Original Article

A novel AIDS/HIV intelligent medical consulting system based on expert systems

Alireza Pour Ebrahimi, Abbas Toloui Ashlaghi[1], Maryam Mahdavy Rad

Department of Information Technology Management, Islamic Azad University, Electronic Campus, Tehran, [1]Department of Industrial Management, Science and Research Branch, Tehran, Iran

ABSTRACT

Background: The purpose of this paper is to propose a novel intelligent model for AIDS/HIV data based on expert system and using it for developing an intelligent medical consulting system for AIDS/HIV. **Materials and Methods:** In this descriptive research, 752 frequently asked questions (FAQs) about AIDS/HIV are gathered from numerous websites about this disease. To perform the data mining and extracting the intelligent model, the 6 stages of Crisp method has been completed for FAQs. The 6 stages include: Business understanding, data understanding, data preparation, modelling, evaluation and deployment. C5.0 Tree classification algorithm is used for modelling. Also, rational unified process (RUP) is used to develop the web-based medical consulting software. Stages of RUP are as follows: Inception, elaboration, construction and transition. The intelligent developed model has been used in the infrastructure of the software and based on client's inquiry and keywords related FAQs are displayed to the client, according to the rank. FAQs' ranks are gradually determined considering clients reading it. Based on displayed FAQs, test and entertainment links are also displayed. **Result:** The accuracy of the AIDS/HIV intelligent web-based medical consulting system is estimated to be 78.76%. **Conclusion:** AIDS/HIV medical consulting systems have been developed using intelligent infrastructure. Being equipped with an intelligent model, providing consulting services on systematic textual data and providing side services based on client's activities causes the implemented system to be unique. The research has been approved by Iranian Ministry of Health and Medical Education for being practical.

Key words: AIDS/HIV, data mining, intelligent system, medical informatics, software engineering, text mining

INTRODUCTION

This paper is about designing a novel AIDS/HIV intelligent medical consulting system based on expert systems. Intelligent medical systems are those which use the knowledge of medical experts in their infrastructure and serve their clients on this basis.[1] Medical fields are exposed to ever-growing information which can be used to enhance healthcare and treatment quality. On the other hand, data mining plays an important role in medical consulting, diagnosing, prevention and treatment methods in regard to patient's problems. The necessity of utilizing intelligent web-based medical consulting systems can be well understood, considering current non-intelligent systems having problems in providing fast and online services to the clients. Since, the answer should go through a long and time-consuming procedure consisting of various methods including revisiting the system, vocal calls and sending email in daily intervals. Furthermore, web-based intelligent medical consulting is up-to-date, which is essential for clients to put their faith and propagate using the system.

Address for correspondence: Mrs. Maryam Mahdavy Rad, Islamic Azad University, Electronic Campus, Tehran, Iran.
E-mail: mmahdavyrad@gmail.com

Access this article online	
Quick Response Code:	Website: www.jehp.net
	DOI: 10.4103/2277-9531.119041

This article may be cited as: Ebrahimi AP, Ashlaghi AT, Rad MM. A novel AIDS/HIV intelligent medical consulting system based on expert systems. J Edu Health Promot 2013;2:54.

These causes to improve responding to the patients who refuse to visit a physician and thus endanger the society of being infected easily because of their insufficient care about their disease. In some religious cultures, people may have bad attitude toward someone having diseases such as AIDS/HIV. As information classification is performed by the system, no human error will ever happen. The outcome of this research can be utilized for medical organizations, specialists, physicians, patients and information seeking people. Also, considering the current methods in prevalent medical systems, using scientific methods for textual information classification can provide some facilities for above-mentioned clients that require information about AIDS/HIV. Medical organizations and specialists can share their knowledge with the system which does not pile them, but keep them in an organized, structured and retrievable manner. Physicians may also use the system whenever they need to consult a colleague. In addition, patients can benefit specialists' views, without having concern about geospatial distance and society's attitude toward them, especially one who is suffering from AIDS/HIV, tuberculosis, hepatitis and etc.

LITERATURE REVIEW

Using expert systems has found its way in medical world. Some software applications in medical domain are: DENDRAL: Describing molecular structures introduced in 1965,[2] MYCIN: Heart diseases diagnosis developed in 1976,[3] PUFF: Lung diseases diagnosis,[4] XBONE: Bone diseases diagnosis,[5] VM: Supervision on ICU patients, CADCUCEUS: Internal diseases diagnosis, BLUE BOX: Diagnosis and treatment of depression, expert systems used to detect acid substances and electrolytes, anesthesia management education, etc,[6] MENINGE: An intelligent diagnosis system for acute children meningitis which is developed using Prolog. This system has been tested for 212 consecutive clinical patients and has led to sensible and efficient diagnosis in 94.8% of the cases. It had been found that the therapeutic error is 11.3% more than the system. Therefore, the system was installed in a medical care unit.[7]

Canfield has conducted a research to present a methodology to use available knowledge in medical centers to primare intelligent split menus for provider data-entry. A simulation has shown that data-entry using the intelligent split menus requires two to five times less effort (user selections or mouse clicks) than menus sorted alphabetically. Canfield has used a set of tests related to echocardiography for simulation. The methodology benefits statistical correlation between set of words such as "anatomy" and "pathology" to train the system to determine sequence of menus frequency. After initial training, choices of a real human client is used to update used frequencies. Then, the system is adjusted with data inout models.[8]

Smith discusses that researchers who develop medical decision support systems are aware of the need to answer real medical cases, especially supervision on specific ethical problems in the medical domain. It means that such intelligent systems should be evaluated for acceptability. Smith mentions cases such as the inherent performance evaluation. For instance, in a Medline search which has taken place using neural networks, it was showed that only 12.5% evaluated inherent performance adequently. The goal of this research was supervision on such cases, using possible evaluation methodology, presenting a framework and specific suggestions for each type of classification problem. Smith provides intelligent system developers to produce evidence of a sufficiency of output performance evaluation.

Using semantic information causes better coverage and performance of intelligent systems. Since, resource classification in intelligent medical systems considering enormous amount of data is vital, a system is required to constantly change the information and update the intelligent model. Yilu Zhou *et al.* state that designing and developing information retrieval techniques for languages other than English is vital, because non-English available resources on the web are increasing rapidly. The research focuses on discovering techniques that search engines use for English that can be used for other languages to facilitate information searching in multi-lingual world. The general structure for the technique consists of four main parts: A focused collection-building technique, a generic language processing ability, an integration of information resources, a post-retrieval analysis module.

Based on this idea, CMedPort was developed as a Chinese web portal in medical world. It not only enables the clients to search web pages through local collections and meta-search engines, but also generates encoding conversions between simplified and traditional Chinese. These capabilities help clients to search locally, summarize and classify documents. Comparing the effectiveness and efficiency of CMedPort with three other significant Chinese search engines, the results indicated that, there were no significant differences among the efficiency measures of CMedPort and benchmark systems. The researchers then evaluate CMedPort's user satisfaction using a questionnaire, and it was shown that none of the three search engines have gained such a level of user satisfaction.[9]

Mohtaram Ne'matollahi has studied several AIDS/HIV national information management systems around the world and recommends such a system for Iran. The goal of the probe was to design an AIDS/HIV surveillance system for Iran to minimize AIDS/HIV epidemic. The recommended system was also based on World Health Organization (WHO) recommendations and similar experiments in USA, Australia, UK and Malaysia. It was presented in six axes: General specifications of the system, minimum data set, data collection and reporting, data processing, data processing rules, data distribution. The results were accepted using Delphi technique and obtained 36.3 of 44 (82.5%). It is noteworthy that machine learning methods to diagnose patients suffering from AIDS/HIV have not been used in Iran.[10]

The Iranian web medical systems can be categorized into three categories: (1) Client asks his/her question from a specialist and revisits the site with a question ID after some days to read the answer, (2) Client searches his/her question in well-organized prevalent FAQs, (3) Client states his/her question in registration process and receives the answer via a phone call.

Foreign medical consulting websites have higher graphical user interface that the client specifies the problematic region on a graphical human body. The system also contains an educational module that shows required operators for physicians graphically. There are also some websites that show required descriptions as films. None of these systems which give services to clients use expert systems, except some which are offline. Some of the systems may not have medical classification and infrastructure and only provide answers for clients' questions. Only a minority of the systems include medical knowledge infrastructure, however responses are not online and systematic.

AIDS/HIV

AIDS is a progressive and preventable disease that causes major asthenia in human's immune system. As the illness progresses, it interferes more and more with the immune system, making different viruses to attack body cells and tissue such as skin, respiratory system, etc., including opportunistic infections, and tumours that do not usually affect people with working immune systems. The dangerous point about HIV virus is that it attacks the immune system that leads to decreasing body ability to confront infections and abnormal cells such as cancer cells. HIV virus involves immune cells in blood (lymphocytes), bone marrow, spleen, liver and lymphatic nodes, which play an important role in producing antibody to confront diseases and cancers. Eventually, it can be said that AIDS is a secondary malfunction of immune system which emerges when body is infected with HIV virus. The disease provides an ideal situation for progress of opportunist infections, which a normal human body can usually defeat and then progress of these infections leads to death. For example, tuberculosis is the most dominant cause of death among AIDS/HIV infected patients. The procedure is invisible and there is no way to detect a human suffering AIDS/HIV, by just looking and this is the reason of the phobia about this disease. However, a mere blood test can reveal the presence of this virus in the body. Some patients may be infected to AIDS/HIV for several years, without being aware of it.

AIDS/HIV is well-known as a developing pandemic disease all over the world. Its characteristics, transmission, development, kill rate, high rate of paralyzing, non-treatability and latent infection regarding the social and economical consequences, are the causes to make it a global dilemma. The number of patients suffering from AIDS/HIV in Iran has increased in rate and has made it to become a sanitary, social and economical problem and a large amount of funds are spent by the government to treat and combat its side effects.[11]

This research focuses on using machinery learning to design an AIDS/HIV intelligent medical consulting system to automate a number of system management tasks and facilitating clients' access to up-to-date information.

SUBJECTS AND METHODS

This descriptive study of applied kind was conducted through a case study about AIDS/HIV. Statistical population of this research is 752 records of FAQs related to AIDS/HIV. The methods used are Crisp method to produce intelligent model and RUP to develop the web-based consulting software.

CRISP-DM method

Data mining is a process of extracting various models, summarizations and values derived from a specific set of data.[12] The analysis step of the "Knowledge Discovery in Databases" process or KDD[13] is a field at the intersection of computer science and statistics, and is a process that attempts to discover patterns in large data sets. Regarding CRISP-DM method, every data mining project consists of six steps, which should be performed continuously and iteratively in the whole data mining procedure. The steps are as:[14] Business understanding, data understanding, data preparation, modelling, evaluation, development. The rest of the subsection is committed to these six steps.

Business understanding

This initial phase focuses on understanding the project objects and needs, then converting this knowledge into a data mining problem.[15] Specialists should contribute in the whole process so as to be able to reach the desired goal.[16] To perform the first step, two brainstorming session was held by Imam Khomeini Hospital specialists to discuss the required data types, and at the end an AIDS/HIV consulting system through FAQ data format was selected. The information was gathered through informing AIDS/HIV websites related to the disease discovery time (about 30 years) until now. The following tasks were also performed: The goal of the system supporter is not only informing, but also providing online treatment services and

Table 1: Predefined classes to determine the border and domain of the system

ID	Class
1	General questions
2	Symptoms-how I can tell you
3	Transmission
4	Testing
5	Laboratory information
6	Background information
7	Living with HIV
8	Medication to fight HIV
9	Prevention
10	Diagnosis
11	STD

HIV = Human Immunodeficiency Virus, STD = Sexually Transmitted Diseases

encouraging the patients to visit a specialist. The ultimate evaluation after implementation were conducted by the Ministry of Health and Medical Education experts, FAQs' categorization is based on experts' ideas and is shown in Table 1.

Data understanding

The data understanding step starts with initial data collection and proceeds with activities that enable you to become familiar with the data, identify data quality problems.[15] Second step consists of two stages: (1) Data gathering:[17] Gathering 752 FAQ records from informing AIDS/HIV websites, (2) Checking each FAQ record and classify it accordingly:[18] Class 5 was merged with class 4, since the number of records were less than expected. Class 5 was omitted, because no FAQ record was classified in this class. Finally, records which cannot be categorized in any of the classes were omitted.

Data preparation

The data preparation or pre-processing phase covers all activities needed to construct the final dataset [data that will be fed into the modeling tool(s)] from the initial raw data.[15] As Pyle has estimated in his book, pre-processing is required in 60% of data mining projects.[19] Pre-processing includes three stages: (1) Preparing raw data:[20] All data were converted into a same format and stored in Microsoft Access DBMS. Furthermore, data records which were irrelevant based on experts' ideas were eliminated from the database. Hence, FAQ records reduced to 372, (2) Three hundred and seventy two questions changed into words (tokenization) and repeated words were omitted. To do so, all the spaces where replaced with enter command. Then, all the words were exported from the database in the form of a Table. Stop words (Words that do not have a meaning and they can only be used in a sentence) were also eliminated using a SQL query. There are 630 keywords to this stage. Then, the frequency of each keyword in FAQ records was calculated using a piece of code. Final keywords were 163, after eliminating those with frequency less than 10, (3) A matrix was formed using a piece of code to determine the frequency of each keyword in FAQ records to be used in data mining and model generation. The matrix contains 163 columns that is the number of keywords. The rows of the matrix are the class of each FAQ. A piece of code was used to fill the Table, which is shown in Table 2.

Modelling

Selection and implementation of the proper data mining algorithm is performed in this step.[21] The model is generated using the matrix introduced in the previous subsection. The aforementioned matrix is formed into a Microsoft Excel file and is introduced to a data mining software namely Clementine 8.0 and using C5.0 method, so that required rules can be achieved.

The first stage is to introduce the matrix to the software. Then, input and output variables should be specified. In this case, all the keywords are inputs and FAQ classes are outputs. The variable type is introduced in the next stage. C5.0 algorithm is used to generate an intelligent tree model and the rules are also generated for tree structure. The output rules were so much to handle. Therefore, by changing the values for support and confidence, the number reduced. These rules are used as "if … then" rules. A subset of the output rules are shown below in Table 3.

Evaluation

In this step, Model/s generated in modelling step is evaluated based on quality and effectiveness.[22] The output of intelligent model which is created by C5.0 algorithm is validated by cross validation method. This validation method has different types. Here we used a special type in which the modelling action had been done for every record. After modelling, train data sets were tested on the model and reached the accuracy of 78.76%. Accuracy of the model using C5.0 algorithm was 78.76%. The output of model accuracy analysis which was created by Clementine software is shown in Figure 1.

Deployment

Models cannot be completely evaluated, except being tested under commercial situation.[23] To do so, the following steps are necessary:[24] Using the produced models, generating a simple report, generating more complicated reports.

Software development using RUP approach

As mentioned before, software production and development using RUP approach consists of four phases: Inception, elaboration, construction and transition.

The steps that should be fulfilled in all the four stages are as follows: Determining organization requirements, determining

Table 2: The output matrix of the third step

Class of FAQ	Keyword 1	Keyword 2	Keyword 163
1	1	0	4
2	0	3	0
3	0	2	2
372	1	1	4

FAQ = Frequently asked question

Table 3: List of the output rules

Rules for 11-contains 1 rule(s)
Rule 1 for 11 (78; 0.897)
if C143>0
and C131≤0
and C122≤0
and C4≤0
and C11≤0
and C110≤3
and C61≤0
and C96≤0
then 11
Rules for 2#- contains 1 rule(s) …
Rules for 3#- contains 4 rule(s) …
Rules for 4#- contains 1 rule(s) …
Rules for 6#- contains 3 rule(s) …
Rules for 8- contains 1 rule (s) …

the goal and subject description, determining operational requirements, determining business actors (Who use the system outside), determining business use cases, use cases belonging to operational units are written under the related package, business units determining in logical view under business object model, object relation are drawn in each business unit, the gathered data analyzed so that the use cases and state/activity diagrams were prepared, creating class diagrams, determining data model and creating system database, coding and software implementation. All the aforementioned steps will be discussed in subsequent subsections.

Inception

(1) Determining organization's requirements: Diagnosis services provider center requires patients suffering from AIDS/HIV, (2) Determining the goal and problem description: Using an intelligent model in the backbone of the system to prepare online responses for clients, (3) Determining operational requirements: To use the system, client types his/her question after entering the website. System will then specify the keywords after eliminating stoppage words. Keywords in client's question will be matched with keywords and Table of semantic keywords. Regarding keywords and using the intelligent model, the class of clients request will be recognized. The possibility of rating was also provided to measure the rate of practicality for each question. Each client can rate every FAQ record from 1 to 10. The client can send the question to the system administer, if a proper answer is not in the FAQs. The administrator will send the question to a specialist. Specialist receives the answer, emails the answer to the individual who had asked the question, puts the FAQ record in its category and updates the system if necessary.

Determining business actors: Everybody who plays a role in the system is a business actor. There are three types of business actors in the system: (1) Patient/client user, (2) System main manager, (3) Specialist. All the activities of the system are performed by one of the aforementioned groups.

Business use case: Regarding what is going to be done, use case diagrams have been reduced to more compact diagrams: Use case diagram - client's log in and log out operations, use case diagram - request and response operations, use case diagram - ranking operation, use case diagram - testing and entertainment operations, use case diagram - content productions operation. Relationships between the five diagrams introduced before hand are shown in Figure 2.

Activity diagram shows all the activities taking place in the system in all conditions. Two activity diagrams are generated in this stage. The former is non-managerial system activity diagram, which is shown in Figure 3.

The later is dedicated to managerial system activities diagram which is shown in Figure 4.

Elaboration

The whole system is classified into three packages: CGS-package, FAQ-package, and user package. Each package containing classes related to a part of system operations. CGS-package contains all classes related to content generation system. Some classes of CGS-package are set-system-management-page class, unanswered-q-management-page class, category-management-page class, FAQ-management-page class and etc., FAQ package contain all classes related to FAQs, which some of them are FAQ-Page class, read-more-page class, spell-checker class, user-ask class and etc., Finally, user-package contains classes related to user's registration and login. Classes of CGS-package are registration-page class, login-page class, use class, home-page class and etc.

Construction

In the construction part, the data model of the system is been generated. The activity diagram shows the structure of system database and it is the infrastructure of the system. To generate the data model in this stage, fields and Tables of system database was specified and then the relations between them were established. And the usage of each field specified.

Transition

System will be implemented and tested in this stage by ASP.net, C#, SQL-Server. A view of the designed and

Results for output field C3

Comparing $C-C3 with C3

Correct	293	78.76%
Wrong	79	21.24%
Total	372	

Coincidence Matrix for $C-C3 (rows show actuals)

	1	11	2	3	4	6	7	8	9
11	0	73	2	0	2	0	0	0	0
2	0	0	16	0	3	0	0	0	0
3	0	2	2	71	1	2	0	0	0
4	1	2	2	0	60	0	0	0	0
6	2	1	1	1	10	37	0	0	0
7	0	0	5	5	3	1	1	5	0
8	1	3	0	2	8	0	0	18	0
9	0	2	0	5	1	0	1	3	17

Figure 1: The output of model accuracy analysis

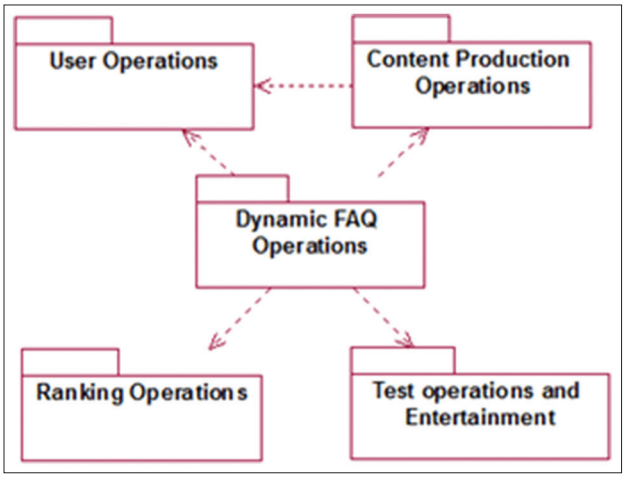

Figure 2: Relational model diagram of use-cases relation

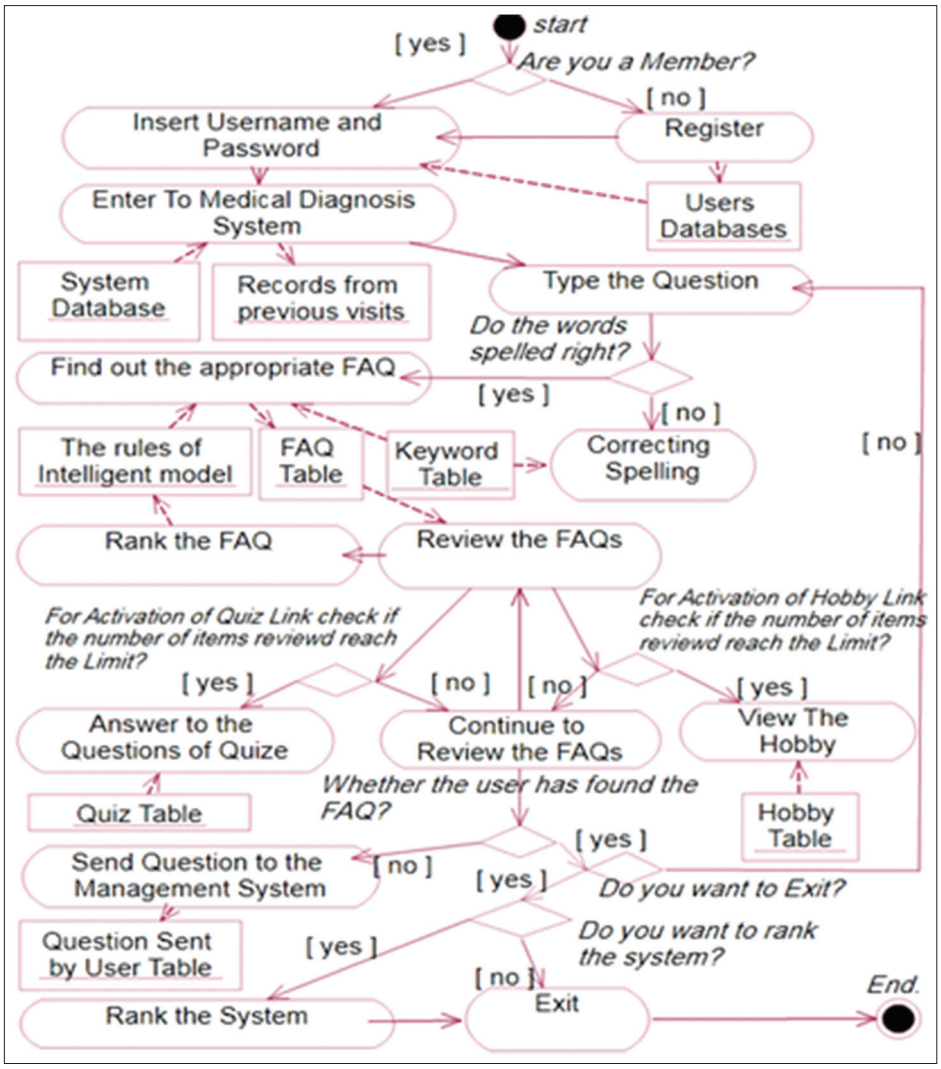

Figure 3: Activity diagram for user/client

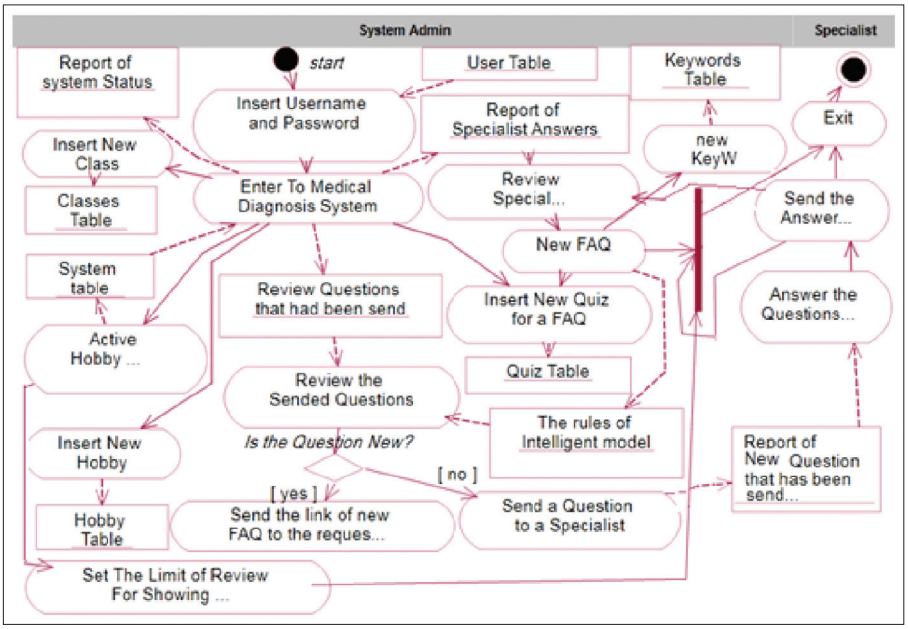

Figure 4: Activity diagram for managerial part

Figure 5: A view of the implemented software according to this research

implemented software is shown in Figure 5. As this software is being developed as an AIDS/HIV virtual clinic, currently it is not been completed yet to upload on the web and unfortunately is not available publicly.

It is noteworthy to mention that experts of Ministry of Health and Medical Education have evaluated the software and licensed this product.

RESULT

The research has been conducted to design an AIDS/HIV intelligent medical consulting system on the web. As informing the society about diseases is one of the most important problems in today's medical domain, the goal of the research was towards this destination. The results of the research consist of: (1) Investigation about text mining and text classification and coding for implementation, (2) Implementation of classification algorithms using C5.0 data mining technique and generating a model to detect the class of input data and using the model to give consultations about AIDS/HIV, (3) Implementation stages which consist of five stages: Study of requirements that lead to five diagram and their respective scenarios, generating system activity diagrams that leads to two diagrams to show all system activities, designing data model that leads to data base design, generating class diagrams, system implementation.

DISCUSSION

In this paper a novel model of an intelligent web-based medical system is presented by focusing on data mining and FAQ data related to AIDS/HIV. First, the goals of the system are specified and literature review and also current

medical systems are detected and analyzed, and the GUI of the system is described. Then, an intelligent model with 78.76% accuracy was generated based on data mining and text mining on pre-categorized FAQs about AIDS/HIV. After that, the system was analyzed and then implemented.

As medical fields are exposed to ever-growing information, which can be used to enhance healthcare and treatment quality, and available evidence indicates that the internet is becoming the main source for delivering information on the prevention, control and treatment of HIV/AIDS.[25] We demonstrate an online consulting system design and development that support the patients and specialist who need help on the AIDS/HIV field. The proposed framework utilizes data mining algorithm for extracting intelligent model to support clients. Experiments do prove that it is feasible to use this method to develop an intelligent web-based medical system for other diseases, which is valuable for further in-depth study.

There are many issue to be taken into account is the development of a web-based intelligent medical consulting system for people living with HIV/AIDS.

First of all, being a web-based system involves global access to it. This may help to improve responding to the patients who refuse to visit a physician and thus endanger the society of being infected easily because of their insufficient care about their disease. Hotline number is the popular way for consulting for AIDS/HIV patients, which could not answer to all needs.

There is another interesting issue about the implemented model that can be used in medical system software for different diseases. As medical domain contains a large volume of data, using data mining model can greatly assist medical goals and prevention of dangerous illnesses. The software facilitates informing the society about various diseases and as a result improves the medical knowledge of them.

Finally because AIDS/HIV medical consulting systems have not been developed using intelligent infrastructure, so being equipped with an intelligent model, providing consulting services on systematic textual data and providing side services based on client's activities cause the implemented system to be unique.

The researchers wish to review the usage of medical information as future work. Also, some traditional formats of information storage can be changed to further enhance the quality of the software.

REFERENCES

1. Shortliffe EH, Perrault LE. Medical informatics: Computer applications in health care and biomedicine. 2nd ed. New York: Springer; 2000.
2. Edward A, Feigenbaum BG, Buchanan D, Meta D. Roots of knowledge systems and expert system applications. Artif Intell 1933;59:233-73.
3. Shorrtliffe EH. Computer-based medical consultations: MYCIN. New York: Elsevier; 1976.
4. Aikins JS, Kunz JC, Shortliffe EH, Fallat RJ. PUFF: An expert system for interpretation of pulmonary function data. Comput Biomed Res 1983;16:199-208.
5. Hatzilygeroudis I, Vassilakos PJ, Tsakalidis A. XBONE: A hybrid expert system for supporting diagnosis of bone diseases. Stud Health Technol Inform 1997;43 Pt A:295-9.
6. Ghazanfari M, Kazemi Z. Expert systems. Tehran: Elmo Sanat; 2004. [Persian].
7. François P, Crémilleux B, Robert C, Demongeot J. MENINGE: A medical consulting system for child's meningitis. Study on a series of consecutive cases. Artif Intell Med 1992;4:281-92.
8. Canfield K. Priming intelligent split menus with text corpora for computerized patient record data-entry. Int J Biomed Comput 1995;39:263-73.
9. Zhou Y, Qin J, Chen H. CMedPort: An integrated approach to facilitating Chinese medical information seeking. Decis Support Syst 2006;42:1432-18.
10. Nematollahi M, Khalesi N, Moghaddasi H, Askarian M. Second Generation of HIV Surveillance System: A Pattern for Iran. Iran Red Crescent Med J 2012;14:309-12.
11. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, *et al*. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 2008;455:661-4.
12. Fayyad U, Piatetsky-Shapiro G, Padhraic S. From data mining to knowledge discovery in databases. American Association for Artificial Intelligence. 1996;17:37-18.
13. Horrocks D. CRISP: An introduction. Md Med 2010;11:20-1.
14. Harper G, Pickett SD. Methods for mining HTS data. Drug Discov Today 2006;11:694-9.
15. Hand D, Mannila H, Smith P. Principles of data mining. India: Prentice Hall; 2005.
16. Fayyad U, Djorgovski S, Weir N. Automating analysis and cataloging of Sky surveys. In: Fayyad U, Piatetsky-Shapiro G., Smyth P, Uthurusamy R, editors. Advances in knowledge discovery and dataminig. Boston: MIT Press; 1996; p. 471-93
17. Kantardzic M. Data mining: Concepts, models and algorithms (Focus on Machine Learning). New York: IEEE-Wiley; 2003.
18. Pyle D. Data preparation for data mining. San Francisco, CA: Morgan Kaufmann; 1999.
19. Emati HR, Barko CD. Key factors for achievinf organizational data mining success. Ind Manage Data Syst 2003;103:282-10.
20. Barson A, Smith S, Thearling K. Building data mining applications for CRM. New York: McGraw Hill; 2000.
21. Han J, Kamber M. Data mining: Concepts and techniques. San Francisco: Morgan Kaufmann; 2000.
22. Lefebure R, Venture G. Data Mining. Paris: Eyroll; 2000.
23. Larose DT. Discovering Knowledge in Data. New York: John Wiley; 2005.
24. Gómez EJ, Cáceres C, López D, Del Pozo F. A web-based self-monitoring system for people living with HIV/AIDS. Comput Methods Programs Biomed 2002;69:75-86.
25. Hull MEC, Taylor PS, Hana JRP, Millar RJ. Software development process – an assessment. Inf Softw Technol 2002;44:1-12.

Source of Support: Nil, **Conflict of Interest:** None declared