SN Computer Science (2020) 1:322
https://doi.org/10.1007/s42979-020-00337-2

REVIEW ARTICLE

Artificial Intelligence-based Learning Techniques for Diabetes Prediction: Challenges and Systematic Review

Surabhi Kaul[1] · **Yogesh Kumar**[2]

Received: 14 August 2020 / Accepted: 17 September 2020 / Published online: 6 October 2020

Abstract

Diabetes is one of the most deadly and chronic diseases which cause an increase in blood sugar. If diabetes remains untreated and unidentified many difficulties may arise due to that. The tedious work is in identifying the process which results in visiting the clinic and consulting the doctor. But this tedious work has been solved with the rise in the approaches used by machine learning. Over recent years a plenty of growth has been seen doing over health informatics by focusing on the technology of the presentation, generation and application of clinical information in health care. With the motive of improving outcomes of health for patients and creating efficiency in health professions a Healthcare Informatics or eHealth solutions, has empowered the accessibility of clinical data through computer networks or cloud computing. Its accessibility and understanding have become easier with language technologies. This paper gives a comprehensive prospect of work accomplished to develop a model that can predict the possibility of diabetes in patients with extreme accuracy. Therefore, various machine learning classification algorithms are used for detecting diabetes. In this paper, the author has studied various machine learning classification algorithms, namely genetic algorithm, decision tree, random forest, Logistic regression, SVM and Naive Bayes. Experiments are carried out on Pima Indians Diabetes Database (PIDD) which is track down from the UCI machine learning repository. Further author has done the comparison among various performances of all the different algorithms. The performances are categories of various measures like Precision, Accuracy, F-Measure and Recall. The paper helps in identifying the algorithm to classify the risk of diabetes. Different techniques were applied to the algorithms for improving the robustness. Additionally, the findings suggest that the best performance of disease risk classification is done with the help of a genetic algorithm.

Keywords Random forest · Diabetes prediction · Artificial neural networks · Support vector machine naive bayes · PIDD

Introduction

Taking care of health or efforts for the betterment of health by way of treatment, diagnosis, salvage from the disease, sickness, injury and other physical and mental deficiencies in people comes under health care. Health specialists provide health care services in different health fields. Doctors, surgeons and their respective associates all are the part of Health experts. Health care includes physical therapy, nursing, pharmacy, Dentistry, athletic training, optometry, etc. All the works come under primary care, tertiary care and secondary care are the part of health care. A well-organized health care plays an important role in the development and progress in a country's economy. Health care proves an important element to maintain the mental and physical health of the people all around the world. In 1980, WHO declares smallpox as the first disease in history which is

This article is part of the topical collection "Computational Statistics" guest edited by Anish Gupta, Mike Hinchey, Vincenzo Puri, Zeev Zalevsky and Wan Abdul Rahim.

✉ Yogesh Kumar
Yogesh.arora10744@gmail.com

Surabhi Kaul
surabhi93kaul@gmail.com

[1] CSE, Chandigarh University, Gharuan, Mohali, Punjab, India

[2] CSE, Chandigarh Group of Colleges, Landran, Mohali, Punjab, India

completely eliminated by deliberate health care was a clear example of efficient health care.

In 2014, 422 million adults all over the world suffering from diabetes mellitus disease. DM enacts a great financial problem on worldwide medical assistance service also the widespread financial situation. The worldwide rate is expected almost 825 billion US dollars in one year, and this cost is arisen because of DM difficulties [2]. For the betterment of the healthiness of the diabetic patient, ailment controlling and taking measures to avoid DM allied problems it is very important and need a deep study about DM. Management of the Glycohemoglobin on the average level is very important for diabetic patient. Glycohemoglobin normal range is less than 7%. For a Diabetic patient to manage this level is difficult and electronic. Digitization [1] and medical assistance had a great effect on all aspects of diabetes. With the help medical facilities information received from the internet and social media and all other allied services related to 'Health' totally transform the scenario of diabetes.

Novelty in the Article

Various researchers [3] have worked on health informatics for diabetes prediction, due to its popularity, but it is difficult for many of the researchers give reviews on things related to diabetes. Different procedures were opted by researchers for this prediction of diabetes through different ways like using data mining [4], mechanical knowledge and deep learning. No survey predicts all the aspects of diabetes in one table and represents with appropriate differentiation between all the prediction and categorizing methods which helped for diabetes prediction. So, the main motive behind this work is to provide a framework [5] for the evaluation between all predictions and various used classification procedures. For best results and examination, authors have systematically emphasized the evaluation and perfection has been calculated.

Motivation

Because of its rising day by day, lots of people suffering from diabetes [8]. Mostly patient has no knowledge regarding their health condition and the difficulties before disease analysis. Efforts to resolve or refine the prediction model's correctness is the major difficulty and to make this design adaptive more than one dataset [9].

The newest statistics on DM describe in the Diabetes Atlas (Seventh Edition) by the International Diabetes Federation (IDF) [6]. Ratio of diabetic patient in 2015 across the world was almost 410 million.

Concentrate on a person having more risk is very necessary to lower the influence of Diabetes [10] mellitus and control the disease. World Health Organization (WHO) defines more possibility of Diabetes Mellitus as given below:

- People of 45 and above age and rarely work out
- Body Mass Index is more than 24 kg/m^2
- Blood glucose is raised beyond normal levels or blood sugar levels during fasting consistently above the normal range (IFG)
- Pregnant women of above 30 age
- High blood pressure
- High blood levels of triglycerides [7]
- Having a family history with diabetes

Background Study

In this section, we expound the framework for diabetes prediction, Importance of diabetes prediction and challenges is being discussed. The section also describes the role of community and behavior risk for diabetes.

The Framework for Diabetes Prediction

Anderson et al. [3] has projected united framework has been produced. Concentrating on the artificial intelligence centered forecast producing a Framework. For diabetes patients, it contributes to give the nearer appearance in the actual phase estimate. Below mentioned figure shows AI [25] focused design for DM prophecy. Inspired from the consequence of Artificial Intelligence centered ailment estimates, it describes allocation schemes and actual period facts estimate service to generate real outcomes for real forecasting, checking, of Artificial Intelligence focused DM prediction [22].

- To learn checking of the device is recommended for DM forecast. For diabetes prediction and monitoring, the recommended structural benefits of effective decision-making technique and helping in good outcome [26].
- Keeping in view the huge development in the ailment, the recommended porotype goal is to deal with efficiently through cloud computing solutions.
- Mostly, research is not reviewing the F-score [21], But some research make a regular estimate of categorizing model with F-score.

Tensor Flow structure used by Xu et al. [34] that run on various CPUs and GPUs. The suggested system using the steps described in Fig. 1, in the process flow diagram.

Put the data in the system is the first step. After that, according to the necessity feature selection is to be

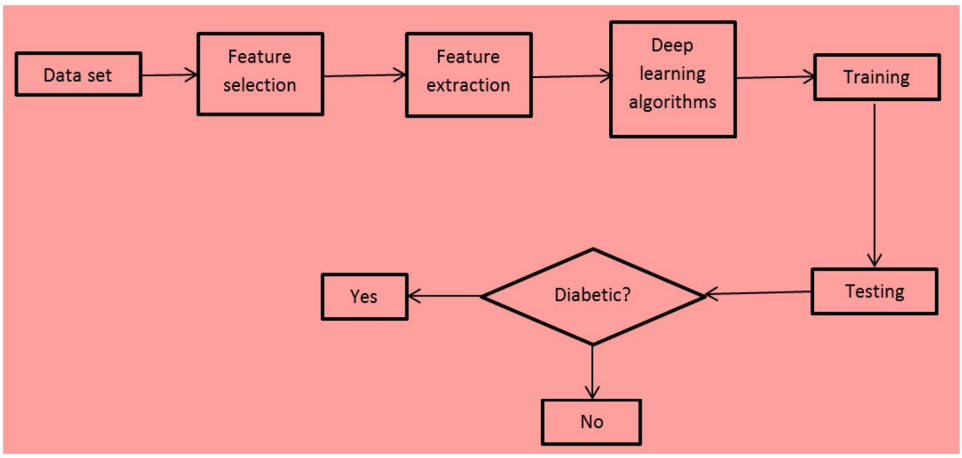

Fig. 1 Flow diagram for diabetes prediction using deep learning model [34]

Table 1 Symptoms of diabetes [31]

	Symptoms
1	Recurring skin, gum, or bladder infections
2	Unusual weight loss
3	Sweet, or alcohol-like smell while breathing
4	Frequent urination
5	Unusual thirst
6	Extreme hunger
8	Quick vision changes
9	Numerous infections
10	Unclear vision
11	Cuts/bruises that are slow to heal
12	Extreme weakness

completed. Then exaction of various factors is completed as per the selected category feature. After completing the extraction procedure many deep learning algorithms have to be applied forgetting the preferred output. Using various classification and prediction procedures available output is then practiced. After the output is examined, testing is executed to check that sample which is present in the output is diabetic [24] or not.

Importance of Diabetes Prediction

To protect and keeping people healthy [34], timely treatment of diabetes and early detection is most important. It will provide relief to reduce the risk of serious heart disease and stroke, blindness, kidney failure and limb amputations.

Table 1 describes some important indications and symptoms of diabetes. Some patients have no signs or symptoms who are suffering from type 2 diabetes, but there has a risk factor [29]. There are so many tests are required to diagnose diabetes for those patients who have a higher risk and found signs and symptoms.

- A fasting plasma glucose test used to detect diabetes, which measures before eating 8 h.
- Another GTT Glucose Tolerance [37] test for detecting DM. This test is conducted after two hours of consuming glucose water.
- Another test is Capillary blood glucose (CBG) in which blood glucose is tested after eating food.

Through fasting blood Sugar test and GTT test one and another day be confirmed positive test result. It is important for the Patients to control diagnose ailment initial stage and so that avoid delay in serious disease complications which cause bad effect on the quality of life [33].

Challenges

If diabetes is not controlled, then they may cause serious health problems such as lower limb amputation, kidney failure, blindness that impact very badly on the quality of life [32]. There are various research issues or challenges that are occurring while predicting the diabetes:

- Loss of eyesight: in 2010 Diabetic retinopathy caused 2.6% of blindness globally and 1.9% of moderate or severe visual loss. Studies show the diabetic patient prevalence of any retinopathy is 35%, but vision-threatening is 7%. People suffering from type 1 diabetes retinopathy rates are higher among them.
- Renal disease: Information [29] received from different sources describes that diabetes is the main cause of 80% cases of kidney-related disease (ESRD). The ratio of kidney problems because of diabetes is 15–55%. Renal disease very much dependent on renal replacement therapy and access to dialysis and—(and in some cases within) in different countries these are highly variable.

- Cardiovascular actions: With an increase in fasting plasma glucose level the danger of cardiovascular disease also increases continuously even before adopting sufficient levels to diagnose diabetes. There is a decrease found in type 1 or type 2 diabetes over the past 20 years among different countries in North America, Scandinavia and United Northern Ireland has shown in the study of the prevalence of cardiovascular events although less decrease in the non-diabetic population [30].
- Below extremity surgical: diabetes became the danger of below extremity surgical due to septic, foot sores. Ratio of surgical extremity in diabetes patient is 20% more than non-diabetic people, and in the last past years, this ratio rising 1.5–3.5 in each year in people with having diabetes. United States of America study showed a reduction in rates of amputations from 40 to 60% among adults with diabetes [19].

Community and Behavior Risk for Diabetes

Diabetes is a prevalent plaguing for the nation, becoming a serious problem which needs to be cured [27]. Almost 29 million in the United States suffering from diabetes and out of these 29 million so many people do not aware that they suffered from the prolonged disease. There is a need to study the main reasons which are responsible for increase, and also find measures to control it.

- Obtain steps for precaution: low-income areas and rural regions in different towns may face obstacles to find proper precaution, like poor transport service, lack of health centers in these towns. Even if they find a health center, many of them cannot bear the expenditure of the treatment.
- Society built-up: structure of the society also has a great impact on the fitness. In those regions where a shortage of sidewalks, bike paths or amusing areas—shortage of all these amenities also causes risk of type 2 diabetes for the residents. Also, in a community the locality of shops/supermarkets can create a problem for the establishment of a fit lifestyle because people having difficulty to find better foodstuff. More than 23 million Americans survive in poor places which are far from a supermarket. A Campaign like Let's Move take initiatives for helping people by providing healthy better foodstuffs and provide clean surroundings for physical activity [28].
- Education: education also affects the risk of diabetes, research shows that society having educated is least affected by diabetes. Education level can define your profession, financial grade, fitness knowledge which helps to take healthiness decisions. Diabetes reduces 7–5.7% and the possibility of obesity reduces from 23 to 18% because of giving additional four years of education, helping more positive health behaviors to the adults over the age of 25.
- Financial constancy: employment status also affects your probabilities diabetes. Shortage of incomes also prevents to take medical services to obtain therapeutic care and buying better food selections. Most, least expensive foods are more calorie-laden and less-nutrient. However, Healthy and rich-nutrient food choices are not always expensive. Food items like Milk and rice having a rich, nutritious protein, and fruits and vegetables also are a good source of nutrients [35].
- Social support: social support increase awareness and prevention of ailments, including type 2 diabetes. If you do not have any social support from family, friends and from members of your community then it will effect your health. A study from the Center for Health Disparities Research shows that social support helps in better health results, mentally and physically, due to being able to express feelings.

Reported Work

Most of the massive topics for Researchers are Data prediction [12]. Numerous researchers have done their research in finding out the technique which is good and helpful for predicting the diabetes more simpler way. Different methods for diabetes predicting [15] have been listed below:

a. Diabetes prediction using machine learning

Analysis of Data [17] in vast data files to separate enclosed and earlier abstruse examples, links, and data difficult to find with conservative assessment methods. Study of vast data sets is also a developing area of great implication in social insurance. Researches using data mining procedures to study the patients' data which are beneficial to find important knowledge which is facilitating medical services and deeply study of disease [36].

Different procedures follow to forecast Diabetes Mellitus, through device learning technique. Many authors used particle swarm optimization [11] (PSO) is a computational method algorithm to forecast type 2 DM researchers offered a scheme For DM prediction, namely linear discriminant analysis. To decrease proportions and extract the features Linear Discriminant Analysis are used (LDA). Forecast algorithms built upon statistical models for diverse onsets of type 2 DM forecast were built to deal with high dimensional data sets. Using support vector regression [16] (SVR) many researchers focused on the glucose in finding the diabetes.

In the diabetes prediction Machine learning [13] methods are broadly used to get superior results. In the field of medical Decision support tool is the most prevalent mechanism procedures, which has grate sorting controls. Another famous machine learning method is a neural network and

gives better results in different aspects. So Random decision forests [14], Decision tree, and artificial neural system mostly uses methods to predict diabetes.

b. Diabetes prediction using data mining

Data mining is the investigation of expansive data sets to separate covered up and beforehand obscure examples, connections and information that are hard to recognize with conventional measurable techniques. The territories where data mining is connected as of late incorporate designing, showcasing, human services and monetary anticipating. Data mining in social insurance is also a rising field of high significance for giving what we can say is high anticipation and a more profound comprehension of restoring data. The amount of accessibility of tremendous measure of patient's data which can be used to extricate valuable information, scientists have been utilizing data mining methods to help medicinal services experts in the analysis of ailments.

In the usually higher part of the paper, the diabetes forecast system chips away at a little dataset, however, our point is to deal with the expansive dataset. The quantity of medicinal test required may influence the execution of the system in this way we additionally concentrate on diminishing the therapeutic test. It relies on upon which parameter or quality is taken in the system for foreseeing diabetes. Our expectation system will take a shot at a bigger dataset and number of therapeutic testing test required will overcome. Our system utilizes two calculations which we will apply to the same data set for anticipating diabetes.

c. Diabetes prediction using deep learning

Deep learning and machine learning are different from each other. Deep learning deals with computational models which are made up of various processing layers built upon the neural networks to study the representation of data with various abstraction levels. Basic difference between ANNs and deep learning are their links, different secret layers and about the meaningful input's abstraction. The limit of traditional artificial neural networks is up to three layers which are used to attain supervised illustrations that are adjusted for the particular duty [20]. Layers of the deep learning system represent the observed designs based upon the data which obtain as input from the previous layer through a local standard [18]. Features of deep these layers are not discovered by human engineers and this is the key aspect of deep learning and made up through general learning technique.

With the help of Machine learning [23] huge research has occurred on diabetes detection. Machine learning based on phases of classification, feature selection and feature extraction. It has been found from the study that the results of machine learning approaches are not its best to solve the problems of object recognition and speech recognition the reason behind this are the dimension of data which is very high. The drawbacks of machine learning are the reason behind the speedy increase in the research process of deep learning. It works in the field of healthcare applications. In the area of health care about anomaly detection has been published in recent years. A deep learning technique is used to configure out diabetes from HRV data with a specified accuracy value which almost matches with the highest accuracy achieved in the automated diabetes detection.

Comparison Analysis

The diabetes prediction using various techniques are presented in Table 2. Table 2 illustrates the purpose of diabetes prediction by exploring various performances of different algorithms presented by different authors on different parameters which may include data set, feature extraction, classifier used and the results obtained by each author.

Conclusion

The main objective of this paper is to improve the correctness of predictive models. The accuracy can be achieved by either refining the performance of the data or with the help of an algorithm. For achieving best results the data can be improved at the earlier phase. PIMA dataset has been taken for performing the accuracy checks on each classifier. Among all, the genetic algorithm leads over others. In this research, the author has concluded one more important factor that the accuracy of a model is highly dependent on the dataset. In our work, the PIMA diabetic dataset has worked every well in providing us with correct results, but the same results cannot be guaranteed on a different dataset. In future work, advanced classifiers such as evolutionary algorithm (EA) for diabetes prediction can be applied along with machine learning algorithms.

Table 2 Comparison among different algorithms for diabetes prediction

Author	Dataset	Feature extraction	Classifier	Result
Bhargava et al. [4]	Real-world male heart disease dataset	Heart diseases	Decision tree algorithm	Accuracy = 79.9%
Asha gowda karegowda et al. [1]	Pima Indians DM statistics	glycemia, hypertension, body mass index	Genetic algorithm	Accuracy = 84%
Dagliati et al. [5]	Real-world medical dataset	Insulin, bmi, diabetes pedigree	Gmm	Accuracy = 81%
Kho et al. [16]	Real-world medical dataset	Glucose, blood pressure, skin problem and age	Ann	Accuracy = 89%
Giri et al. [8]	Delta elevators	Skin problem, blood pressure	Elm (extreme learning machine)	Accuracy = 82%
Bhargava et al. [4]	School of medicine, University of Virginia	Age, weight, waist, hip	Random forest algorithm	Accuracy = 84%
Forecast et al. [7]	Uci machine repository standard dataset	Body mass index, triceps skin-fold thickness	Hoeffding tree algorithm	Accuracy = 77%
Shana et al. [28]	Pima Indian Diabetes Dataset	Number of times pregnant, age	Decision tree j48	Accuracy = 73.75%
Naiarun et al. [20]	26 Primary Care Units (PCU) in sawanpracharak regional hospital	Age, weight, waist	Logistic regression	Accuracy = 82.35%
Mir et al. [18]	26 Primary Care Units (PCU) in sawanpracharak regional hospital	Body mass index, height, weight	Random forest	Accuracy = 85.55%
Deshmukh et al. [22]	The National Inpatient Sample (NIS) data	Age, race, sex	Random forest	Accuracy = 89.55%
Ramesh et al. [24]	The Pima Indian Diabetic Database	Skin, blood, age, height	Svm	Accuracy = 78%
Khalilia et al. [15]	https://www.stanford.edu/~hastie/papers/lars/diabetes.data	Body mass index, blood pressure, cholesterol level	Knn	Accuracy = 70%
Aljarullah et al. [2]	Pima Indian Diabetic Database (PIDD)	Plasma glucose concentration, triceps skin fold, age	Knn	Accuracy = 82%
Sisodia et al. [29]	Real-world medical dataset	Plasma glucose concentration,blood pressure, triceps skin fold, serum insulin	Ann	Accuracy = 83%
Sisodia et al. [29]	Pima indian diabetic database (PIDD)	All features skin color, age, waist size etc	Svm	Accuracy = 72%
Kaur et al. [13]	Pima Indians Diabetes Dataset	Plasmaglucose, blood pressure, triceps skin fold thickness	Decision tree	Accuracy = 78.17%
Thirugnanam et al. [31]	Pima Indians Diabetes Dataset	Bp, skin, age, plasma glucose	Naïve bayes	Accuracy = 77.86%
Zia et al. [36]	Pima Indians Diabetes Dataset	Age, height, weight, smoking habit, alcohol habit, family history	Genetic algo	Accuracy = 90%
Joshi et al. [10]	Pima Indian Diabetes Dataset	Glucose, blood pressure, skin thickness	Knn, naïve bayes	Hybrid model provides best accuracy

Compliance with Ethical Standards

Conflict of Interest The authors declare that they do not have any conflict of interests that influence the work reported in this paper.

References

1. Huang GB, Zhu QY, Siew CK, et al. Extreme learning machine: theory and applications. Neurocomputing. 2006;70:489–501.
2. Al Jarullah AA. Decision tree discovery for the diagnosis of type II diabetes. In: 2011 International conference on innovations in information technology. IEEE; 2011, pp. 303–7.
3. Anderson JP, et al. Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: an application of machine learning using electronic health records. J Diabetes SciTechnol. 2016;10:6–18.
4. Bhargava N, Dayma S, Kumar A, Singh P. An approach for classification using simple CART algorithm in WEKA. In: 2017 11th international conference on intelligent systems and control (ISCO). IEEE; 2017. pp. 212–6.
5. Dagliati A, et al. Machine learning methods to predict diabetes complications. J Diabetes SciTechnol. 2018;12:295–302.
6. Fagherazzi G, et al. Science direct digital diabetes: perspectives for diabetes prevention, management and research. Diabetes Metab. 2019;45(4):322–9.
7. Forecast M., et al. (2005) Diabetes Mellitus Forecast Using Artificial Neural Network (ANN), In 2005 Asian Conference on Sensors and the International Conference on New Techniques in Pharmaceutical and Biomedical Research. 135–139.
8. Giri D, et al. Knowledge-based systems automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform. Knowl-Based Syst. 2013;37:274–82.
9. Jegan C. Classification of diabetes disease using support vector machine. Int J Eng Res App. 2017;3(2):1797–801.
10. Joshi R, et al. Analysis and prediction of diabetes diseases using machine learning algorithm: ensemble approach. Int Res J EngTechnol. 2020;10:426–35.
11. Karegowda AG, et al. Cascading K-means clustering and K-nearest neighbor classifier for categorization of diabetic patients. Int J EngAdvTechnol. 2012;1(3):147–51.
12. Karegowda AG, et al. Application of genetic algorithm optimized neural network connection weights for medical diagnosis of pima Indians diabetes. Int J Soft Comput. 2011;2(2):15–23.
13. Kaur G. Improved J48 classification algorithm for the prediction of diabetes. Int J Comput App. 2014;98(22):13–7.
14. Kavakiotis I, et al. Machine learning and data mining methods in diabetes research. ComputStructBiotechnol J. 2016;15:104–16.
15. Khalilia M, et al. Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform DecisMak. 2011;11(1):51.
16. Kho AN, et al. Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. J Am Med Inform Assoc. 2012;19(2):212–8.
17. Mercaldo F, et al. Diabetes mellitus affected patients classification and diagnosis through machine learning techniques. ProcediaComputSci. 2017;112:2519–28.
18. Mir A., et al. (2018) Diabetes Disease Prediction using Machine Learning on Big Data of Healthcare. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). 1–6.
19. Monica M, et al. Investigating structure and dynamics of proteins in amorphous phases using neutron scattering. ComputStructBiotechnol J. 2017;15:117–30.
20. Naiarun N, et al. Comparison of classifiers for the risk of diabetes prediction. ProcediaComputSci. 2015;69:132–42.
21. Patil N. (2017) Diabetes disease prediction using data mining. In 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS). 1–5.
22. Deshmukh JS, et al. Social factors influencing diabetes mellitus in adults attending a tertiary care hospital in Nagpur: a cross sectional study. Int J Res Med Sci. 2017;5(11):4988–92.
23. Pavate A. (2015). Risk prediction of disease complications in type 2 diabetes patients using soft computing techniques. In 2015 Fifth International Conference on Advances in Computing and Communications (ICACC). 371–375.
24. Ramesh S, et al. A deep learning approach to identify diabetes. AdvSci Tech Lett. 2017;145:44–9.
25. Rashed S., Noori H., (2018) Machine Learning Based Unified Framework for Diabetes Prediction. In Proceedings of the 2018 International Conference on Big Data Engineering and Technology. 46–50.
26. Saxena K, et al. Diagnosis of diabetes mellitus using k nearest neighbor algorithm. Int J ComputSci Trends Technol (IJCST). 2014;2(4):36–433.
27. Sadhana SS, et al. Analysis of diabetic data set using hive and r. Int J EmergTechnolAdvEng. 2014;4(7):626–9.
28. Shana J, Venkatachalam T. An improved method for counting frequent itemsets using bloom filter. ProcediaComputSci. 2015;47:84–91.
29. Sisodia D, Sisodia DS. Sciencedirect prediction of diabetes using classification algorithms. ProcediaComputSci. 2018;132:1578–85.
30. Swapna G, et al. Diabetes detection using ECG signals: an overview. Deep learning techniques for biomedical and health informatics. Cham: Springer; 2020. p. 299–327.
31. Thirugnanam M, et al. Improving the prediction rate of diabetes diagnosis using fuzzy, neural network, case based (FNC) approach. ProcediaEng. 2012;38:1709–18.
32. Thiyagarajan C. A survey on diabetes mellitus prediction using machine learning techniques. Int J ApplEng Res. 2016;11(3):1810–4.
33. Weng SF, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE. 2017;12(4):e0174944.
34. Xu W., et al. (2017) Risk prediction of type II diabetes based on random forest model. In 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 382–386, IEEE.
35. Yu W, et al. Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. BMC Med Inform DecisMak. 2010;10(1):16.
36. Y. Kumar, et al. (2020) Machine Learning Aspects and its Applications Towards Different Research Area. International Conference on Computation, Automation and Knowledge Management (ICCAKM), Dubai, United Arab Emirates. 150–156.
37. Zia UA, et al. Predicting diabetes in medical datasets using machine learning techniques. Int J SciEng Res. 2017;8(5):1538–51.