Risk Analysis of Diabetes using IoT and Deep Learning

Sujaritha.M, Monica Murugesan, Bhuvana MK, Saleekha

Abstract: Diabetes mellitus is a most common disease faced by most of patients can have uncontrollable glucose level can lead to chronic disease to prevent this risk of higher chances of chronic diseases which can be implemented using Internet of Things(IoT) which is applied in various areas for solving problems of healthcare involved in monitoring and diagnosis of different parts of body using wearables or biosensors. In the proposed system, IoT devices and cloud technologies are connected to transfer data and execute the decisions on well-defined rules and deep learning technique is applied on diabetes data to decide the risk of diabetic patient which is solved by defining rules, system can understand the which data lies under which partition and knowledge representation can be made using the result the system can decide whether to suggest lifestyle modifications or proper in-take medication for improving their health and reduce adverse reactions in other parts of body or preventing to cause psychological effects.

***Index Terms*: Diabetes Mellitus, Deep Learning, Internet of Things, lifestyle modifications**

I. INTRODUCTION

Almost 382 million people has diabetes over the world. Diabetes known as diabetes mellitus a cluster of chronic illness which is faced due to increase in level of blood glucose level and decreased insulin level in body. Symptoms are polyuria-frequent urination, polyphagia high appetite, polydipsia-increased intake of water. 3 types of Diabetes names are Type -1, Type-2 and Gestational Diabetes.

Type-1 Diabetes is also called juvenile diabetes mellitus where patient suffers right from childhood as pancreas cannot produce sufficient insulin hence intake of insulin and diabetic medication as per doctor's advice is must. In rare cases, people may suffer from secondary diabetes which is same as type-1 which doesn't affect beta cells but affects immune system by some disease which affects pancreas. The destruction of beta cells prevents entry of glucose into blood without insulin hence it piles up in blood causes rise in blood sugar levels. Patients of Type -1 Diabetes go through Diabetes Ketoacidoisis in which body cannot store glucose hence converts fat cells in the form of ketones.

Hence it is combination of Ketones and glucose in the blood. Weight Loss: Patients constantly urinates or flushes calories from body hence their body weight gets reduced.

Revised Manuscript Received on May 05, 2019.

Sujaritha. M, Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India

Monica Murugesan, Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India

Bhuvana MK, Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India

Saleekha, Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India

Damage to other parts: Over the age, Patients suffering from this type of diabetes can affect various parts of body like heart mostly thickens the arteries ultimately causing to artherosclerosis, kidney, nerves, even affects eyes.

Diabetes of Type-2 is diabetes caused due to sedentary habits, improper eating etc. All over the world almost 95% suffer from Type-2 Diabetes and it is a non-insulin diabetes. But in this type the insulin production is high as compared to rather one. Patients can suffer from Stroke, Retinal problems-Glaucoma or Cataract, Sepsis, Artherosclerosis.

Gestational Diabetes is a diabetes caused during pregnancy and disappears after child-birth. Usually it occurs in second but appears mostly in later phases. Patient faces this type of diabetes as their body cannot generate insulin which ensures balancing glucose levels in blood during pregnancy.

Sometimes this diabetes can harm baby also hence it is necessary to diagnose the risk as earlier as possible. Pregnant women can have gestational diabetes with possible criterias-BMI above 30, baby whose weight exceeds 4.5, it can be heredity as if patient's mother had means patient can also get diabetes. if patient is of Asian, African or Arabic etc.

Sometimes these types of diabetes can lead to chronic diseases like sepsis, cerebrovascular diseases and cardiovascular diseases, kidney diseases, retinal diseases etc. If the patient's sugar level was constantly high like within range of 200-500 mg/dl or in terms of HbA1C along with various factors like improper diet management and hygiene, heavy smoking/alcohol intake, frequent urination etc. Person is likely to face one of the chronic illnesses Hence the system will detect risk associated to chronic illness with sensor values and will decide the patient undergo diet modification treatment or diabetic medication.

II. LITERATURE SURVEY

Aljumah, Abdullah A et.al, developed system which generate patterns associated with the existence of diabetes using regression technique of data mining which is done using Oracle Data Mining .The source of dataset is from NCD(Non Communicable Disease) World Health Organization in Saudi Arabia also used in analyzing different treatment which can be applied and effective for various ranges of age groups based on different factors .Age groups were classified as younger and older. The result obtained as drug treatment is ineffective for younger groups as compared to older groups[1].

Poorejbari, S.et al; had designed a system which aims to treat and manage diabetes by using cloud computing approach.

Is designed for patients for making it usable at home as well as hospital for diagnosing type-2 diabetes. Survey suggests that type-2 diabetes can get cured with help of lifestyle changes .It uses three main components:- The necessary information from patients gathered by Home Context Manager along with responds patient's requests, Hospital Environment-accessed by nurse or doctors in hospital and Cloud Infrastructure[2].

Abdalraouf H et al; had a system which performs analysis on unstructured data using Recurrent Neural Network and Convolutional Neural Network, the earlier neural network can extract high features but for huge images it requires multiple convolution and stacking of these pixels hence RNN is used along with it which is possibly done by using unsupervised neural language which helps training deep learning algorithm then parameterizes then framework parameterizes with previous information combines with set of maps learned by convolutional layer via long short term memory. The performance varied over different datasets for Standford Review it was 93.3% and Sentiment Trademark 48.8% compressed but accuracy was 89.2%. It has a significant role in terms of parameters, generating convolutional layer as well as pooling layer[3].

Jaana L et al; developed a methodology of predicting risk of diabetes placed on diabetes risk score to prevent risk associated with type -2 diabetes which causes various chronic diseases in patients. A sample was produced which were of 35 years and 64 years diagnosed with type-2 diabetes and analysis were done using multi-variant logistic regression ultimately assigns a parameter with score. This Risk Score was nothing but sum of various variables proposed 2 years before. The risk score ranged between 0 to 20 drug treated diabetes >=9, sensitivity of 0.78 and 0.81 and specificity of 0.77 and 0.76 predictive values 0.13.[4].

Alper Kursat U et al; developed a system using genetic algorithm oriented semantic features which are used for generalizing text in text classification algorithms and comprises of selecting features from text. First stage of computation is carried out using state of the art algorithm. Second stage is carried out using latent semantic indexing entitled by genetic algorithm but outcome of these where corresponded to larger values[5].

Yu Cheng et al; developed a risk prediction system which just gathers and stores in cloud i.e. preserving records in electronic form and such data preserving techniques, deriving how we can collect patterns from data out of which EHR(Electronic Health Record) is the successful technique. Deep learning technique has been applied over the data of EHR by assigning one dimension in a matrix and another dimension in another. Second layer is derived as a result of convolution of first layer. Third layer is max pool layer which defines sparsity and Fourth layer is softmax prediction layer over which one can extract patterns[6].

Gauri D et al developed a model called predictive risk which integrates various data analysis techniques like data mining, machine learning and statistics as it predicts future with current data in the current case it might predict risk associated with diabetes. The dataset used is Pima-Indian Dataset and algorithm which has been used is machine learning in Hadoop MapReduce in order to find missed values and prediction is done on it[7].

Sneha J et al performed prediction on Diabetes dataset using Back propogation Neural Network which has 8 input and 1 output layer, one hidden layer with 10 neurons which was developed using MATLAB can get accurate results with help of GUI of patient's diagnosis. The accuracy rate is 81%[8].

Kanimozhi Selvi et al developed a system which first trains the data then later classifies it. The dataset used was diabetes. The dataset is loaded into Hive and then classified using Naïve Bayes algorithm. They had just analysed but further information prediction was further carried out[9].

Kazemi, M et al; designed system details of health status of patients using convenience sampling methods. Later on NDS(Neuropathy Disability Score) was administered on six hundred patients. Based upon the score severity was classified. With SVM method, one-against-all, one- against-one, three-kernel functions and radial basis were used to predict class of dataset. Synthetic minority oversampling technique is used to improve the outcome[10].

Parisa P et al; a system is designed which ensures balanced diet to be followed by a person in order to prevent high calories causing obesity leads to chronic diseases so it will accept the image of food and performs convolutional neural network on it recognizes food portion and its content and its accuracy came out as 99% [11].Senthil Kumar et al, has designed system for decision making based on medical datasets-heart disease, diabetes and hepatitis diabetes retrieved from UC-Irvine and objective of system is to generalize and differentiate the values provided by various algorithms such as CART, C4.5 and ID3.In terms of complexity, C4.5 performs better than ID3 and CART performs better than all in terms of efficiency[12].

William S et al performed analysis using Artifical Neural Networks which is done using past training datasets with which one can predict the future risk of the disease also along can gather and predict the person's diet, exercise etc. Prediction with ANN is done using insulin, diet, exercise as well as glucose vector and X which contains data of stress or something vector. In this algorithm it works based on trial and error method and has 97 events [13].

Diego C et al developed a device which gathers reading of events performed by humans like eating, sitting jogging and running and analysis is performed over it using machine learning technique and its accuracy rate was 95.53%[14].

Mithun M et al(2017); developed a diet suggestion system for diabetic patients by acquiring list of food items in a definite ratio as a result of arranging a healthy diet which assures least variations in blood glucose levels which helps in managing their glucose levels. The system uses two important factors Glycemic Index and Glycemic Load. Glycemic Load is defined in terms of food items which are available to individual then framed in terms of linear programming to optimize the Glycemic Load. The function is fed to algorithm which suggests food items and predictions are discussed[15].

Irshad F et al(2014); designed a diet recommendation system for diabetic patients also has studied about semantic web details that are used in other systems OWL base ontology is base ontology whereas SWRL ontology derive rules for diet and exercise out of these recommendations it allows us to define new knowledge and it is integrated into system. These recommendations make a person's life more reliable[16].

III. MATH

There are many systems designed for diabetic patients nowadays from checking out blood glucose till providing an assistance for leading better life. In this system, data will be gathered using blood pressure sensor, ECG sensor and Foot pressure sensor attached with temperature sensor. The above sensor values will be used in symptom assessment which contains queries about whether patients have any other symptoms or bad habits. The symptom assessment data will be stored in a cloud server. A decision is made on the data stored in the cloud and retrieved values are proceed over analysis in order to make patients to visit doctor or need to visit dietician.

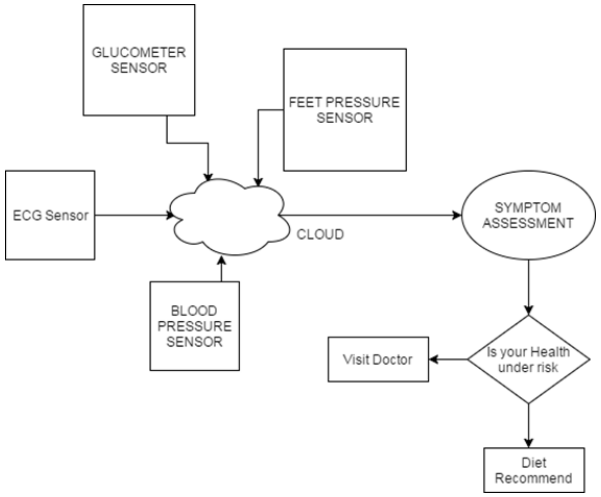


Fig. 1.1 Diabetes Risk Flow

The system is designed using Raspberry-Pi 3 with OS enabled 1.4GHz 64-bit quad-core processor, dual-band wireless LAN, Bluetooth 4.2 and Power-over-Ethernet support.

A. Sensors

Sensors used are glucometer sensor, blood pressure sensors, ECG sensors, Feet Pressure Sensors.

- Glucometer Sensor: Glucometer Sensor allows user to monitor their sugar levels either in fasting or post-lunch.
- Blood Pressure Sensor: Blood Pressure Sensor allows to gather reading blood pressure of user in order to determine existence of high blood pressure.
- Feet Pressure Sensor: Feet Pressure Sensor allows to determine pressure in feet with respect to increased level of swelling and it is embedded with temperature sensor which monitors temperature of feet.
- ECG Sensor: ECG sensor monitors the heartbeat of diabetic patient.

B. Softwares

The system is programmed using python scripts to maintain records of tests, recommending for dietary changes or making an appointment for doctor.

- Python: Python is a interpreter object oriented programming language also allows to run the code in different platforms it is easy to execute as the number of lines is less. With python we can develop web application, embedded and mobile application too.
- Thingspeak: Thingspeak is an opensource free cloud where we can upload, view and download the data.
- Bootstrap: Bootstrap is the most famous and easiest web development framework which allows one to develop a stylish website using HTML, CSS, JavaScript .It is a powerful front end framework and is divided into Basic, CSS, Layout, Components and Plugins and at Twitter Jacob Thornton and Mark Otto developed it and on GitHub in August 2011 it was released as an open source product.
- PHP: PHP is a Hypertext Preprocessor and it is a server scripting language it is an open source used for web development and the main difference between PHP and client-side HTML code is nothing but it is executed on server and the generated HTML then sends to the client.
- MySQL: It is also known as RDBMS i.e. relational database management system. It is also open source and free. The Swedish company MySQL AB sponsored MYSQL and after that Sun Microsystems bought it.

C. Modules

- Data Collection and Storage: The system involves gathering data from sensors like and later stored in cloud over raspberry-pi.
- Algorithm Formulation: A model is designed for risk prediction with which can analyse danger in diabetic condition so a deep learning with optimal layer is necessary
- Need a Diet Recommendation: If the risk factor is low then the system will suggest patient to go for diet recommendation.
- Feedback: The system will ask patient about their improvement in health, if the status still remain same then system will provide certain tips to patient.

IV. RISK ANALYSIS

Risk analysis is nothing but helping the patients to manage their glucose levels with respect to diet and other routine ultimately can prevent patient from facing chronic diseases CVD, stroke, cellulits, sepsis. The dataset is real time data. This risk analysis model also helps in reduction of diseases that may arise in later stages of diabetes. Hence it is very importance to predict diabetes in early stage and take precautionary measures such that the blood glucose level and other related diseases are controlled which are more commonly seen in aged people but there are various other techniques used which are developed with Genetic Algorithm. In comparison with second one, first one is selected as it has capability to predict diseases.

There are many methods like data mining, neural networks, regression techniques etc.

So using Deep learning we can foretell the risk depends upon the nodes so here the nodes are sensors like blood pressure, ECG, glucose levels, feet pressure etc. Hence we can use Neural Network with minimal layers.

V. DEEP LEARNING

It is a part of machine learning which imitates the simulation of Human Brain. It uses this imitation for processing data and generating patterns which can be used in decision making and or unsupervised or unlabeled. Deep Learning comprises of neural network and it has layers of neurons or neural units which has behavior of human brain.

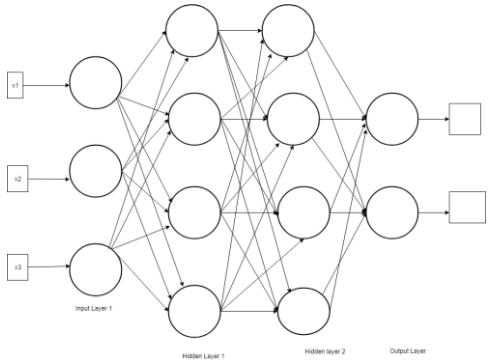


Fig. 1.2 Deep learning process

Deep Learning Algorithm's Types

1. Perceptron

Perceptron is the first generation of neural network which has one input layer of some appearance and output layer along with weights. The disadvantage its ability to handle linearly separable functions but unable to handle non-linear separable functions.

2. Multilayer Perceptron:

Multilayer Perceptron is next generation of neural network which has input layer and a output layer and hidden layer. It is a type of feedforward neural network In this algorithm every node works on activation function for training data it uses backpropogation technique. It distinguishes in such a way as if it appears as non-linearly separable.

y(vi)=tanh(vi) and y(vi)=(1+e-vi)-1….(1)

3. Back-propagation

Back propagation is an artificial neural network based on supervised learning approach. It has three layers: Input, Output in between layers and output present in invisible layer. Weights between input to Output in between layers are updated so that error occurs at each computation will be small. The limitations of back-propagation are: local minima, network paralysis and slow confluence.

4. Recurrent Neural Network

Recurrent Neural Network is another deep learning technique which generates edges of objects in an image by feeding the input in the within same layer but in another step rather than loading it in another layer. It is also used in identifying patterns from speech input or text files. The thought behind RNNs is to utilize successive data. RNNs are called intermittent in light of the fact that they play out a similar undertaking for each component of a grouping, with the yield being relied upon the past calculations. Another approach to consider RNNs is tied in with having a memory which catches data about what has been determined up until this point. A recurrent neural network and the unfurling in time of the calculation associated with its forward calculation.

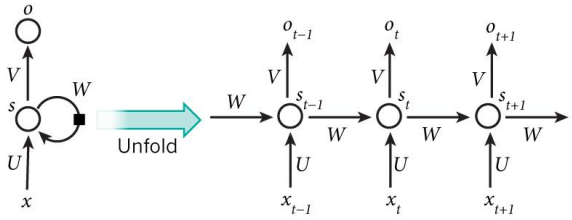


Fig. 1.3 RNN Computation

xt is the input at time step t. For example, x_1 could be a one-hot vector corresponding to the second word of a sentence.st is the hidden state at time step t. It's the "memory" of the network. s_t is calculated based on the previous hidden state and the input at the current step: s_t=f(Ux_t+ Ws_{t-1}). The function f usually is a nonlinearity such as tanh or ReLU. s_{-1}, which is required to calculate the first hidden state, is typically initialized to all zeroes. ot is the output at step t. For example, if we wanted to predict the next word in a sentence it would be a vector of probabilities across our vocabulary. o_t= softmax(Vs_t).

VI. TYPES OF RNN

- Bidirectional Recurrent Neural Network

Bidirectional RNNs depend on the possibility that the result at time t may not just rely upon the past components in the succession, yet in addition future components. For instance, to anticipate a missing word in a grouping you need to take a gander at both the right and the left context. Bidirectional RNNs are very basic. They are only two RNNs stacked over one another. The yield is then processed dependent on the hidden condition of both RNNs.

- Deep Bidirectional RNN

Deep (Bidirectional) RNNs are similar to Bidirectional RNNs, only then can have multiple layers per time step. In practice it can have higher learning capacity generally need a lot of training data.

- LSTM(Long Short Term Memory) Networks

LSTM systems are very mainstream nowadays and we quickly discussed them above. LSTMs don't have an in a general sense diverse design from RNNs, however they utilize an alternate capacity to process the hidden state. The memory in LSTMs are called cells and you can consider them black boxes that take as information the past state h_{t-1} and current input x_t. Inside these cells choose what to keep in (and what to eradicate from) memory. They at that point join the past express, the present memory, and the info. For reasons unknown, these sorts of units are effective at catching long haul conditions.

4. Artificial Neural Network

Artificial neuron network (ANN) is a kind of neural system whose calculation depends on the elements and structure of natural neural systems In this neural system the data flows through the system hence changes the structure of the ANN.

These associations are weighted; the higher the number the more prominent impact one unit has on another, like a human cerebrum. As the information experiences every unit the system is becoming familiar with the information. On the opposite side of the system is the yield units, and this is the place the system reacts to the information that it was given and handled.

VII. RESULT

The device has been designed and sensor data are stored and viewed on Thingspeak cloud. Using Weka tool observation has been got using Multilayer Perceptron Algorithm.

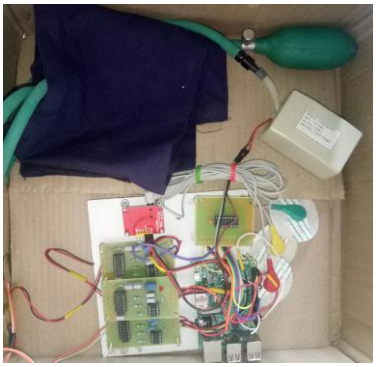


Fig. 1.4 Bio-sensor Connection

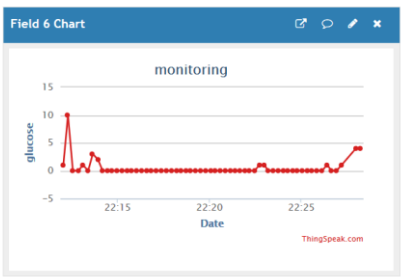


Fig. 1.5 Data in Cloud

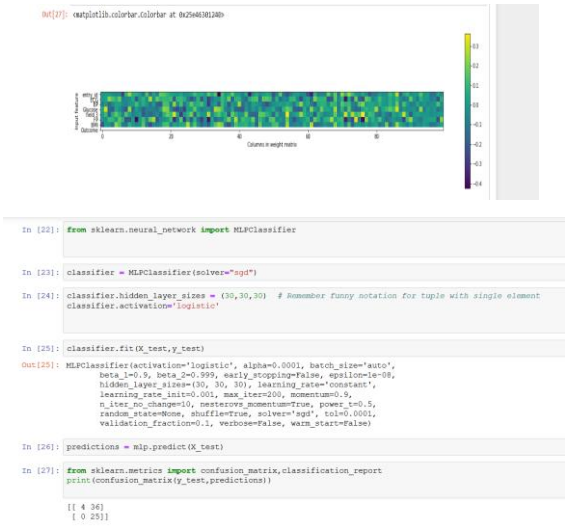
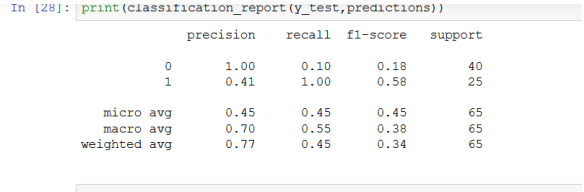


Fig. 1.6 Risk Analysis in Python



In [28]: `print(classification_report(y_test,predictions))`

	precision	recall	f1-score	support
0	1.00	0.10	0.18	40
1	0.41	1.00	0.58	25
micro avg	0.45	0.45	0.45	65
macro avg	0.70	0.55	0.38	65
weighted avg	0.77	0.45	0.34	65

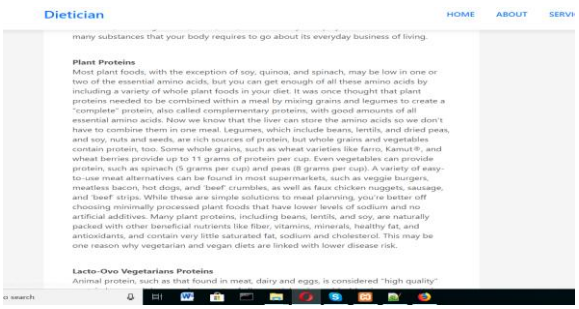


Fig. 1.7 Diet WebPage

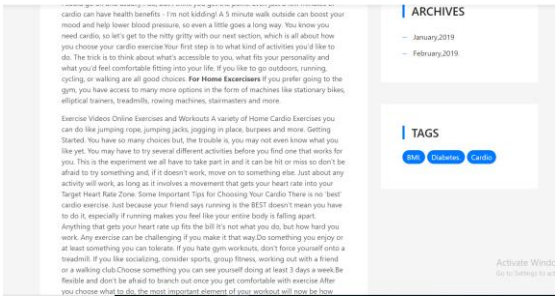


Fig. 1.8 Exercise

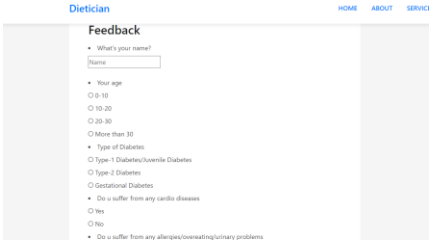


Fig. 1.9 Feedback

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The dataset has seven attributes –ECG, Blood Pressure, Feet Pressure, Glucose and BMI values. It was analysed using multilayer perceptron. The algorithm had one linear with 195 instances its accuracy was 0.65 for test 0.63. Total Number of instances are 260.During classification fit the average(micro average):0.45 and macro average is 0.70. The risk assessment is carried out using Python Programming in Jupyter notebook.

Risk is compared with BMI i.e.if a person is obese or overweight (>=25 or >=25 and <=30) then they have chances of high sugar and ultimately can have high risk of neurological problems like stroke, paralysis etc or can have cardiovascular problems. Hence it is necessary to take precautions hand in hand that is diet as well as medications properly.

But for those who are overweight and whose blood glucose is less or within range like (5.0 or less than 5.0 as per HbA1C) they can undergo dietary modifications and lifestyle changes.

REFERENCES

1. Aljumah, Abdullah A., Mohammed Gulam Ahamad, and Mohammad Khubeb Siddiqui. "Application of data mining: Diabetes health care in young and old patients."Journal of King Saud University-Computer and Information Sciences 25.2 (2013):127-136.
2. Poorejbari, S., Vahdat-Nejad, H., &Mansoor, W.(2017). Diabetes Patients Monitoring by Cloud Computing.In Cloud Computing Systems and Applications in Healthcare (pp. 99-116).IGI Global.
3. Hassan, A., & Mahmood, A. (2018). Convolutional Recurrent Deep Learning Model for Sentence Classification. IEEE Access, 6, 13949-13957.
4. Lindström, J., & Tuomilehto, J. (2003). The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes care, 26(3), 725-731.
5. Uysal, A. K., & Gunal, S. (2014). Text classification using genetic algorithm oriented latent semantic features. Expert Systems with Applications, 41(13), 5938-5947.
6. Cheng, Y., Wang, F., Zhang, P., & Hu, J. (2016, June). Risk prediction with electronic health records: A deep learning approach. In Proceedings of the 2016 SIAM International Conference on Data Mining (pp. 432-440). Society for Industrial and Applied Mathematics.
7. Kalyankar, G. D., Poojara, S. R., & Dharwadkar, N. V. (2017, February). Predictive analysis of diabetic patient data using machine learning and Hadoop. In I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2017 International Conference on (pp. 619-624). IEEE.
8. Joshi, S., & Borse, M. (2016, September). Detection and Prediction of Diabetes Mellitus Using Back-Propagation Neural Network. In Micro-Electronics and Telecommunication Engineering (ICMETE), 2016 International Conference on (pp. 110-113). IEEE.
9. D. Jayaprakash C.S. Kanimozhi Selvi, S.V. Kogilavani, S. Malliga, "Classification and Prediction of Diabetics using Weka and Hive Tool." International Journal of Advance Engineering and Research Development on pp(105-111) Vol 5
10. Kazemi, M., Moghimbeigi, A., Kiani, J., Mahjub, H., &Faradmal, J. (2016). Diabetic peripheral neuropathy class prediction by multicategory support vector machine model: a cross-sectional study. Epidemiology and health, 38.
11. Pouladzadeh, P., Kuhad, P., Peddi, S. V. B., Yassine, A., & Shirmohammadi, S. (2016, May). Food calorie measurement using deep learning neural network. In IEEE International Instrumentation and Measurement Technology Conference Proceedings (I2MTC) (pp. 1-6).
12. Kumar, D. S., Sathyadevi, G., & Sivanesh, S. (2011). Decision support system for medical diagnosis using data mining. International Journal of Computer Science Issues (IJCSI), 8(3), 147.
13. Sandham, W., Nikoletou, D., Hamilton, D., Paterson, K., Japp, A., & MacGregor, C. (1998, September). Blood glucose prediction for diabetes therapy using a recurrent artificial neural network. In Signal Processing Conference (EUSIPCO 1998), 9th European (pp. 1-4). IEEE.
14. Castro, D., Coral, W., Rodriguez, C., Cabra, J., & Colorado, J. (2017). Wearable-Based Human Activity Recognition Using an IoT Approach. Journal of Sensor and Actuator Networks, 6(4), 28.
15. Mohan Mithun K, Poonkuzhali S."A Diet Suggestion System for Diabetic Patients Based on Linear Programming." International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), on IEEE, 2017.
16. Faiz, I., Mukhtar, H., & Khan, S. (2014, October). An integrated approach of diet and exercise recommendations for diabetes patients. In 2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom) (pp. 537-542). IEEE.
17. Lee, Malrey, Thomas M. Gatton, and Keun-Kwang Lee. "A monitoring and advisory system for diabetes patient management using a rule-based method and KNN." Sensors 10.4 (2010): 3934-3953.