International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 4
ISSN: 2454-4248
747 – 750

A Review on Predictive Analysis for DiabeticBlood Glucose and Reductionof over fitting inDiabetes using Deep Learning Neural Network

Ankitha D S[1], Deepika D V[2]
Department of Computer Science and Engineering
K.S. Institute of Technology, Bengaluru, India.
[1]*naga.ankitha@gmail.com*,[2] *deepika13.vijay@gmail.com*

Pradeep K R[3]
Assistant professor, Department of Computer Science and Engineering, K.S. Institute of Technology, Bengaluru, India
[3]Pradeepkr22@gmail.com

***Abstract*—** In this research, a prediction system is developed for the illness of diabetes and dropout strategy is made use to minimize the issues of overfitting. The key idea is arbitrarily drop unit from neural network during preparing. Expectation of blood glucose levels Measured by continuous glucose observing gadgets, by utilizing clinical information. The certain rate of a patients in the data set take as a training data and test on the left-over portion of the patients, i.e., the machine need not re-calibrate on other patients in the data set.

Keywords- *Dropout, Data overfitting, Diabetes Prediction, Neural Network, Deep Learning (DL).*

*****

I. INTRODUCTION

Diabetes is a chronic disease caused due to abnormally high levels of sugar glucose in the blood. Diabetes is usually referred as Diabetes mellitus. Diabetes is due to one of two mechanisms, Insufficient production of insulin (which is made by pancreas and brings down blood glucose), or Insufficient sensitivity of cells to the activity of insulin.

Diabetes is grouped into two types namely[1], Type I and Type II diabetes. In Type I diabetes, is a chronic condition in which the pancreas produce little or no insulin, which is also known as insulin-dependent diabetes. In type II diabetes, the human body cannot use insulin the right way, which is also termed as non insulin-dependent diabetes.

The idea of deep learning (DL) is a quickly developing one which is overflowing with thoughts as of recent years. Deep learning techniques are used in various fields, including medical field and optimal character acknowledgment [2]. The strategies of deep learning, in particular - deep learning neural system, to propose a model for diabetes forecast with high exactness. Deep neural systems contain various non-straight concealed layers and this makes them extremely expressive models that can learn exceptionally convoluted connections between their data sources and yields. With limited training information, however, many of this complicated blood relationship will be the resultant role of sampling noise, so they will exist in the training set but not in real test data even if it is drawn from the same distribution [3]. This leads to overfitting and many method acting have been developed for reduction.

II. RELATED WORK

In this paper, the outcome of a previous methods as well as outcome of proposed methods are discussed. The outcome of the proposed method is more accurate and précised when compared to outcome of previous method. Smith et al. utilized the perceptron based calculation called ADAPtive learning routine (ADAP), which is an early neural system demonstrate, to build up a diabetes expectation display for estimating the entry of diabetes mellitus. The framework's execution measures were finished utilizing standard clinical benchmarks as specificity and affectability. The outcomes acquired were then contrasted and those secured from applying direct per-captor models and calculated relapse [3].

The three neural network structures, such has multilayer perceptron (MLP), general regression neural network (GRNN), and radial basis function (RBF) are proposed by Kayaer and Yıldırım and they utilized the same data set to evaluate these three models. The performance gained by employing MLP was better than that of RBF method for all spread values tried. Among the them, GRNN was able to provide the finest result on the test data [4].

III. PROPOSED METHOD

Square graph of the proposed strategy is laid out in Figure. 1. Here, the process begins by entering information into the Input layer. At that point there are two completely associated Layer which is followed by dropout layer. At last result is yeild from yeild layer (i.e. output layer) with a single hub[5]. Together these layer build a multilayer perceptron. These method is used to reduce the overfitting in diabetes predication.

 ISSN: 2454-4248
747 – 750

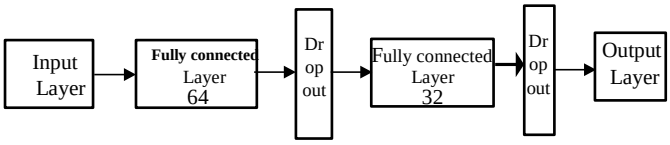
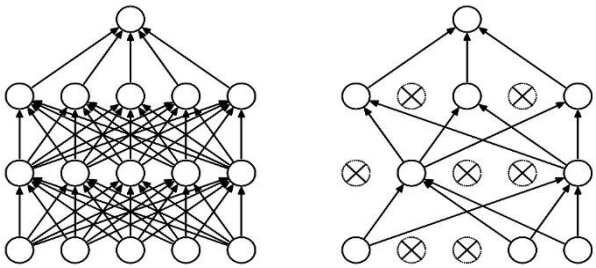


Figure 1:- Description of Proposed Method.

A. *DROPOUT*

Dropout is a shape of regularization. The term dropout means dropping out units (hidden and visible) in a neural network [5]. By dropping a unit out, it means temporarily removing it from the network, along with all its incoming and outgoing connections, as shown in Figure 2.



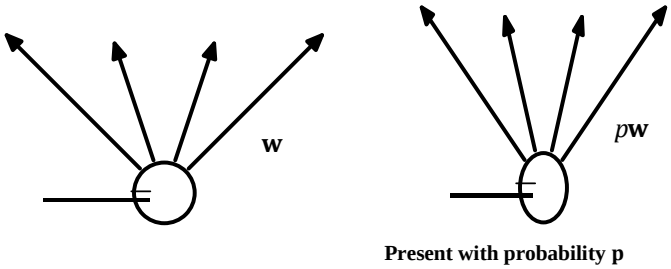
(a) Standard Neural Net (b) After Applying Dropout

Figure 2:- A demonstration of dropout, adapted from the method of Srivastava et.al. [5].

The decision of which units to drop is arbitrary. In the least difficult case, every unit is held with a fixed probability p free of different units, where p can be picked utilizing a validation set or can basically be set at 0.5, which is by all accounts near ideal for a wide range of networks and undertakings [5]. For the information units, be that as it may, the ideal likelihood of maintenance is usually nearer to 1 than to 0.5.

Applying dropout to a neural system adds up to testing a thinned" organize from it. The diminished system comprises of the considerable number of units that survived dropout (Figure 2b). A neural net with n units, can be viewed as an accumulation of 2n conceivable diminished neural networks. These organizes all offer weights so the aggregate number of parameters is still O (n2), or less [6ss]. For every introduction of each preparation case, another diminished system is tested and trained. So preparing a neural system with dropout can be viewed as preparing a gathering of 2n thinned systems with broad weight sharing, where each diminished system gets trained very infrequently, if by any stretch of the imagination.

At test time, it is not practical to unequivocally normal the forecasts from exponentially numerous diminished models. Notwithstanding, an extremely basic surmised averaging technique functions admirably in rehearse. The thought is to utilize a solitary neural net at test time without dropout. The weights of this system are downsized adaptations of the Prepared weights. In the event that a unit is held with likelihood p amid preparing, the active weights of that unit are duplicated by p at test time as appeared in Figure 3. This guarantees for any concealed unit the normal yield (under the conveyance used to drop units at preparing time) is the same as the genuine yield at test time [6]. By doing this scaling, 2n systems with shared weights can be joined into a solitary neural system to be utilized at test time. It is found that preparation a system with dropout and utilizing this surmised averaging technique at test time prompts signicantly bring down speculation mistake on a wide assortment of classication issues contrasted with preparing with other regularization techniques.



a) At training time (b) At test time

Figure3:- Left: A unit at training time that is present with probability *p* and is connected to units in the next layer with weights **w**. **Right**: At test time, the unit is always present and the weights are multiplied by *p*. The output at test time is same as the expected output at training time.

IV. BLOOD GLUCOSE PREDICTION

The proposed strategy speaks to semi-directed discovering that takes after an altogether different come nearer from those depicted previously. It is not a traditional measurement-based approach. Rather, it depends on work estimate on information defined manifolds, utilizing diffusion polynomials [7]. In this area deep learning technique which comprises of two layers are described. Given the time series $\{s_p(t_j)\}$ of BG levels at time t_j for each patient p in the patient set P, where $t_j - t_{j-1}$ = 5 min, it is started by formatting the data into the form $\{(x_j, y_j)\}$, where

$$x_j = \left(s_p(t_{j-d+1}), \ldots\ldots, s_p(t_j)\right) \in R^d \text{ and}$$

$$y_j = s_p(t_{j+m}) \in R \text{ for all patients } p \in P \quad \textbf{(1)}$$

The notation used are

$$P := \{ x_j = \left(s_p(t_{j-d+1}), \ldots, s_p(t_j)\right) : p \in P \} \quad \textbf{(2)}$$

Likewise, build the diffusion framework from P. This is finished by normalizing the lines of the weight grid W_N^{ε} .

 ISSN: 2454-4248
747 – 750

Having defined the information P and the comparing diffusion lattice, the strategy proceeds as follows.

A. First Layer: Three Networks in Different Clusters

To begin the first layer training, frame the preparation understanding set TP by haphazardly choosing (as indicated by a uniform likelihood circulation) M% of the patients in P [7]. The training information are currently defined to be every one of the information (x_j, y_j) comparing to the patients in TP. The notations utilised are

$$C: = \{ x_j = (s_p(t_{j-d+1}),\ldots,s_p(t_j)): p \in TP \} \quad (3)$$

And

$$C^*: = \{(x_j,y_j) = ((s_p(t_{j-d+1}),\ldots, s_p(t_j)), s_p(t_{j+m})) : p \in TP\} \quad (4)$$

Next, the short-term prediction $L_{Xj}(t_{j+1})$ of the BG level $s_p(t_{j+1})$ after 5min, for all the given measurements $x_j \in C$, by applying the linear indicator technique. In view of these 5-min predictions, the measurements in C are partitioned into three groups C_o, C_e and C_r,

$C_o = \{ x_j \in C: 0 \le L_{Xj}(t_{j+1}) \le 70\}$ (hypoglycaemia (5)

$C_e = \{ x_j \in C: 70 < L_{Xj}(t_{j+1}) \le 180\}$ (euglycemia) (6)

$C_r = \{ x_j \in C: 180 < L_{Xj}(t_{j+1}) \le 450\}$ (hyperglycaemia) (7)

With

$$C_l^* = \{(x_j, y_j): x_j \in C_l\}, \ell \in \{o, e, r\} \quad (8)$$

The inspiration of this progression is to assemble more data concerning the training set to guarantee more precise predictions in each BG range—as noted beforehand, results of prediction error in the different BG ranges are extremely different.

After obtaining the three clusters C_o^*, C_e^* and C_r^*, the three predictors are computed

$$f_o(x): = S(C_o^*,x), \quad f_e(x): = S(C_e^*,x) \quad (9)$$

$$f_r(x): = S(C_r^*, x), \text{ for all } x \in P \quad (10)$$

As well as the "judge" predictor, based on the entire training set C^*,

$$f_J(x): = S(C^*,x), x \in P \quad (11)$$

B. Second Layer (Judge): Final Output

In the last training layer, a final yield is delivered in view of which f_ℓ, $\ell \in \{o, e, r\}$ gives the best situation in the PRED-EGA lattice, utilizing f_J as the reference esteem. The PRED-EGA network grid is developed by utilizing correlations of f_o (resp., f_e and f_r) with the reference esteem f_J—specifically, it includes comparing

$$f_o(x_j) \ (\text{resp.}, f_e(x_j) \text{and} f_r(x_j)) \text{ with } f_J(x_j) \quad (12)$$

as well as the rates of change

$$\frac{f_o(x_j+1) - f_o(x_j-1)}{2(t_j+1-t_j-1)} \left(\text{resp.}, \frac{f_e(x_{j+1}) - f_e(x_{j-1})}{2(t_{j+1}-t_{j-1})} \text{ and } \frac{f_r(x_{j+1}) - f_r(x_{j-1})}{2(x_{j+1}-t_{j-1})} \right)$$

$$\text{with } \frac{f_e(x_{j+1}) - f_e(x_{j-1})}{2(t_{j+1}-t_{j-1})} \quad (13)$$

C. Algorithm Deep Network for BG prediction

Input:

Time series $\{s_p(t_j)\}$, $p \in P$, formatted as $P = \{x_j\}$ with

$x_j = (s_p(t_{j-d+1}),\ldots\ldots,s_p(t_j))$ and $y_j = s_p(t_{j+m})$, and corresponding diffusion matrix

$d \in N$ (specifies sampling horizon), $m \in N$ (specifies prediction horizon)

$M \in (0,100)$ (percentage of data used for training).

Let TP contain M% of patients from P (drawn according to uniform prob. distr.)

Set $C = \{x_j\}$ and $C^* = \{(x_j, y_j)\}$ for all patients $p \in TP$

Output:

Prediction $f(x_j) \approx s_p(t_{j+m})$ for $x_j \in P$.

First layer:

For $x_j \in C$ do

Make 5-min prediction $L_{Xj}(t_{j+1})$

end

Set $C_o = \{ x_j \in C: 0 \le L_{Xj}(t_{j+1}) \le 70\}$

Set $C_e = \{ x_j \in C: 70 < L_{Xj}(t_{j+1}) \le 180\}$

Set $C_r = \{ x_j \in C: 180 < L_{Xj}(t_{j+1}) \le 450\}$

Set $C_l^* = \{(x_j, y_j): x_j \in C_l\}$, $\ell \in \{o, e, r\}$.

for $x_j \in P$ do

for $\ell \in \{o, e, r\}$ do

Compute $f_l(x_j) = S(C_l^*, x_j)$

end

Compute $f_J(x_j) = S(C^*, x_j)$

end

Second layer:

for $x_j \in P$ do

for $\ell \in \{o, e, r\}$ do

Construct PRED-EGA grid $f_l(x_j)$ for using $f_j(x_j)$ as reference value

end

Let $\ell \in \{o, e, r\}$ denote the subscript for which f_{lj} produces the best PRED-EGA placement

 ISSN: 2454-4248
747 – 750

Set final output f ($\mathbf{x}_j$) = $\mathbf{f}_{lj}$

End

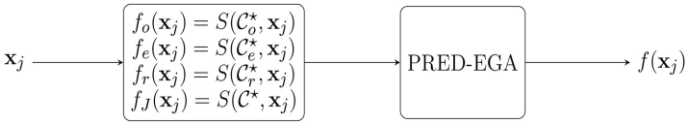


Figure 4:-Flowchart diagram: Deep Network for BG prediction.

A summary of the procedure is given in Algorithm. A flowchart diagram of the algorithm is shown in figure 4.

V. CONCLUSIONS

In this paper the utilization of dropout technique is proposed so as to decrease data overfitting in prescient model. This model is utilized for estimating the sickness of diabetes. A novel type of deep neural network for diabetes visualization with expanded exactness is examined. In this way, the accuracy achieved is 88.41% over the PID Data Set. By reducing the impact of overfitting in the proposed demonstrate, expanded Precision is accomplished by means of experimentation. The prediction of blood glucose levels based on continuous glucose monitoring system using deep learning neural network is also suggested. This also illustrate a case where domain information can be utilized to manufacture a proper compositional structure, prompting a miserly profound learning plan.

ACKNOWLEDGEMENT

We would like to express our great attitude to R&D, Department of CSE, KSIT, Bengaluru for their constant guidance and encouragement.

REFERENCES

[1] Alberti, K. G. M. M., and Zimmet, P. F.: “Definition, Diagnosis and Classification of Diabetes Mellitus and Its Complications. Part 1: Diagnosis and Classification of Diabetes Mellitus,” In: Provisional report of a WHO consultation. Diabetic medicine, 15(7), 539-553 (1998).

[2] Nielsen, M. A.”Neural Networks and Deep Learning,” http://neuralnetworksanddeeplearning.com,last

[3] Smith, J. W., Everhart, J., Dickson, W., Knowler W., and Johannes, R.” Using the Adap Learning Algorithm to Forecast the Onset of Diabetes Mellitus In: Proceedings of the An-nual Symposium on Computer Application in Medical Care,” American Medical Informatics Association, p. 261 (1988).

[4] Kayaer, K. and Yıldırım, T s” Medical Diagnosis on Pima Indian Diabetes Using General Regression Neural Networks” In: Proceedings of the International Conference on Artificial Neural Networks and Neural Information Processing, p (2003).

[5] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.”Dropout: A Simple Way to Prevent Neural Networks from Overfitting” In: Journal of Machine Learning Research, vol. 15, no. 1,(2014).

[6] Wager, S. Wang, and P. Liang. “Dropout training as adaptive regularization” In Advancesin Neural Information Processing Systems 26, pages 351{359, 2013.

[7] MontufarGF,PascanuR,ChoK,BengioY”Onthenumberofline arregionsof deepneural networks” In: Advances in Neural Information Processing Systems. Red Hook, NY: Curran Associates, Inc. (2014) p

[8] Telgarsky M.”Representation benefits of deep feedforward networks” arXiv preprint (2015) arXiv:150908101.

[9] Naumova V, Pereverzyev SV, Sivananthan S. “A meta-learning approach to the regularized learning - Case study: blood glucose prediction. Neural Network” (2012) 33:181–93. doi: 10.1016/j.neunet.2012.05.004

[10] Sivananthan S, Naumova V, Man CD, Facchinetti A, Renard E, Cobelli C, et al. “Assessment of blood glucose predictors: the prediction-error grid analysis” Diabet Technol Ther. (2011) 13:787–96. doi: 10.1089/dia. 2011.0033 Sivananthan S, Naumova V, Man CD, Facchinetti A, Renard E, Cobelli C, et al.” Assessment of blood glucose predictors: the prediction-error grid analysis” Diabet Technol Ther. (2011) 13:787–96. doi: 10.1089/dia. 2011.0033

[11] Sparacino G, Zanderigo F, Corazza S, Maran A, Facchinetti A, Cobelli C.”Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series” IEEE Trans Biomed Eng. (2007) 54:931–7. doi: 10.1109/TBME.2006.889774

[12] Mhaskar HN, Naumova V, Pereverzyev SV.”Filtered Legendre expansion method for numerical differentiation at the boundary point with application to blood glucose predictions” Appl Math Comput. (2013) 224:835–47. doi: 10.1016/j.amc.2013.09.015

[13] http://direcnet.jaeb.org/Studies.aspx DirecNetCentralss Laboratory. (2005).