

© 2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Genetic Algorithm based Fuzzy Multiple Regression for the Nocturnal Hypoglycaemia Detection

Sai Ho Ling, *Member, IEEE*, Hung Nguyen, *Senior Member, IEEE*, and Kit Yan Chan

Abstract—Low blood glucose (Hypoglycaemia) is dangerous and can result in unconsciousness, seizures and even death. It has a common and serious side effect of insulin therapy in patients with diabetes. We measure physiological parameters (heart rate, corrected QT interval of the electrocardiogram (ECG) signal, change of heart rate, and the change of corrected QT interval) continuously to provide detection of hypoglycaemic. Based on these physiological parameters, we have developed a genetic algorithm based multiple regression model to determine the presence of hypoglycaemic episodes. Genetic algorithm is used to determine the optimal parameters of the multiple regression. The overall data were organized into a training set (8 patients) and a testing set (another 8 patient) which are randomly selected. The clinical results show that the proposed algorithm can achieve predictions with good sensitivities and acceptable specificities.

I. INTRODUCTION

EPISODES of hypoglycaemia, especially at night, are common for Type 1 diabetes mellitus patients (T1DM), largely because usual insulin preparations do not adequately mimic the normal patterns of endogenous insulin secretion [10]. In the DCCT, patients assigned to intensive therapy experienced a threefold increase incidence of severe hypoglycaemic episodes over those receiving conventional therapy [2-3]. In that report, hypoglycaemic episodes are defined as those in which the patient has blood glucose levels < 3.33 mmol/l (60mg/dl). Severe hypoglycaemic episodes are defined as those in which the patient has blood glucose levels < 2.80 mmol/l (50mg/dl). Thus hypoglycaemia was proved to be a limiting factor in achieving improved diabetes control. In this paper, the hypoglycaemic episodes are defined as those in which the patient has blood glucose levels < 3.33 mmol/l.

In this paper, we have developed a genetic algorithm based multiple regression with fuzzy inference system for the detection of hypoglycaemic episodes in T1DM children using physiological parameters such as heart rate and corrected QT interval of electrocardiogram signal. Multiple regression [5] is a statistical method which can be used to

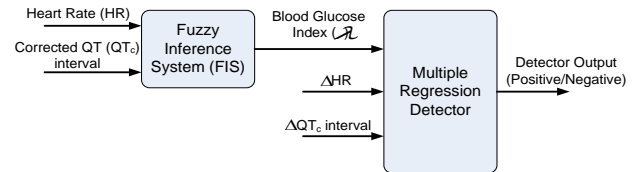


Fig. 1. Hypoglycaemic Episodes Detection System

generate the relationship between a set of independent or predictor variables and a set of dependent or criterion variable. Regression approach can be used to develop models in parametrical polynomial form by fitting data. Fuzzy inference system (FIS) [6] is good in representing some expert knowledge and experience in some linguistic rules which can be easily understood by human being. In the proposed system, FIS is used to pre-process the physiological parameters. By introduce FIS, the overall performance in terms of sensitivity and specificity of the detection system is significantly improved. To optimize the model parameters of multiple regression and the rules of fuzzy inference system, a global learning algorithm, Genetic Algorithm (GA) [1, 8], is introduced. GA is a powerful random global search technique to handle optimization problem. It can find the globally optimal solution over a domain. A case study is given to show the proposed detector on detect the hypoglycaemic episodes successfully in T1DM children.

This paper is organized as follows: an overview of the methods used for non-invasive and continuous detection of hypoglycaemia is discussed in Section II. Section III presents the development and results of the genetic algorithm based multiple regression with fuzzy inference system used for the identification of nocturnal hypoglycaemic episodes in T1DM children. A conclusion for this study is drawn in Section IV.

II. METHODS

A. Hypoglycaemic Episodes Detector (Genetic Algorithm based Multiple Regression with Fuzzy Inference System)

To realize the detection of hypoglycaemic episodes in T1DM, a genetic algorithm based multiple regression with fuzzy inference system is developed and the system block diagram is shown in Fig. 1 and it is a 4 inputs and 1 output system. The inputs are the heart rate (HR), corrected QT interval of the electrocardiogram signal (QT_c), change of heart rate (ΔHR), and the change of corrected QT interval

S.H. Ling is with the Centre for Health Technologies, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW, Australia. (e-mail: Steve.Ling@uts.edu.au; phone: 61-2-9514 2390; fax: 61-2-9514 2868).

H.T. Nguyen is with the Centre for Health Technologies, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW, Australia. (e-mail: Hung.Nguyen@uts.edu.au).

K.Y. Chan is with the Digital Ecosystems and Business Intelligence Institute, Curtin University of Technology, WA, Australia. (e-mail: Kit.Chan@curtin.edu.au).

(ΔQT_c), and the output is the binary status of hypoglycaemia (+1 represent non-hypoglycaemia and -1 represents hypoglycaemia). The hypoglycaemic episodes detector system mainly consists of two sub-systems, namely fuzzy inference system (FIS) and multiple regression detector. The FIS plays a main role to approximate the correlation between the physiological parameters (HR and QT_c) and the blood glucose index (λ) by using some fuzzy rules. Experimental results will be given in the next section to illustrate the performance in terms of the sensitivity and specificity of the hypoglycaemic episodes detection system and they show that it is increased significantly when FIS is introduced. The other sub-system is called multiple regression detector which is a model to classify the status of hypoglycaemia. This model is estimated by the model parameters. In this proposed system, the model parameters are trained by a genetic algorithm. The details of the genetic algorithm, fuzzy inference system and multiple regression detector are discussed as follows:

1) Genetic algorithm

To optimize the parameters of the fuzzy inference system and the multiple regression detector, a global learning algorithm, Genetic Algorithm (GA), is used. The GA process is shown in Fig. 2. First, a set of population of chromosomes P is created. Each chromosome \mathbf{p} contains a set of genes p_{ij} , where $i = 1, 2, \dots, n_p, j = 1, 2, \dots, n_g, n_p$ and n_g denote the population size (number of chromosome) and the number of gene respectively. Second, the chromosomes are evaluated by a defined fitness function which is written as,

$$\text{fitness} = f(\mathbf{p}_i) \quad (1)$$

The form of the fitness function depends on the application. The better chromosomes are return higher fitness values in this process. Third, some of the chromosomes are selected to undergo genetic operations for reproduction by the method of normalized geometric ranking [7]. It is a selection based on a non-stationary penalty function which is a function of the generation number. As the number of generation increases, the penalty increases that puts more and more selective pressure on the GA to find the feasible solution. In general, a higher-rank chromosome will have a higher chance to be selected. Fourth, genetic operation of crossover is performed. The crossover operation is mainly for exchanging information from the two parents, chromosomes \mathbf{p}_1 and \mathbf{p}_2 , obtained in the selection process with a defined probability of crossover μ_c . This probability gives an expected number of chromosomes that undergo the crossover. In this paper, blend- α [4] is selected as the operation of crossover which has a good searching ability in handling multimodal and separability problem effectively. For Blend- α crossover, the resulting offspring is chosen randomly from the interval $[X_j^1, X_j^2]$ following the uniform distribution, where

$$X_j^1 = \min(p_{1j}, p_{2j}) - \alpha d_j \quad (2)$$

$$X_j^2 = \max(p_{1j}, p_{2j}) + \alpha d_j \quad (3)$$

```

=====
begin
   $\tau \rightarrow 0$  //  $\tau$ : iteration number
  initialize  $P(\tau)$  //  $P(\tau)$ : population for iteration  $\tau$ 
  evaluate  $f(P(\tau))$  //  $f(P(\tau))$ : fitness function
  while (not termination condition) do
    begin
       $\tau \rightarrow \tau + 1$ 
      select 2 parent  $\mathbf{p}_1$  and  $\mathbf{p}_2$  from  $P(\tau-1)$ 
      perform crossover and mutation operations
      reproduce a new  $P(\tau)$ 
      evaluate  $f(P(\tau))$ 
    end
  end
end
=====

```

Fig. 2. Procedure of the Genetic Algorithm

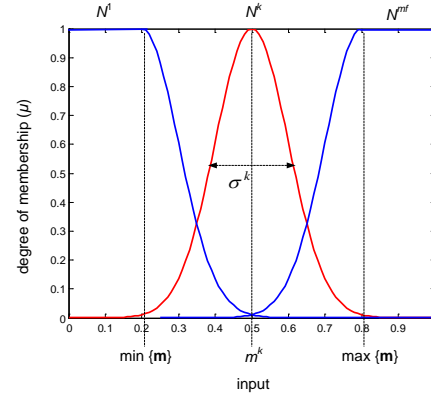


Fig. 3. Fuzzy inputs

where $d_j = |p_{1j} - p_{2j}|$, p_{1j} and p_{2j} are the j -th elements of \mathbf{p}_1 and \mathbf{p}_2 , respectively, and α is a positive constant.

After the crossover operation, the mutation operation follows. The mutation operation is to change the genes of the chromosomes in the population such that the features inherited from their parents can be changed. A probability of mutation μ_m is defined to govern the operation and it gives an expected number of genes that undergo the mutation. Non-uniform mutation [8-9] is investigated in this paper and it is an operation with a fine-tuning capability. The mutated gene is given by,

$$\hat{p}_{ij} = \begin{cases} p_{ij} + \Delta(\tau, \delta_{\max}^j - p_{ij}) & \text{if } r_d = 0 \\ p_{ij} - \Delta(\tau, p_{ij} - \delta_{\min}^j) & \text{if } r_d = 1 \end{cases}, \quad (4)$$

where r_d is a random number equal to 0 or 1 only. The function $\Delta(\tau, y)$ return a value in the range $[0, y]$ approaches 0 as τ increases. It is defined as follows.

$$\Delta(\tau, y) = y \left(1 - r \left(1 - \frac{\tau}{T} \right)^\zeta \right), \quad (5)$$

where r is a random number in the range of $[0, 1]$, τ represents the current generation number, T represents the total iteration number and ζ is a system parameter that determines the degree of non-uniformity. After going through the mutation operation, the new offspring will be evaluated using the fitness function. The new population will be formed when the new offspring replaces the chromosome

with the smallest fitness value. After the operations of selection, crossover and mutation, a new population is generated. This new population will repeat the same process. Such an iterative process will be terminated when a defined condition is met.

This GA is used to tune (optimize) the fuzzy rules of FIS and the model parameters of the multiple regression model.

2) Fuzzy inference system

Referring to Fig. 1, the fuzzy inference system is used to realize the approximated correlation between the physiological parameters (HR and QT_c) and blood glucose index (λ). This system pre-processes the physiological parameters by using a set of fuzzy rules and reasons the approximated blood glucose index. The approximated index, λ then is one of the inputs of the multiple regression model. The sensitivity and specificity can be enhanced by using the pre-processing system. The detail of FIS is given below.

a) Fuzzification and reasoning

The first step is to take the inputs and determine the degree of membership to which they belong to each of the appropriate fuzzy sets via membership functions. In this study, there are two inputs, heart rate (HR) and corrected QT interval (QT_c). The degree of the membership function shows in Fig. 3 for input HR $\mu_{N_{HR}^k}(HR(t))$ is a bell-shaped function as given by:

when $m_{HR}^k \neq \max\{\mathbf{m}_{HR}\}$ or $\min\{\mathbf{m}_{HR}\}$

$$\mu_{N_{HR}^k}(HR(t)) = e^{-\frac{(HR(t)-m_{HR}^k)^2}{2\sigma_{HR}^k}}, \quad (6)$$

where $\mathbf{m}_{HR} = [m_{HR}^1, m_{HR}^2, \dots, m_{HR}^k, \dots, m_{HR}^{m_f}]$, $k=1, 2, \dots, m_f$, m_f denotes the number of membership function, $t=1, 2, \dots, n_d$, n_d denotes the number of input-output data pair, parameter m_{HR}^k and σ_{HR}^k are the mean value and the standard deviation of the member function, respectively.

If $m_{HR}^k = \min\{\mathbf{m}_{HR}\}$ then

$$\mu_{N_{HR}^k}(HR(t)) = \begin{cases} 1 & \text{if } HR(t) \leq \min\{\mathbf{m}_{HR}\} \\ e^{-\frac{(HR(t)-m_{HR}^k)^2}{2\sigma_{HR}^k}} & \text{if } HR(t) > \min\{\mathbf{m}_{HR}\} \end{cases} \quad (7)$$

If $m_{HR}^k = \max\{\mathbf{m}_{HR}\}$ then

$$\mu_{N_{HR}^k}(HR(t)) = \begin{cases} e^{-\frac{(HR(t)-m_{HR}^k)^2}{2\sigma_{HR}^k}} & \text{if } HR(t) \geq \max\{\mathbf{m}_{HR}\} \\ 1 & \text{if } HR(t) < \max\{\mathbf{m}_{HR}\} \end{cases} \quad (8)$$

Similarly, the degree of the membership function for input QT_c ($\mu_{N_{QT}^k}(QT(t))$) is same as HR.

With these fuzzy inputs and fuzzy output λ , the behavior of the FIS is governed by a set of fuzzy if-then rules in the following format:

Rule γ : IF $HR(t)$ is $N_{HR}^k(HR(t))$ AND $QT(t)$ is $N_{QT}^k(QT(t))$ THEN $\lambda(t)$ is w_γ (9)

where $N_{HR}^k(HR(t))$ and $N_{QT}^k(QT(t))$ are fuzzy terms of rule γ , $\gamma=1, 2, \dots, n_r$; n_r denotes the number of rules and n_r is equal

to $(m_f)^{n_{in}}$ where n_{in} represents the number of inputs of the FIS; $w_\gamma \in [0, 1]$ is the fuzzy singleton to be determined.

b) Defuzzification

Defuzzification is the process of translating the output of the fuzzy rules into a scale. The final value of $\lambda(t)$ is given by:

$$\lambda(t) = \sum_{\gamma=1}^{n_r} m_\gamma(t) w_\gamma \quad (10)$$

where

$$m_\gamma(t) = \frac{\mu_{N_{HR}^k}(HR(t)) \times \mu_{N_{QT}^k}(QT(t))}{\sum_{\gamma=1}^{n_r} (\mu_{N_{HR}^k}(HR(t)) \times \mu_{N_{QT}^k}(QT(t)))} \quad (11)$$

c) Training with GA

To learn the input-output relationship of the HR, QT_c and λ by using FIS, GA is employed to optimize the fuzzy rules by finding out the best parameters $m_{HR}^k, \sigma_{HR}^k, m_{QT}^k, \sigma_{QT}^k, w_\gamma$ of the FIS. The input-output relationship is described by,

$$\lambda^d(t) = \mathbf{g}(\mathbf{z}^d(t)), \quad \mathbf{z}^d(t) = [HR^d(t) \quad QT^d(t)] \quad (12)$$

where $\mathbf{z}^d(t)$ and $\lambda^d(t)$ are the given physiological inputs and the desired BGI output of a nonlinear function $\mathbf{g}(\cdot)$ respectively. If blood glucose levels, $\lambda^d(t) < 3.33 \text{ mmol/L}$, $\lambda^d(t)$ will be set to 1, otherwise, it will be set to 0.

Referring to (1) of the genetic algorithm, the fitness function is defined as,

$$\text{fitness} = -\frac{1}{n_d} \sum_{t=1}^{n_d} (\lambda^d(t) - \lambda(t))^2 \quad (13)$$

The objective is to maximize the fitness function of (13) (minimize the mean square error between the desired $\lambda^d(t)$ and the $\lambda(t)$ from FIS) using the GA by setting the chromosome to be $[m_{HR}^k, \sigma_{HR}^k, m_{QT}^k, \sigma_{QT}^k, w_j]$ for all j, k .

3) Multiple regression detector

In this section, a multiple regression detector is introduced to model the relationship between the system's inputs and the presence of hypoglycaemic episodes. Referring to Fig. 1, the detector inputs are: 1) blood glucose index λ which is estimated by the FIS; 2) change of heart rate (ΔHR) and 3) the change of corrected QT interval (ΔQT_c).

ΔHR and ΔQT_c are useful information to enhance the sensitivity of the testing results. Some experimental results will be given to explain on it.

a) Multiple regression model

In general, multiple regression model procedures will estimate as the following form:

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{X}_i + \beta_2 \mathbf{X}_i^2 + \dots + \beta_j \mathbf{X}_i^j, \quad (14)$$

where the \mathbf{Y}_i is output to represent the binary status of hypoglycaemic. When $\mathbf{Y}_i \geq 0$, which represents positive

TABLE I
TRAINING AND TESTING RESULTS (MEAN)-SET SPECIFICITY (TRAINING) = 54-59%

Method	Training		Testing	
	Sen	Spec	Sen	Spec
PR-FIS2	70.24%	58.62%	70.45%	55.14%
PR-FIS3	71.27%	58.68%	68.13%	56.76%
LR-4	65.08%	54.51%	43.39%	67.70%
LR-2	68.89%	54.86%	38.21%	71.50%

TABLE II
TESTING RESULTS (BEST)-SET SPECIFICITY (TRAINING) = 54-59%

Method	With Best Sensitivity		With Best Specificity	
	Sen	Spec	Sen	Spec
PR-FIS2	75.57%	50.23%	64.29%	63.85%
PR-FIS3	75.57%	48.36%	64.29%	64.79%
LR-4	46.43%	68.08%	42.86%	70.89%
LR-2	39.29%	70.89%	37.50%	71.83%

(sick), appositively, when $\mathbf{Y}_i < 0$, which represents negative (healthy). \mathbf{X}_i is the inputs of the system, $\mathbf{X}_i = [\lambda_i \Delta\text{HR}_i \Delta\text{QT}_i]$, $i = 1, 2, \dots, n_d$; $\boldsymbol{\beta}$ denotes the parameters of the regression model and η denotes the number of order. In order to find the optimal set of the parameters $\boldsymbol{\beta}$ of the model, the GA is used.

b) Trained with GA

The objective of the multiple regression model is to detect the hypoglycaemic episodes accurately. To measure the performance of the biomedical classification test, sensitivity and specificity are introduced. The sensitivity measures the proportion of actual positives which are correctly identified and the specificity measures the proportion of negatives which are correctly identified. The definitions of the sensitivity (ζ) and the specificity (κ) are given as follows:

$$\zeta = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad \text{and} \quad \kappa = \frac{N_{TN}}{N_{TN} + N_{FP}},$$

where N_{TP} is number of true positive which implies the sick people correctly diagnosed as sick; N_{FN} is number of false negative which implies the sick people wrongly diagnosed as healthy; N_{FP} is number of false positive which implied healthy people wrongly diagnosed as sick; and N_{TN} is number of true negative which implied healthy people correctly diagnosed as healthy.

The objective of the system is to maximize the sensitivity and the specificity, where the fitness function is defined as follow:

$$fitness = v\zeta + (1-v)\kappa \quad (15)$$

where $v \in [0, 1]$ is a constant value to control a balance of the sensitivity ζ and specificity κ . A larger value of the v gives a strong force to the system to maximize the sensitivity; however, it will reduce the performance of the specificity. For clinical study, the sensitivity of the detection system is most significant than the specificity due to the sensitivity represent the sick peoples. It is important to find out the abnormal condition for peoples accurately. The choosing of the v is trial by error through experiment. In this clinical study, v is set at 0.52 where the mean training sensitivity is larger than the mean training specificity around 10–20%, i.e. ~70% (Sensitivity) vs ~60% (Specificity).

TABLE III
TRAINING AND TESTING RESULTS (MEAN)- SET SPECIFICITY (TRAINING) = 84-89%

Method	Training		Testing	
	Sen	Spec	Sen	Spec
PR-FIS2	88.98%	43.68%	75.57%	40.31%
PR-FIS3	87.70%	45.60%	73.75%	42.16%
LR-4	87.78%	37.65%	47.77%	50.92%
LR-2	84.21%	38.00%	48.75%	52.70%

TABLE IV
TESTING RESULTS (BEST)- SET SPECIFICITY (TRAINING) = 84-89%

Method	With Best Sensitivity		With Best Specificity	
	Sensitivity	Specificity	Sensitivity	Specificity
PR-FIS2	80.36%	38.03%	67.86%	50.23%
PR-FIS3	80.36%	35.68%	67.86%	50.23%
LR-4	53.57%	58.22%	46.43%	61.50%
LR-2	50.00%	55.87%	44.64%	67.61%

In this system, the objective is to maximize the fitness function of (15) using the GA by optimizing the parameters $\boldsymbol{\beta}$ represented by the chromosome.

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

Sixteen children with T1DM (14.6 ± 1.5 years) volunteered for the 10-hour overnight hypoglycaemia study at the Princess Margaret Hospital for Children in Perth, Australia. Each patient is monitored overnight for the natural occurrence of nocturnal hypoglycaemia. We measure the required physiological parameters, while the actual blood glucose (BG) levels are collected as reference using Yellow Spring Instruments. The main parameters used for the detection of hypoglycaemia are the heart rate and corrected QT interval [11-12]. The responses from 16 T1DM children exhibit significant changes during the hypoglycaemia phase against the non-hypoglycaemia phase. Normalization is used to reduce patient-to-patient variability and to enable group comparison by dividing the patient's heart rate and corrected QT interval by his/her corresponding values at time zero.

The detection of hypoglycaemic episodes ($\text{BG} < 3.3 \text{ mmol/l}$) using these variable is based on a genetic algorithm based multiple regression with fuzzy inference system developed from the obtained clinical data. In effect, it estimates the presence of hypoglycaemia at sample period k_s based on the basic of the data at sampling period k and the previous data at sampling period $k_s - 1$. In general, the sampling period is 5 minutes and approximately 35-40 data points are collected from each patient.

The overall data set consisted of a training set and a testing set, each with 8 patients randomly selected. For these, the whole data set which included both hypoglycaemia data part and non-hypoglycaemia data part are used.

For comparison and analysis purposes, 4 approaches are used to tackle the T1DM problem and they are: i) Proposed second order ($\eta=2$) multiple Regression Fuzzy Inference System (PR-FIS2); ii) Proposed third order ($\eta=3$) multiple Regression Fuzzy Inference System (PR-FIS3); iii) Linear multiple regression with 4 inputs (LR-4), the 4 inputs are

HR, QT_c , ΔHR , ΔQT_c , and iv) Linear multiple regression with 2 inputs (LR-2), the 2 inputs are HR and QT_c only. Eight patients are selected for training. By using GA which is used to find the optimized fuzzy rules and membership functions, and the model parameters of the multiple regression model, the basic settings of the parameters of the GA are shown as follows.

Population size n_p : 50;

Probability of crossover μ_c : 0.8;

Probability of mutation μ_m : 0.1;

The constant value of crossover α : 0.5;

The shape parameter of the mutation ζ : 2;

Number of iteration T : 1000 for FIS tuning, 200 for regression tuning.

There are two steps to develop the proposed multiple regression with fuzzy inference system to T1DM problem. The first step is to determine the optimized fuzzy rules and membership function of the FIS to approximate the relationship between the inputs HR, QT_c and the output blood glucose index λ . Once the optimized FIS is developed, the second step is to develop a multiple regression model to detect the nocturnal hypoglycaemic episodes in T1DM.

In the fuzzy inference system, the number of the membership function is set at 5 ($m_f = 5$), thus, the total number of fuzzy rules is equal to 25. Referring to (13), the fitness function of the GA for tuning the parameters [m_{HR}^k , σ_{HR}^k , m_{QT}^k , σ_{QT}^k , w_γ] of FIS is defined where $k = 1, 2, \dots, 5$ and $\gamma = 1, 2, \dots, 25$. The total number of input-output data pair n_d is 320 for 8 patients.

After the training process, the optimized fuzzy membership functions are found. Once an optimized fuzzy inference system is developed, a multiple regression is used to model the T1DM problem. For comparison and analysis purpose, second order and third order of the multiple regression models are used. According to (14), η is set at 2 and 3 respectively. To train the regression model, we use GA to maximize the ζ and κ by optimizing the chromosomes represented by [$\beta_0 \beta_1 \dots \beta_\eta$]. In this study, 8 testing patients are used. Referring to (15), when we set the constant value v at 0.52, the specificity of the training process for all approaches are keep around 54% to 59%. The average training and testing results in term of the sensitivity and specificity are tabulated in Table I. All results are averaged over 20 runs. Among these 20 set of results, the results with the best sensitivity and with the best specificity are tabulated in Table II. From Table I, we can see that the overall performance of multiple regression with fuzzy inference system (PR-FIS2, PR-FIS3) outperform the multiple regression without FIS (LR-4, LR-2). The average testing sensitivity is around 70% which is better than LR-4 and LR-2 (43% and 38% only). Compare with PR-FIS2 and PR-FIS3, the PR-FIS2 gives a better sensitivity. On the other hand, compare with LR-4 and LR-2, the average testing sensitivity of LR-4 performs better, which is affected that

ΔHR ; ΔQT_c are the useful physiological information to enhance the sensitivity. Similarly, Table III and Table IV show the training and testing results with the training specificity set to around 84% to 89%, where the v is set at 0.47. From these tables, PR-FIS2 also gives a better sensitivity compared with other approaches. In short, from the Table II and Table IV, we can see that the best sensitivity is 75.5% with an acceptable specificity, 50.2%, using by the method PR-FIS2 with $\alpha=0.52$.

IV. CONCLUSION

In this paper, a genetic algorithm based multiple regression with fuzzy inference system detection algorithm is developed to recognize the presence of hypoglycaemic episodes. Genetic algorithm is proposed to optimize the fuzzy rules and the regression model. A clinical study is given to illustrate the proposed algorithm outperforms the others in terms of the sensitivity and specificity. The performance of the proposed algorithm for detection of hypoglycaemic episodes for T1DM is satisfactory, as the best sensitivity is 75.5% and specificity is over 50%.

ACKNOWLEDGMENT

The authors would like to thank Dr. Nejhdeh Ghevondian and Assoc. Prof. Timothy Jones for their contribution. This research is supported by the grant of Juvenile Diabetes Research Foundation.

REFERENCES

- [1] L. Davis, *Handbook of Genetic Algorithms*. NY: Van Nostrand Reinhold, 1991.
- [2] DCCT Research Group, "The effect of intensive treatment of diabetes on the development and progression of long-term complications in IDDM", *N. Eng. J. Med.*, no. 329, pp. 977-986, 1993.
- [3] DCCT Research Group, "Adverse events and their association with treatment regimens in the Diabetes Control and Complications Trial", *Diabetes Care*, no. 18, pp. 1415-1427, 1995.
- [4] L.J. Eshelman and J.D. Schaffer, "Real-coded genetic algorithms and interval-schemata," *Foundations of Genetic Algorithms 2*, pp. 187-202, 1993.
- [5] D.A. Freedman, *Statistical Models: Theory and Practice*, Cambridge University Press, 2005.
- [6] J.S. Jang and C.T. Sun, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, 1997.
- [7] J. Joines and C. Houck, "On the use of non-stationary penalty functions to solve constrained optimization problems with genetic algorithm," in Proc. of 1994 International Symposium evolutionary computation, Orlando, 1994, pp. 579-584.
- [8] Z. Michalewicz, *Genetic Algorithm + Data Structures = Evolution Programs*, 2nd extended ed. Springer-Verlag, 1994.
- [9] A. Neubauer, "A theoretical analysis of the non-uniform mutation operator for the modified genetic algorithm," in *IEEE Int. Conf. Evolutionary Computation*, 1997, Indianapolis, pp. 93-96.
- [10] J.C. Pickup, "Sensitivity glucose sensing in diabetes", *Lancet*, no. 355, pp. 426-427, 2000.
- [11] H.T. Nguyen, "Intelligent technologies for real-time biomedical engineering applications", *Int. J. Automation and Control*, vol. 2, no.2, pp. 274-285, 2008.
- [12] H.T. Nguyen, N. Ghevondian, T. Jones, "Detection of nocturnal hypoglycemic episodes (natural occurrence) in children with type 1 diabetes using an optimal Bayesian neural network algorithm", in

30th Ann. Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Vancouver, Canada, Aug 2008, pp. 1311-1314.