

# A MODEL FOR PREDICTING TYPE-II DIABETES USING MACHINE LEARNING APPROACH

NETRA PATIL

Research Scholar, Department of Computer Engineering  
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune.

DR. NAVEENKUMAR JAYAKUMAR

Associate Professor, Department of Computer Engineering  
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune.

**Abstract-** Diabetes Mellitus (DM) is a metabolic diseases group where the person will have high blood sugar due to the pancreas unable to produce sufficient insulin or the cell's which are not responding to the insulin produced. Diabetes is a chronic disease and a major public health challenge worldwide. The main drawback is that there is lack of awareness of the people on eating habits. In our country, diabetes patient counts have increased steadily due to this reason. Diabetes is a metabolic disease affecting a multitude of people worldwide. Its incidence rates are increasing alarmingly every year. If untreated, diabetes-related complications in many vital organs of the body may turn fatal. Early detection of diabetes is very important for timely treatment which can stop the disease progressing to such complications. RR-interval signals known as heart rate variability (HRV) signals (derived from electrocardiogram (ECG) signals) can be effectively used for the non-invasive detection of diabetes. This research paper presents a methodology for classification of diabetic and normal HRV signals using deep learning mechanism. We employ long short-term memory (LSTM), convolutional neural network (CNN) and its combinations for extracting complex temporal dynamic features of the input HRV data. These features are passed into support vector machine (SVM) for classification. We have obtained the performance improvement of 0.03% and 0.06% in CNN and CNN-LSTM architecture respectively compared to our earlier work without using SVM. The classification system proposed can help the clinicians to diagnose diabetes using ECG signals with a very high accuracy of 95.7%.

**Keywords:** Deep learning, Diabetes, Heart rate variability, ECG, CNN, LSTM.

## 1. INTRODUCTION

Diabetes mellitus (DM), commonly known as diabetes, is a group of metabolic disorders characterized by high blood sugar levels over a prolonged period. Symptoms of high blood sugar include frequent urination, increased thirst, and increased hunger. Machine learning is one of the most essential domains in the field of research with the goal of predicting and conducting a systematic review. According to Official World Health Organization data, India has the highest number of diabetes [1]. The total number of diabetics in India in the year 2017 is positioned as 31.7 million and by the year 2030, it is presumed to ascend to 79.4 million [2]. The disease, diabetes is chronic and has become one of the leading lifestyle ailments, characterized by prolonged elevated blood sugar levels. Failure of organs like liver, heart, kidneys, stomach, etc. is caused in the long run due to the effect of diabetes.

Classification of diabetes mellitus can be described as follows:

- i. Type I diabetes: The diabetic condition that depends on insulin occurs mainly in children and adolescents because of the genetic disorders.
- ii. Type II diabetes: Generally occurs in adults during the age of 40 years discernible by high blood sugar level.
- iii. Pregnancy diabetes: The diabetes that occur during the pregnancy period.
- iv. Diabetic retinopathy: This type of disorder leads to eye blindness.
- v. Diabetic neuropathy: Nerve disorder is the cause of this type of diabetes.

### 1.1 Factors Responsible for Diabetes:

- i. Combination of genetic susceptibility and environmental factor can cause diabetes.
- ii. Overweight may lead to cause diabetes in the long run.
- iii. If a parent or sibling has diabetes, then the risk is supposed to be increased.
- iv. Aging increases risk of diabetes.
- v. More than 140/90 mm of Hg is linked to an increased risk of diabetes.

vi. Low levels of high-density lipoprotein (HDL) are also the cause of occurring the risk.

### 1.2 Complications arise Due to Diabetes:

The complications progress moderately. Possible complications those are included for arising:

- i. Cardiovascular disease: Diabetes vividly increases the risk of various cardiovascular problems
- ii. Damage in the nerves (Neuropathy)
- iii. Damage in the kidneys (Nephropathy)
- iv. Damage in eyes (Retinopathy);
- v. Damage in foot: Deficient blood flow to the feet increases the risk;
- vi. Acute skin condition: Bacterial and fungal infections may happen;
- vii. Impairment of hearing: The problems of hearing are common;
- viii. Alzheimer's disease: Increases the chance of Alzheimer's disease.

Our aim to focus on Type-II diabetes, which is a lifestyle related disease. It is caused due to high level of insulin which causes resistance to insulin by the body cells which leads to increased sugar level in the blood. The research idea is not focusing on Type-I diabetes, where sufficient insulin is not secreted from pancreases of an individual.

### 1.3 Scope and Objective

#### Objectives:

Early prediction, providing alerts to non-diabetic human being and management for Type -II diabetic patients with proper care about drugs and lifestyle for adults throughout world, under proper guidance ,monitoring by preventive medicine practitioners. It improves the accuracy of predicting diabetes.

#### Scope:

Now days, diabetes is an immense challenge to the human health on earth. Improper lifestyle, eating habits like junk food, unawareness and/or ignorance about the health has contributed to the lifestyle diseases like Obesity, Type -II diabetes, B.P. etc The scope is limited only for adults throughout the world. Proposed system is applicable to all mankind for management of type-II diabetes, obesity and control. The proposed system can be used in real life applications like-- All medical centers, health care, Clinicians and Doctors.

### 1.4 Problem Definition

Diabetes Mellitus (DM) is a metabolic diseases group where the person will have high blood sugar due to the pancreas unable to produce sufficient insulin or the cell's which are not responding to the insulin produced. Diabetes is a chronic disease and a major public health challenge worldwide. The main drawback is that there is lack of awareness of the people on eating habits. In our country, diabetes patient counts have increased steadily due to this reason. Diabetes is a metabolic disease affecting a multitude of people worldwide. Its incidence rates are increasing alarmingly every year. If untreated, diabetes-related complications in many vital organs of the body may turn fatal. Early detection of diabetes is very important for timely treatment which can stop the disease progressing to such complications.. This research proposal presents a methodology for classification, progression modeling and effective drug administration of diabetic through machine learning mechanism. . The classification system proposed can help the clinicians to diagnose diabetes and provide model for administering drug. "Progression Modeling and Effective Drug Administration for Diabetes through Machine Learning".

## 2. CONCEPTUAL BACKGROUND

The data samples are partitioned into the target class in classification technique and the same is predicted for each and every data point. For instance, we may classify a patient as "higher risk" or "lower risk" on the premise of their diseases using the data classification methods. Some of habitually used techniques are discussed in the following.

### 2.1 Support Vector Machine (SVM)

Support vector machine [9] is a supervised learning technique and represents the dataset as points in n-dimensional space, n being the number of features. The purpose of SVM is to establish a hyper plane which partitions the datasets in different sorts and the hyper plane should be at utmost margin from the various sorts. For robustness, the hyper plane is needed to be chosen in such a way that it is having high margin and maximizes the distances between the nearest data point of either class.

Advantages:

- SVM provides better accuracy and can easily handle complex nonlinear data points.
- It removes the over fitting nature of the samples.

Disadvantages:

- It is difficult to use in large datasets.
- Execution is comparative slow.

### 2.2 k-Nearest Neighbors (k-NN)

In this classification technique, the anonymous data points are discovered using the familiar data points which are known as nearest neighbors. k-Nearest neighbors (k-NN) [5] is conceptually simple and is also called as lazy learning, where “k” is the nearest neighbor. In k-NN algorithm, the aim is to vigorously recognize k samples in the training dataset which are identical to a new sample.

Advantages:

- It is easy to implement.
- Training is done in a faster manner

Disadvantages:

- Time becomes prohibitive for finding the nearest neighbor in the training data which is of huge size, thus making it slow.
- It requires large storage space.
- The transparency of knowledge representation is very poor.

### 2.3 Decision Tree

In the decision tree technique [6], on the premise of parameters, a tree- or graph-like shape is constructed and it contains a predefined target variable. Traversing from root to leaf is done for the decision to take, and the traversing is done till the criteria are met.

Advantages:

- Domain knowledge is not required for the decision tree construction.
- Inexactness of complex decision is minimized which results to assign exact values to the outcome of various actions.
- Easy interpretations and it can handle both numerical and categorical data.

Disadvantages:

- One output attribute is restricted for decision tree.
- Decision tree is an unstable classifier.
- Categorical output is generated.

### 2.4 Random Forest

Random forest [7] is a classifier that constitutes a few decision trees and considered as one of the dimensionality reduction methods. It is one of the ensemble methods for classification, regression, and other terms. It can be used to rank the importance of variables.

Advantages:

- Random forest improves classification accuracy.
- It works well with the dataset of large number of input variables.

Disadvantages:

- Random Forest is fast to train but once trained, it becomes slow to create predictions.
- It is slow to evaluate.
- Interpretation is very hard.

## 2.5 Naïve Bayes Approach

Based on the Bayes' theorem, naïve Bayes is a supervised learning technique. To classify the text documents, it is one of the most successful known algorithms for learning because of its better outcomes in multi-class problems and rules of independence.

Advantages:

- Simple and easy to implement.
- More accuracy in result due to higher value of probability

Disadvantages:

- Strong assumption on the shape of data distribution.
- Loss of accuracy.

## 2.6 Artificial Neural Network

The artificial neural network [9] is aroused by the neural network of human being, and it is a combination of three layers, i.e., input layer, hidden layer, and output layer, which is also called as MLP (Multilayer Perceptron). The hidden layer is similar to neuron, and each hidden layer consists of probabilistic behavior.

Advantages:

- Ability to learn and model nonlinear and complex relationships.
- Ability to generalize the model and predict the unseen data.
- Resistant to partial damage.

Disadvantages:

- Optimizing the network can be challenging because of the number of parameters to be set in.
- For large neural networks, it requires high processing time.

## 2.7 Logistic Regression

Logistic regression is also known as log it model for dichotomic output variables and was comprised for classification prediction of diseases. It is a statistical method for analyzing. Here, one or more than one independent variables as certain the consequences.

Advantages:

- Easy to implement and very efficient to train.
- Can handle nonlinear effect and interaction effect.

Disadvantages:

- Cannot predict continuous outcomes.
- Vulnerable to overconfidence, i.e., the models can appear to have more predictive power than they actually do, resulting in over fitting.
- Requires quite a large sample size to achieve stable results.

## 3. RELATED WORKS

A lot of research has happened on the non-invasive automated detection of diabetes using machine learning techniques. Machine learning was employed based on steps of feature extraction, feature selection and classification. There were a variety of works which differed in what type of features was extracted and what classifiers were tried upon. It was further observed that the performance of traditional machine learning algorithms is not up to the acceptable level in crucial artificial intelligence problems of speech recognition and object recognition mainly because of the fact that the dimension of the data handled is high. The shortcomings of machine learning boosted the deep learning research. Deep learning also has its applications in healthcare. Lot of works has recently been published mainly in anomaly detection in the area of healthcare. Related to diabetes detection, used deep learning techniques to detect diabetes from the input HRV data with an accuracy value that closely matches with the maximum accuracy achieved for automated diabetes detection. Deep learning is a form of machine learning. Unlike in machine learning, feature extraction and classification are not explicitly done in deep learning networks. The hidden layers of the deep learning network do all these implicitly within itself without involving the external researcher. A short description of deep learning networks is given below.

### 3.1 Recurrent neural network (RNN)

Recurrent neural network (RNN) is capable of extracting dynamic temporal behaviour from an input time sequence. Basic RNNs are a network of nodes emulating neurons, each with a directed (one-way) connection to every other node. Each node has a time varying realvalued activation. Each connection (synapse) has a realvalued weight which can be modified in every iteration. Nodes are either input nodes to receive data from outside of the network or output nodes that yields results, or hidden nodes that modify the data which passes through them via their route from input to output. The difference

from the traditional feed forward neural networks is that RNN is capable of using its internal state, otherwise known as memory, to process sequences of inputs.

### 3.2. Long short-term memory (LSTM)

Long short-term memory (LSTM) units are a special type of building units for RNN. It can analyse, classify and predict temporal data sequences of time lags of any size. A typical LSTM network is made up of memory, input, output and forget gates. The memory in LSTM can remember values over arbitrary time intervals. Each of the three gates is a form of neuron (which computes an activation function of a weighted sum). More than that, these gates control the passage of values in LSTM layers; hence these special neurons are named as gates. By long short-term, the fact underlined is that LSTM's memory can really last for large time duration. LSTM tackles the issue of exploding and vanishing gradient problem which is an important issue while training traditional RNNs.

### 3.3. Convolutional neural network (CNN)

Convolutional neural network (CNN) is an improvised variant of multilayer perceptron. CNN is generally made up of an input, an output layer and many hidden layers. The hidden layers of a CNN typically are made up of convolutional, pooling, and fully connected layers.

## 4. LITERATURE SURVEY

Li Ching Ng et al. The main purpose of drug delivery. The transdermal drug delivery of diabetes. This review presents the recent advances and developments in transdermal research to achieve better diabetes management. Different technologies and approaches have been explored and applied to the transdermal systems to optimize diabetes management. Studies have shown that these transdermal systems demonstrate higher bioavailability compared to oral administration due to the avoidance of first-pass hepatic metabolism and a sustained drug release pattern. Besides that, transdermal systems have the advantage of reducing dosing frequency as drugs are released at a predetermined rate and control blood glucose level over a prolonged time, contributing to better patient compliance. drug administration and glucose level monitoring, a breakthrough innovation in treatment strategies for management of the disease would be welcome. If the challenges in manufacturing the transdermal patch could be overcome, we will expect a tremendous shift in diabetes management as transdermal drug delivery systems in diabetes management are deemed to be a promising approach in providing a better clinical outcome compared to conventional dosage forms.

Brandon Fan et al. Achieves state-of-the-art results and can be applied to multiple other healthcare and information extraction tasks including medical entity extraction and entity recognition. According to the U.S. Department of Health and Human Services, adverse drug events (ADEs), or harmful side effects, account for 1/3 of total hospital admissions each year. The goal of this research is to utilize novel deep learning methods for accurate detection and identification of professionally unreported drug side effects using widely available public data (open data). This study also shows the feasibility of BERT word embedding's over other types of non-retrained word embedding's and retrained embedding's by comparing ADE detection results with our proposed model achieving an AUC of 0.94. Not only can the model be used for ADE extraction, but it can also be used for general information extraction tasks including named entity recognition and content extraction, generalizing the model to a wider array of tasks. Despite the success in the ADE extraction task, further research should analyze ways to classify drug side effects into manageable categories for doctor viewing and drug professionals. a method for recognizing semantic similarity between drug side effects is an important consideration in order to map similar drug side effects into one "formal definition".

Saima Amjad et al. The integration of nanotechnology with herbal remedies for the effective management of diabetes. The herbs are rich resources of phytoconstituents which were found to be useful for the management of DM as well as wound healing. Plant derived drugs may be considered to be safe in their traditional form, but when they come in Nano form or with Nano carriers may induce toxicity. More studies are therefore needed to evaluate the potential side effects, mechanism and mode of action of NHDs for anti-diabetic and insulin mimetic activities. The applications of nanotechnology in the formulation of Nano-herbal drugs have shown promise in addressing the challenges and improving the safety and efficacy of plant products for treating DM

Ambika Choudhury et al. This paper mainly targets the review of diabetes disease detection using the techniques of machine learning. Further, PIMA Indian Diabetic dataset is employed in machine learning techniques like artificial neural networks, decision tree, random forest, naïve Bayes, k-nearest neighbors, support vector machines, and logistic regression and discussed the results with their pros and cons. a four-layer artificial neural network is designed. Back propagation method and Bayesian regulation (BR) algorithms are used to train and avoid over fitting the dataset. The training of data is done in such a way that it forms a single and accurate output displaying in the regression graphs

Sujaritha.M et al. 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. 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. 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.

Jatin N Bagrecha et al. in assisting doctors to predict diabetes disease in prior based on present health parameters like blood plasma, age ,insulin level, pregnancy, body mass, skin thickness, pediatric conditions and blood pressure. Machine learning tools have huge impact in medical field day by day. It is quite difficult to make correct decisions in future prediction of disease. But this project uses convolution neural network system to make efficient classification. Here classification happens as Diabetic or Non-Diabetic based on the health parameters neural networks are prime machine learning algorithms for accurate prognosis of disease in medical field. It uses health parameters as input for neurons; here 8 input parameters are given as input for the neural network. Initially, weights are assigned randomly for neuron and later on during the training phase weights are updated accordingly. Iteration continues, until the output of neural network matches with target output. This paper depicts the implementation and development of software tool built in MATLAB. The algorithm used is Back-Propagation Neural Network (BPNN) and performance ranges around 81 percent.

Quan Zou et al. This system used decision tree, random forest and neural network to predict diabetes mellitus. The dataset is the hospital physical examination data we used principal component analysis (PCA) and minimum redundancy maximum relevance (MRMR) to reduce the dimensionality. Five-fold cross validation was used to examine the models. In order to verify the universal applicability of the methods, we chose some methods that have the better performance to conduct independent test experiments. We randomly selected 68994 healthy people and diabetic patients' data, respectively as training set. to exactly predict and diagnose this disease by using machine learning is worthy studying. According to the all above experiments, we found the accuracy of using PCA is not good, and the results of using the all features and using MRMR have better results.

Mohamed Chetoui et al. In this paper we introduce the use of different texture features for DR, mainly Local Ternary Pattern (LTP) and Local Energy-based Shape Histogram (LESH). We show that they outperform LBP extracted features. Support Vector Machines (SVM) are used for the classification of the extracted histogram. A histogram binning scheme for features representation is proposed. The proposed approach is suitable even for small datasets. New techniques based on deep learning are data hungry but show impressive performances in different classification tasks including DR. Future work includes benchmarking the performance of deep learning techniques and the proposed texture-based features in a small dataset such as MESSIDOR.

Swapna G. et al. The proposed system diabetes is diagnosed by the analysis of Heart Rate Variability (HRV) signals obtained from ECG signals. We employed deep learning networks of Convolutional neural network (CNN) and CNN-LSTM (LSTM = Long Short Term Memory) combination to automatically detect the abnormality. Unlike the conventional analysis meth-ods so far followed, deep learning techniques do not require any feature extraction. We initially performed classification splitting the database into separate training and testing data. Our system can assist the clinicians to diagnose diabetes accurately. As explained under Results section, further improvement in accuracy can be explored by feeding into the proposed architecture large sized input dataset compared to the dataset size used in this work.

Gauri D.Kalyankar et al. In this system machine learning algorithm in Hadoop Map Reduce environment are implemented for Pima Indian diabetes data set to find out missing values in it and to discover patterns from it. This work will be able to predict types of diabetes are widespread, related future risks and according to the risk level of patient the type of treatment can be provided. Suggests that implemented algorithms are able to impute missing values and to recognize patterns from the data set. In future work pattern matching will be employed by applying these discovered patterns on testing data set to predict diabetic prevalent and risk levels associated with it.

Priyanka Indoria et al. This paper focuses on recent developments in machine learning which have made significant impacts in the detection and diagnosis of diabetes. Type 1 and type 2 diabetes are the most common forms of the disease, but there are also other kinds, such as gestational diabetes, which occurs during pregnancy, as well as other forms. Insulin is a hormone that is produced by the beta cells in the pancreas. Insulin works like a key to a door. Insulin attaches itself to 'doors' on the cell, opening the door to allow glucose to move from the blood stream, through the door, and into the cell. If the pancreas is not able to produce enough insulin (insulin deficiency) or if the body can - not use the insulin it produces (insulin resistance), glucose builds up in the bloodstream (hyperglycemia) and diabetes develops.

Basharat Naqvi et al. In this paper description of chosen classification models and dataset. Next, makes evaluation and performs a comparison of performance of 5 classification techniques on chosen dataset. Then, it provides results by considering evaluation metrics such as accuracy, precision and recall. This work finds that the decision tree is the best technique for prediction of disease in diabetic patients. This work considers datasets related to diabetic patients and applies various techniques by considering evaluation metrics such as accuracy, recall, precision in order to carry out performance evaluation and comparison of chosen classifiers.

Minyechil Alehegn et al. In this system most known predictive algorithms are applied SVM, Naïve Net, Decision Stump, and Proposed Ensemble method (PEM). An ensemble hybrid model by combining the individual techniques/methods into one we made Proposed Ensemble method (PEM). The proposed ensemble method (PEM) provides high accuracy of 90.36% we concentrated only Diabetes disease for future it can be extended to apply this method in another diseases Small amount sample data used on this study. it can be apply in large amount of data for future extension .on this study also only a single data set used therefore for future multiple data set can be used for prediction .in this study only limited base classifier used .for future it is possible to use another base classifier like ANN, Nave Bayes, KNN ,Random tree ,and other.

Tejas N. Joshi et al. This paper Machine learning is an emerging scientific field in data science dealing with the ways in which machines learn from experience. The aim of this project is to develop a system which can perform early prediction of diabetes for a patient with a higher accuracy by combining the results of different machine learning techniques. This project aims to predict diabetes via three different supervised machine learning methods including: SVM, Logistic regression, ANN. This project also aims to propose an effective technique for earlier detection of the diabetes disease. This work has described a machine learning approach to predicting diabetes levels. The technique may also help researchers to develop an accurate and effective tool that will reach at the table of clinicians to help them make better decision about the disease status.

D.Ramana kumar et al. This system attempt is made to develop a novel approach to find the appropriate solution using probabilistic and Machine Learning models which have the ability to predict whether the patient is having diabetes or not. Predicting the disease in early stages leads in treating the patient before it becomes critical. The proposed model has ability to extract hidden knowledge from a huge amount of diabetes-related data - collected from Web services data repository. Data mining techniques are used for variety of applications like health care domain which is the most considerable factor in human life. There are many classifications of methods which analyze the medical data to find the features and take the decisions on its own based on the given data. The real time medical decision support system with stream data classifier is needed to analyze medical data streams and make real time predictions.

## CONCLUSION

Considerable part of human population is under the grip of diabetes which is incurable. If not managed well, diabetes can lead to health hazards. Hence, early detection of diabetes is extremely crucial. The proposed progression modeling system can serve as a reliable tool to clinicians to detect early diabetes. Further depending upon the progression modeling based on classification, the model for administering drug is provided for diabetics patient through machine on large sized input data for research. The predicted information can serve as a warning signal for the patient as well as the doctor to take sufficient control and precautionary measures.

## REFERENCES

- [1] Li Ching Ng "Transdermal drug delivery systems in diabetes management: A review" Asian Journal of Pharmaceutical Sciences Available online at [www.sciencedirect.com](http://www.sciencedirect.com) journal homepage: [www.elsevier.com/locate/AJPS](http://www.elsevier.com/locate/AJPS) August 12, 2019
- [2] Brandon Fan "Adverse drug event detection and extraction from open data: A deep learning approach" Information Processing and Management 57 (2020) 102131 Science Direct

- [3] Saima Amjad “A novel strategy of anodized herbal drugs and their delivery in the treatment of diabetes: Present status and future prospects” *Journal of Herbal Medicine*, <https://doi.org/10.1016/j.hermed.2019.100279> 2019
- [4] Ambika Choudhury “A Survey on Medical Diagnosis of Diabetes Using Machine Learning Techniques” Springer Nature Singapore Pvt. Ltd. 2019 [https://doi.org/10.1007/978-981-13-1280-9\\_6](https://doi.org/10.1007/978-981-13-1280-9_6)
- [5] Sujaritha.M, “Risk Analysis of Diabetes using IOT and Deep Learning” *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8, Issue-7S, May 2019
- [6] Jatin N Bagrecha “Diabetes Disease Prediction using Neural Network” *International Journal for Research in Applied Science & Engineering Technology (IJRASET)* ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177 Volume 7 Issue IV, Apr 2019- Available at [www.ijraset.com](http://www.ijraset.com)
- [7] Quan Zou “Predicting Diabetes Mellitus with Machine Learning Techniques” ORIGINAL RESEARCH published: 06 November 2018 DOI: 10.3389/fgene.2018.00515
- [8] Mohamed Chetoui “Diabetic Retinopathy Detection Using Machine Learning and Texture Features” 2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)
- [9] Swapna G, “Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals” *International Conference on Computational Intelligence and Data Science (ICIDS 2018)* *Procedia Computer Science* 132 (2018) 1253–1262 0.1016/j.procs.2018.05.041
- [10] Gauri D. Kalyankar “Predictive Analysis of Diabetic Patient Data Using Machine Learning and Hadoop” *International conference on I-SMAC (IOT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017)*
- [11] Priyanka Indoria “A Survey: Detection and Prediction of Diabetes Using Machine Learning Techniques” *International Journal of Engineering Research & Technology (IJERT)* <http://www.ijert.org> ISSN: 2278-0181 IJERTV7IS030167 (This work is licensed under a Creative Commons Attribution 4.0 International License.) Published by: [www.ijert.org](http://www.ijert.org) Vol. 7 Issue 03, March-2018
- [12] Basharat Naqvi “Prediction Techniques for Diagnosis of Diabetic Disease: A Comparative Study” *IJCSNS International Journal of Computer Science and Network Security*, VOL.18 No.8, August 2018
- [13] Minyechil Alehegn “Analysis and Prediction of Diabetes Mellitus using Machine Learning Algorithm” *International Journal of Pure and Applied Mathematics* Volume 118 No. 9 2018, 871-878 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version) URL: <http://www.ijpam.eu> Special Issue
- [14] Tejas N. Joshi “Diabetes Prediction Using Machine Learning Techniques” *Int. Journal of Engineering Research and Application* [www.ijera.com](http://www.ijera.com) ISSN: 2248-9622, Vol. 8, Issue 1, (Part II) January 2018, pp.-09-13
- [15] D.Ramana kumar “Health care System: Stream Machine Learning Classifier for features Prediction in Diabetes Therapy” *International Journal of Applied Engineering Research* ISSN 0973-4562 Volume 13, Number 1 (2018) pp. 59-65 © Research India Publications. <http://www.ripublication.com>