

Treatment of Diabetes Type II Using Genetic Algorithm

<https://doi.org/10.3991/ijoe.v15i11.10751>

Majdoleen Al Switi, Bahaaldeem Alshraideh (✉), Abedalrhman Alshraideh,
Abudalla Massad, Mohammad Alshraideh
The University of Jordan, Amman, Jordan
b2alshraideh@yahoo.com

Abstract—Chronic diseases is an important research field because of the growth of the number of affected people around the world. When someone has diabetes, the body either does not make enough insulin or cannot use its own insulin as well as it should. This causes sugar to build up in the blood leading to complications like heart disease, stroke, and neuropathy. Poor circulation leading to loss of limbs, blindness, kidney failure, nerve damage, and death. Diagnosis plays a vital role in diabetes treatment otherwise it leads to long term complications in terms of costs of the treatment of the patients and leads to many risks over the patient himself as mentioned above. In this research, we propose a new methodology to extract the best testing sequence evaluation mechanism for helping doctors to evaluate their patient's cases and make the best decisions about the medicine being given. We managed to create chromosomes population each of which consists of a binary decision tree, as this implementation considered being the best scenario of our problem. The system proves its efficiency by applying it on 50 patients and the results show an accuracy percentage of 95.4%.

Keywords—Clinical decision support system, Genetic algorithms, Diabetes Management.

1 Introduction

Diabetes has become the fourth leading cause of death in developing countries especially in the third world and there is a significant indication that it is reaching wave proportions in many developments and newly developed countries, with evidence pointing to preventable factors such as inactive lifestyle and poor diet [1, 10, 11].

According to the World Health Organization, it affects around 194 million people worldwide, and that number is expected to increase to at least 300 million by 2025 [2, 8,9,14]. Extra researches show that in 2012, about half of all adults about 117 million people had one or more chronic health conditions. The main objective of this research is to find the best testing criteria depending on patient lab test results metrics which in return leads to best drug suggestion assessment with high speed and accuracy outputs, using the genetic algorithm technique. For this purpose, we proposed to implement a genetic algorithm which consists of a population (Genes) set, each of them which contains the full 8 lab test metrics (Body Mass Index, fasting blood sugar test, Low-

density lipoprotein, High-density lipoprotein, HBA1C, Diastolic blood pressure, Tri-glyceride, Previous Treatment) that describe the patient situation and direct the expert to the best diagnoses and best medicine and portion. We propose a new methodology to allow each gene to have its own thresholds and suggestion data flow conditions. Also considered the evaluation of each gene to be the average of correct patients' suggestions, from a set of 50 real patients test and decisions made by specialized doctors. The algorithm starts by initiating the population with the random ordering of the 8 lab parameter evaluation and random binary path sequences, and continue to test the best candidates to be the parents to the next generation. Crossover and mutation for each generation are done randomly until a generation meets the desired accuracy threshold in the evaluation phase.

2 Background and Related Works

2.1 Diagnosis of diabetes

There are many factors must be taken into our consideration in order to classify the patient if he or she diabetic, so that the diagnosing of diabetes mellitus is not an easy task to physicians. Diagnosing of diabetes mellitus is not depended only over the knowledge of physicians about the disease but it also depends on the experience of the physicians with previous same cases and same condition of patients [15].

The decision of physicians is done by comparing the results of the patient with another result which has been done before to another patient who had the same case and the same conditions. According to what we mentioned before that the physicians need to have a look to previous results, so they need to have a very important system which is called classifier system, that this system helps to list all previous decision which has been done over the patient who has same conditions [15].

2.2 Diabetes mellitus treatments

The goal of diabetes management is to keep blood glucose levels as close to normal as safely possible. Since diabetes may greatly increase the risk for heart disease and peripheral artery disease, measures to control blood pressure and cholesterol levels are essential parts of diabetes treatment as well.

People with diabetes must take responsibility for their daily care. This includes monitoring blood glucose levels, dietary management, maintaining physical activity, keeping weight and stress under control, monitoring oral medications and, if required, insulin use via injections or pump. To prove the importance of patient self-care, many societies such as UCSF (University of California San Francisco) Diabetes Teaching Center offers self-management education programs that emphasize individualized diabetes care. The program enables patients to make more consistent and appropriate adjustments in their therapy and lifestyle [1].

Do not try to structure your paper by lists, do not misuse list-items as headings. A list item (a bullet, a dash) contains maximum one paragraph. If there is more than one

paragraph in one list item then it's most likely a sub-section. Consider using a run-in heading level 3 or 4.

2.3 Diabetes type 2 treatment

Although type 2 of diabetes has less risk than type 1, Sometimes blood sugar levels remain high in people with type 2 diabetes even though they eat in a healthy manner and exercise. Therefore doctors treat these cases using one of 2 methods:

Insulin therapy: Using an insulin injection or its alternative solution the insulin pump. The pump is about the size of a pager and is usually worn on a patient belt. Insulin is delivered through a small tube (catheter) that is placed under the skin (usually in the abdomen). There are four major types of insulin(Rapid-acting, Short-acting, Intermediate-acting, and Long-acting) [1].

Oral medications: When a patient case does not require direct insulin supplement medications taken in pill form may be prescribed. The medications work in several different ways. These include improving the effectiveness of the body's natural insulin, reduce blood sugar production, increase insulin production and inhibit blood sugar absorption. Oral diabetes medications are sometimes taken in combination with insulin [1].

2.4 Genetic algorithms overviews

Genetic algorithms are based on the ideas of natural selection and genetics. Non-figurative representations also called 'chromosomes' or 'genotype' of candidate solutions and 'individuals' or 'phenotypes' are subject to Genetic Algorithm optimization problems [3, 12]. To simplify the idea, a Genetic Algorithm search space is comprised of genes corresponding with letters, chromosomes equal to words and the genotype with a word family [4, 13].

Genetic algorithms mimic "the survival of the fittest" of individuals dominating over the weaker ones over successive generations which is analogous with the behaviour of DNA chromosomes within a population of individuals [5].

Genetic Algorithms Operations

Crossover: In crossover (or recombination), two programs are selected from the population, both are then copied to a mating pool. A crossover point is randomly chosen in each program, and the subtrees below the crossover points are swapped. The two programs, with swapped subtrees, are then copied to the new population [3].

Mutation: In mutation, a single program is selected from the population and copied to a mating pool. A mutation point is chosen randomly, somewhere in the program, and the subtree below the mutation point is replaced with a new, randomly generated subtree. The new program is then copied into the new population [3].

Selection: The selection process determines which individual should be passed to the next generation. The fitness value is the selection key of the individuals. However, a number of methods help us to select the individuals based on the fitness values.

Basic terms of genetic algorithms

Function pool: It is a set of functions that will be used by the intermediate nodes in the structure of the tree. The function pool can contain different types of functions which are problem dependent. All these functions may have a different number of inputs but always have a single output e.g. for logical problems logical functions like AND, OR, etc. are used. The function pool that will be used is (+, -, *, /, square, $\sqrt{\quad}$, sin, cos, asin, acos, log, abs, reciprocal) [6].

Fitness function: The most significant concept of genetic programming is the fitness function. Genetic Programming can solve a number of problems; it is the fitness function which connects GA to the given problem. Each individual is assigned a fitness value by this function. It determines how well a solution is able to solve the problem. It varies greatly from one type of problem to another. It is chosen in such a way that highly fitted solutions have high fitness value. Fitness function is the only index to select a chromosome to reproduce for the next generation [7].

Testing criteria: The testing in our system is based on two metrics: Accuracy, and ROC.

Accuracy: Let P_c be the number of correctly treated patients, P_r be the number of wrong treated patients. Then accuracy is defined as:

$$\text{Accuracy} = \frac{P_c}{P_c + P_r} * 100\%$$

2.5 Using genetics algorithms in medical problem

Genetic algorithm has already been used by a lot of authors to classify 2-class problems [8, 16]. In [17] Karegowda et al. used neural networks and presented a hybrid model which uses Genetic Algorithms (GA) and Back Propagation Network (BPN) for classification of diabetes among Pima Indians.

In [18] Polat et al. proposed two different approaches for diabetes data classification - principal component analysis and neuro-fuzzy inference and Generalized Discriminant Analysis (GDA) and least square support vector machine (LS-SVM). They achieved an accuracy of 89.47% and 79.16% respectively. Muni, et al [19] proposed a method for multiclass classifier and introduced a new concept of unfitness for improving genetic evolution. Hasan Temurtas et al [20] proposed a neural approach for classification of diabetes data and achieved 82.37% accuracy.

Pradhan et al. used Comparative Partner Selection (CPS) along with GP to design a 2-class classifier for detecting diabetes [9]. Cheung used C4.5, Naive Bayes, BNND and BNNF algorithms and reached the classification accuracies 81.11%, 81.48%, 81.11% and 80.96%, respectively [21]. Ephzibah [22] used a fuzzy rule-based classification system for feature subset selection in diabetes diagnosis. This approach proves to be cost-effective. Arcanjo et al. proposed a KNN-GP (KGP) algorithm, a semi-supervised transductive one, based on the three basic assumptions of semi-supervised learning. This system was implemented on 8 datasets of UCI repository but inferior results were obtained for diabetes dataset [23].

Kishore et al. [24] proposed an interesting method which considers a class problem as a set of two-class problems [19]. Smart investigated the multi-class approach

using modified genetic operators and concluded that GP can be used to improve multi-class problems [26]. Lim et al. presented an excellent comparison of 33 classification algorithms in [25]. They used a large number of benchmark data sets for comparison. None of these 33 algorithms uses GP [19]. In [27] Sonu Kumar and Archana Singh proposed an effective methodology for the automated detection of Diabetes Mellitus. It states that the user sitting at home can diagnose whether he/she is suffering from Diabetes Mellitus or not. In [28] Alexis Marcano-Cedeno and Diego Andina proposed the artificial meta plasticity on multilayer perceptron (AMMLP) as a data mining (DM) technique for the diabetes disease diagnosis.

In [29] Tamilarasi et al developed a system to predict Diabetes using Genetic Algorithm. Aishwarya and Ano used the genetic algorithm and Extreme Learning Machine (ELM) for diabetes disease diagnosis [30]. Fayssal et al [31] have designed a diagnosis system for diabetes using fuzzy classifier and modified Artificial Bee Colony algorithm. Polat et al. [32] also developed a cascade learning system to diagnose diabetes. Chang and Lilly [33] developed an evolutionary approach to derive a compact fuzzy classification system.

Kahramanli and Allahverdi [34] designed a hybrid neural network system for the classification of the diabetes database. For instance, Campos-Delgado et al [35] developed a fuzzy-based controller that incorporates expert knowledge to regulate the blood glucose level. Magni and Bellazzi [36] devised a stochastic model to extract variability from a self-monitoring blood sugar level time series. Lavanya and Pravin Kumar [38] developed an expert system based on PSO and Fuzzy Technique for Medical Diagnosis.

3 Proposed Approach

3.1 Attributes used for defining chromosomes.

For the attributes used inside our generated chromosomes, we first start to study all the medical tests and parameters related to the diabetes disease. Later we filtered these parameters depending on how each parameter interacts as a complication with high levels of glucose in the patient's blood. Also the importance of its high values which indicates that the medical treatment being used is sufficient or not. Moreover, we classified the attributes into three main categories; the first one is the Basic information of a patient (Age, Gender, BMI, Current weight, Height, Diabetes Duration, Triceps skinfold thickness). The second category is the patient's previous status and test (Previous weight, Previous Treatment, Previous HBA1c, Previous FBS), The third category is the Lab results and physical examination (Current HBA1c, Current FBS, Blood Pressure, Cholesterol, Triglyceride, Low-Density lipoprotein, High-Density lipoprotein, and 2-hour oral glucose tolerance test). Later for type 2 diabetes, we extract the most demanding attributes shown in Table 1 below:

Table 1. Selected attributes used in the chromosomes construction.

NO	Lab Test
1	Body mass index (weight in kg/(height in m) ²)
2	HBA1C
3	Fasting blood sugar test (FBS)
4	Diastolic blood pressure (mm Hg)
5	Triglyceride
6	Low-density lipoprotein
7	High-density lipoprotein
8	Previous Treatment

The main reason for choosing only 8 attributes is to reduce the overall computational time of the genetic algorithm. On the other hand, the selected 8 attributes cover all the previously mentioned ones.

3.2 Proposed algorithm.

There were a few types of research about diabetes using genetic algorithm while diabetes diagnostic has many types of research and developed systems using other expert systems techniques.

The main focus in this research is not to diagnostic the existence of the disease, as the patient already infected with diabetes, but to focus on helping the doctor to make the decision wither the supplement drug is working well and no need to change for extra doses, or if there is a need to change the dose even the medicine type. Moreover, patients with the same diagnosis and almost the same lab tests still need deferent treatment types, this can be explained by the fact that each body act in a deferent way with the supplied medicines.

3.3 Generate the initial random population (chromosomes).

In this research, the chromosomes considered to be an object which represents a selection tree that checks each of the main eight factors with a special sequence and unique test values. The initial set of chromosomes are randomly generated and tested using the fitness function starts with 200 chromosomes as in real population. It is worth to mention here, that each chromosome may use the same attribute many times with deferent values; also it may contain up to 50 nodes in the selection tree. This way of implementation allows us to achieve a very high accurate test sequence since we do not limit the mutation and crossover phase. Figure 1 and 2 below show a sample generated chromosome:

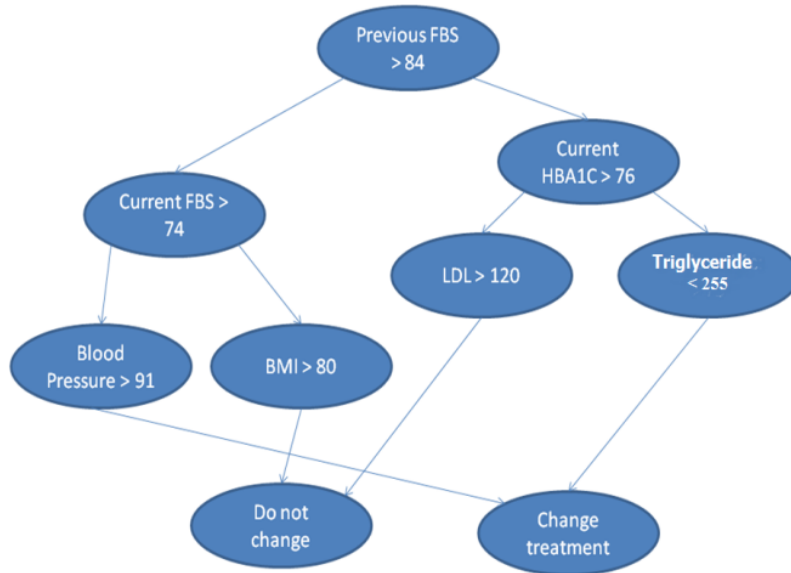


Fig. 1. A selection of binary tree sample chromosome.

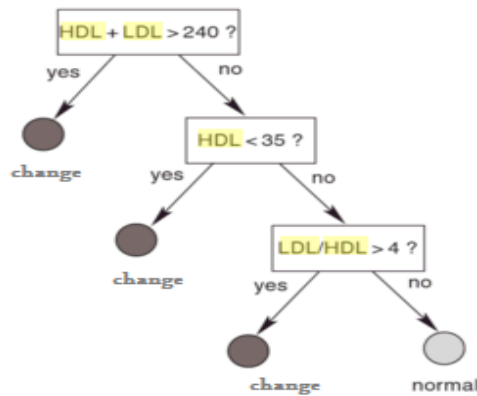


Fig. 2. The subtree of the chromosome implementation.

3.4 Roles used in chromosome generation

For the generation of the chromosomes and the mutation, we had to constrain the ranges of each parameter from the minimum value to the highest value. We depend on the standard average of each parameter from labs forms. Table 2 shows the parameter with their ranges.

Table 2. Attributes allowed ranges to be selected inside the chromosome

NO	Lab Test	Minimum value	Maximum value	Range Acceptable
1	BMI	16-18.5	30-35	18.5-30
2	HBA1C	3.5-5.6	6.5- 7.9	5.6-6.5
3	FBS	50-100	126-800	100-126
4	Diastolic blood pressure	50-80	89-120	80-89
5	Triglyceride	155-199	500-650	199-500
6	Low density lipoprotein	90-110	190-220	110-190
7	High density lipoprotein	20-40	60-70	40-60

The algorithm starts to select one of the attributes randomly, later its select one of the condition tests like (greater than, less than, and equal to) number selected from the ranges in Table 2. Moreover, the condition may contain some equation like LDL + HDL is greater than. After that, another randomly selected attribute is chosen and assigned a testing condition. It is worth to mention that the same attribute may be selected but with constraint not to have the same parent attribute. This mechanism initially generates 11 nodes inside each chromosome in the first generation. Later it is allowed to the chromosome in cross over to change subtree with a lower level with another one with a higher level. Als, we have constraints of not exceeding 50 attribute test.

3.5 Calculation of fitness values

In this phase, each chromosome is being evaluated using the defined fitness function. After that, upon the evaluation values, the proposed algorithm decides wither to proceed to the next phase (generating the new chromosomes generation) or stop if one of the evaluated chromosomes achieved the desired threshold which in our system considered being correct treatment suggestion accuracy of 95.4%.

For that we first need to define our fitness function to be 50 patients real test results, which is evaluated using each generated chromosome, later the suggestion results decided by each chromosome when compared to the real doctor decision, and then calculate the accuracy for each chromosome results. The accuracy results then evaluated, and if any chromosome has reached the desired accuracy threshold, the process stops and the selected chromosome considered being the best-generated chromosome. On another hand, if the desired accuracy is not achieved, the fitness function finds the average accuracy from the calculated chromosomes, and consider it as the threshold for chromosomes to be either selected as one of the best chosen chromosomes to be a parent for the next generation or just consider it as sub-parent in the next phase (here we implement the crossover and mutation between the parents and the sub-parents). Also, the fitness function contains the accuracy threshold which is considered to be the value where the whole operation stops if any chromosome reaches it. But still, have a restriction on the allowed number of created generations (200 generations only). To prevent the system from executing all forever if the desired accuracy does not obtain. Later the output chromosome considered being the best sequence testing algorithm for diabetes medicine efficiency.

3.6 Mutation and crossover phase.

To generate new chromosomes from the best-preselected ones in the fitness function phase, in order to achieve better results in the accuracy test. We concatenate and replaces parts from the parents' chromosomes with part of sup parent chromosomes (sup parent are considered to be chromosomes below the average of the fitness function) in a random matter, although we still have some constraints like the crossover mechanism which can be only done on the second root or higher, also most crossover weight is made over main nodes and the rest is made on the leaves. Also, the same attribute cannot be repeated more than 5 times (achieved experimentally). On the other hand, the mutation is held on the attributes level like changing the BMI test to FSB test and can be done also to the values of each attribute like changing the value of diastolic blood pressure from 60 to 65. It is worth to mention here that the mutation also can be done on the condition test of each attribute, as for an example, a node of $LDL > 78$ may mutate to $LDL \geq 78$. Also, the mutation can be implemented combined with crossover or separated. This found to be more effective from changing only the sequence order.

The number of generated chromosomes is fixed to 200 chromosomes; this can be implemented by allowing the same chromosome to be used with many other ones in the crossover process. By adopting this rule, we enhance the chance of having good results in the early generated generations as our results show that we could achieve the desired accuracy on the 167th generation. Figure 3 shows the three phases of the proposed algorithm and its work principle. The figure illustrates the procedure of each phase and when the process terminates, by meeting the desired accuracy or by the finishing constraints of 200 generations.

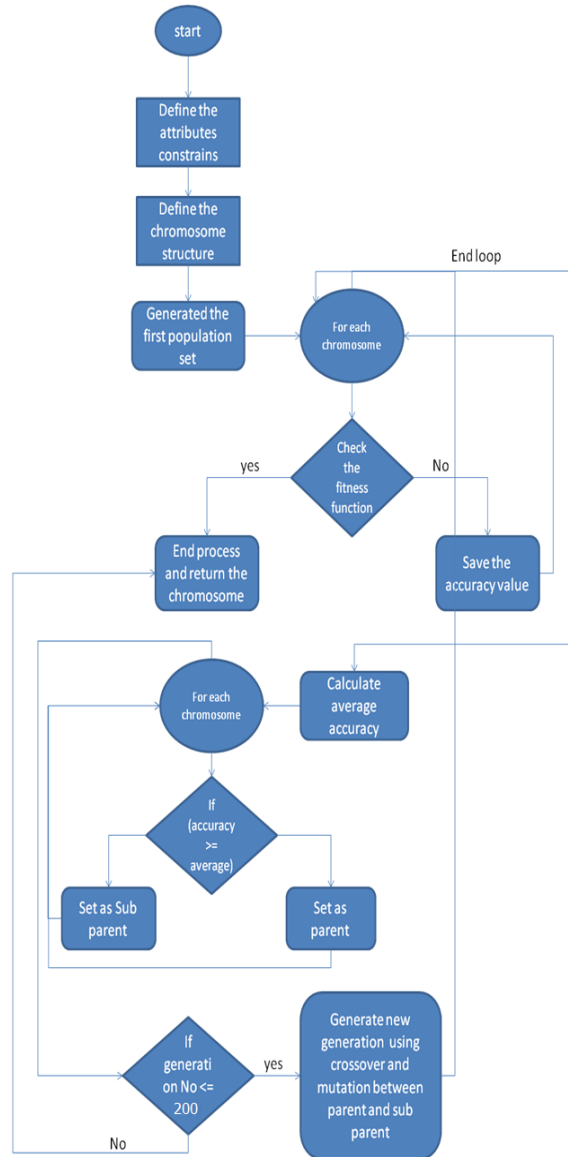


Fig. 3. Flow diagram of the proposed algorithm

4 Experimental Settings and Dataset

4.1 Dataset used

In our research, we adopted a dataset of 50 patients taken from many labs in Jordan, with the subjective opinion of the responsible doctor, and later test the output

testing criteria with other patients’ cases dataset to prove our generated algorithm efficiency. Table 3 shows a 25 patients lab tests used in our system, the table contains the initial treatment used and the new treatment decided by the doctors responsible for each patient.

Table 3. Sample of the patient's lab tests used in our thesis.

ID	BMI	HBA1C	FBS	Blood pressure	Triglyceride	LDL Mg/dl	HDL Mg/dl	Previous treatment	Decision
1	27.7	6.7	195	145	110	60	114		Insulin
2	28.1	6.9	181	140	105	93.2	48	insulin	Insulin
3	27.4	7.1	179	150	130	103.2	70	insulin	Insulin
4	26.9	6.2	142	130	110	93.6	51	2 oral	Insulin
5	28.7	6.0	149	140	170	131.4	50	2 oral	Insulin
6	28.3	6.5	173	130	100	125.4	52	insulin	Insulin
7	27.9	6.2	196	110	155	64.8	61	insulin	Insulin
8	39.4	7.1	178	96	198	99.6	57	insulin	Insulin
9	23.2	6.4	130	72	156	185	56	insulin	Insulin
10	22.2	6.8	165	50	192	119.4	42	insulin	Insulin
11	34.1	7.4	90	84	160	93.4	43	insulin	Insulin
12	36.0	7.7	125	74	139	144.8	32	insulin	Insulin
13	27.6	5.9	129	62	160	115.4	72	2 oral	Insulin
14	32.9	5.4	158	66	255	184.2	39	2 oral	2 oral
15	40.2	6.1	189	84	180	110.6	56	insulin	Insulin
16	22.7	5.9	129	90	184	81.2	66	2 oral	Insulin
17	24.2	6.7	88	94	186	112.8	86	insulin	Insulin
18	37.7	7.2	196	82	202	162.2	43	insulin	Insulin
19	46.8	7.4	158	92	254	116.4	47	insulin	Insulin
20	25.4	6.3	103	85	207	126.8	64	insulin	Insulin
21	32.5	5.5	129	87	162	144	57	2 oral	2 oral
22	42.7	5.2	146	88	187	146.6	65	2 oral	2 oral
23	27.9	7.6	147	58	156	128.2	53	insulin	Insulin
24	33.7	6.2	99	74	167	125.8	76	2 oral	Insulin
25	31.2	6.1	81	97	187	189	68	2 oral	2 oral

4.2 Experimental results.

For the experimental results, we manage to gather 80 samples to be the dataset used for both the genetic algorithm evolution and for the evaluation and efficiency tests. First of all, the genetic algorithm fitness function was adopted to be the decision of the doctor on 50 patient cases. Later the algorithm was established by initial chromosomes number of 200, and the experimental results show that after the 167 generations the genetic algorithm could achieve a fitness function of 95.4% accuracy. The results were achieved by applying each of the 50 patient inputs into the new chromosome and compare the results with the doctor decision and calculate the results then calculate the overall average between the system and the doctor. Table 4 shows the lab tests results for one patient in Jordan and the doctor decision.

It is worth to mention here that we implement two types of our model; the first one enabled only the mutation and crossover between the sub-parent chromosomes (chromosomes under the average accuracy). The second one enables only the mutation and crossover between sub-parents and parents. The first implementation achieved only 54.3% after 200 generations while the second implementation achieved 95.4% which made us adopt the mutation and crossover between chromosomes from the over average accuracy and lower ones.

Table 4. Lab test result for diabetic type 2 patient.

Lab result	values
BMI	28.1
HBA1C	6.8
FBS	189
Blood pressure	96
Triglyceride	109
Low density lipoprotein	65
High density lipoprotein	185

Doctor recommendation was to move the patient from two tablets per day to on insulin injection which was the same result of our generated system. For the system evaluation, we later run a test for the next 30 patient cases to get a real evaluation of data that the system never has seen before. The results came as shown in the table below with accuracy of 96.6%, which approves the system efficiency. Table 5 below shows the results for the system test on 30 patients in Jordan and the decision taken also the overall accuracy of 96.6%.

Table 5. The results for the system test on 30 patients in Jordan.

Patient ID	Doctor decision	System decision	Evaluation
1	1 oral medication	1 oral medication	1
2	1 oral medication	1 oral medication	1
3	2 oral medication	2 oral medication	1
4	2 oral medication	2 oral medication	1
5	1 oral medication	1 oral medication	1
6	2 oral medication	2 oral medication	1
7	2 oral medication	2 oral medication	1
8	Insulin	Insulin	1
9	Insulin	Insulin	1
10	1 oral medication	1 oral medication	1
11	Insulin	Insulin	1
12	Insulin	Insulin	1
13	Insulin	Insulin	1
14	2 oral medication	2 oral medication	1
15	Insulin	Insulin	1
16	Insulin	Insulin	1
17	Insulin	Insulin	1

18	Insulin	2 oral medication	0
19	2 oral medication	2 oral medication	1
20	1 oral medication	1 oral medication	1
21	Insulin	Insulin	1
22	Insulin	Insulin	1
23	Insulin	Insulin	1
24	Insulin	Insulin	1
25	2 oral medication	2 oral medication	1
26	2 oral medication	2 oral medication	1
27	Insulin	Insulin	1
28	Insulin	Insulin	1
29	1 oral medication	1 oral medication	1
30	Insulin	Insulin	1
Match count			29

For the system evaluation, we later run a test for the next 10 patient cases to get a real evaluation of data that the system never has seen before. The results came as shown in the table below with an accuracy of 100%, which approves the system efficiency. Table 6 below shows the results for the system test on 10 patients in Jordan and the decision taken also the overall accuracy of 100%.

Table 6. The results for the system test on 10 patients in Jordan.

Patient ID	Doctor decision	System decision	Evaluation
1	Insulin	Insulin	1
2	Insulin	Insulin	1
3	1 oral medication	1 oral medication	1
4	2 oral medication	2 oral medication	1
5	1 oral medication	1 oral medication	1
6	Insulin	Insulin	1
7	Insulin	Insulin	1
8	Insulin	Insulin	1
9	Insulin	Insulin	1
10	2 oral medication	oral medication	1

5 Conclusion

This research presents a system to help doctors to find a treatment for diabetes disease type 2, in this system we use real lab test to specify the suitable treatment to the patient. New diabetes type 2 treatment assessment system based on genetic algorithm was implemented, we managed to create chromosomes population each of which consists of the binary decision tree, as this implementation considered being the best scenario of our problem. The system proves its efficiency by applying it on 50 patients and the results show an accuracy percentage of 95.4%.

6 References

- [1] UCSF Medical Center https://www.ucsfhealth.org/conditions/diabetes_mellitus/.
- [2] American Diabetes Association <http://www.diabetes.org/diabetes-statistics>.
- [3] Al Adwan F., AlShraideh M., AlSalem M. R. (2015), A Genetic Algorithm Approach for Breaking of Simplified Data Encryption Standard. *International Journal of Security and Its Applications*, vol. (9), issue (9), pp. 295-304. <https://doi.org/10.14257/ijisia.2015.9.9.26>
- [4] Qbea'h M., Alshraideh M., Sabri K. E. (2016), Detecting and preventing SQL injection attacks: a formal approach, 2016 Cybersecurity and Cyberforensics Conference (CCC), IEEE, pp.123-129. <https://doi.org/10.1109/ccc.2016.26>
- [5] Manaseer S., Manaseer W., Alshraideh M., Abuhashish N., Adwan O. (2016), Automatic Test Data Generation for Java Card Applications Using Genetic Algorithm, *Journal of Software Engineering and Applications*, 8 (12), pp. 603-616. <https://doi.org/10.4236/jsea.2015.812057>
- [6] Alshraideh M., Qatawneh M., AlMobaiden W., Sleit A. (2010). Program-operators to improve test data generation search, *WSEAS Transactions on Computers* 9 (8), pp. 799-811.
- [7] Allawi H., Al Manaseer W., Al Shraideh M. (2018). A greedy particle swarm optimization (GPSO) algorithm for testing real-world smart card applications, *International Journal on Software Tools for Technology Transfer*, pp.1-12, DOI <https://doi.org/10.1007/s10009-018-00506-y>.
- [8] Aslam M.W., Nandi A. K. (2010). Detection of Diabetes Using Genetic Programming, 18th European Signal Processing Conference(EUSIPCO-2010), Aalborg, Denmark.
- [9] Pradhan M. A., Acharya P., Gawade, R., Pateria A. (2011). Design of Classifier for Detection of Diabetes using Genetic Programming, *International Conference on Computer Science and Information Technology (ICCSIT'2011)* Pattaya, pp.125-130.
- [10] Mendes R. R. F., Voznika B., Freitas A., Nievola J. (2001). Discovering fuzzy classification rules with genetic programming and co-evolution, in *Lecture Notes in Artificial Intelligence*, vol. 2168, Proc. 5th Eur. Conf. PKDD, pp. 314–3 25. https://doi.org/10.1007/3-540-44794-6_26
- [11] Aishwarya S., Anto S. (2014). A Medical Expert System based on Genetic Algorithm and Extreme Learning Machine for Diabetes Disease Diagnosis, *International Journal of Science, Engineering and Technology Research (IJSETR)*, Volume 3, Issue 5, May 2014.
- [12] Sapna S., Tamilarasi A., Pravin Kumar M. (2012). Implementation of Genetic Algorithm in Predicting Diabetes, *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 1, No 3.
- [13] Yang S., Zhang M. (2018). Application of Brain Neural Network in Personalized English Education System, *International Journal of Emerging Technologies in Learning (iJET)* 13(10). <https://doi.org/10.3991/ijet.v13i10.9488>
- [14] Thinnukool O., Khuwuthyakorn P., Wientong P., Suksati B., Waisayanand N. (2019). Type 2 Diabetes Mobile Application for Supporting for Clinical Treatment: Case Development Report, *International Journal of Online and Biomedical Engineering (iJOE)*, 15(2). <https://doi.org/10.3991/ijoe.v15i02.9769>
- [15] Pradhan, Abdul Rahman M. A., Pushkar Acharya, Ravindra Gawade, Ashish Pateria (2011). Design of Classifier for Detection of Diabetes using Genetic Programming I, *International Conference on Computer Science and Information Technology (ICCSIT'2011)*, pp.125-130.
- [16] Bojarczuk C. C., Lopes H. S., Freitas A. (2000). Genetic Programming for knowledge discovery in chest pain diagnosis, I *IEEE Eng.Med. Mag.*, vol. 19, no. 4, pp. 38–44. <https://doi.org/10.1109/51.853480>

- [17] Asha Gowda Karegowda, Manjunath A.S., Jayaram M.A. (2011). Application of Genetic Algorithm Optimized Neural Network Connection Weights for Medical Diagnosis of Pima Indians Diabetes I, *International Journal on Soft Computing (IJSC)*, Vol.2, No.2, pp. 15-23. <https://doi.org/10.5121/ijsc.2011.2202>
- [18] Polat K., Gunes S., Aslan A. (2008). A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine", *Expert systems with applications*, vol.34(1), pp. 214-221. <https://doi.org/10.1016/j.eswa.2006.09.012>
- [19] Durga Prasad Muni, Nikhil R. Pal, Jyotirmoy Das (2004). A Novel Approach to Design Classifiers Using Genetic Programming, *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, VOL. 8, NO. 2, pp. 183-196. <https://doi.org/10.1109/tevc.2004.825567>
- [20] Hasan T., Nijat Y., Feyzullah T. (2009). A comparative study on diabetes disease using neural networks", *Expert system with applications*, vol.36, pp.8610-8615.
- [21] Kemal Polata, Salih Güne,sa Sülayman Tosunb. (2007). Diagnosis of heart disease using artificial immune recognition system and fuzzy weighted pre-processing, *ELSEVIER, PATTERN RECOGNITION*. <https://doi.org/10.1016/j.patcog.2010.12.007>
- [22] Ephzibah E.P. (2011). Cost-Effective Approach on Feature Selection Using Genetic Algorithms and Fuzzy Logic for Diabetes Diagnosis. I, *International Journal on Soft Computing (IJSC)*, Vol.2, No.1, pp. 1-10. <https://doi.org/10.5121/ijsc.2011.2101>
- [23] Filipe de L. Arcanjo, Gisele L. Pappa, Paulo V. Bicalho, Wagner Meira Jr., Altigran S. da Silva. (2011). Semi-supervised Genetic Programming for Classification, *GECCO'11*, Dublin, Ireland. <https://doi.org/10.1145/2001576.2001746>
- [24] Kishore J. K., Patnaik L. M., Mani V., Agrawal V. K. (2000). Application of genetic programming for multicategory pattern classification, *IEEE Trans. Evol. Comput.*, vol. 4, pp. 242–258. <https://doi.org/10.1109/4235.873235>
- [25] Lim T.-S., Loh W.-Y., Shih Y.-S. (2000). A comparison of prediction accuracy, complexity and training time of thirty-three old and new classification algorithms, *Mach. Learning J.*, vol. 40, pp. 203–228.
- [26] Richmond Smart W. (2005). Genetic Programming for Multiclass Object Classification, A thesis submitted to the Victoria University of Wellington.
- [27] SonuKumari S., Singh A. (2012). A Data Mining Approach for the Diagnosis of Diabetes Mellitus, 978-1-4673-4603-0112/\$31.00 ©2012 IEEE.
- [28] Pradhan M. A., Bamnote G.R., Tribhuvan V. (2012). A Genetic Programming Approach for Detection of Diabetes *IJCER*, Vol.2, Issue 6.
- [29] Aishwarya S., Anto S. (2014). A Medical Expert System based on Genetic Algorithm and Extreme Learning Machine for Diabetes Disease Diagnosis, *International Journal of Science, Engineering and Technology Research (IJSETR)*, Volume 3, Issue 5.
- [30] Fayssal BChikh., M.A. (2013). Design of fuzzy classifier for diabetes disease using Modified Artificial Bee Colony algorithm, *Computer methods and programs in biomedicine*, No.1, pp.92-103. <https://doi.org/10.1016/j.cmpb.2013.07.009>
- [31] Polat K., Gunes S., Arslan A. (2008). A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine, *Expert Syst. Appl.*, vol. 34, no. 1, pp. 482–487. <https://doi.org/10.1016/j.eswa.2006.09.012>
- [32] Chang X., Lilly J. H. (2004). Evolutionary design of a fuzzy classifier from data, *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 4, pp. 1894–1906. <https://doi.org/10.1109/tsmcb.2004.831160>

- [33] Kahramanli H., Allahverdi N. (2008). Design of a hybrid system for diabetes and heart diseases, *Expert Syst. Appl.*, vol. 35, no. 1, pp. 82–89. <https://doi.org/10.1016/j.eswa.2007.06.004>
- [34] Campos-Delgado D. U., Hernandez-Ordóñez M., Femat R., Gordillo-Moscó A. (2006). Fuzzy-based controller for glucose regulation in type-1 diabetic patients by the subcutaneous route, *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2201–2210. <https://doi.org/10.1109/tbme.2006.879461>
- [35] Magni P., Bellazzi R. (2006). A stochastic model to assess the variability of blood glucose time series in diabetic patients self-monitoring,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 977–985. <https://doi.org/10.1109/tbme.2006.873388>
- [36] MLavanya N., Praveen Kumar M. (2012). Expert system based on PSO and Fuzzy Technique for Medical Diagnosis, *IJART*, Vol.2, Issue 4, pp. 66-71.

7 Author

Majdoleen Al Switi works for The University of Jordan, in the Department of computer science at Jordan. majdoleen_majed2012@yahoo.com

Bahaaldeem Alshraideh works in the Medicine department at The University of Jordan, Jordan. b2alshraideh@yahoo.com

Abedalrhman Alshraideh is with The University of Jordan at the Medicine department, in Jordan. Abed_1997@yahoo.co.uk

Abudalla Massad works at the Medicine department in The University of Jordan, Jordan. a.massad97@yahoo.com

Mohammad Alshraideh is with The University of Jordan in the Department of computer science at Jordan. mshridah@ju.edu.jo

Article submitted 2019-04-27. Resubmitted 2019-05-26. Final acceptance 2019-05-26. Final version published as submitted by the authors.