.8 Software Project Modules	41
5	EXPERIMENTAL RESULTS	43
	5.1 Results	44
	5.2 Testing and test cases	52
	5.3 Testing for unit testing and Validation Testing	55
6	CONCLUSION AND FUTURE WORK	57
	REFERENCES	59
	APPENDICES	60
A	SOURCE CODE	60
B	SCREENSHOTS	70

List Of Figures

3.1	Steps for system design	12
3.2	DFD elements	14
3.3	Dataflow diagram for proposed application	15
3.6.1	Use case diagram	17
3.6.2	Class Diagram	18
3.6.3	Sequence Diagram	19
3.6.4	Collobration Diagram	19
3.6.5	Activity Diagram	20
3.6.6	Component Diagram	21
3.6.7	Deployment Diagram	22
4.1	Software Requirement Analysis	29
4.5	Spiral Model	32
4.7.4	Steps For Java	37

List Of Tables

3.7.3	Natural Disease Table	25
4.6	Hardware Specifications	34

CHAPTER - I

INTRODUCTION

1. INTRODUCTION

People care deeply about their health and want to be, now more than ever, in charge of their health and health care. Life is more hectic than has ever been, the medicine that is practiced today is an Evidence-Based Medicine (here after, EBM) in which medical expertise is not only based on years of practice but on the latest discoveries as well. Tools that can help us manage and better keep track of our health such as Google Health and Microsoft Health Vault are reasons and facts that make people more powerful when it comes to health care knowledge and management. The traditional health care system is also becoming one that embraces the Internet and the electronic world. Electronic Health Records (here after, EHR) are becoming the standard in the healthcare domain. Researches and studies show that the potential benefits of having an EHR system are

- Health information recording and clinical data repositories— immediate access to patient diagnoses, allergies, and lab test results that enable better and time-efficient medical decisions.
- Medication management—rapid access to information regarding potential adverse drug reactions, immunizations, supplies, etc.
- Decision support—the ability to capture and use quality medical data for decisions in the workflow of healthcare
- Obtain treatments that are tailored to specific health needs—rapid access to information that is focused on certain topics.
- First the symptoms provided by the user are processed by a expert system for identifying the diseases. If the rules required for processing the data by the above are not present in the database, then the system automatically calls the machine learning algorithm technique.

1.1 Aim Of The Project

In order to embrace the views that the EHR system has, we need better, faster, and more reliable access to information. In the medical domain, the richest and most used source of information is Medline, a database of extensive life science published articles. All research discoveries come and enter the repository at high rate, making the process of identifying and disseminating reliable information a very difficult task. The work that we present in this paper is focused on two tasks: automatically identifying sentences published in medical abstracts (Medline) as containing or not information about diseases and treatments, and automatically identifying semantic relations that exist between diseases and treatments, as expressed in these texts. The second task is focused on three semantic relations: Cure, Prevent, and Side Effect

The tasks that are addressed here are the foundation of an information technology framework that identifies and disseminates healthcare information. People want fast access to reliable information and in a manner that is suitable to their habits and workflow. Medical care related

Information is a source of power for both healthcare providers and laypeople. Studies reveal that people are searching the web and read medical related information in order to be informed about their health. Show how a new outbreak of the influenza virus can be detected from search engine query data. Our objective for this work is to show what Natural Language Processing (NLP) and Machine Learning (ML) techniques—what representation of information and what classification algorithms—are suitable to use for identifying and classifying relevant medical information in short texts.

We acknowledge the fact that tools capable of identifying reliable information in the medical domain stand as building blocks for a healthcare system that is up-to-date with the latest discoveries. In this research, we focus on diseases and treatment information, and the relation that exists between these two entities. Our interests are in line with the tendency of having a personalized medicine, one in which each patient has its medical care tailored to its needs. It is not enough to read and know only about one study that states that a treatment is beneficial for a certain disease. Healthcare providers

need to be up-to-date with all new discoveries about a certain treatment, in order to identify if it might have side effects for certain types of patients.

We envision the potential and value of the findings of our work as guidelines for the performance of a frame work that is capable to find relevant information about diseases and treatments in a medical domain repository. The results that we obtained show that it is a realistic scenario to use NLP and ML techniques to build a tool, similar to an RSS feed, capable to identify and disseminate textual information related to diseases and treatments. Therefore, this study is aimed at designing and examining various representation techniques in combination with various learning methods to identify and extract biomedical relations from literature.

1.2 Methodology

The present work deals with the concepts of Machine learning algorithms and the development of Web based online expert systems. An expert system follows the methodology of task-based specification and independent problem solving techniques. Here Machine learning technique is considered to find a good matching for the symptoms in the database. The new algorithm mainly focuses on the determination of the diseases affected to the Natural plants. First the symptoms provided by the user are processed by a rule based expert system for identifying the diseases. If the rules required for processing the data by the above are not present in the database, then the system automatically calls the machine learning algorithm technique. As a whole, the system results global solution for recognizing the diseases in Natural plants. And corresponding treatments to the diseases may also be suggested to the users.

1.3 Significance Of The Work

In proposed System, we try to implement ML Approach in data mining to identify the plant diseases. The proposed method tries to create a medical dictionary for the end users in order to predict and identify the diseases very accurately on plant.

1.4 Report Organization

2. LITERATURE SURVEY

2.1 Introduction

Literature survey is the most important step in software development process. Before developing the tool, it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, the next steps are to determine which operating system and language used for developing the tool. Once the programmers start building the tool, the programmers need lot of external support. This support obtained from senior programmers, from book or from websites. Before building the system the above considerations are taken into account for developing the proposed system.

2.2 Related Work

In 2011, an innovative approach was presented[1] to automatically grade the disease on plant leaves. According to that, plant pathologists mainly rely on naked eye prediction and a disease scoring scale to grade the disease. That leads some problems associated with manual grading. This manual grading is not only time consuming but also not feasible. Hence an image processing-based approach to automatically grade the disease spread on plant leaves by employing Fuzzy Logic had been proposed. The results are proved to be accurate and satisfactory in contrast with manual diseases are inevitable