



وزارة التعليم العالي والبحث العلمي

جامعة القادسية

كلية علوم الحاسوب وتكنولوجيا المعلومات

قسم الحاسوب

DIAGNOSIS OF DISEASES USING ASSOCIATION RULE

اعداد الطالبات

نور علي مهدي

اسراء كمال عبد

اشراف

م. م. الاء عبد مسلم عبد علي

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

سبح لله ما في السموات والأرض وهو العزيز الحكيم، له ملك السموات والأرض يحيي ويميت وهو على كل شيء قدير، هو الأول والآخر والظاهر والباطن وهو بكل شيء عليم، هو الذي خلق السموات والأرض في ستة أيام ثم استوى على العرش يعلم ما يلج في الأرض وما يخرج منها وما ينزل من السماء وما يعرج فيها وهو معكم أين ما كنتم والله بما تعملون بصير.

صدق الله العلي العظيم

(سورة الحديد. الآية ١)

كلمة شكر

بسم الله الرحمن الرحيم

"وما توفيقي الا بالله عليه توكلت واليه أنيب"

صدق الله العظيم (سورة هود. الآية ٨٨)

أحمد الله انه وفقنا ومنحنا القوة والشجاعة والصبر على تحمل أعباء هذا البحث.
أتقدم بالشكر الجزيل الى الاستاذة المشرفة الست ألاء عبد مسلم، التي أكن لها كل
التقدير والاحترام التي لم تبخل علينا بتوجيهاتها ونصائحها القيمة ..
أتقدم بالشكر الجزيل والامتنان الكثير الى عائلتي الكريمة خاصة والدي العزيزين..
أشكر كل من ساعدني من قريب او بعيد ، اخص بالذكر الدكتور صالح الجنابي جزاه الله
كل خير.

الاهداء

أليكِ يا منبع الامل الصافي الحنون..
والامل المشرق الذي لا يغيب ضوئه كالشمس والقمر..
اليكِ اهدي عباراتي .. ورسالتي.. وأزكى تحياتي
أمي
اليكِ يا من غمرتني بعطفك وحنانك وزرعت في نفسي حب الخير..
اليكِ اهدي حبي وقلبي ورسالتي وجهدي وعمري..
والدي العزيز
الى شموع اضاءت لي دربي وأخواني واخواتي.
الى وطني العراق العزيز ... ارضاً وشعباً.

SUPERVISOR CERTIFICATION

We certify that this thesis "DIAGNOSIS OF DISEASE USING ASSOATION RULE" was prepared under my supervision at the computer science department to science college as a partial fulfillment of the requirement for the degree of B.s.c of Science in computer science.

SIGNATURE:

NAME:M.S.C. ALAA ABID MUSLAM

DATE:21/5/2017

ABSTRACT

In this project is to collect medical information related to the patient electronically where information is entered and stored in the computer. Then design computer programs are helping to make medical decisions for the patient based on the information provided by the doctor about the situation and symptoms. That is, these programs can make decisions using a set of laws for the analysis of input data and interfaces to the final result .

The goal of the project is to design a system that has a great educational interest as a doctor skilled in his work he can pass his experience and convey his knowledge to future generations If the cooperation with the world of the computer to put the medical program includes salvation experience this doctor and so medical students and young doctors benefit from this experience doctor in any country and at any time they lived.

List Of Content

Chapter 1	Review
1.1	Introduction
1.2	Computer uses in the medical field
1.3	Objective
1.4	Theses layout
Chapter 2	Data mining
2.1	Introduction
2.2	Continuous Innovation
2.3	Data, Information and Knowledge
2.4	Data Warehouses
2.4.1	What can data mining do?
2.4.2	How does data mining work?
2.4.3	What technological infrastructure is required
2.5	Association Rule Mining
2.6	The a-priori algorithm
3.1	The interfaces
4.1	Conclusions
4.2	Future Work
4.3	Languages

CHAPTER ONE

REVIEW

1.1 INTRODUCTION

The history and origin of the word (diagnosis): medical diagnosis date ; The history of medical diagnosis in earnest from the days of ancient Egypt Irnhotep and Hippocrates in ancient Greece, in: traditional Chinese medicine, there are four ways to diagnosis Inspection, hearing, smell, interrogation, and palpation, The book Babylonian medical, diagnostic and guide written by Esagil-Ken-apli (Florida from 1069 to 1046 BC) introduced the experimental logic and rationality in the diagnosis of disease or the use of, the book relied on the analysis of logical rules to combine the symptoms observed on the patient's body with diagnosis and prognosis.

Since the medical or the actual process diagnostics in the formation of diagnosis constitute cognitive process; where the doctor several sources of data used and put the puzzle pieces together to form an impression of the situation and the formation of the diagnosis, the initial impression of the diagnosis can be a broad term that describes a class of diseases, rather than a particular disease or condition, after the initial diagnostic impression.. The doctor follow-up tests and procedures to get more data to support or reject the original diagnosis and will try to narrow the House of Commons to a more specific level, diagnostic procedures are the special things that doctors needed to narrow the diagnostic possibilities.

Medical diagnosis which is to identify any disease or condition that causes signs and symptoms appear on the person, and it is called simply the diagnosis in the context of the implicit medicine, they are each working to identify the disease or the existing situation and called the result to be reached through this process, “diagnosis1’ Based diagnosis is always due to information from the history and physical examination, but often may include one or more of the diagnostic procedures, such as diagnostic tests, which are also ‘in the process.

Diagnosis is often difficult, because many of the signs and symptoms are non-specific, for example, redness of the skin (flare), in itself, is a sign of many disorders, Where to find help in the differential diagnosis so, where

they are compared to several possible explanations and contradiction with the case, and this has the potential of different symptoms of the disease, followed by tests to distinguish between patterns, Sometimes make the diagnosis easily through the sign or symptom (or a group of several) be clear and come back together to a single cause, get a diagnosis is one of the main things that drive us to visit a doctor, and so from the standpoint of statistics, has involved classification also to conduct diagnostic tests.[2]

1.2 COMPUTER USES IN THE MEDICAL FIELD

Computer and Medicine: Has become in this day of computer mail considerable importance in all areas, The ability of the computer boundless data storage, and speed, which does not trade in processed, reformatted, enabled him deserved to become a key element in all of science, became so plays an active role in various fields of life, and different aspect, Which concerns us to talk about this project is the role of the computer in the medical field and the delivery of health care, and this is what Snaakcch in this research, God willing. Since the computer was invented, which plays an active role in hospitals, it is used to store patient data and medical records, print reports, and other familiar things. But As the life and evolution, and the start of digital time plays a Machine a strong role in easing the burden on humans, strong ambition grew up in the computer plays far beyond this role, to be a partner to the doctor in determining treatment, and diagnosis of the disease, and even in a surgical procedure too.

Where The diagnosis is a collection of medical information on the patient electronically where the information is entered and stored in the computer. Where are designing computer programs are helping to make medical Alorteurat for the patient based on the information provided by the doctor about the status of that and symptoms In other words, programs can make decisions using a set of laws for the analysis of input data and interfaces to the final result (such as software used to diagnose based on symptoms) These are called Expert system software specialist system.[4]

1.3 OBJECTIVE

The objective will be to System has the advantage of designing a great learning as a doctor skilled in his work he can pass his experience and convey his knowledge to future generations if they cooperate with the computer scientist to develop a medical program includes salvation experience this

doctor and that medical students benefit and junior doctors from experience that a doctor in any country they were in anytime lived Has already been used computerized diagnostic method in some cases in hospitals, and even exceeded to predict what might be construed a situation in the future, for example: in the heart department in Christaana hospital in NewYork in the United States, has been to predict what will accrue to the health condition of the patient based on ECG analysis by computer and thus determine the treatment plan for the patient .The development tool used will be Microsoft Visual Basic 6.0, and emphasis will be on the software.[6]

1.4 THESIS LAYOUT

This thesis shows the design of diagnosis of diseases using association rules system by obtaining the select item and applies the association rules algorithm. The system consist of three main stages: initial association rules, read item, and finally the implement algorithm stage.

The thesis consists of four chapters

Chapter one this chapter gives a brief introduction ‘and the related works done.

Chapter two it describe the data mining and the definition. It also gives their knowledge discovery.

Chapter three system design and implementation results: it illustrates the proposed system (diagnosis of diseases using association rules) and its implementation results and discussion.

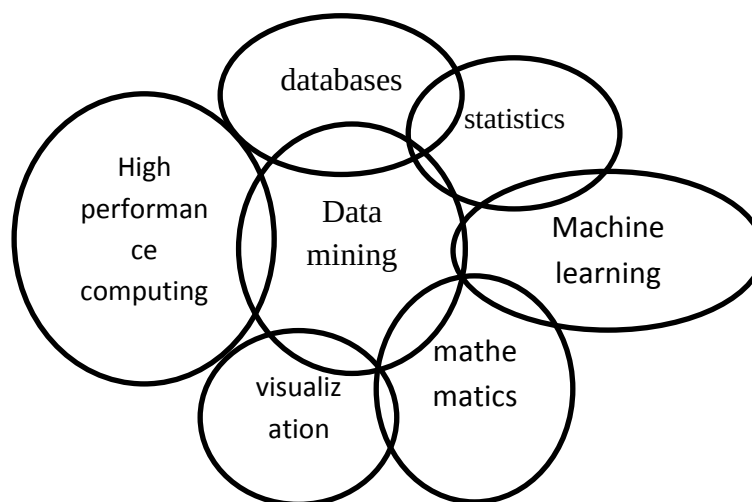
Chapter four contains the conclusion and gives some suggestion for future works.[1]

CHAPTER TWO

DATA MINING

2.1 INTRODUCTION

Data Mining Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.[7]



2.2 CONTINUOUS INNOVATION

Although data mining is a relatively new term, the technology is not. Companies have used powerful computers to sift through volumes of supermarket scanner data and analyze market research reports for years.

Flow ever, continuous innovations in computer processing power, disk storage, and statistical software are dramatically increasing the accuracy of analysis while driving down the cost.

For example, one Midwest grocery chain used the data mining capacity of Oracle software to analyze local buying patterns. They discovered that when men bought diapers on Thursdays and Saturdays, they also tended to buy beer. Further analysis showed that these shoppers typically did their weekly grocery shopping on Saturdays. On Thursdays, however, they only bought a few items. The retailer concluded that they purchased the beer to have it available for the upcoming weekend. The grocery chain could use this newly discovered information in various ways to increase revenue. For example, they could move the beer display closer to the diaper display. And, they could make sure beer and diapers were sold at full price on Thursdays.[2]

2.3 DATA, INFORMATION, AND KNOWLEDGE

Data are any facts, numbers, or text that can be processed by a computer. Today, organizations are accumulating vast and growing amounts of data in different formats and different databases. This includes:

- operational or transactional data such as, sales, cost, inventory, payroll, and accounting,
- nonoperational data, such as industry sales, forecast data, and macro-economic data,
- meta data - data about the data itself, such as logical database design or data dictionary definitions.[3]

INFORMATION

The patterns, associations, or relationships among all this data can provide information. For example, analysis of retail point of sale transaction data can yield information on which products are selling and when.[1]

KNOWLEDGE

Information can be converted into knowledge about historical patterns and future trends. For example, summary information on retail supermarket sales can be analyzed in light of promotional efforts to provide knowledge of consumer buying behavior. Thus, a manufacturer or retailer could determine which items are most susceptible to promotional efforts.[1]

2.4 DATA WAREHOUSES

Dramatic advances in data capture, processing power, data transmission, and storage capabilities are enabling organizations to integrate their various databases into data warehouses.

Data warehousing is defined as a process of centralized data management and retrieval. Data warehousing, like data mining, is a relatively new term although the concept itself has been around for years. Data warehousing represents an ideal vision of maintaining a central repository of all organizational data. Centralization of data is needed to maximize user access and analysis.

Dramatic technological advances are making this vision a reality for many companies. And, equally dramatic advances in data analysis software are allowing users to access this data freely. The data analysis software is what support data mining.[5]

2.4.1 WHAT CAN DATA MINING DO?

Data mining is primarily used today by companies with a strong consumer focus - retail, financial, communication, and marketing organizations. It enables these companies to determine relationships among “internal” factors such as price, product positioning, or staff skills, and “external” factors such as economic indicators, competition, and customer demographics. And, it enables them to determine the impact on sales, customer satisfaction, and corporate profits. Finally, it enables them to “drill down” into summary information to view detail transactional data.

With data mining, a retailer could use point-of-sale records of customer purchases to send targeted promotions based on an individual’s purchase history. By mining demographic data from comment or warranty cards, the retailer could develop products and promotions to appeal to specific customer segments.

For example, Blockbuster Entertainment mines its video rental history database to recommend rentals to individual customers. American Express can suggest products to its cardholders based on analysis of their monthly expenditures.

Wal-Mart is pioneering massive data mining to transform its supplier relationships. Wal-Mart captures point-of-sale transactions from over 2,900 stores in 6 countries and continuously transmits this data to its massive 7.5

terabyte Teradata data warehouse. Wal-Mart allows more than 3500 suppliers, to access data on their products and performs data analyses. These suppliers use this data to identify customer buying patterns at the store display level. They use this information to manage local store inventory and identify new merchandising opportunities. In 1995, Wal-Mart computers processed over 1 million complex data queries.

The National Basketball Association (NBA) is exploring a data mining application that can be used in conjunction with image recordings of basketball games. The Advanced Scout software analyzes the movements of players to help coaches orchestrate plays and strategies. For example, an analysis of the play-by-play sheet of the game played between the New York Knicks and the Cleveland Cavaliers on January 6, 1995 reveals that when Mark Price played the Guard position, John Williams attempted four jump shots and made each one! Advanced Scout not only finds this pattern, but explains that it is interesting because it differs considerably from the average shooting percentage of 49.30% for the Cavaliers during that game.

By using the NBA universal clock a coach can automatically bring up the video clips showing each of the jump shots attempted by Williams with Price on the floor, without needing to comb through hours of video footage.

Those clips show a very successful pick-and-roll play in which Price draws the Knick's defense and then finds Williams for an open jump shot.[2]

2.4.2 HOW DOES DATA MINING WORK?

While large-scale information technology has been evolving separate transaction and analytical systems, data mining provides the link between the two. Data mining software analyzes relationships and patterns in stored transaction data based on open-ended user queries. Several types of analytical software are available: statistical, machine learning, and neural networks. Generally any of four types of relationships are sought:

- **Classes:** Stored data is used to locate data in predetermined groups. For example, a restaurant chain could mine customer purchase data to determine when customers visit and what they typically order. This information could be used to increase traffic by having daily specials.
- **Clusters:** Data items are grouped according to logical relationships or consumer preferences. For example, data can be mined to identify market segments or consumer affinities.

- Associations: Data can be mined to identify associations. The beer diaper example is an example of associative mining.
- Sequential patterns: Data is mined to anticipate behavior patterns and trends. For example, an outdoor equipment retailer could predict the likelihood of a backpack being purchased based on a consumer's purchase of sleeping bags and hiking shoes.[6]

DATA MINING CONSISTS OF FIVE MAJOR ELEMENTS:

- Extract, transform, and load transaction data onto the data warehouse system.
- Store and manage the data in a multidimensional database system.
- Provide data access to business analysts and information technology professionals.
- Analyze the data by application software.
- Present the data in a useful format, such as a graph or table.[5]

DIFFERENT LEVELS OF ANALYSIS ARE AVAILABLE:

- Artificial neural networks: Non-linear predictive models that learn through training and resemble biological neural networks in structure.
- Genetic algorithms: Optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of natural evolution.
- Decision trees: Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). CART and CHAID are decision tree techniques used for classification of a dataset. They provide a set of rules that you can apply to a new (unclassified) dataset to predict which records will have a given outcome. CART segments a dataset by creating 2-way splits while CHAID segments using chi square tests to create multi-way splits. CART typically requires less data preparation than CHAID.
- Nearest neighbor method: A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar

to it in a historical dataset (where $k \geq 1$). Sometimes called the k-nearest neighbor technique. Rule induction: The extraction of useful if-then rules from data based on statistical significance.

- Data visualization: The visual interpretation of complex relationships in multidimensional data. Graphics tools are used to illustrate data relationships.[7]

2.4.3 WHAT TECHNOLOGICAL INFRASTRUCTURE IS REQUIRED?

Today, data mining applications are available on all size systems for mainframe, client/server and PC platforms. System prices range from several thousand dollars for the smallest applications up to \$1 million a terabyte for the largest. Enterprise-wide applications generally range in size from 10 gigabytes to over 11 terabytes. NCR has the capacity to deliver applications exceeding 100 terabytes. There are two critical technological drivers:

- Size of the database: the more data being processed and maintained, the more powerful the system required.
- Query complexity: the more complex the queries and the greater the number of queries being processed, the more powerful the system required.

Relational database storage and management technology is adequate for many data mining applications less than 50 gigabytes. However, this infrastructure needs to be significantly enhanced to support larger applications. Some vendors have added extensive indexing capabilities to improve query performance. Others use new hardware architectures such as Massively Parallel Processors (MPP) to achieve order-of-magnitude improvements in query time. For example, MPP systems from NCR link hundreds of high-speed Pentium processors to achieve performance levels exceeding those of the largest supercomputers.[4]

2.5 ASSOCIATION RULE MINING

Association rule learning is a popular and well researched method for discovering interesting relations between variables in large databases. It is Mining intended to identify strong rules discovered in databases using different measures of interestingness.(1] Based on the concept of strong rules, Rakesh Agrawal et al.[2] introduced association rules for discovering regularities between products ‘n large-scale transaction data recorded by

point-of-sale(P05) systems in supermarkets. For example, the rule found in the sales data of a supermarket would indicate that if a customer buys onions and potatoes together, they are likely to also buy hamburger meat. Such information can be used as the basis for decisions about marketing activities such as, e.g., promotional pricing or product placements. In addition to the above example from market basket analysis association rules are employed

today in many application areas including Web usage mining, intrusion detection, Continuous production, and bioinformatics. In contrast with sequence mining, association rule learning typically does not consider the order of items either within a transaction or across transactions.

Association Rule Mining searches for interesting relationships among items in a given data set.[1]

THE MAIN EQUATION FOR FINDING ASSOCIATION RULES IN SUPER MARKETS IS:

“Which items are frequently purchased together by my customer?”

An association rule is an implication of the form: $X \rightarrow Y$.

Where X and Y are extracted from set of items.

The support and confidence of an association rule are calculated by the following equations:

Support = the number of transactions that contains X and Y / total number of transactions.

Confidence = the number of transactions that contains X and Y / the number of transactions that contains X.[3]

2.7 THE APRIORI ALGORITHM

A-priori is an influential algorithm for mining frequent item sets for Boolean association rules.

The name is based on the fact that the algorithm uses prior knowledge of frequent item set properties.

A-priori employs an iterative approach known as a level-wise search, where K-item sets are used to explore (K+1)-item sets.

A-priori property contains two-sets process consisting of joins and prune

actions.[2]

Here is example illustrate the A-priori algorithm

WE SUPPOSE THE MINIMUM SUP-COUNT IS 2

D	
TID	list of items
T100	I1,I2,I5
T200	I2,I4
T300	I2,I3
T400	I1,I2,I4
T500	I1,I3
T600	I2,I3
T700	I1,I3
T800	I1,I2,I3,I5
T900	I1,I2,I3
ALL ELECTRONICS BRANCH	

SCAN D FOR COUNT OF EACH CANDIDATE →

C1		→	L1	
ITEMSET	SUP-COUNT		ITEMSET	SUP-COUNT
I1	6		I1	6
I2	7		I2	7
I3	6		I3	6
I4	2		I4	2
I5	2		I5	2

GENERATE C2 AND SCAN D FOR COUNT OF EACH CANDIDATE

C2			L2	
ITEMSET	SUP-COUNT		ITEMSET	SUP-COUNT
I1,I2	4	→ COMPARE AND USING PRUNE ACTION	I1,I2	4
I1,I3	4		I1,I3	4
I1,I4	1		I1,I5	2
I1,I5	2		I2,I3	4
I2,I3	4		I2,I4	2
I2,I4	2		I2,I5	2
I2,I5	2			
I3,I4	0			
I3,I5	1			
I4,I5	0			

GENERATE C3 AND SCAN D FOR COUNT OF EACH CANDIDATE →

C3			L3	
ITEMSET	SUP-COUNT		ITEMSET	SUP-COUNT
I1,I2,I3	2	→	I1,I2,I3	2
I1,I2,I5	2		I1,I2,I5	2

Supposes the data contain the frequent itemset

$$L = \{I1, I2, I5\}.$$

What are the association rules that can be generated from L ?

The nonempty subsets of L are

$$\{I1, I2\}, \{I1, I5\}, \{I2, I5\}, \{I1\}, \{I2\} \text{ and } \{I5\}.$$

The resulting association rules are as follows.

$$I1 \text{ and } I2 \rightarrow I5 \quad \text{Confidence} = 2/4 = 50\%$$

$$I1 \text{ and } I5 \rightarrow I2 \quad \text{Confidence} = 2/2 = 100\%$$

$$I2 \text{ and } I5 \rightarrow I1 \quad \text{Confidence} = 2/2 = 100\%$$

$$I1 \rightarrow I2 \text{ and } I5 \quad \text{Confidence} = 2/6 = 33\%$$

$$I2 \rightarrow I1 \text{ and } I5 \quad \text{Confidence} = 2/7 = 29\%$$

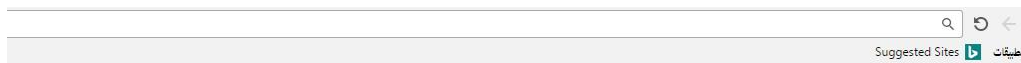
$$I5 \rightarrow I1 \text{ and } I2 \quad \text{Confidence} = 2/2 = 100\%$$

CHAPTER THREE

System design and implementation

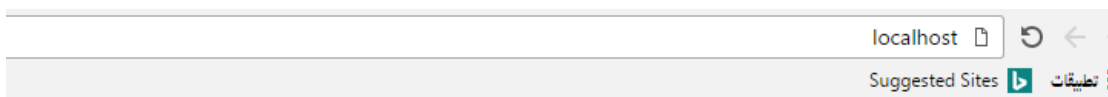
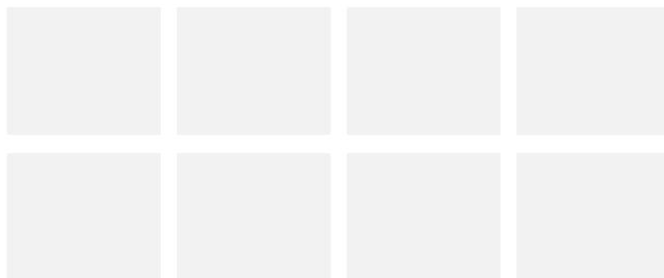
In this chapter we review the interfaces main application that was built and how they are accessible to the data base..

1-from the main browser we enter the local server.



Google

ابحث عن Google أو اكتب عنوان URL



2-from the local server enter the data base.

localhost (C
Suggested Sir

The AppServ Open Project - 8.4.0 for Windows
Now you running on PHP 5.6.22

 **phpMyAdmin Database Manager Version 4.6.2**
 **PHP Information Version**

About AppServ Version 8.4.0 for Windows

AppServ is a merging open source software installer package for Windows includes :

- **Apache Web Server** Version **2.4.20**
- **PHP Script Language** Version **5.6.22 & 7.0.7**
- **MySQL Database** Version **5.7.13**
- **phpMyAdmin Database Manager** Version **4.6.2**
- [ChangeLog](#)
- [README](#)
- [AUTHORS](#)
- [COPYING](#)

Official Site : <http://www.AppServ.org>

Hosting by : <http://www.AppServHosting.com>

3-in this interface there is a user name and password entered by the administrator of application or use.



Welcome to phpMyAdmin

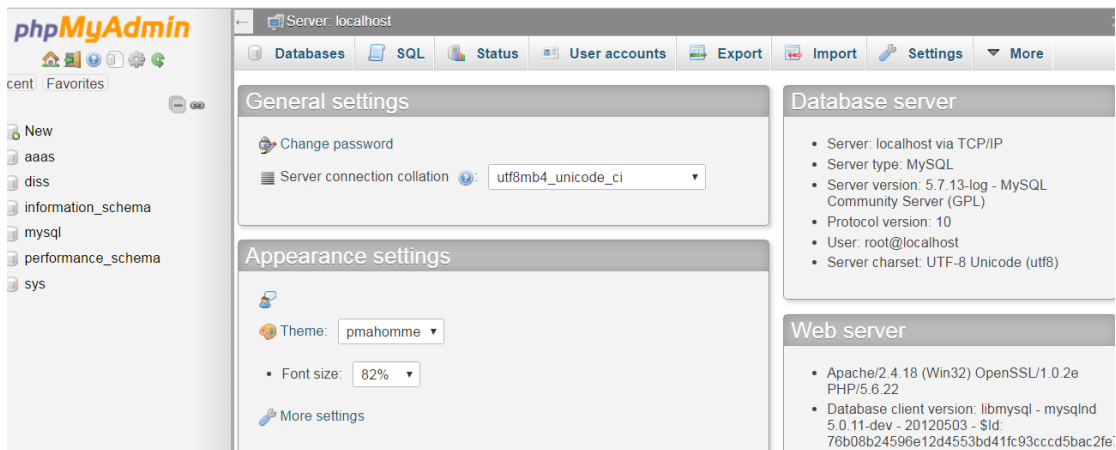
Log in 

Username:

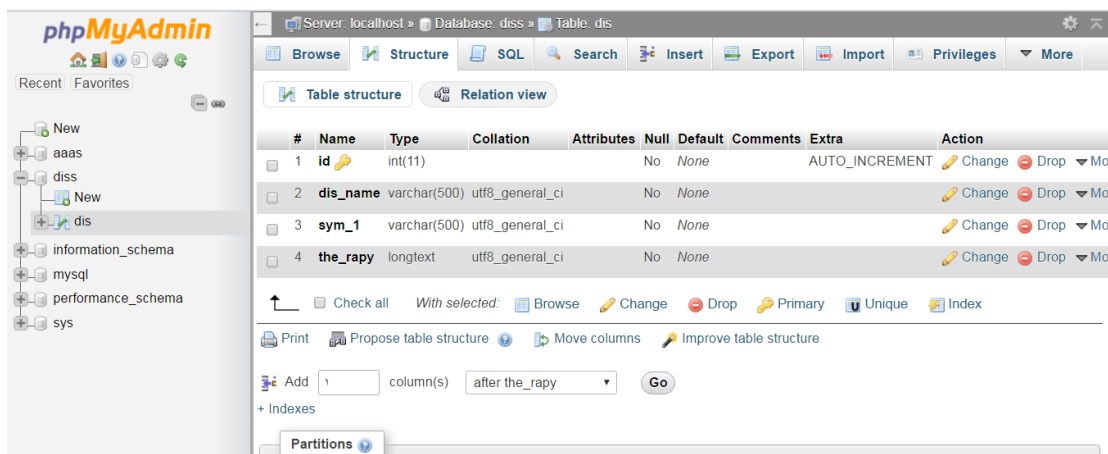
Password:

Go

4-this interface contain the data base server, web server, sitting...etc.



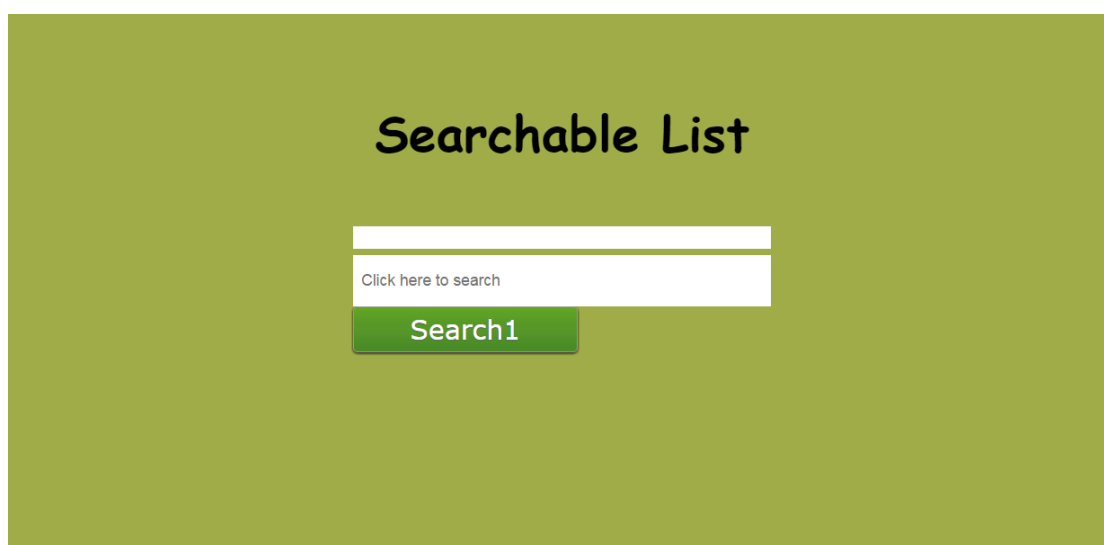
5-when we enter the table dis (shortcut for the word disease)which contain four rows :the ID, dis name, symptoms and the rapy.



6-in this interface we add the diseases, the methods of treatment and prevention.

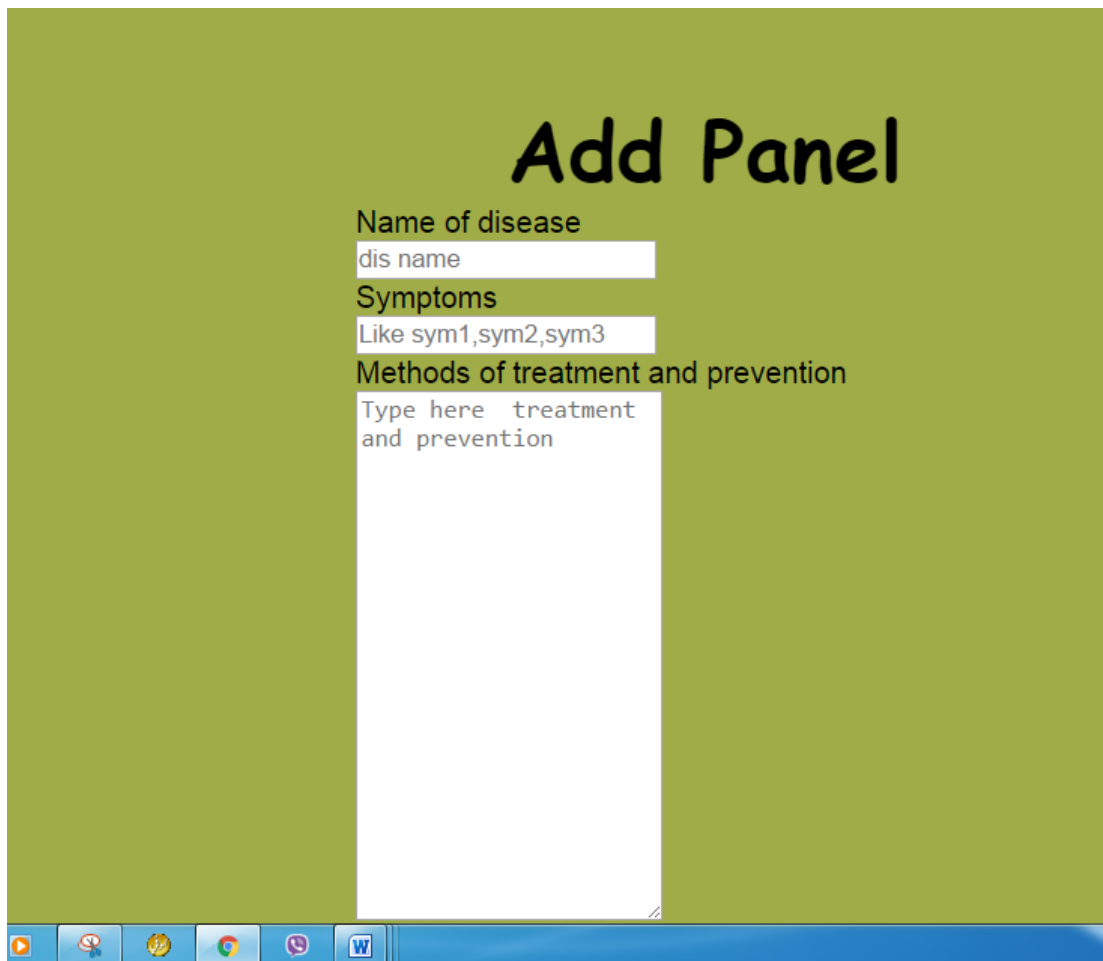
Column	Type	Function	Null	Value
id	int(11)			
dis_name	varchar(500)			
sym_1	varchar(500)			
the_rapy	longtext			

7- Searchable List (index .php): Through the interface search we can find the name of the disease and methods of treatment and prevention by writing the symptoms of the disease in place of the input and then press the key to search as shown in the picture below.



The image shows a web interface with a solid olive-green background. At the top center, the text "Searchable List" is displayed in a large, black, sans-serif font. Below this title, there is a search form. The form consists of a white rectangular input field with a thin green border. Inside the input field, the text "Click here to search" is written in a small, grey, sans-serif font. Directly beneath the input field is a green rectangular button with a thin green border. The button contains the text "Search1" in a white, sans-serif font.

8- Add Panel (add .php) :interface of adding data to data base .



The image shows a web browser window with a green background. The title 'Add Panel' is displayed in large black font. Below the title, there are three input fields:

- Name of disease**: A text input field containing the placeholder text 'dis name'.
- Symptoms**: A text input field containing the placeholder text 'Like sym1,sym2,sym3'.
- Methods of treatment and prevention**: A larger text input field containing the placeholder text 'Type here treatment and prevention'.

The browser's taskbar is visible at the bottom, showing icons for various applications including a media player, a folder, a clock, a web browser, a messaging app, and a word processor.

9- Searchable Results (search .php): After search in the main interface for the symptoms of the disease and pressure on the key research showing us interface declaration.



CHAPTER FOUR

Conclusions and Future Work

4.1 conclusions

- . It is unable to predict the future of the disease by providing part of the symptoms that are of a resident weight through a number of people .**
- . Whenever increasing number of persons or samples increases diagnostic accuracy .**

4.2 Future Work

- 1.The use of neural networks to improve the results of rules of engagement .**
- 2.Use fuzzy logic to find the percentage of the disease.**
- 3. Find the true value of each disease rather than binary.**

PHP

PHP is a server-side scripting language designed primarily for web development but also used as a general-purpose programming language. Originally created by Rasmus Lerdorf in 1994, the PHP reference implementation is now produced by The PHP Development Team.

PHP originally stood for *Personal Home Page*, but it now stands for the recursive acronym *PHP: Hypertext Preprocessor*.

PHP code may be embedded into HTML or HTML5 markup, or it can be used in combination with various web template systems, web content management systems and web frameworks.

PHP code is usually processed by a PHP interpreter implemented as a module in the web server or as a Common Gateway Interface (CGI) executable. The web server software combines the results of the interpreted and executed PHP code, which may be any type of data, including images, with the generated web page. PHP code may also be executed with a command-line interface (CLI) and can be used to implement standalone graphical applications.

The standard PHP interpreter, powered by the Zend Engine, is free software released under the PHP License.

PHP has been widely ported and can be deployed on most web servers on almost every operating system and platform, free of charge.

The PHP language evolved without a written formal specification or standard until 2014, leaving the canonical PHP interpreter as a *de facto* standard. Since 2014 work has gone on to create a formal PHP specification.

HTML

HTML is a computer language devised to allow website creation. These websites can then be viewed by anyone else connected to the Internet. It is relatively easy to learn, with the basics being accessible to most people in one sitting; and quite powerful in what it allows you to create. It is constantly undergoing revision and evolution to meet the demands and requirements of the growing Internet audience under the direction of the » W3C, the organization charged with designing and maintaining the language.

The definition of HTML is Hypertext Markup Language.

- *Hypertext* is the method by which you move around on the web by clicking on special text called hyperlinks which bring you to the next page. The fact that it is *hyper* just means it is not linear i.e. you can go to any place on the Internet whenever you want by clicking on links , there is no set order to do things in.
- *Markup* is what HTML tags do to the text inside them. They mark it as a certain type of text (*italicized* text, for example).
- HTML is a *Language*, as it has code-words and syntax like any other language.

How does it work?

HTML consists of a series of short codes typed into a text-file by the site author , these are the tags. The text is then saved as a html file, and viewed through a browser, like *Internet Explorer* or *Netscape Navigator*. This browser reads the file and translates the text into a visible form, hopefully rendering the page as the author had intended. Writing your own HTML entails using tags correctly to create your vision. You can use anything from a rudimentary text-editor to a powerful graphical editor to create HTML pages.

What are the tags up to?

The tags are what separate normal text from HTML code. You might know them as the words between the <angle-brackets>. They allow all the cool stuff like images and tables and stuff, just by telling your browser what to render on the page. Different tags will perform different functions. The tags themselves don't appear when you view your page through a browser, but their effects do. The simplest tags do nothing more than apply formatting to some text, like this:

These words will be bold, and these will not.

In the example above, the tags were wrapped around some text, and their effect will be that the contained text will be bolded when viewed through an ordinary web browser.

If you want to see a list of a load of tags to see what's ahead of you, look at this tag reference. Learning the tags themselves is dealt with in the next section of this website, My First Site.

Is this going to take long?

Well, it depends on what you want from it. *Knowing* HTML will take only a few days of reading and learning the codes for what you want. You can have the basics down in an hour. Once you know the tags you can create HTML pages.

However, *using* HTML and designing good websites is a different story, which is why I try to do more than just teach you code here at HTML Source , I like to add in as much *advice* as possible too. Good website design is half skill and half talent, I reckon. Learning techniques and correct use of your tag knowledge will improve your work immensely, and a good understanding of general design and the audience you're trying to reach will improve your website's chances of success. Luckily, these things can be researched and understood, as long as you're willing to work at it so you can output better websites.

The range of skills you will learn as a result of running your own website is impressive. You'll learn about aspects of graphic design, typography and computer programming. Your efficiency with computers in general increases. You'll also learn about promotion and your writing will probably improve too, as you adapt to write for certain audiences.

SQL

SQL is a standard language for storing, manipulating and retrieving data in databases.

Our SQL tutorial will teach you how to use SQL in: MySQL, SQL Server, MS Access, Oracle, Sybase, Informix, Postgres, and other database systems.

Examples in Each Chapter

With our online SQL editor, you can edit the SQL statements, and click on a button to view the result.

Example

SELECT * FROM Customers;

References

1. Fayyad, U.G., Piatetsky A, Shapiro, P., Smyth, & R. Uthurusamy, Advance in knowledge discovery & Data Mining. Cambridge, MA (The AAAI Press/The MIT Press), 1996.
2. Siraj, F (2007). Data Mining Models and Tasks, data mining lecture.
3. Dunham, M. H. (2003). Data Mining; Introductory and Advanced Topics, Prentice Hall.
4. Data Mining, <http://www.uqucs.com/vb/showthread.php?t=3308> Access data: 19/11/2010.
5. Data Mining, <http://www.alhasebat.net/vb/showthread.php?t=6494> Access data: 19/11/2010.
6. Adriaan, P. and D. Zanting. Data Mining. Addison-Wesley :Harlow, England, 1996.
7. Al-Hamami, Alaa H, abass F Kader, Hussein K. AL-Khefaji, "Design and Implementation of Generate of large Dense, or sparse Database to test Association rules Miners" (selected researchers papers), Scientific journal of Federation of Arab Scientific Research Council, 2002.