Artificial Intelligent - Based diabetes prediction system

Reported Work

Most of the massive topics for Researchers are Data predic- tion. Numerous researchers have done their research in finding out the technique which is good and helpful for predicting the diabetes more simpler way. Diferent methods for diabetes predicting have been listed below:

a. Diabetes prediction using machine

- learning Analysis of Data in vast data files to separate enclosed and earlier abstruse examples, links, and data dif- fcult to fnd with conservative assessment methods.
- Study of vast data sets is also a developing area of great implication in social insurance. Researches using data mining proce- dures to study the patients' data which are benefcial to fnd important knowledge which is facilitating medical services and deeply study of disease.
- Diferent procedures follow to forecast Diabetes Melli- tus, through device learning technique. Many authors used particle swarm optimization (PSO) is a computational method algorithm to forecast type 2 DM researchers ofered 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 (SVR) many researchers focused on the glucose in fnding the 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 feld of high signifcance 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.

Comparison Analysis

- The diabetes prediction using various techniques are pre- sented in Table 2.
- Table 2 illustrates the purpose of diabetes prediction by exploring various performances of diferent algorithms presented by diferent authors on diferent param- eters which may include data set, feature extraction, classi- fer used and the results obtained by each author.

Conclusion

- The main objective of this paper is to improve the correct- ness of predictive models.
- The accuracy can be achieved by either refning 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 classi- fer. Among all, the genetic algorithm leads over others.
- In this research, the author has concluded one more impor- tant 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 diferent dataset. In future work, advanced classifers 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 dis- ease 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, Univer- sity 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 Sam- ple (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 Data- base (PIDD)	Plasma glucose concentra- tion, triceps skin fold, age	Knn	Accuracy = 82%
Sisodia et al. [29]	Real-world medical dataset	Plasma glucose concentration,blood pres- sure, 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 con- fict of interests that infuence the work reported in this paper