RESEARCH ARTICLE

SAMRIDDHI Volume 13, Special Issue 1, 2021 Print ISSN : 2229-7111 Online ISSN : 2454-5767

A Review of Methodology and techniques for analysis of diabetes Prediction

Janhavi R Raut*[1], Priyanka Banger[2]

[1] Jagdish Prasad Jhabarmal Tibrewala, University, Rajasthan, India; e-mail : jau.raut09@gmail.com,
[2] Dept. of Comp. Science, Sonopant Dandekar College, Palghar, India; e-mail : bangarpriyanka20@gmail.com

ABSTRACT

In almost all nations, diabetes mellitus was among the most prevalent persistent diseases and continues to boom in numbers and significance, as economic growth and urbanisation contribute to evolving lifestyles marked by decreased physical hobby and weight issues.In this paper, based on different predictive analytics algorithms, we analysed different cases related to diabetes mellitus and concluded that a single approach is not appropriate for predictive analytics.

Key Words: Diabetes, KNN, SVM, Random Forest, Naïve Bayes, Decision Tree

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, (2021); DOI : 10.18090/samriddhi.v13iS1.9

INTRODUCTION

Diabetes is a disease that happens when the hormone insulin is not developed or used properly by your body. This allows too much glucose (sugar) to accumulate in the blood.Insulin is a blood sugar-regulating hormone. According to World Heath Organisation, 1.6millions deaths due to diabetes in 2016 and around 2.2 million deaths in 2012. Unregulated diabetes disease may be damaged by nerves, kidney and other body organs. Increased Hunger, Increased thirst, suddenly weight loss, frequent urination are some symptoms of diabetes disease.So early prediction is required to save the human life and money. Basically, there are three types of diabetes i.e., Type 1 Diabetes where human body does not make insulin. Type 2 Diabetes where body is not sufficient to use the insulin and third type is Gestational diabetes where high blood sugar present during pregnancies.

EXISTING PREDICTIVE TECHNIQUES

Machine Learning Techniques

Machine learning is an increasing technique that lets machine to learn from existing records efficiently. For constructing mathematical models and making predictions using statistical data or knowledge, machine learning uses various techniques. It is currently used in the various tasks, including image recognition,auto-tagging for Facebook, speech recognition, email filtering, recommendation system, and many others.

Data Mining Techniques

Data mining was among the most valuable tools to help developers, researchers, and people obtain valuable data from large data sets. Information Discovery in the Database is also called data mining. Data cleaning, data integration, data collection, data transformation, data mining, pattern evaluation, and knowledge presentation are part of the knowledge discovery process.

Corresponding Author : Janhavi R Raut, Shri. JJT University, Jhunjhunu, Rajasthan, India; e-mail : jau.raut09@gmail.com,

How to cite this article : Raut, J.R, Banger, P. (2021). A Review of Methodology and techniques for analysis of diabetes Prediction.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, Volume 13, Special Issue (1), 38-41.

Source of support : Nil

Conflict of interest : None

Table-:1 Review of methodology and techniques to analysis of diabetes prediction

Author	Year	Methodology and Algorithm	Dataset	Research Objective	Research Conclusion
Priya B. Patel et. al [1]	2017	Gaussian Naive Bayes, KNN, SVM and Decision Tree	Pima Indian Diabetic Set	To overcome the limitation of traditional diabetes system	KNN algorithm best suitable for application
Saman H. et. al.[2]	2017	Naïve bayes, MLP, J.48, ZeroR, Random Forest, and Logistic Regression	Pima Indians Diabetes Database	To extract knowledge from dataset andgenerate comprehensive and intelligent results.	MLP give low error rate, ZeroR determine better performance, J4.8 give accurate result
S.Selvakumar et. al. [3]	2017	Binary Logistic Regression, Multilayer Perceptron and K-Nearest Neighbor	A multi-dimensional healthcare dataset	classification accuracy was compared	k-Nearest Neighbor is higher accuracy
MinyechilAleh egn et. al. [4]	2018	SVM, Naïve Net,DecisionStump , and Proposed Ensemble method	Pima Indian Diabetes Data Set	An ensemble hybrid model by combining the individual Methods	proposed ensemble method (PEM) provides highest accuracy
Mr. R. Sengamuthu et. al. [5]	2018	J48, C4.5, GA, KNN, MLP, Naïve Bayes, ANN, Homogenity, Generic GA, PLS-LDA, Bayesian	PIMA Indian Dataset	Compared Modified J48 with other classifiers	J48 Classifier gives highest accuracy using WEKA & MATLAB tool.
Nandhini.M, Kavitha.R[6]	2017	Naive Bayes, Multilayer Perceptron and IBK	Diabetes dataset	To predict diabetes using bestclassification algorithm	Naive Bayes is best algorithm
Deepti Sisodia, Dilip S. Sisodia [7]	2018	Decision Tree, SVM and Naive Bayes	Pima Indians Diabetes Database (PIDD)	Compared algorithm to find out suitable classifer	Naive Bayes
Amina Azrar et. al. [8]	2018	KNN, Decision Tree, Naïve Bayes	Pima Indians Diabetes Data set	Comparison for different data mining algorithms for diabetes prediction	Decision Tree
R. Manimaran and Dr.M.Vanitha [9]	2017	Multilayer Perceptron (MLP), BayesNet, JRip, C4.5, Fuzzy Lattice Reasoning (FLR)	MV dataset	use of Decision Tree algorithm for classification and predict Diabetes	C4.5 and JRip
Tejas N. Joshi, Prof. Pramila M. Chawan [10]	2018	SVM, Logistic regression, ANN		To propose an effectivetechnique for earlier detection of the diabetes disease	machine learning is best approach to predicting diabetes

Ahmad Ashari et. al. [11]	2013	Naïve Bayes, Decision Tree and k-Nearest Neighbor	PIMA	Propose a novel method in searchingalternative design	Decision Tree is the fastest
Harsha Sethi et al. [12]	2017	ANN, KNN, Naïve Bayes, SVM	400 datasets from diverse section of the society	To accurate diagnosis of	ensemble technique assured an accuracy
IoannisKavakiotis et al. [13]	2017	Machine Learning along with Data Mining process	Electronic Health Records	Data Mining identify through Machine Learning along with DataMiningtechniques	Data Mining techniques
S. R. Priyanka Shetty,Sujata Joshi [14]	2016	Data mining technique	Laboratories data	For diabetes prediction design a model	Tool for diabetes Prediction
J.Omana and Dr.M.Moorthi [15]	2018	Adaboost, Bagging, Random Forest, Decision Stump, KNN, Apriori		Comparative evaluation on the performance of algorithm	Decision Stump show higher accuracy when boosted with Adaboost algorithm

CONCLUSION

The various review papers more focus on predictive analysis process and techniques to estimate and prediction of diabetes disease. Several approaches utilised for diabetes dataset to predict diabetes disease treat in better way. The comparative analysis of various techniques helps for deciding which approach is much suitable for prediction purpose in future. So, hybrid techniques are more effective and give more accurate result for diagnosis and predicting diabetes disease.

BIBLIOGRAPHY

[1] P. P. S. H. D. P. Priya B. Patel, "Analyze Data Mining Algorithms For Prediction Of Diabetes," International Journal of Engineering Development and Research, vol. 5, no. 3, pp. 466-473, 2017.

[2] A. S. S. A. S. Saman Hina, "Analyzing Diabetes Datasets using Data Mining," Journal of Basic & Applied Sciences, vol. 13, pp. 466-471, 2017.

[3] K. K. S. S.Selvakumar, "Prediction of Diabetes Diagnosis Using Classification Based Data Mining Techniques," International Journal of Statistics and Systems, vol. 12, no. 2, pp. 183-188, 2017.

[4] "Analysis and Prediction of Diabetes Mellitus using Machine Learning Algorithm," International Journal of Pure and Applied Mathematics, vol. 118, no. 9, pp. 871-878, 2018.

[5] M. R. A. ,. M. D. K. Mr. R. Sengamuthu, "VARIOUS DATA MINING TECHNIQUES ANALYSIS TO PREDICT DIABETES MELLITUS," International Research Journal of Engineering and Technology, vol. 5, no. 5, pp. 676-679, 2018.

[6] K. Nandhini.M, "Prediction of diabetes using Classification algorithms," International Journal of Science, Engineering and Management, vol. 2, no. 12, pp. 287-291, 2017.

[7] D. S. S. Deepti Sisodia, "Prediction of Diabetes using Classification Algorithms," International Conference on Computational Intelligence and Data Science, p. 1578–158, 2018.

[8] Y. A. K. Z. M. A. Amina Azrar, "Data Mining Models Comparison for Diabetes Prediction," International Journal of Advanced Computer Science and Applications, vol. 9, no. 8, pp. 320-323, 2018.

[9] D. M. R. Manimaran, "Prediction of Diabetes Disease Using Classification Data Mining Techniques," International Journal of Engineering and Technology, vol. 9, no. 5, pp. 3610-3614, 2017.

[10] P. P. M. C. Tejas N. Joshi, "Diabetes Prediction Using Machine Learning Techniques," S. Dewangan.et.al. Int. Journal of Engineering Research and Application, vol. 8, no. 1, pp. 09-13, 2018.

[11] I. P. A. M. T. Ahmad Ashari, "Performance Comparison between Naïve Bayes, Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool," International Journal of Advanced Computer Science and Applications, vol. 4, no. 11, pp. 33-39, 2013.

[12] Sethi, H., Goraya, A., & Sharma, V. (2017). Artificial Intelligence based Ensemble Model for Diagnosis of Diabetes. International Journal of Advanced Research in Computer Science, 8(5).

[13] O. T. ,. A. S. ,. N. M. I. V. ,. I. C. Ioannis Kavakiotis, "Machine Learning and Data Mining Methods in Diabetes Research," Computational and Structural Biotechnology Journal, p. 104–116, 2017.

[14] S. J. S. R. Priyanka Shetty, "A Tool for Diabetes Prediction and Monitoring Using Data Mining Technique," I.J. Information Technology and Computer Science, pp. 26-32, 2016.

[15] D. J.Omana, "Estimation of Association Summarization Techniques Performance in Prediction of Diabetes Mellitus," Journal of Network Communications and Emerging Technologies (JNCET), vol. 8, no. 3, pp. 46-51, 2018.