668 International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.3 Issue No.5, pp : 668-670 1 May 2014

Prediction of Hyperglycemia in Diabetic Patients Using Data Mining Techniques

K.Jothi, M. Ramya, S.Kauser, L.Thomas Robinson, S.Raj Anand[5]

Department of Master of Computer Applications, Vel Tech Multi Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Avadi,Chennai,Tamil Nadu

jothikarunagaran@gmail.com, ramyamano123@gmail.com, kauserzainab@gmail.com, son.mca@gmail.com, rajsra_mca@rediff.com

Abstract: ***Diabetes mellitus is a group of metabolic diseases in which a person hyperglycemia, either because the pancreas does not produce enough insulin, or because body's cell does not respond to the insulin that is produced. In diabetics patients the improper control of blood glucose level will lead to many complications. Prediction of hyperglycemia events is by Continuous Glucose monitoring. In this paper prediction of hyperglycemia uses ARIMA model and METABO system. The prediction of blood glucose is done every 10 or 15 minutes. The METABO system is which aims at monitoring of diabetics and recording, interpreting. The results are processed with proper system modeling.***

Keywords: *Hyperglycemia, Data mining, METABO system, ARIMA model.*

1. Introduction: While 33 million, men are diabetic, 29 million women are affected by high blood sugar. The latest global figures on diabetes, released by the International Diabetes Federation (IDF), has raised a serious alarm for India by saying that nearly 52% of Indians aren't aware that they are suffering from high blood sugar[1]. India is presently home to 62 million diabetics .Diabetes mellitus is metabolic disorder by inability of pancreas to produce blood glucose. Hyperglycemia is a condition that occurs when blood glucose level goes high. Early signs of hyperglycemia in diabetics include Headaches, Difficulty concentrating, Blurred vision, frequent urination, Fatigue, (weak tired feeling), Weight loss, Blood sugar more than 180 mg/dl. Prolonged hyperglycemia in diabetes may result in: neuropathies, nephropathies and retinopathies, Damage of Nerve, Damage to eyes, blood vessels or kidneys. The improper blood glucose level leads to diseases such as cardio vascular, Retinal, Renal nervous disorder.[2] When a person with diabetes has hyperglycemia frequently or for long periods of time as indicated by a high HbA1c blood test, damage to nerves, blood vessels, and other body organs can occur. Hyperglycemia may also lead to more serious conditions including ketoacidosismostly in people with type 1 diabetes and hyperglycemic hyperosmolar nonketotic syndrome (HHNS) in people with type 2 diabetes or in people at risk for type 2 diabetes.

2. State of Predicting Hyperglycemia:

Hyperglycemia prediction in diabetic patients is attracted by researchers. Signs of hyperglycemia in diabetics may include: increased thirst, blurred vision, fatigue, weight loss[3]. Hyperglycemia predicted by best knowledge has used continuous glucose monitoring using data mining techniques. Few techniques have been developed to predict hyperglycemia in diabetic patients. ARIMA model is used for the prediction of future blood glucose[1]. METABO is a diabetes monitoring system which aims at recording and interpreting patient condition as well as provides decision support for patient. METABO system provides a multi parametric monitoring system which facilitates the efficiency and systematic recording of information[4]

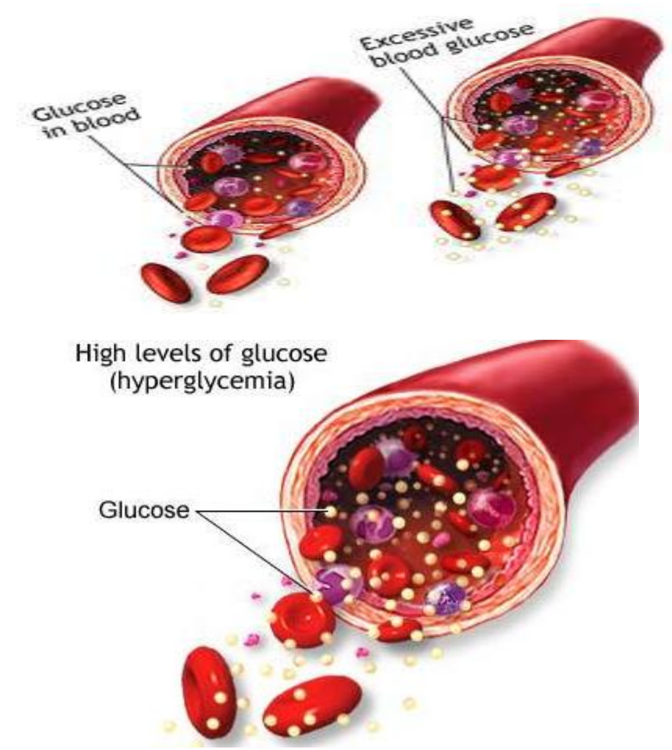


Fig: Normal blood glucose and Hyperglycemia diabetic in patient body.

3. Methodology:

A stochastic model that is extremely useful in the representation of certain practically occurring series is the Auto Regressive model. In this model, the current value of the process is expressed as a linear aggregate of previous values of the process. Another kind of model is the Moving Average model which depends on the previous deviations. To achieve greater flexibility in fitting of actual time series, it is advantageous to include both Auto Regressive and Moving Average terms in the model. Many time series data obtained practically are of non

International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.3 Issue No.5, pp : 668-670 1 May 2014

stationary in nature. ARIMA models are the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing and logging. ARIMA models are fine tuned versions of random walk and random trend models[2]. The fine tuning consists of adding lags of the differenced series and/or lags of the forecast errors to the prediction equation. The first step in fitting an ARIMA model is the determination of the order of differencing needed to stationarize the series. The optimal order of differencing is often the differencing at which the standard deviation is minimum.

4.Time series data:

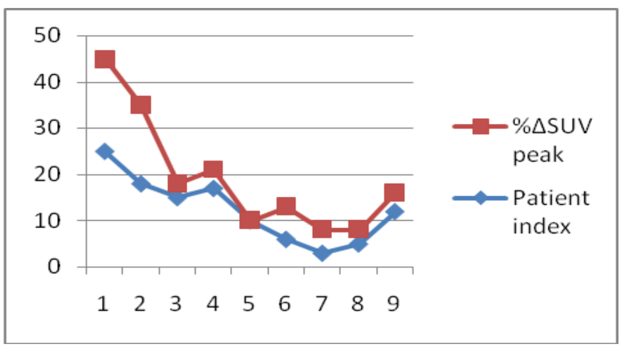
A **time series** is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals.Time series analysis includes methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series is the use of a model to predict future values based on previously observed values.

5. Auto regressive

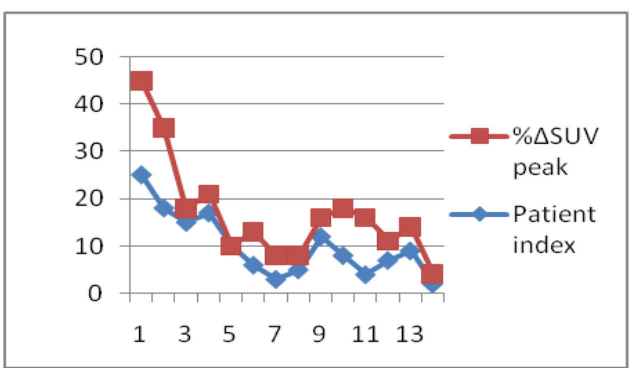
The glucose time series is described locally by an auto-regressive model of first-order (AR (1)), Corresponding to the following time-domain difference equation:

$$y_i = ay_{i-1} + w_i$$

Testing Results Using One Prediction Model



Leave-one-out cross validation testing results



Graph1 & 2:This model predicts patient glucose level by patient index.

The prediction strategy is as follows. Let θ denote the vector of the parameters of the model employed to describe the glucose time-series, i.e., θ = (a, σ2). At each sampling time tn, a new value of θ is first determined by fitting the model against past glucose data yn, yn-1,yn-2,... by weighted linear least squares. [6] The model is used to calculate the prediction of glucose level T stepsOnce θ is determined ahead, i.e., $\theta n{+}T$. Fora sampling interval of 3 min, a value of T equal to 10 or 15 corresponds to a PH equal to 30 or 45, respectively. The value $\theta n{+}T$is calculated iteratively for i = n+1, n+2, ...,n+T with wi=0.In determining the model parameters θ at a given time, all the past data yn, yn1,yn2,....,y1 participate, with different qualified weights[9][10][11]. A typical choice is to employ exponential weighting, i.e., μkis the weight of the sample taken k instants before the actual sampling time i.e., μkis the weight of the sample taken at time tn-k(k= 0,1,...,n-1). With μ, taken in the range (0,1), acts as a forgetting factor. If a forgetting factor is not used (which is equivalent to letting μ= 1), glucose samples collected tens of hours, if not days, before the actual sampling time would influence the prediction, with a possible corrosion of the algorithm capability to promptly track changes in the signal, in particular those due to perturbations, e.g., meals. From an algorithmic point of view, recursive least squares (RLS) implementations are possible in order to estimate the unknown model parameters θ in a computationally efficient manner.

6. ARIMA Modeling:

After stationarizing the data by preprocessingi.e., through regularization, the next step is to fitting in an ARIMA model. The more systematic way to do this through Auto correlation andPartial Auto correlation plots of the regularized data.[7] ACF plot is merely a bar chart of the coefficients of correlation between the time series and lags of itself. PACF plot is a plot of partial correlation coefficient between the series and lags of itself. The terms corresponding to exponential decline in ACF and peak in PACF would contribute to AR processes and Peak in ACF and exponential decline in PACF would contribute for MA processes. The next step is to determine the coefficients of model parameters by Maximum likelihood estimation. A conditional likelihood function is selected in order to get good starting point to contain an exact likelihood function.

Then the diagnosis check is carried out to validate the model. In successive trials the observation of the residuals obtained can helpto refine the structure of the functions in the model.An ARIMA model is generally given by

$$\Phi(B)\ g(t) = \theta(B)\varepsilon(t)$$

Where g(t) is the glucose level at time 't',Φ(B) and θ(B) are the parameters of AR and MA processes involved and ε(t) is the error term.Φ(B) and θ(B) are functions of backward shift operator

The Third order ARIMAmodel has been selected first through empirical approach and then confirmed with optimization. The prediction efficiency of the model has been validated initially with simulated data and then with five real life subjects" data who were using the Minimed MedtronicCGMdevice.

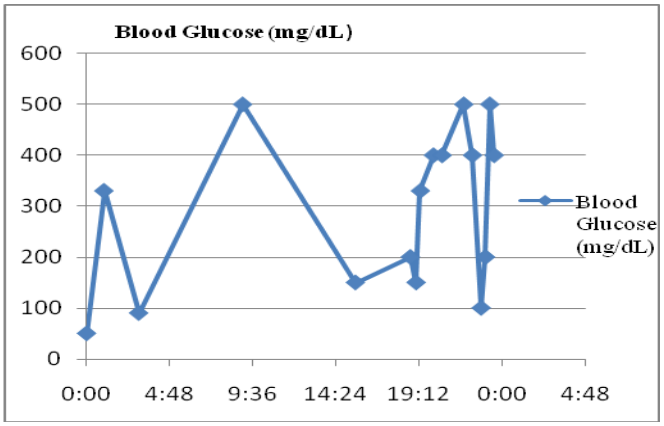


Fig: the prediction of glucose level on different days using this model

7. METABO system:

METABO is a diabetes monitoring system which aims at recording and interpreting patient condition as well as provides decision support for patient. METABO system provides a multi parametric monitoring system which facilitates the efficiency and systematic recording of information.

This system records continuous and discontinuous glucose emeasurements. This system is used to monitor hyperglycemia diabetes.[3]METABO modeling is used to predict blood glucose values and provide a decision support to patients.

The METABO system is responsible for analysis of data related to patient that may either directly affect the glucose levels or influence glucose control in long term. The metabolic modeling captures glucose with the aim to predict blood glucose values and provide alert and decision support to patient.

The glucose prediction model is able to predict high blood glucose levels over specific time horizon.

8. Conclusion:

The continuous glucose monitoring will improve in prediction of hyperglycemia diabetes and the alert the patients. A clinically important task in diabetes is the prevention of hyperglycemia events. ARIMA model is used for the prediction of future blood glucose so that the awaiting dangerous hyperglycemia can be incidental in advance and preventive measures can be taken.

9. References:

i. *C. Gireesh, V. Punna Rao,"Blood Glucose Prediction Algorithms for Hypoglycemicand/or Hyperglycemic Alerts", of IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 3, September 2012.*

ii. *International Diabetes federation, Public health foundation of India.http://ccebdm.org/news.php.*

iii. *Giovanni Sparacino, Andrea Facchinetti, and Claudio Cobell,"Smart Continuous Glucose Monitoring Sensors: On-Line Signal Processing Issues"Published online 2010 July 12. doi: 10.3390/s100706751*

iv. *S.Shanthi,Dr.D.Kumar,Prof.S.Varatharaj,S.Santhana Selvi,"Prediction of Hypo/Hyperglycemia through System*

v. *Identification, Modelingand Regularization of Ill-Posed Data", of International Journal of Computer Science & Emerging Technologies (E-ISSN: 2044-6004)171Volume 1, Issue 4, December2010.*

vi. *Eleni Georga, Vasilios Protopappas, Member, IEEE, Alejandra Guillen, Giuseppe Fico, Diego Ardigo, Maria Teresa Arredondo, Themis P. Exarchos, Student Member, IEEE, Demosthenes Polyzos,and Dimitrios I. Fotiadis, Senior Member, IEEE", Data Mining for Blood Glucose Prediction and KnowledgeDiscovery in Diabetic Patients: The METABO Diabetes Modelingand Management System", of 31st Annual International Conference of the IEEE EMBSMinneapolis, Minnesota, USA, September 2-6, 2009.*

vii. *D. Takahashi, Y. Xiao, F. Hu, and M. Lewis, "A survey of insulin dependent diabetes-part I: therapies and devices", J. of Telemedicine and Applications, Article ID 639019, vol. 2008, pp. 1-15, 2008.*

viii. *E. D. Lehmann, "Application of Information Technology in ClinicalDiabetes Care", Medical Informatics, vol.21, pp. 255-378, 1996.*

ix. *E. D. Lehmann, "Application of Information Technology in ClinicalDiabetes Care", Medical Informatics, vol. 22, pp. 1-120, 1997.*

x. *D. Takahashi, Y. Xiao, F. Hu, and M. Lewis, "A survey of insulindependent diabetes-part II: control methods", J. of Telemedicine and Applications, Article ID 739385, vol. 2008, pp. 1-14, 2008.*

xi. *E. D. Lehmann, and T. Deutch, "A physiological model of glucoseinsulin interaction in type 1 diabetes mellitus", J. Biomed. Eng., vol.14, pp. 235-242, May, 1992.*

xii. *W. A. Sandham, D. J. Hamilton, A. Japp, and K. Patterson, "Neuralnetwork and neuro-fuzzy systems for improving diabetes therapy", inProc. 20th Annu. Conf. IEEE Engineering in Medicine and BiologySociety, Hong Kong, China, 1998, vol. 20, no. 3, pp. 1438–1441.*

xiii. *Z. Trajanoski, and P. Wach, "Neural Predictive Controller for InsulinDelivery Using the Subcutaneous Route", IEEE Trans. BiomedicalEngineering, vol. 45, no. 9, pp. 1122-1134, Sept., 1998.*

xiv. *S. Mougiakakou, A. Prountzou, and K. Nikita, "A Real TimeSimulation Model of Glucose-Insulin Metabolism for Type 1 DiabetesPatients", in Proc. 27th Annu. Conf. IEEE Engineering in Medicineand Biology Society, Shanghai, China, 2005, pp. 298-301.*

xv. *H.T. Nguyen, N. Ghevondian, and T. Jones, "Neural-networkDetection of Hypoglycemic Episodes in Children with Type 1Diabetes using Physiological Parameters", in Proc. 28th Annu. Conf.*

xvi. *IEEE Engineering in Medicine and Biology Society, New York City,USA, 2006, pp. 6053-6056.*

xvii. *V. Tresp, T. Briegel, and J. Moody, "Neural Network Models for the Blood Glucose Metabolism of a Diabetic", IEEE Trans. NeuralNetworks, vol. 10, no. 5, pp. 1204-1213, Sept., 1999.*

xviii. *G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, and C. Cobelli, "Glucose Concentration can be Predicted Ahead in Time from Continuous Glucose Monitoring Sensor Time-Series", IEEETrans. Biomedical Engineering, vol. 54, no. 5, pp. 931-937, May,2007.*

xix. *P. Barnes, H. Tindall, C. Baynes, D. Hicks, K. McAuley, and M.Shukla "Diabetes Care Pathway", NHS, version 3, April 2008.*