UNIVERSITI SAINS MALAYSIA

Laporan Akhir Projek Penyelidikan Jangka Pendek

Development of a Web-Based Medical Information Repository Integrated With an Artificial Intelligence-based Medical Decision Support System

By
Dr. Lim Chee Peng

A0837

PAN ASIA ICT R&D GRANT

Final Report

Tutup Geran

PROF. DR LEE KEAT TEONG
Pengarah
Pejabat Pengurusan & Kreativiti Penyelidikan
Universiti Sains Malaysia

10/6/06

Project Title:

Development of a Web-based Medical Information Repository Integrated with an Artificial Intelligence-based Medical Decision Support System

Project No.:

0402A13_S11

Name of Researcher:

Dr. Lim Chee Peng

University of Science Malaysia

Summary

This project is focused on the design and development of a web-based Medical Information Repository (MIR) and an Artificial Intelligence (AI)-based system for support of medical decision making. The MIR contains medical records pertaining to suspected stroke patients that are collected from a teaching hospital for evaluating the usefulness of the developed system. Applicability of the system to the early prognosis and diagnosis of acute stroke patients is demonstrated.

A total of more than 1000 patient records have been collected and stored electronically in the MIR. The information collected comprises patients' medical history, physical examination and observation results, laboratory test results, evaluation of medical specialists.

An AI-based Decision Support System (DSS) is also developed. Specifically, the Fuzzy ARTMAP (FAM) and Fuzzy Min-Max (FMM) neural network models and other complementary techniques are used to develop intelligent medical decision support tools. A series of experiments is performed to evaluate the performance of the AI-based DSS in analysing and classifying the stroke patients into different categories. These include prediction of (i) the condition on discharge, (ii) the Rankin scale category of patients upon discharge; and (iii) the length of hospital stay by patients. Three performance indicators, i.e., Accuracy, Sensitivity, and Specificity, are calculated. Logistic Regression (LR) models, a technique that is commonly used in clinical data analysis, are used for performance comparison.

In addition, fuzzy if-then rules are extracted from the AI-based DSS in order to provide explanation for the predictions. The if-then rules allow domain experts (medical specialists) to understand the rationale behind the predictions made by the neural-network-based decision support tools. The empirical study reveals the potential of deploying the AI-based DSS as a usable and useful computerized tool in supporting the decision making process in the medical domain.

Table of Contents

1. Introduction

Computer-based decision support (CBDS) systems are increasingly used by professionals in many diversify fields, e.g., manufacturing, bio-medicine, finance, and economy. This trend is driven by the need to support and enhance the human decision making processes by extracting and deducing information/knowledge from data. Apart from data-driven analytical techniques such as decision and classification theory, a CBDS system can utilize Artificial Intelligence (AI) and knowledge-based approaches to increase the capability of decision making.

In general, medical prognosis and diagnosis is the process where physicians attempt to identify the disease a patient is suffering from based on the patient's physical symptoms, clinical tests data, and other related signs and information. In some cases, physicians may require additional opinion for disease diagnosis since some diseases have common symptoms and some are related to others. However, it is not always easy to obtain a second opinion or to statistically analyze the symptoms. This problem has inspired researchers to develop CBDS tools for medical prognosis and diagnosis.

Since medical prognosis and diagnosis tasks can be viewed as a pattern classification problem, artificial neural network (ANN) models have shown good potentials as a tool to diagnose (classify) patients based on their symptoms. Indeed, ANNs have the ability to integrate data-based analytical techniques such as decision and classification theory, and to use knowledge-based approaches to provide useful information to support the decision making process [1-4]. From the literature, a lot of successful ANN applications to medical problems can be found. In [1], a pulmonary disease diagnostic system was proposed. The system used real clinical data to attempt to treat a whole category of distressed body organs. A model selection method that used the self-organizing map (SOM) for breast cancer diagnosis was demonstrated in [2]. In [3], the performances of a number of ANN models for breast cancer diagnosis were compared and analyzed. A neuromuscular disorder diagnosis system that employed two different ANN models (supervised and unsupervised) for analyzing features selected from electromyography (EMG) was presented in [4]. From the literature review, it is obvious that CBDS tools are useful to assist in medical prognosis and diagnosis tasks.

2. Research Problem

The aim of the research is to develop and implement a web-based Medical Information Repository (MIR) integrated with an Artificial Intelligence (AI)-based medical Decision Support System (DSS) to promote advanced, quality healthcare information and services for all people, especially the rural communities, through the use of Information and Communication Technology (ICT) as well as AI-based methodology.

The main objectives of the research are as follows.

1. to establish a web-based repository for storage and analysis of patient records;
2. to devise an intelligent software system based on the Adaptive Resonance Theory (ART) neural networks and other complementary techniques for medical decision support;
3. to evaluate the effectiveness of the proposed MIR integrated with the DSS for early diagnosis of stroke patients

3. Research Findings

There are two main aspects of work in this research, i.e., the establishment of a medical information repository (MIR) and the development of an AI-based decision support system (DSS). For the research conducted, it is found that the establishment of the MIR needs a long period. The main effort is geared towards collection and analysis of medical records pertaining to acute stroke patients. As the medical records are collected from real patient information, there are a number of challenges faced. The collected records often are incomplete, and contain missing data. As such, extra effort is needed to cleanse the medical records that that they can be useful for the development of the AI-based DSS.

For the AI-based DSS, the main challenge is on the development of an intelligent system that is able to learn autonomously in stationary as well as non-stationary environment. The ability to adapt and to continue learning without erasing old information is an essential characteristic where training data samples are limited and new knowledge is continuously encountered. Adaptation also provides a degree of robustness by compensating for minor variability in characteristics of processing elements.

4. Fulfillment of Objectives

In general, there are three objectives as set out in the proposal. First, a web-based MIR is established for storing and analyzing patient records. The main tasks accomplished include

(i) collecting medical records of acute stroke patients from hospital;

(ii) transforming the medical records into an electronic (stroke) database;

(iii) cleansing the missing/incomplete information in the stroke database;

After the data cleansing process, useful statistics and facts pertaining to the stroke database are extracted for evaluation of healthcare professionals.

Second, the backbone of the AI-based DSS is based on a family of ANN models called the Adaptive Resonance Theory (ART) networks [5-7]. Fuzzy ARTMAP (FAM) [7], a specific ANN developed from ART theory, and its variant are developed and deployed as the intelligent DSS in this study. This is because FAM-based models are able to absorb novel information without corrupting previously learned knowledge. Besides FAM, the Fuzzy Min-Max (FMM) [8] neural network is developed to classify stroke patient records. FAM and FMM share a common characteristic, i.e., both networks possess the same incremental learning property. They are able to learn autonomously through a single pass of a set of training data, and to absorb information into their knowledge bases continually without resorting to the re-training problem. In addition, a useful data visualization tool is developed to identify important features from the medical records for pattern classification.

Third, a series of empirical studies is conducted to evaluate the AI-based DSS (FAM and FMM) using the medical records contained in the MIR. The MIR and DSS are integrated into a web-based platform. Useful performance indicators including accuracy, sensitivity, and specificity rates for diagnosis of acute stroke patients are obtained. Furthermore, fuzzy if-then rules are extracted from the DSS to help explain how the predictions are made. The results demonstrate that the AI-based are useful for providing useful decision support for medical diagnosis of acute stroke patients.

Based on the above explanation, the research conducted has successfully fulfilled the three objectives initially set out in the proposal.

5. Project Design and Implementation

The project design and implementation is divided into two parts: web-based MIR and AI-based DSS.

5.1 The Web-Based MIR

Figure 1 show the context diagram of the system developed. In general, the overall architecture of the MIR built in accordance with the client/server model. This architecture allows users to access the system using a standard web browser, such as Internet Explorer. Information contained in the system can be shared among medical practitioners at various locations, subject to the accessibility to the internet.

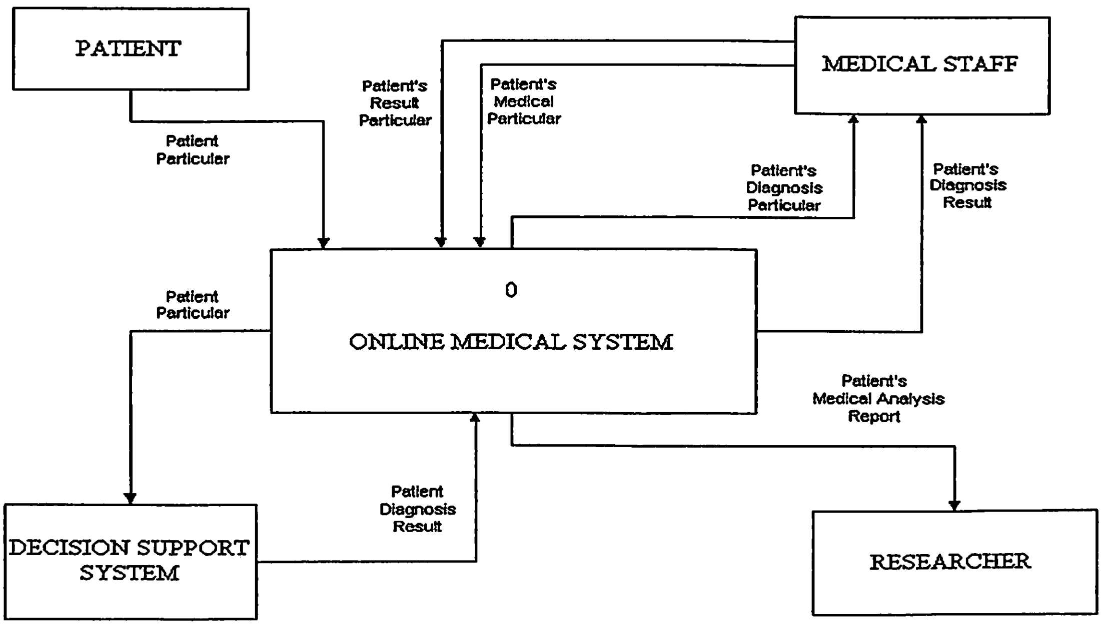

Figure 1 Context diagram of the MIR and DSS integrated system.

A database comprising medical records of more than 1000 acute stroke patients has been established. A proforma was first designed to record information of suspected acute stroke patients admitted to a number of participating hospitals. A total number of 70 symptoms/features of patients were collected. These comprised patients' medical history, physical examination and observation results, laboratory test results, evaluation of medical specialists. Some observations were recorded for a period of twenty days after the occurrence of stroke. Some examples of the information collected are shown in Table 1.

Age	Sex
Underlying Hypertension	Alcoholic Consumption
History of Diabetes Mellitus	Cigarette Smoking
Glasgow Coma Scale Score	Rankin Scale on Admission
Systolic Blood Pressure	Diastolic Blood Pressure
Type of Ward Admitted	Initial Blood Sugar Level
Cholesterol Level	HDL Cholesterol Level
Triglyceride Level	Body Temperature on Admission
Subtype of stroke	First or Recurrent Stroke
Epileptic Seizure	Urinary Tract Infection
Chest Infection	Other Infection
Pressure Sore/Skin Break	Deep Vein Thrombosis
Pulmonary Embolism	Depression
Symptomatic Intracranial Hemorrhage	Constipation
Congestive Heart Failure	Conductive Arrhythmia
Arthritis	Retention Urination

Table 1 Examples of the symptoms/features collected

A lot of effort is spent in "cleansing" and "filtering" the proforma. Follow-ups with patients to collect the missing information are conducted. Very often, the record is purged from the database owing to non-availability of a lot of information. As a result, an extended period is spent to collect an adequate number of patient records for analysis and evaluation. The main aim is to ensure that a representative MIR of acute stroke patients can be established, and an appropriate analysis using the AI-based DSS can be performed.

5.2 The AI-Based DSS

For the AI-based DSS, the activities accomplished include

(i) evaluating and validating the performance of the AI-based DSS and other statistical methods for medical decision support using the MIR of acute stroke patients

(ii) developing a data visualization tool for analyzing and selecting important data features;

(iii) extracting fuzzy if-then rules from AI-based DSS for explaining predictions;

Three experimental studies using the stroke patient database are conducted. The first problem is the prediction of the Condition on Discharge (either Class 1-Alive or Class 2-Dead) of patients. A total number of 1014 data samples (patient records), with 940

Alive cases and 74 Dead cases are used. The second problem is prediction of Rankin Scale category of patients on discharge (either Class 1-Rankin scale between 0 and 1 or Class 2-Rankin scale between 2 and 6). After eliminating some missing data, a total number of 661 data samples, with 141 in Class 1 and 520 in Class 2, are used. The third problem is prediction of the number of hospital days stayed, either Class 1-fewer than or equal to 15 days; or Class 2-more than 15 days. In this problem, 1002 cases, with 710 cases in Class 1 and 292 cases in Class 2, are used.

A total number of 18 most representative features/symptoms are extracted from the database for experimentation. Table 2 summarizes the three data sets used in the experimental studies.

Features/Symptoms	Data Sets		
	COD	Rankin	HDay2
Age	*	*	*
Sex	*	*	*
Underlying Hypertension	*	*	*
Alcoholic Consumption	*	*	*
History of Diabetes Mellitus	*	*	*
Cigarette Smoking	*	*	*
Glasgow Coma Scale Score	*	*	*
Rankin Scale on Admission	*	*	*
Systolic Blood Pressure	*	*	*
Diastolic Blood Pressure	*	*	*
Initial Blood Sugar Level	*	*	*
Cholesterol Level	*	*	*
HDL Cholesterol Level	*	*	*
Triglyceride Level	*	*	*
Body Temperature on Admission	*	*	*
Subtype of Stroke	*	*	*
First or Recurrent Stroke	*	*	*
Type of Ward Admitted	*	*	*
Condition on Discharge (COD)	#		
Rankin Scale on Discharge (Rankin)		#	
Number of Hospital Days (HDay2)			#

Table 2: Features (denoted as *) used in the each dataset where # denoted as the target class. Underlying hypertension is measured in either "no hypertension", "uncertain" or a series of range in years. Alcoholic consumption is evaluated in either "yes" or 'no" while history of diabetes is deliberated in either "yes", "no" or "uncertain". Subtype of stroke consists of five categories (LACI, PACI, TACI, POCI, and ICH). Type of wards admitted is divided into general neurological wad, organized stroke unit, and general ward.

Two variants of FAM network models are employed, i.e. single FAM and multiple FAM models. In addition, logistic regression (LR) models, which are commonly used in medical diagnostic applications, are established. The LR outcomes serve as a yardstick for comparison with those from FAM.

In general, multiple FAM is able to achieve better results than those from single FAM and LR. For the COD problem, multiple FAM obtained the same Accuracy rate as LR and slightly higher than that of single FAM. For the Rankin and Hday2 problems, multiple FAM obtained an accuracy rate which is 4.35% and 0.74% higher than those from LR, and 1.37% and 4.19% higher than those from single FAM, respectively.

In pattern classification, the main task is to learn and construct an appropriate function that is able to assign a set of input features into a finite set of classes. If noisy and irrelevant input features are used, the learning process may fail to formulate a good decision boundary that has a discriminatory power for data samples of various classes. As a result, feature selection has a significant impact on classification accuracy.

The circle-segments method and the Principal Component Analysis (PCA) are used to select a set of important data features for pattern classification with four different machine learning systems, i.e., Multilayer Perceptron (MLP), Support Vector Machine (SVM), Fuzzy ARTMAP (FAM) and k-Nearest Neighbour (kNN). The input features comprise patients' medical history, physical examination, laboratory test results, etc. The problem is the prediction of the Rankin Scale category of patients upon discharge, either Class 1-Rankin scale between 0 and 1 (141 samples) or Class 2-Rankin scale between 2 and 6 (520 samples).

Based on the circle-segments, three features, i.e., V8, V16, and V18, are identified as important input features for classification. The PCA method is also used to analyse the data set. For a real data set, five or six principal components may be required to account for 70% to 75% of total variation, and the more principal components that are required, the less useful each one becomes. From the PCA method, eight features, i.e., V2, V3, V4, V6, V7, V8, V16, and V17, are identified as the important input features for classification.

In general, the classification performances improve with feature selection either using PCA or circle-segments. The circle-segments method yields the best accuracy rates for all the four machine learning systems despite the fact that only three features are used for classification (a reduction of 83% of the number of input features). In terms of sensitivity, the circle-segments method also shows improvements for MLP, SVM, and kNN from 17%-35% as compared with those before feature selection. The specificity rates are better than those before feature selection using MLP and FAM, but inferior for SVM and kNN. There is a trade-off between improvements in sensitivity and specificity rates with the use of the circle-segments method, i.e., more substantial improvement in sensitivity with marginal degradation in specificity for SVM and kNN while less substantial improvement in sensitivity with marginal improvement in specificity for MLP and FAM. This observation is interesting as it is important for a medical diagnostic system to have high sensitivity as well as specificity rates so that patients with and without the diseases can be identified accurately.

By comparing the results of circle-segments and PCA, the accuracy and sensitivity rates of circle-segments are better than those of PCA. The trade-off is that PCA generally yields better specificity rates. Again, it is interesting to note the substantial improvement in sensitivity with marginal inferior performance in specificity of both the circle-segment and PCA results. Another observation is that the PCA results do not show substantial improvements in terms of accuracy, sensitivity, and specificity as compared with those from the original data set.

When it comes to safety-critical applications such as medical prognosis and diagnosis, the domain experts often hesitate, if not totally reluctant, to use ANN models as a solution owing to the lack of a comprehensive explanation facility on the predictions from ANN models. Indeed, an ANN without an explanation capability is called a "black-box" model. Nevertheless, a "black-box" can be turned into a "white-box" by translating the internal contents of the ANN into an understandable and meaningful form of knowledge. This can be accomplished by extracting symbolic rules from ANN models.

In an attempt to incorporate rule extraction into the FMM model, a Genetic Algorithm (GA)-based rule extractor is applied. The GA is then used to select the minimum number of rule antecedents with the maximum accuracy. The selected rules are quantized, and translated into fuzzy if-then format. Each rule is tagged with a confidence factor, and has "don't care" antecedents to simplify its interpretation. An empirical study with the problem of predicting the Rankin Scale of patients upon discharge, either Class 1-Rankin scale between 0 and 1 (141 samples) or Class 2-Rankin scale between 2 and 6 (520 samples) is conducted. An example of the rules extracted is as follows:

IF	Age is *High*
	Sex is *Don't Care*
	Underlying Hypertension is from *Medium to High*
	Alcoholic Consumption is from *Medium to Very High*
	History of Diabetes Mellitus is *High*
	Cigarette Smoking is from *Medium to Very High*
	Glasgow Coma Scale Score is *High*
	Rankin Scale on Admission is from *Very Low to Low*
	Systolic Blood Pressure is *Don't Care*
	Diastolic Blood Pressure is *Don't Care*
	Initial Blood Sugar Level is *Don't Care*
	Cholesterol Level is *Don't Care*
	HDL Cholesterol Level is *Don't Care*
	Triglyceride Level is *Don't Care*
	Body Temperature on Admission is *Don't Care*
	Subtype of Stroke is *Don't Care*
	First or Recurrent Stroke is *Don't Care*
	Type of Ward admitted is *High (One)*
THEN	Output is Class 1 for RS, with confidence factor of 0.8

The rule can be simplified by eliminating the "don't care" antecedents, as follows.

IF	Age is *High*
	Underlying Hypertension is from *Medium to High*
	Alcoholic Consumption is from *Medium to Very High*
	History of Diabetes Mellitus is *High*
	Cigarette Smoking is from *Medium to Very High*
	Glasgow Coma Scale Score is *High*
	Rankin Scale on Admission is from *Very Low to Low*
	Type of Ward admitted is *High (One)*
THEN	Output is Class 1 for RS, with confidence factor of 0.8

From the above rule example, it is clear that the GA-based rule extractor is able to extract easily comprehensible rules for domain users (medical practitioners).

6. Project Outputs and Dissemination

The main output of the project is a web-based MIR integrated with AI-based DSSs.

A database of more than 1000 anonymous stroke records is established. This electronic, web-based MIR enables administration and management of patient records in an efficient manner. Useful statistics related to stroke diagnosis as well as patient demographic information are also established.

During the course of design and development the AI-based DSSs, a number of AI-based algorithms and complementary (statistical) methods are developed, which include FAM, FMM, PCA, circle segments. In addition, a set of diagnostic rules is extracted from FAM and FMM systems to help provide explanation and knowledge for medical practitioners.

Some snapshots of the web-based MIR integrated with the AI-based DSSs are shown in Figures 2-7. Figure 2 shows the login page of the web-based system. Figure 3 depicts some example of the anonymous stroke patient records. Figures 4 and 5 present the parameter settings and the training sessions of FAM. Figure 6 gives a set of rules extracted from FAM for explaining its predictions. Figure 7 shows how the FAM can be used to give a prediction of Rankin Scale category of a suspected stroke patient.

In terms of dissemination of the knowledge gained, three scientific articles are produced. One of them is published in an international conference proceedings, and the other two are submitted to international journals and are currently under review.

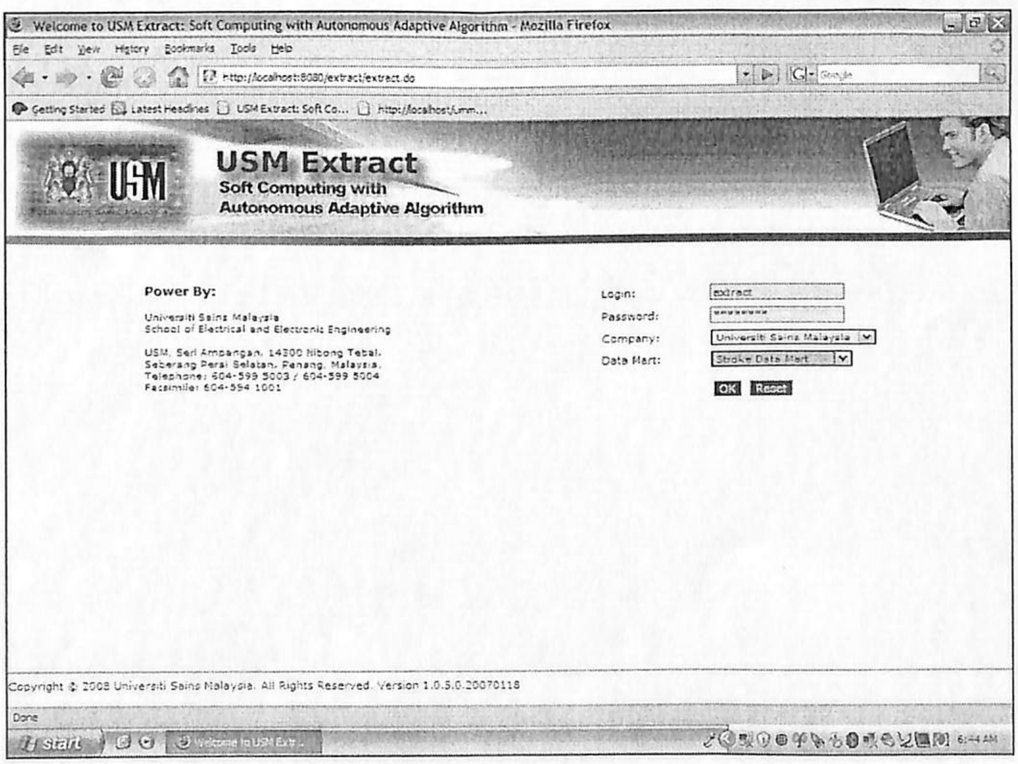

Figure 2 The login page of the developed system

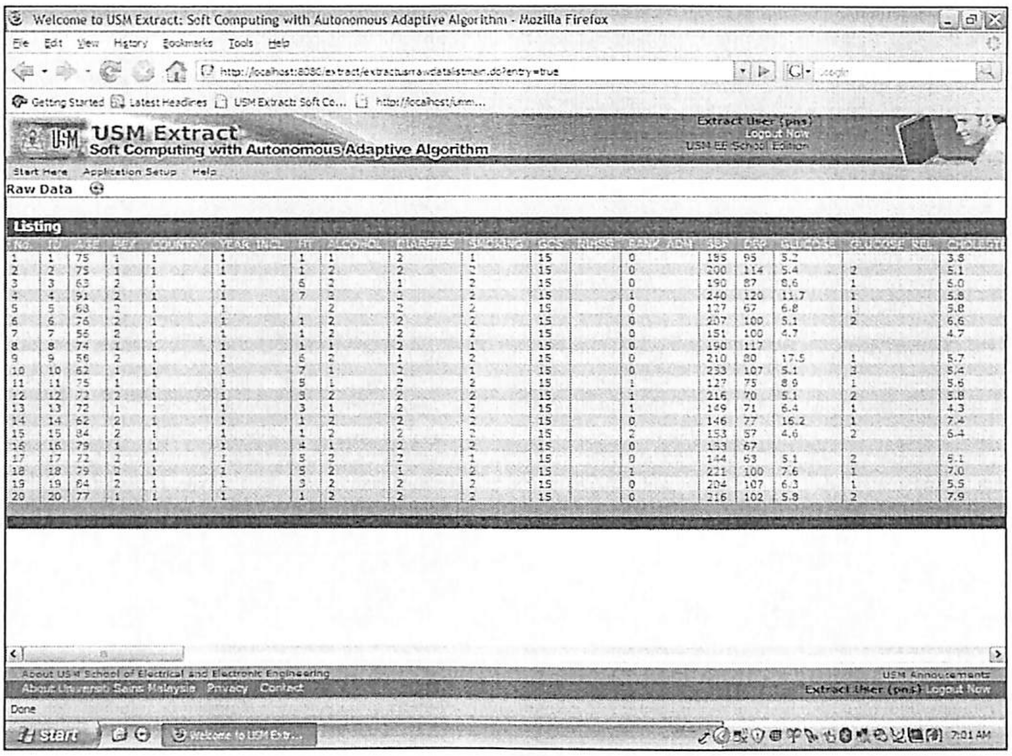

Figure 3 The web-based MIR

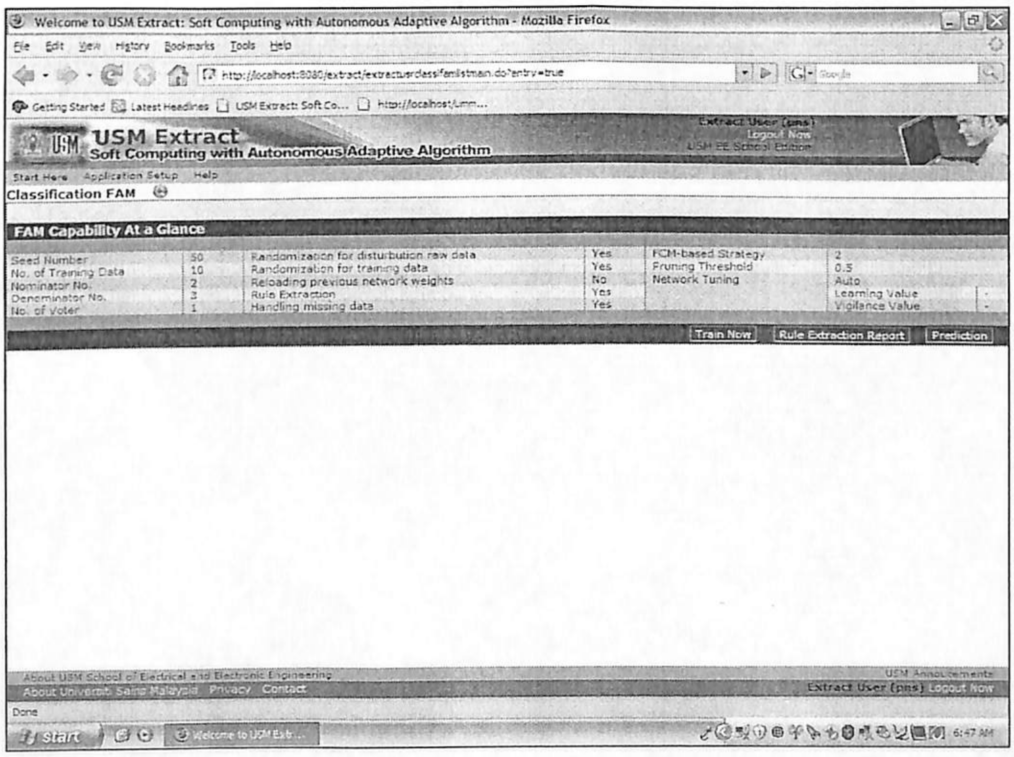

Figure 4 Parameter settings of the FAM system

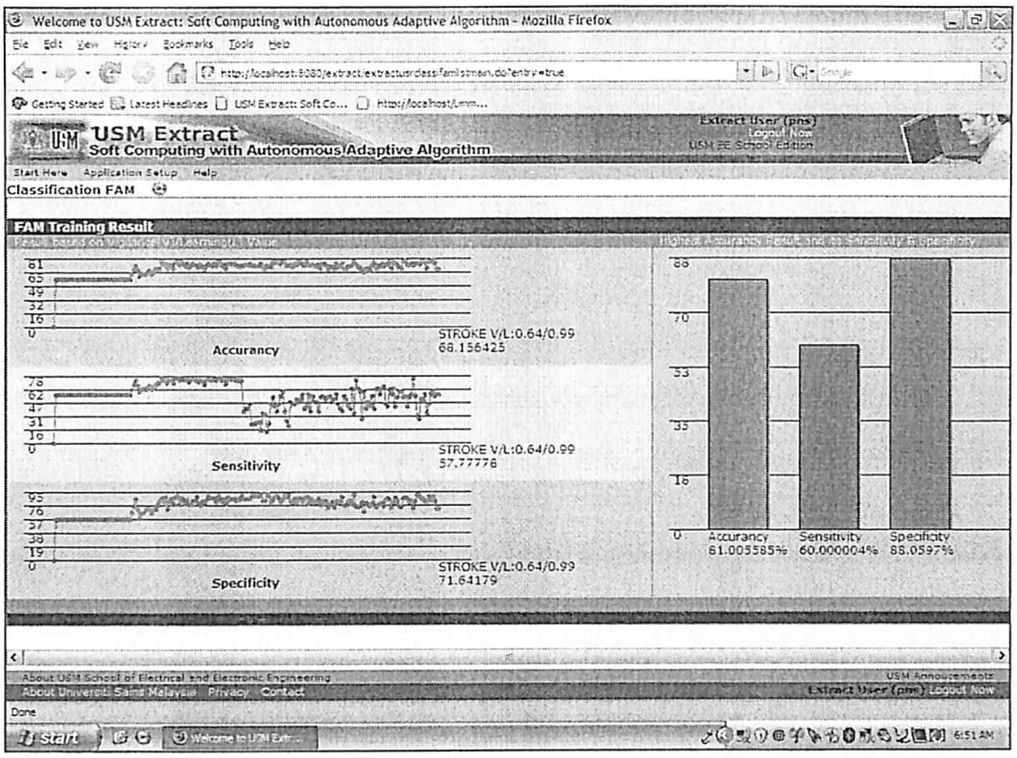

Figure 5 Training sessions of the FAM system

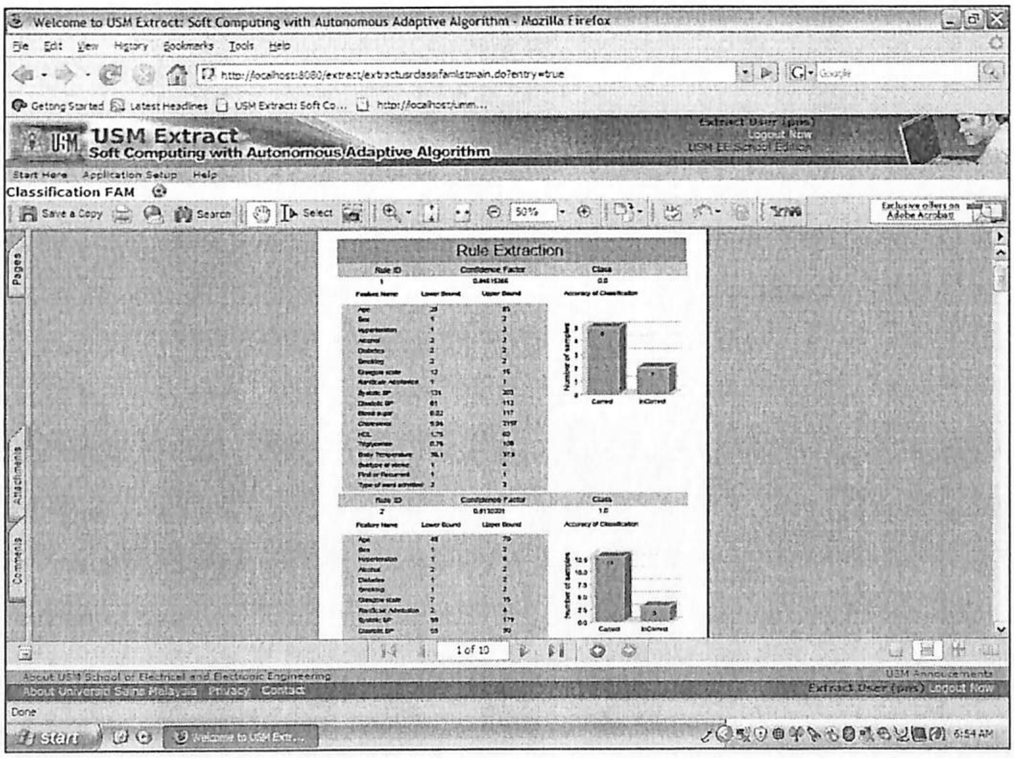

Figure 6 Rules with confidence factors extracted from FAM

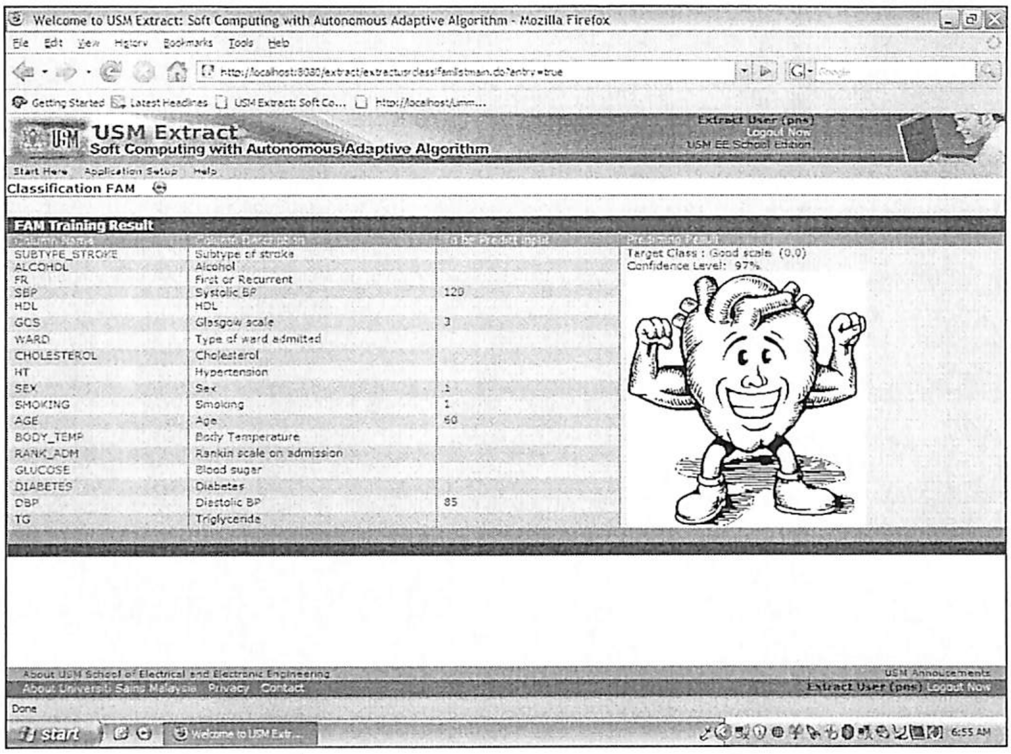

Figure 7 Use of the FAM-based DSS for Rankin Scale prediction

7. Capacity Building

The research team has gained a lot of experience in terms of designing web-based MIR, cleansing and analyzing real medical data, developing AI-based, especially FAM-based and FMM-based, DSSs for medical diagnosis, and extracting rules from AI-based systems for enhancing their practical applicability.

A number of postgraduate students (2 masters and 2 PhD) in the area of AI and medical decision support systems are trained during the course of this project. In addition, valuable experience and knowledge in building MIR, especially in the aspect of tackling the issue on missing/incomplete information and on cleansing information from databases, are gained.

The research findings are submitted for publication in international learned journals. This enhances the credibility and reputation of the research team as well as the institutions involved, i.e., University of Science Malaysia (USM) and University Malaya (UM), in international standing.

8. Project Management

The project is conducted in two institutions, USM and UM. The USM pole is responsible for the development of the AI-based DSSs, while the UM pole is accountable for medical data collection and interpretation of the AI-based DSSs predictions. Quarterly project meetings to review work conducted, to enhance the work done, and to derive any necessary improvements for the project are held.

The research funding is handled by a team of professionals from the Bursary Department of USM. They are specialized in administrating research funding from either local or overseas organizations. An effective account keeping system is also established to manage and keep track of progressive expenses of the project.

In summary, all aspects of the project are managed efficiently.

9. Impact

The project has direct impact to use of ICT and AI for medical decision support. This output of this project, i.e., a web-based MIR integrated with AI-based DSSs is an effort toward the development and implementation of a telehealth/telemedicine system. According to [38], medical repositories for telehealth applications are still limited at present, and the diffusion of electronic medical records is still at its infancy. On the other hand, according to [39], intelligent agents and AI have been identified as emerging technologies to support a rapidly expanding scope of telemedicine/telehealth applications.

This project, therefore, contribute towards overcoming the above two pressing issues by (i) developing an electronic medical repository to assist the diffusion of electronic medical records, (ii) devising an AI-based decision support system that is able to learn autonomously and continuously to cater for future telehealth applications, as well as a continuous education tool for further training of medical practitioners.

In Malaysia, the government has embarked on a series of initiatives to realize Vision 2020, i.e., the national agenda to transform Malaysia into a knowledge-based society. One of the initiatives is the MSC (Multimedia Super Corridor) flagship applications, and telehealth is one of them. The output of this project could also be incorporated into the MSC flagship applications, with the aim of providing equitable, accessible, and high quality healthcare services through the use of ICT.

10. Overall Assessment

The main difficulty faced during the course of the research is the collection and cleansing of the medical database. As the medical records are collected from a hospital, a number of unavoidable issues have occurred, resulting in medical records with incomplete information, missing data, etc.

The quality of the medical database is of utmost importance for the successful development of the AI-based DSSs. This is because the FAM and FMM systems learn from data samples, and noisy data samples would lead to compromised performances. Therefore, the research team has spent a lot of effort and time in cleansing and analyzing the medical database. This has resulted in an extended period of time in completing the activities related to medical data collection and analysis.

Other than the timing issue, the research has been conducted and progressed in a smooth manner. A number of useful outputs in terms of software and scientific publications have been produced.

In summary, the research has fulfilled all the objectives set up in the proposal.

References

[1] G. P. K. Economou, C. Spiropoulos, N. M. Economopoulos, N. Charokopos, D. Lymberopoulos, M. Spiliopoulou, E. Haralambopulu, and C. E. Goutis, Medical diagnosis and artificial neural networks: a medical expert system applied to pulmonary diseases, *Proceedings of the 1994 IEEE Workshop Neural Networks for Signal Processing*, pp. 482-489, 1994.

[2] D. West and V. West, Model selection for a medical diagnosis decision support system: A Breast Cancer Detection Case, *Artificial Intelligence in Medicine*, vol. 20, pp. 183-204, 2000.

[3] T. Kıyan and T. Yıldırım, Breast cancer diagnosis using statistical neural networks, *Proceedings of the International XII Turkish Symposium on Artificial Intelligence and Neural Networks*, pp. 754-760, 2003.

[4] C. S. Pattichis, C. N. Schizas, and L. T. Middleton, Neural network models in EMG diagnosis, *IEEE Transactions on Biomedical Engineering*, vol. 42, pp. 486-496, 1995.

[5] G.A. Carpenter, and S. Grossberg, A massively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision, Graphics and Image Processing*, vol. 37, pp. 54-115, 1987.

[6] G.A. Carpenter, and S. Grossberg, The ART of adaptive pattern recognition by a self-organizing neural network, *IEEE Computer*, vol.21, pp. 77-88, 1988.

[7] G.A. Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds, and D.B. Rosen, Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps, *IEEE Transactions on Neural Networks*, vol. 3, pp. 698-712, 1992.

[8] P. K. Simpson, Fuzzy min-max neural networks-Part 1: Classification, *IEEE Transactions on Neural Networks*, vol. 3, pp. 776-786, 1992