ISSN 2278-3091
Volume 9, No.5, September - October 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse160952020.pdf
https://doi.org/10.30534/ijatcse/2020/160952020

A UI Based Co-relational Analysis for Diabetic Prognosis using SVM Techniques

Challa Sundeep Babu[1], Dr G Siva Nageswara Rao[2]
[1]PhD Scholar, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram,AP, India, sundeepchalla789@gmail.com
[2] Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram,AP, India, sivanags@kluniversity.in

ABSTRACT

Diabetes is one among the common and growing concern disease in several countries and every one of them are working to stop this disease at early stage .The typical recognizing process is that patients got to visit the diagnostic center, consult specialist and stay for each day or more to urge their reports.

The objective of this model is to predict whether someone is diabetic or not, based upon input attributes like : age, body mass index(BMI) and blood pressure(BP) .This model uses support vector machine (SVM) which may be a supervised classification machine learning algorithm .

Key words: Diabetes, Correlation, SVM, Superior label, Hyperplane, AI, CBR

1.INTRODUCTION

Diabetes is one among the common and rapidly increasing diseases within the world. it's a serious ill health in most of the countries. Diabetes may be a condition during which your body is unable to supply the specified amount of insulin needed to manage the quantity of sugar within the body. This results in various diseases including heart condition, renal disorder, blindness, nerve damage and blood vessels damage. There are two general reasons for diabetes:[1][2][3]

(1) The pancreas doesn't make enough insulin or the body doesn't produce enough insulin. Only 5-10 you look after people with diabetes have this type of the disease (Type-1).[4][5]

(2) Cells don't answer the insulin that's produced (Type-2).

Diabetes disease diagnosis and interpretation of the diabetes data is a crucial classification problem. A classifier is required and to be designed that's cost efficient, convenient and accurate. Support Vector Machines (SVM) have shown remarkable success within the area of employee computer aided diagnostic systems to enhance diagnostic decisions.[6][7]

AI provides an excellent deal of human ideologies and is involved in human related fields of application. These systems find an area within the diagnosis.[8][9]

Now a days, the report for diabetes test may be a very lengthy process. the standard recognizing process is that patients got to visit the diagnostic centre, consult specialist and stay for each day or more to urge their reports. [10][11]

To make the diagnosis a way easy process , the clients are needed a system which outputs most accurate results and doesn't consume individual's time and deviate him/her from his day to day routine. [12][13]

A diagnosis may be a classification process. A physician has got to analyze lot of things before diagnosing the diabetes which makes physician's job difficult. So so as to form the procedure much easy and price efficient one, we are using machine learning technique for predicting the status of diabetes in a person.[14][15][16].

A function are often specified by identifying the state at which data is to be input to the system, its input file domain, the output domain, and therefore the sort of processing to be carried on the input file to get the output.

Following are the requirements:

- Should enter the small print like Name, Age, BMI and vital sign (BP).
- Should press the PREDICT button to understand the result.

Following are the non functional requirements:

- Should be available for installation on computer.
- Crashing of application shouldn't occur.
- reaction time for every query sent by the user should be minimum.
- Application should be updated properly.

2. LITERATURE SURVEY

Machine learning may be a branch of AI that aims at solving real world engineering problems. It provides the chance to find out without being explicitly programmed and it's supported the concept of learning from data. it's such a lot ubiquitously used dozen a times each day that we might not even realize it. The advantage of machine learning (ML) methods is that it uses mathematical models, heuristic learning, knowledge acquisitions and decision trees for deciding. Thus, it provides controllability, observability and stability. It updates easily by adding a replacement patient‘s record

The application of machine learning models on human disease diagnosis aids doctors supported the symptoms at an early stage, albeit some diseases exhibit similar symptoms. one among the important problems in multivariate techniques is to pick relevant features from the available set of attributes. The common feature selection techniques include wrapper subset evaluation, filtering and embedded models. Embedded models use classifiers to construct ensembles, the wrapper subset evaluation method provides ranks to features supported their importance and filter methods rank the features supported statistical measurements.

A computer aided diagnosis system generally consists of a knowledge domain and a way for solving an intended problem. On the idea of the query posted to the system, it provides assistance to the physicians in diagnosing the patients accurately. The knowledge domain of such medical systems relies on the inputs that originate from the clinical experience of field experts. Knowledge acquisition is that the process of remodeling human expert knowledge and skills acquired through clinical practice to software. it's quite time consuming and labour intensive task.

Common methods like Case Based Reasoning (CBR) solves the knowledge acquisition problem to some extent because the past records are maintained during a database, including possible remedies, past clinical decisions, preventive measures and expected diagnostic outcome measures. During patient diagnosis, the clinical database is matched for analogous past patient‘s record for taking suitable decisions.

Some of the main problems faced during the event of an expert diagnosis system are: doctors are less interested to share their knowledge with others, experience knowledge (called common sense) is practically impossible to be separated and designing a singular expert system for diagnosing all diseases is difficult.

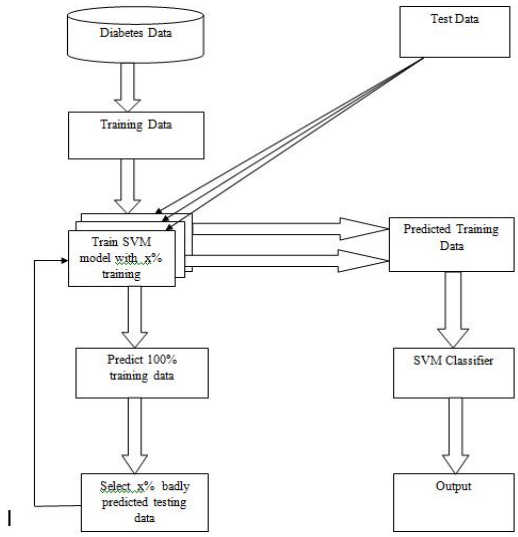


Figure 1: Architecture Design

3. SYSTEM DESIGN

Initially we'd like to load data of the person into application.

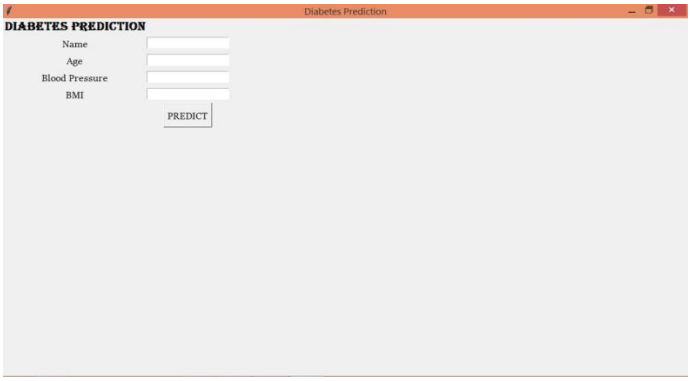
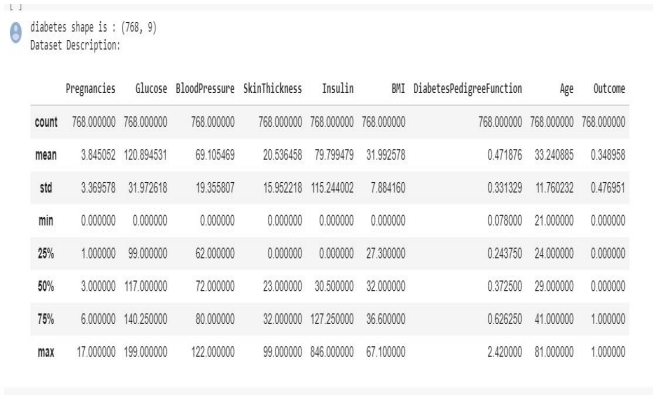


Figure 2: Data Load Page

The following figure displays the info set description.



diabetes shape is : (768, 9)
Dataset Description:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000
mean	3.845052	120.894531	69.105469	20.536458	79.799479	31.992578	0.471876	33.240885	0.348958
std	3.369578	31.972618	19.355807	15.952218	115.244002	7.884160	0.331329	11.760232	0.476951
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.078000	21.000000	0.000000
25%	1.000000	99.000000	62.000000	0.000000	0.000000	27.300000	0.243750	24.000000	0.000000
50%	3.000000	117.000000	72.000000	23.000000	30.500000	32.000000	0.372500	29.000000	0.000000
75%	6.000000	140.250000	80.000000	32.000000	127.250000	36.600000	0.626250	41.000000	1.000000
max	17.000000	199.000000	122.000000	99.000000	846.000000	67.100000	2.420000	81.000000	1.000000

Figure 3: Dataset Description

Next we can visualize the dataset using plot

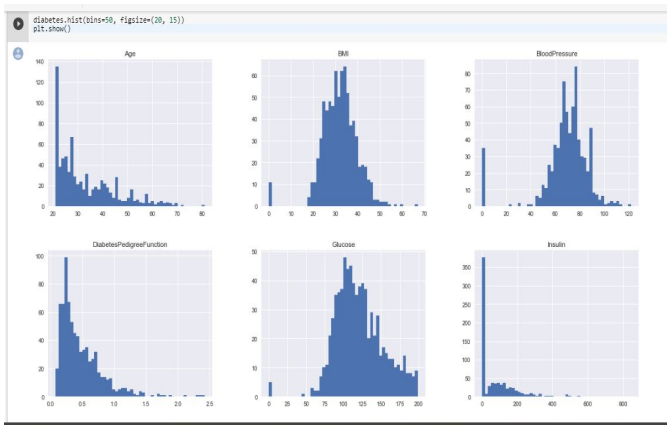


Figure 4: Dataset Visual using plot

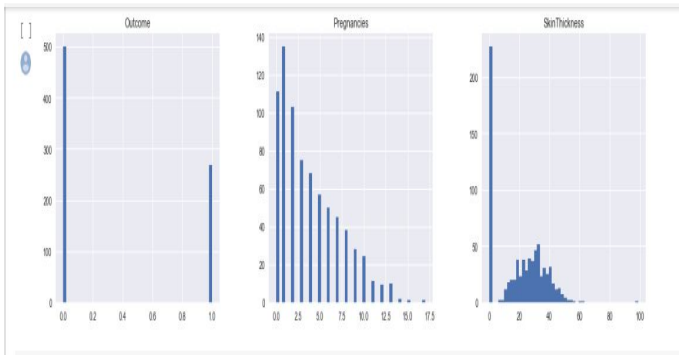
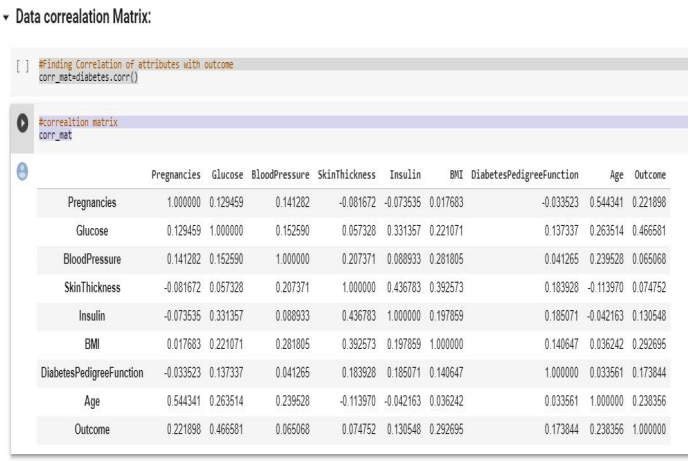


Figure 5: Dataset Visual using plot

We need to next print the result count as shown as below:

Figure 6: Outcome count

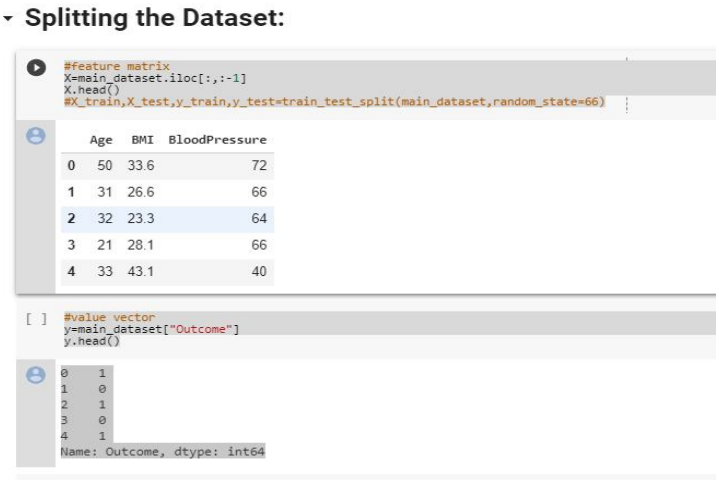
Finding Correlation of attributes with outcome



	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
Pregnancies	1.000000	0.129459	0.141282	-0.081672	-0.073535	0.017683	-0.033523	0.544341	0.221898
Glucose	0.129459	1.000000	0.152590	0.057328	0.331357	0.221071	0.137337	0.263514	0.466581
BloodPressure	0.141282	0.152590	1.000000	0.207371	0.088933	0.281805	0.041265	0.239528	0.065068
SkinThickness	-0.081672	0.057328	0.207371	1.000000	0.436783	0.392573	0.183928	-0.113970	0.074752
Insulin	-0.073535	0.331357	0.088933	0.436783	1.000000	0.197859	0.185071	-0.042163	0.130548
BMI	0.017683	0.221071	0.281805	0.392573	0.197859	1.000000	0.140647	0.036242	0.292695
DiabetesPedigreeFunction	-0.033523	0.137337	0.041265	0.183928	0.185071	0.140647	1.000000	0.033561	0.173844
Age	0.544341	0.263514	0.239528	-0.113970	-0.042163	0.036242	0.033561	1.000000	0.238356
Outcome	0.221898	0.466581	0.065068	0.074752	0.130548	0.292695	0.173844	0.238356	1.000000

Figure 7: Finding Correlation of Attributes

Spliting the Dataset



	Age	BMI	BloodPressure
0	50	33.6	72
1	31	26.6	66
2	32	23.3	64
3	21	28.1	66
4	33	43.1	40

Figure 8: Splitting the Dataset

4. CONCLUSION

Accuracy is up-to 64.38%. Results are supported Correlation analysis of Pima Indians Diabetic Dataset from the machine learning laboratory at University of California, Irvine. The app would contains details of the near by diagnostic centres supported one's location along side their list. A clean record of user's history is maintained along side credentials.

Would accompany medical shops for delivery of ordered medicines from the associated pharmaceutical stores.

All the patients data are trained by using SVM. Choosing the simplest kernel for SVM plays a crucial role within the outcome. within the proposed work, SVM with Radial basis function kernel is employed for classification. The performance parameter like classification accuracy of the SVM and RBF have found to be high, thus making it an honest option for the classification process.

REFERENCES

1.https://www.medicalnewstoday.com/info/diabetes/type2diabetes.php

2.http://scikitlearn.org/stable/modules/preprocessing.html

3. https://towardsdatascience.com/machine-learning-for-diabetes-562dd7df4d42

4. https://www.geeksforgeeks.org/ml-handling-missing-values/amp/

5. https://www.svm-tutorial.com/2014/11/svm-understanding-math-part-1/

6.. Venkata Vara Prasad Padyala , Dr K V D Kiran "**Improved handoff mechanism for infiltrating user equipments in composite networks**" IJECE Vol. 10, No. 3, pp. 2600~2606 ISSN: 2088-8708, June 2020.

7. Venkata Vara Prasad Padyala , Dr K V D Kiran**" A Prediction Schemme of mobility of cognitive FEMTOCELL LTE-A/LTE_UE under different speed scenarios**" IJET, 2018.

8. Vara Prasad P V , Sowmya N**"Introduction to dynamic malware analysis for cyber intelligence and forensics** " International Journal of Mechanical Engineering and Technology 2018

9. Guru Prasad, S., Srikanth, K. S., & Rajanna, B. V. (2018**). Advanced active power filter performance for grid integrated hybrid renewable power generation systems.**Indonesian Journal of Electrical Engineering and Computer Science, 11(1), 60-73

10. Swain, G. (2018).**High capacity image steganography using modified LSB substitution and PVD against pixel difference histogram analysis**. Security and Communication Networks, 2018 doi:10.1155/2018/1505896

11. Madhav, B. T. P., Reddy, V. S., Reddy, D. R., Sankar, K. R., Ramanujan, E. V. S. H., Prakash, V. V. S., & Rao, M. V. (2018).**Tree shaped fractal antenna with multiband characteristics**. International Journal of Engineering and Technology(UAE),

12. Madhav, B. T. P., Venkateswara Rao, M., & Anilkumar, T. (2018).**Conformal band notched circular monopole antenna loaded with split ring resonator. Wireless Personal**Communications, 103(3), 1965-1976. doi:10.1007/s11277-018-5890-3

13. Venkata Ramana, N., Nagesh, P., Lanka, S., & Karri, R. R. (2019). Big data analytics and iot gadgets for tech savvy cities doi:10.1007/978-3-030-03302-6_12.

14. Venkata Ramana, N., Seravana Kumar, P. V. M., Nagesh, P. (2018). Analytic architecture to overcome real time traffic control as an intelligent transportation system using big data. International Journal of Engineering and Technology(UAE), 7(2.18 Special Issue 18),

15. Nagesh, P., Radha Madhavi, M., Surendra Prasad Babu, M. Reusable object oriented design for stereographic projections of tetragonal point groups . International Journal of Innovative Technology and Exploring Engineering,v0lume.8,No.8,177-180,2019

16. G Siva Nageswara Rao "**Food Waste Protein sequence analysis using clustering and classification techniques "**in International Journal of Advanced tends in computer science and Engineering (ISSN 2278-3091) volume 8, Issue 5, pp 2289-2298, October 2019.