International Journal of Scientific Research in Computer Science and Engineering **Research Paper**

Vol.8, Issue.1, pp.93-97, February (2020) **E-ISSN:** 2320-7639

Reasoning with Certainty Factor for Prediction of Diabetes Disease on Machine Learning Platform

M.M. Mastoli[1*], U.R. Pol[2], Rahul D. Patil[3]

[1,2]Department of Computer Science Shivaji University Kolhapur, Maharashtra, India
[3]Quality Assurance Engineer, Menon Bearing Ltd, Kolhapur, Maharashtra

[*]*Corresponding Author: it.manjiri@gmail.com, Tel.: 02312609347*

Available online at: www.isroset.org

Received: 18/Jan/2020, Accepted: 25/Jan/2020, Online: 28/Feb/2020

***Abstract*—** Diabetes Mellitus commonly called as diabetes, is one of the common and growing endocrine diseases. AI provides ability to prevent these type of disease at an early stage by predicting the symptoms using several methods. Two areas which may benefit from the application of Machine Learning techniques in the medical field are diagnosis and outcome prediction. Expert systems as a branch of AI can incorporate with machine learning tools, this technology can be used as a solve the problem where there is a shortage of physician in the field of healthcare. The researcher has designed and developed a machine learning model for predicting the diabetes and its types like expert systems. This present reasoning with certainty factor and considered when designed the knowledge base with the patient's level of belief. The main purpose of this paper is to study certainty factor for prediction of diabetes disease for more accurate prediction and diagnosis of diseases and application of machine learning for heath care systems.

Keywords— Artificial intelligence, Machine Learning Model,Certainty Factor, Expert System,Diabetes Mellitus.

I. INTRODUCTION

Motivated by an increase in computation, storage and the generation of amazing volumes of data,where computers are being utilized to perform a many of complex tasks with remarkable accuracy. Machine learning (ML) is the name given to both the academic discipline and technology which allow computers to perform complex tasks. As an academic discipline, machine learning comprises elements of, statistics, mathematics, computer science and electronics. Machine learning is the technology which is helping to drive advances in the development of artificial intelligence. It is impressively employed in both academia and industry to drive the development of intelligent solutions with the ability to make accurate predictions using diverse sources of data [1].

The increasingly growing number of applications of artificial intelligence, machine learning and deep Learning in healthcare allows us to foretaste at a future where data, analysis, and innovation work hand-in-hand to help countless patients without them ever realizing it. Soon, it will be quite common to find ML&AI-based applications embedded with real-time patient data available from different healthcare systems in multiple countries, thereby increasing the efficacy of new treatment options which were unavailable before.

Healthcare covers detailed processes of the diagnosis, treatment and prevention of disease,[2]. The medical industry in most countries is evolving at a rapid space. The healthcare industry with rich data as they generate massive amounts of data, including electronic medical records, administrative reports and other findings [3].

Health informatics are becoming a very research-intensive field and the largest consumer of public funds. With the occurrence of computers and new algorithms, health care has seen an increase in computer tools and could no longer ignore these emerging tools. This resulted in uniting of healthcare and computing to form health informatics. This is expected to create more efficiency and effectiveness in the health care system, while at the same time, improve the quality of health care and lower cost. [4]

The advantages of artificial intelligence and machine learning in medicine are vast. Where the expert system can be able to convert at machine learning platform which will predict more accurate results.

Expert System is a computer program with collection of facts, rule and knowledge about a particular domain. This emulates the decision making ability of a human expert and acts in all respects like a human expert. Uncertainty is a challenging part in human's everyday life. There is a high level of uncertainty management in intelligent systems. This is because human reasoning and decision making is fuzzy, involving a high degree of vagueness in evidence, concept utilization and mental model formulation.thi present resrech will utilize the power of certainty factor for prediction of diabetes disease for more accurate prediction and diagnosis

II. DIABETES MELLITUS AND MEDICAL EXPERT SYSTEMS

When human body is unable to store and use glucose (sugar) properly, the disease is referred as Diabetes Mellitus commonly known as diabetes [5].

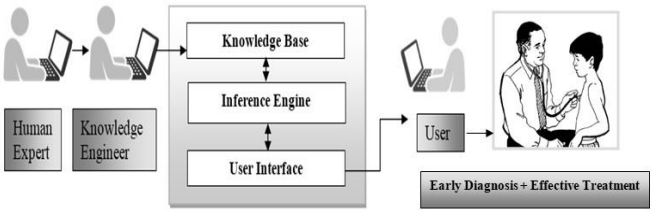

Fig-1: Expert Systems in Healthcare

Many expert systems have been developed in the area of healthcare. Following are the table-1 represents summary for expert system in diabetes healthcare.

Table 1: Summary of expert system in diabetes healthcare.

Sr. No.	Author	Approaches	Output
1	Tawfik Saeed Zeki et al [6]	VP-Expert (Khashavi-Najafabadi, Bayat, 2010), (VP_Expert Primer)	Indications, diagnosis and primary treatment advices to the diabetics
2	Dilip Kumar Choubey et al [7]	A rule based expert fuzzy expert system	Diagnosis of diabetes.
3	Anindito Yoga Pratama et al [8]	Android base mobile application expert system	Determines a person's diabetes risk.
4	M. Kalpana et al[9]	Fuzzy expert system	Diagnosis of diabetes.
5	Danijela TADIC et al [10]	Fuzzy model	Type 2 diabetes for evaluation and choice of optimal therapeutic procedure on individual level.
6	Hanslal Prajapati et al [11]	Fuzzy Inference System (FIS)	Diagnosis of Diabetes
7	Abdulla Al-Malaise Al-Ghamdi et al [12]	Cloud computing based expert system	Better treatment of diabetes.
8	Ibrahim M. Ahmed et al [13]	Prolog Expert system	treatment of diabetes type 2
9	Seyedeh Talayeh Tabibi et al [14]	Expert system	advices for diabetics' treatment
10	Meysam Rahmani Katigari et al [15]	Expert system	Detect and diagnose the severity of diabetic neuropathy.
11	A.D.Dhivya et al [16]	Fuzzy Expert System	Early diagnosis of the disease and proceed the treatment
12	K.Vijaya Lakshmi et al [17]	Expert system	Diagnose gestational diabetes mellitus
13	Onuiri Ernest E et al [18]	Web based application expert system	Help medical practitioners in the recording of patients' glucose level and measuring the right insulin injection for a patient
14	R. Radha et al [19]	Fuzzy Expert system	Forming the relations for prediction
15	Vishali Bhandari et al [20]	Mamdani-type and Sugeno-type Fuzzy expert systems	Comparative FIS used for prediction

III. CERTAINTY FACTOR (CF)

Uncertainty with the likelihood of the consequent.the main method for dealing with uncertainty uses certainty factors. Certainty factor important way of combining belief and disbelief into a single number and used to rank hypotheses in order of importance.

Certainty factors are defined on the interval: where

$$-1 \leq CF(H,E) \leq 1$$

Where, 0 means no evidence, values greater than 0 favor the hypothesis, and less than 0 favor the negation of the hypothesis. The above CF values might be elicited by asking questions to paitients:

How much do you believe that feeling 'Inceased_Thirst' being diabetic? If the evidence is to confirm the hypothesis

or

How much do you disbelieve that making a 'Slurred_Speech' being diabetic?

An answer of 80 percent to each question will set
CF(H | E) = 0.80 and
CF (H' | E) = -0.80.

The certainty factor is used to accommodate the uncertainty of an expert who is often thought of analyzing information with expressions such as ***May, likely, almost certainly*** etc.

Notation certainty factor is:

$$CF[h,e]=M[h,e] - MD[h,e] \quad \ldots\ldots\ldots\ldots (I)$$

Where,

CF [h, e] represents the certainty factor and

MB [h, e] defines the size of the confidence in the hypothesis h, if given evidence e (between 0 and 1),

MD [h, e] defines the size of the distrust of hypothesis h, if given evidence e (between 0 and 1).

Arising from existing evidence there are three things that might happen to the hypotheses. First one, the rules with *single* evidence *e* and *a single* hypothesis of *h*, where the notation certainty factor is:

$$CF[h,e]=CF(e) \times CF(rule) \quad \ldots\ldots\ldots\ldots (II)$$

In practice, certainty factor, i.e. CF (rule) is determined by experts, whereas certainty factor, CF (e) is determined by the user in our case these are patients, while consulting with the system. Next second rules are with *multi evidence e* and *single* hypothesis *h*. Notation for certainty factor is determined by the interface used, could use a disjunction or conjunction. Where notation on connective disjunction is written as follows:

IF e1 AND e2 ... AND en THEN h (CF rule)

.....(III)

Where calculation of the value of CF combination, defined by:

CF(h,e)=min[CF(e1),CF(e2),...,CF(en)]*CF(rule) .....(IV)

In the meantime, for connecting conjunction, can be written as follows:

IF e1 OR e2 ... OR en THEN h (CF rule) .....(V)

So the calculation of the value of the certainty factor(CF) combination, defined by the following notation:

CF[h,e] = max[CF(e1),CF(e2),...,CF(en)]*CF(rule) .....(VI)

Last third, the combination of two rules with different *evidence e1 and e2* but refers to the *same* hypothesis. The equations for combining two certainty factor are as follows:

$$CF(CF1, CF2) = \begin{cases} CF_1 + CF_2(1 - CF_1) & if\ CF_1 > 0\ and\ CF_2 > 0 \quad \text{VII.A} \\ \frac{CF_1 + CF_2}{1 - [|CF_1|, |CF_2|]} & if\ CF_1 < 0\ or\ CF_2 < 0 \quad \text{VII.B} \\ CF_1 + CF_2(1 + CF_1) & if\ CF_1 < 0\ and\ CF_2 < 0 \quad \text{VII.C} \end{cases}$$

In the current study, the combination formulation certainty factor calculations, using the formula VII.A, to calculate the value of certainty factor dual premise, while for a single premise, we use the formula 4, the terms entered by the user certainty factor ranges between (0 and 1).

IV. PROPOSED METHODOLOGY

The decision tree is a machine learning technique which used by researchers for for the forecasting and classification of the diabetes. This supervised learning algorithm that follows the greedy approach and works in a top down manner. It can handle a large database and works well with both numerical and categorical variables and uses a white box model approach and classifies the data in a hierarchical structure. It makes data easy to understand and represent too.

This research methodology is divided into several steps, start with primary and secondary data collection of symptoms of each type of diabetes. Once the symptoms of the patients collected from a physician, researcher has classified the symptoms, according to the type of diabetes. The designing dataset is a complex phenomenon which needs expert knowledge which is input for machine learning model. With the help of expert researcher has designed the Machine Learning Model. After creating the machine learning model for prediction of diabetes the model is then tested with some patients for accuracy calculation[21].

V. RESULTS AND DISCUSSION

During data collection regarding the symptoms by patients, they also requested to write their certainty factor, according to their beliefs for respective symptoms. This certainty factor value will be calculated according to the value of the certainty factor between -1 to 1, there are some symptoms with binary value with a definite value either 0 or 1. (*values not applicable)[22-23]

Columns represented as

A- Type2Diabetes, B-Type1Diabetes, and
CFa- Certainty factor for Type2Diabetes,.
CFb- Certainty factor for Type1Diabetes,

Table 1: Dataset as Knowledge Base List of Symptom with CF values entered by Patients

Sr. No.	Symptoms	A	CFa	B	CFb
1	Family_History	0	*	1	*
2	Age	11	*	12	*
3	Obesity	22	.90	23	.10
4	Previous_IFG_IGT	0	*	1	*
5	Hypertension	0	.90	5	.10
6	HDL_Cholesterol	0	*	3	*
7	Triglyceride	0	*	5	*
8	Inceased_Thirst	1	.95	1	.30
9	Increased_Urinate	1	.95	1	.30
10	Increased_Appetite	1	.90	1	.10
11	Weight_Variation	0	*	1	*
12	Impaired_Vision	0	.70	1	.50
13	Tiredness	1	.60	1	.40
14	Impatience	0	.50	1	.40
15	Infection	0	.30	1	.30
16	Itchy_Skin	0	.40	0	.40
17	Depression_Stress	0	.60	0	.60
18	Tingling_Sensation	0	.70	0	.70
19	Fruity_Breath_Odour	1	.40	0	.20
20	Bed_Wetting	1	.30	0	.30
21	Slow_Healing_Wound	0	.70	1	.70
22	FamilyHis_Pregnancy	0	*	0	*
23	Previous_Pregnancy	0	*	0	*
24	BabyOver_9Pd_PrePreg	0	*	0	*
25	Sleeplessness	1	.40	1	.30
26	Trembling	1	.40	1	.20
27	Sweating	1	.30	1	.20
28	Anxiety	1	.60	1	.50
29	Confusion	1	.20	1	.20
30	Weakness	1	.50	0	.30

31	Mood_Swings	1	.20	0	.10
32	Nausea	1	.10	0	.10
33	Vomiting	1	*	0	*
34	Dry_Skin	0	.10	1	*
35	Aches&Pains	0	.30	1	.30
36	Recurresnt_fungal_infectn	0	*	1	*
37	Nightmares	1	.40	1	.10
38	Seizures	1	*	1	*
39	Sadness	1	.10	1	.20
40	Unconsciousness	1	.10	1	*
41	Numbness	1	*	1	*
42	VaginalMycoticInfectn	1	*	1	*
43	Rapid_Heart_Beat	0	.20	1	.10
44	Recurring_Gum_Infe	0	.10	0	*
45	Impotency	1	*	1	*
46	high blood Pressure	0	.60	0	.60
47	Sleep_Walking	1	*	1	*
48	Makeg_unusual_noises	1	.10	1	.10
49	Leg_Cramps	1	.50	1	.20
50	Slurred_Speech	1	.10	1	.20
51	Flushed_face	1	.10	1	.10
52	Pale_Skin	1	*	0	*
53	LossofMenstruation	1	*	1	*
54	Stomach_Pain	1	.30	1	.10
55	Deep_Breathing	1	*	1	*
56	Areas_Darked_Skin	0	*	1	*
57	Difficult_Concentrating	1	*	1	.10
58	Dehydration	1	.30	1	.20
59	LackofCoordination	1	.10	1	.20
60	Hist_Heart_Stroke	0	.10	0	*
61	Poly_Ovary_Syndrome	0	*	0	*
62	LowbloodSugar_NewbornBaby	0	*	0	*
63	WaistSize02cmM88cmF	0	*	0	*
64	WaistHipRatio.9M.85F	0	*	0	*

Based on the symptoms are selected by patients, the system detects types of diabtes diseases that meet these symptoms.

Table 2: Symptom with CF values entered by Patients with predicted type of diabetes notify by (✓)

CFno	Symptoms	A	CFa	B	CFb	A	B
CF8	Inceased_Thirst	1	.95	1	.30	✓	
CF9	Increased_Urinate	1	.95	1	.30	✓	
CF10	Increased_Appetite	1	.90	1	.10	✓	

For example,from table 2

The symptom number 8 i.e. CF8 = 0.95 and symptoms 9 CF9 =0.95, model prediction is that the patient suffering for type2diabetes.There for the CF value for the type2diabetes must be high.

Table 3: Prediction using CF value.

Predicted type of diabtes	CF value	Accuracy
Type2Diabetes	0.90	100%

VI. CONCLUSION AND FUTURE SCOPE

The certainty factors in the expert system are very useful, but those are valued only after completion with the expert systems. This approach is necessary for the accuracy of the model.The machine learning model is able to predict the type of diabetes and works like expert system. Use of methods of certainty factor in machine learning platform will also enhance the processing and predition accuracy.

REFERENCES

[1] Jordan MI, Mitchell TM. “Machine learning: Trends, perspectives, and prospects”. Sci (NY). 2015; 349(6245):255–60.

[2] J.-J. Yang, J. Li, J. Mulder, Y. Wang, S. Chen, H. Wu, Q. Wang, and H. Pan, “Emerging information technologies for enhanced healthcare,” Comput. Ind., vol. 69, pp. 3–11, 2015.

[3] Nilmini Wickramasinghe,Rajeev K. Bali, Brian Lehaney , Jonathan L. Schaffer , M. Chris Gibbons “Knowledge Management in Healthcare Primer” Routledge 270 Madison Ave, New York, pp 7-15,2009

[4] Shortliffe, EH.,Perrault, LE., (Eds.). “Medical informatics: Computer applications in health care and biomedicine” (2nd Edition). New York: Springer, 2000.

[5] M. M. Mastoli,U. R. Pol,Rahul D. Patil, "AI for Diabetic Retinopathy",ISROSET Publisher,India,International Journal of Scientific Research in Computer Science and Engineering, Vol.7, Issue.6, pp.29-34, 2019

[6] Tawfik Saeed Zeki, Mohammad V. Malakooti, Yousef Ataeipoor, S.Talayeh Tabibi ,“An Expert System for Diabetes Diagnosis”, American Academic & Scholarly Research Journal Vol. 4, No. 5, Sept 2012

[7] Dilip Kumar Choubey, Sanchita Paul, Vinay Kumar Dhandhenia “Rule based diagnosis system for diabetes” Biomedical Research 2017; 28 (12): 5196-5209,

[8] Anindito Yoga Pratama, Dewi Agushinta R., and Remi Senjaya ,“Design Of Mobile Expert System For Diabetes Risk Diagnosis And Information”, Journal of Information Systems, Volume 9, Issue 1, April 2013, 32-36

[9] M. Kalpana, Dr. A.V Senthil Kumar,“Diagnosis of Diabetes using Correlation fuzzy logic in Fuzzy Expert System”, International Journal of Advanced Research in Computer Science, 3 (1), Jan –Feb, 2012,244-250

[10] D., Tadić, P., Popović, A., Đukić,“A Fuzzy Approach to Evaluation” Yugoslav Journal of Operations Research Volume 20 (2010), Number 1, 99-116 , 10.2298/YJOR1001099T

[11] Hanslal Prajapati1, Anurag Jain2, Sanjay Kumar Pal3 “An Enhance Expert System for Diagnosis of Diabetes using Fuzzy Rules over PIMA Dataset”, International Journal of Advance Engineering and Research Development (IJAERD) Volume 4, Issue 9, September-2017, e-ISSN: 2348 - 4470, print-ISSN: 2348-6406

[12] Abdulla Al-Malaise Al-Ghamdi et al, “An Expert System of Determining Diabetes Treatment Based on Cloud Computing

Platforms", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (5) , 2011, 1982-1987
[13] Alfonese, Marco et al. "Daily Meal Planner Expert System for Diabetics Type-2." (2015).
[14] Seyedeh Talayeh Tabibi , Tawfik Saeed Zaki , Yousef Ataeepoor," Developing an Expert System for Diabetics Treatment Advices", International Journal of Hospital Research 2013, 2(3):155-162
[15] Meysam Rahmani Katigari, Haleh Ayatollahi, Mojtaba Malek, and Mehran Kamkar Haghighi, "Fuzzy expert system for diagnosing diabetic neuropathy", World J Diabetes. 2017 Feb 15; 8(2): 80–88.
[16] A.D.Dhivya , A.Felix ,"A Fuzzy Rule based Expert System for T2DM Diagnosis" International Journal of Engineering & Technology, 7 (4.10) (2018) 432-435
[17] K.Vijaya Lakshmi,E.Sreedevi, Prof.M.Padmavathamma "Modeling An Expert System For Diagnosis Of Gestational Diabetes Mellitus Based On Risk Factors" Publications Of Problems & Application In Engineering Research – Paper, Vol 04, Special Issue01; 2013
[18] Onuiri Ernest E , Ndukwe Victor U , Igwe Nkechinyere , Olise Saviour C. "Simulation of an Expert System for Diabetes Diagnosis" International Journal of Advanced Computing, ISSN:2051-0845, Vol.48, Issue.1 1614-1621
[19] R. Radha and S.P. Rajagopalan,"Fuzzy Logic Approach for Diagnosis of Diabetics", Information Technology Journal Volume 6 (1): 96-102, 2007
[20] Vishali Bhandari, Rajeev Kumar ,"Comparative Analysis of Fuzzy Expert Systems for Diabetic Diagnosis", International Journal of Computer Applications Volume 132 - Number 6.
[21] Tb. Ai Munandar , Suherman,Sumiati, "The Use of Certainty Factor with Multiple Rules for Diagnosing Internal Disease", International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 1, Issue 1, September 2012 ISSN 2319 – 4847
[22] M.M.Mastoli,U.R.Pol,R.D.Patil,"Machine Learning Model For Prediction of Diabetes Mellitus", International Journal of Recent Technology and Engineering, Blue Eyes Intelligence Engineering & Sciences Publication, Volume-8 Issue-5, January 2020
[23] M.M.Mastoli, U.R.Pol, R.D.Patil, "Machine Learning Classification Algorithms for Predictive Analysis in Healthcare", International Research Journal of Engineering and Technology, Volume-6 Issue-12, pp.1225-1229,2019.

AUTHORS PROFILE

Ms. Manjiri Mahadev Mastoli pursed Master of Computer Application from CSIBER, Shivaji University, Kolhapur, Maharashtra. She is currently pursuing Ph.D. and currently working as research Scholar in Department of Computer Sciences, Shivaji University, Kolhapur, Maharashtra. Her main research work focuses on Artificial Intelligence and Machine Learning, Data Science, She has 6 years of teaching experience and 5 years of Research Experience.

Dr.Urmila Pol pursed Bachelor of Science from Shivaji University and Master of Compter Application from CSIBER, Shivaji University, Kolhapur, Maharashtra. She is currently working as Assitant Professor in Department of Computer Sciences, Shivaji University, Kolhapur, Maharashtra. She is Member of International Association of Engineer. Currently, she is guiding four Ph.D and one M.Phil student. Her main research work focuses on Artificial Intelligence, Data Science, Big Data Analytics, LMS and Open Source Technologies, She has 20 years of teaching experience and 15 years of Research Experience.

Mr. R.D. Patil pursed Bachelor of Enginerring Shivaji University, Kolhapur and Master of Business Administration from IGNOU. He is currently working as Quality Assurance Engineer at Menon Bearing Ltd, Kolhapur, Maharashtra. His main research work focuses on Oprational Research, Artificial Intelligence and Machine Learning. He has 8 years of Quality Assurance experience and 4 years of Research Experience.

