International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

Privacy Preserving Data Mining using Jaya based Genetic Algorithm

Srutipragyan Swain[1], Prasant Kumar Pattnaik[2], Banchhanidhi Dash[3]
[1]School of Computer Engineering
KIIT,University
Bhubaneswar,India
sruti56@gmail.com
[2]School of Computer Engineering
KIIT,University
Bhubaneswar,India
patnaikprasantfcs@kiit.ac.in
[3]School of Computer Engineering
KIIT,University
Bhubaneswar,India
banchhanidhi.dashfcs@kiit.ac.in

Abstract—Privacy protection has emerged as a key concern in the field of data mining because of the rise in the sharing of sensitive data across networks among organizations, governments, and other parties. For knowledge extraction from these huge set of data, association rule mining is used to analyze the patterns of data. For a variety of optimization issues encountered in the real world, evolutionary algorithms (EAs) offer efficient solutions. The available EA solutions in the privacy-preserving area are limited to specific issues like cost function evaluation. Here, a JAYA based Genetic algorithm has been proposed for privacy preservation.“JAYA” means victory a word from Sanskrit origin**..**This algorithm doesn't need any parameters that are specific to it and it moves towards best solution avoiding the worst. Hence the name Jaya. Jaya based genetic algorithm is applied to original dataset of chromosomes. Privacy preservation is achieved by comparing the support of original dataset of chromosome with the simulation output.

Keywords- PPDM(Privacy Preserving Data Mining),Association Rule Hiding,Jaya Algorithm; GA (Genetic Algorithm).

I. INTRODUCTION

With the increase in number of computing devices, the data is collected with a dramatic pace. However huge amount of data is of no use unless necessary information or knowledge can be extracted by mining the data. Since the availability of information is less in comparison to the availability of data, knowledge extraction from database has emerged one of the important areas of research in the computer science domain. For knowledgeextraction, association rule mining is used to analyze the patterns of data. However; data privacy has to be considered due the presence of sensitive data.Again, redundant data diminishes the performance of this system. All these lead to Privacy Preserving Data Mining (PPDM) [1]. Since, there is increase in sharing of sensitive data among organizations, business sectors, and healthcare industry over the internet it ismore vulnerable to public exposure. But these data have to be analyzed for knowledge extraction.

II. LITERATURE REVIEW

The balance between data mining and exposure is maintained by PPDM.Privacy can be achieved at two different levels. It may be input privacy where the data enters the system or output privacy i.e., during the pattern representation. l-diversity [2, 3],K-anonymity [4, 5] and *t*-closeness [6] are the techniques by which Input privacy can be achieved.In real life data generated may not be crisp always. The generated data may contain uncertainties. PPDM techniques aim at securing the data by removing the sensitive data or by applying masking technique over original data. Sensitive attributes can be identified dynamically based on the threshold limit of their respective characteristics [7].To maintain the privacy the data owner use swapping method to modify the value of sensitive attributes identified above, in such a way that their original properties remain unaltered.

The Privacy-Preserving Record Linkage (PPRL) [8], which protects privacy while enabling the linking of databases to organizations. In order to analyze them in 15 dimensions, a taxonomy based on PPRL approaches is suggested. For privacy preservation Cryptography techniques [9] such as homomorphic encryption, multiparty computation and secret sharing methods are applied on business and medical data. Simulation results show that secret sharing method out performs homomorphic encryption. Fuzzy based technique like triangular membership function is used to change the original dataset for privacy preserving data mining [10]. The methods

Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

like suppression and perturbation [11] are used over quasi-identifier for protection of privacy. This method overcomes the information loss during the privacy preservation process. The solutions for privacy preservation with respect to cloud services can be classified on the basis of advanced cryptographic components [12].The solution provides access which is anonymous, ability to unlink and maintaining confidentiality of data those are transmitted.Further, Confidence and support are the two important factors which are to be taken care of while mining association rule, where support is the transaction's percentage involving AUB for any association rule A=>B and confidence for same association rule is the ratio of transactions number by A [13]. In order to prevent modification of a given database depending upon support and confidence of sensitive rules can be reduced by introducing new parameters namely M confidence, hiding counter and M support which are based on the original concept of confidence and support [14].

Identification of sensitive item can also be done by hiding vulnerable association rules by identifying the most used item sets and accordingly creating the association rules. The concept of representative association rules is used to find out sensitive attributes. In case of distributed k-means clustering privacy-preservation can be done using multiparty k-means clustering in which k-means clustering is evenly applied on each data site which is vertically partitioned [15].

III. AN EMPERICAL STUDY ON DIABETES PATIENT DATASET

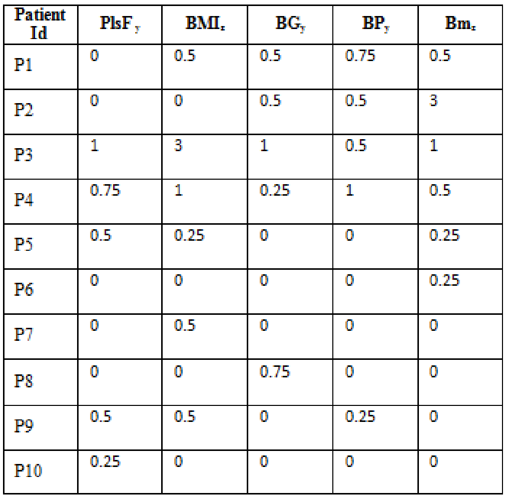
A few parameters of diabetes patient and some potential range of symptom values have been considered. Association rule mining is the concept of finding association rules from huge datasets which helps in decision making. Further, Sensitive association rules can be hidden using fuzzy association rule hiding model.

Table.1 shows the sample information system of diabetes patient. The FAR (Fuzzified Association Rule) of the Sample information system is computed by using the concept of FPR (Fuzzy proximity relation). The sensitive rules are mined by computing the confidence and support of association rules. The items in the critical regions are input to the chromosome dataset. Here we have considered the attribute PlasmaF and its FIS (Fuzzy Information System) have been computed in Table2.

TABLE I. SAMPLE INFORMATION SYSTEM.

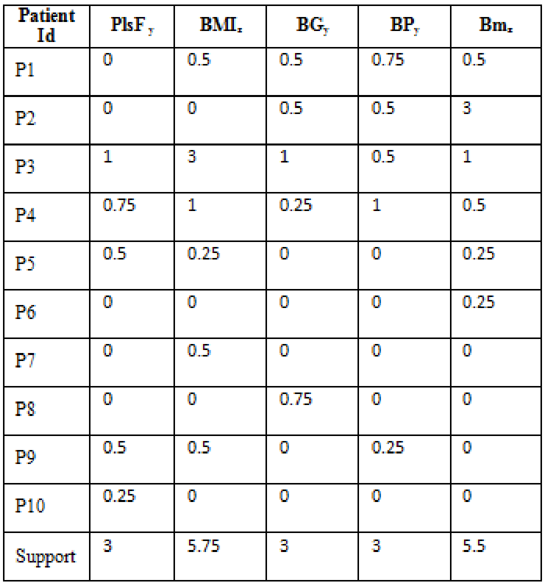
Patient Id	PlasmaF	BMI	Blood Glucose min	Blood Pressure	Blood Glucose max
P1	3.2	73.1	41.2	196.5	263.4
P2	3.3	65.5	66.3	221.8	196.3
P3	4.1	83.6	19.9	156.3	203.5
P4	4.5	80.9	24.5	149.5	187.2
P5	3.7	72.3	16.6	232.8	179.4
P6	5.8	67.4	17.9	229.7	181.6
P7	6.7	79.5	33.2	198.4	153.3
P8	3.6	63.1	21.5	116.8	166.8
P9	4.8	85.2	27.8	212.5	134.5
P10	5.3	61.7	16.8	157.8	157.9

TABLE II. FUZZY VALUES OF SENSITIVE ATTRIBUTES



Patient Id	PlsF $_y$	BMI$_z$	BG$_y$	BP$_y$	Bm$_z$
P1	0	0.5	0.5	0.75	0.5
P2	0	0	0.5	0.5	3
P3	1	3	1	0.5	1
P4	0.75	1	0.25	1	0.5
P5	0.5	0.25	0	0	0.25
P6	0	0	0	0	0.25
P7	0	0.5	0	0	0
P8	0	0	0.75	0	0
P9	0.5	0.5	0	0.25	0
P10	0.25	0	0	0	0

TABLE III. SUPPORT VALUES OF CHROMOSOMES



Patient Id	PlsF $_y$	BMI$_z$	BG$_y$	BP$_y$	Bm$_z$
P1	0	0.5	0.5	0.75	0.5
P2	0	0	0.5	0.5	3
P3	1	3	1	0.5	1
P4	0.75	1	0.25	1	0.5
P5	0.5	0.25	0	0	0.25
P6	0	0	0	0	0.25
P7	0	0.5	0	0	0
P8	0	0	0.75	0	0
P9	0.5	0.5	0	0.25	0
P10	0.25	0	0	0	0
Support	3	5.75	3	3	5.5

IV. PROPOSED MODEL

In the proposed hybrid model, to hide sensitive association rule the combination of Jaya and genetic algorithm has been implemented.

A. *Jaya Algorithm*

- One of the simple yet powerful optimization algorithms known as Jaya algorithm which has wide application in solving optimization problems. The two phases on which the Teaching Learning Based Optimization algorithm operates are the teacher phase and the learner phase. The learner phase is absent here. It requires only the teacher phase.
- The algorithm avoids the worst solution and moves towards best solution. The nature of striving towards success to achieve victory explains the name *Jaya* from *Sanskrit*. It can be defined as follows.
- Step-1 The number of design variables, parameters and termination criteria are initialized.
- Step-2 Until the termination condition is not met; Steps 3 through 5 are repeated.
- Step-3 The best solution is computed. The worst solution is also computed in this step.
- Step-4 The modified solution is calculated as follows

$K'_{b,c,a}= K_{b,c,a}+d_{1,b,a}(K_{b,best,a}- K_{b,c,a}|)-d_{2,b,a}(K_{b,worst,a}-| K_{b,c,a}|)$

- $K_{b,best,a}$is the best candidate's value for the variable b.
- $K_{b,worst,a}$is the worst candidate's value for the variable b.
- $K'_{b,c,a}$ is the updated value of $K_{b,c,a}$.
- Random numbers $d_{1,b,a}$ and $d_{2,b,a}$ are used for the bth variable in the ath iteration.
- Step-5 If $K'_{b,c,a}>K_{b,c,a}$then the old solution is updated.
- Else
- The previous solution is kept.
- Jaya uses two standard regulating parameters, m and n in contrast to other heuristic algorithms which is population based. Here, n is the number of generations and m is the number of candidate solutions. Let f(x) is the objective function to be minimized or maximized [9, 10]. At any iteration a, assume that there are m number of design variables i.e. b = 1, 2,..., m, and n number of candidate solutions i.e. population size, c = 1,2,...,n.
- Let $K_{b,c,a}$is the value of the *b*th variable for the *c*th candidate during the *a*th iteration, then this value is modified as per the following Equation.

$$K'_{b,c,a}= K_{b,c,a}+d_{1,b,a}(K_{b,best,a}- |K_{b,c,a}|)-c_{2,b,a}(K_{b,worst,a}-| K_{b,c,a}|) \quad (1)$$

Equation (1) indicates that $d_{1,b,a}(K_{b,best,a}- |K_{b,c,a}|)$moves towards best solution and the term $-d_{2,b,a}(K_{b,worst,a}-| K_{b,c,a}|)$ avoids worst solution. The term $K'_{b,c,a}$ is accepted if it gives the value which is better than the previous.

B. *Genetic Algorithm*

Inspired by the natural selection process, genetic algorithms are evolutionary algorithms which aim to solve a defined problem by a group of individuals called chromosomes. The genetic algorithm starts with a population made up of individuals chosen at random.

A new population is created in each iteration to replace every member of the previous population. In the process of elitism best individuals are kept. The next population is created using the best chromosome from the previous one. The crossover and mutation are the two processes used to generate new population called offspring from the initial population [17].

C. *Fitness Function Evaluation of chromosome dataset using Jaya Algorithm*

The fitness of a chromosome is evaluated as per the fitness equation in JAYA algorithm using Python. For fitness function evaluation sphere function is considered. The values of chromosome found in critical region are considered as the design variables x_i *for i=1.2,...,n* and $f\ (x_i)$ is the objective function. The sphere function can be demonstrated as follows:

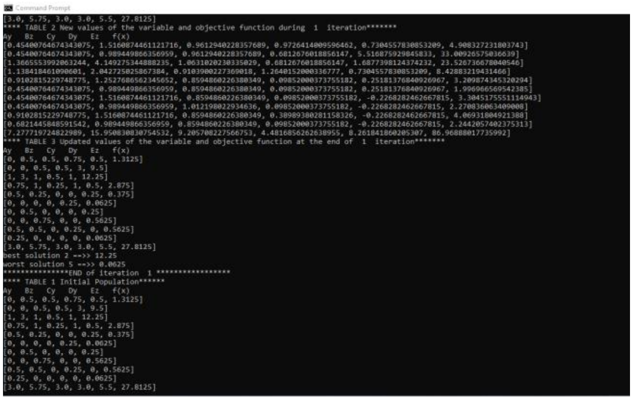
$$\max f(xi) = \sum_{i=1}^{n} x_i^2 \quad (2)$$

Table.Vshows the updated value of chromosomes after second iteration of Jaya algorithm. The parents are selected according to the fitness value of the chromosomes.

***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

TABLE IV. INITIAL POPULATION INPUT TO JAYA ALGORITHM.

Patient Id	$PlsF_y$	BMI_z	BG_y	BP_y	Bm_z
P1	0	0.5	0.5	0.75	0.5
P2	0	0	0.5	0.5	3
P3	1	3	1	0.5	1
P4	0.75	1	0.25	1	0.5
P5	0.5	0.25	0	0	0.25
P6	0	0	0	0	0.25
P7	0	0.5	0	0	0
P8	0	0	0.75	0	0
P9	0.5	0.5	0	0.25	0
P10	0.25	0	0	0	0



```
[3.0, 5.75, 3.0, 3.0, 5.5, 27.8125]
**** TABLE 2 New values of the variable and objective function during  1  iteration*******
Ay   Bz   Cy   Dy   Ez   f(x)
[0.45400764674343075, 1.5160874461121716, 0.9612940228357689, 0.9726414009596462, 0.7304557830853209, 4.908327231803743]
[0.45400764674343075, 0.989449866356959, 0.9612940228357689, 0.6812676018856147, 5.516875929845833, 33.00926575036639]
[1.3665553992063244, 4.149275344888235, 1.0631020230335029, 0.6812676018856147, 1.6877398124374232, 23.526736678040546]
[1.138418461090601, 2.042725025867384, 0.9103900227369018, 1.2640152000336777, 0.7304557830853209, 8.42883219431466]
[0.9102815229748775, 1.2527686562345652, 0.8594860226380349, 0.09852000373755182, 0.25181376840926967, 3.209874345320294]
[0.45400764674343075, 0.989449866356959, 0.8594860226380349, 0.09852000373755182, 0.25181376840926967, 1.996966569542385]
[0.45400764674343075, 1.5160874461121716, 0.8594860226380349, 0.09852000373755182, -0.2268282462667815, 3.3045175551114943]
[0.45400764674343075, 0.989449866356959, 1.012198022934636, 0.09852000373755182, -0.2268282462667815, 2.270836063409008]
[0.9102815229748775, 1.5160874461121716, 0.8594860226380349, 0.38989380281158326, -0.2268282462667815, 4.06931804921388]
[0.6821445848591542, 0.989449866356959, 0.8594860226380349, 0.09852000373755182, -0.2268282462667815, 2.2442057402375313]
[7.277719724822989, 15.950830830754532, 9.205708227566753, 4.4816856262638955, 8.261841860205307, 86.96888017735992]
**** TABLE 3 Updated values of the variable and objective function at the end of  1  iteration*******
Ay   Bz   Cy   Dy   Ez   f(x)
[0, 0.5, 0.5, 0.75, 0.5, 1.3125]
[0, 0, 0.5, 0.5, 3, 9.5]
[1, 3, 1, 0.5, 1, 12.25]
[0.75, 1, 0.25, 1, 0.5, 2.875]
[0.5, 0.25, 0, 0, 0.25, 0.375]
[0, 0, 0, 0, 0.25, 0.0625]
[0, 0.5, 0, 0, 0, 0.25]
[0, 0, 0.75, 0, 0, 0.5625]
[0.5, 0.5, 0, 0.25, 0, 0.5625]
[0.25, 0, 0, 0, 0, 0.0625]
[3.0, 5.75, 3.0, 3.0, 5.5, 27.8125]
best solution 2 ==>> 12.25
worst solution 5 ==>> 0.0625
***************END of iteration  1 *****************
**** TABLE 1 Initial Population******
Ay   Bz   Cy   Dy   Ez   f(x)
[0, 0.5, 0.5, 0.75, 0.5, 1.3125]
[0, 0, 0.5, 0.5, 3, 9.5]
[1, 3, 1, 0.5, 1, 12.25]
[0.75, 1, 0.25, 1, 0.5, 2.875]
[0.5, 0.25, 0, 0, 0.25, 0.375]
[0, 0, 0, 0, 0.25, 0.0625]
[0, 0.5, 0, 0, 0, 0.25]
[0, 0, 0.75, 0, 0, 0.5625]
[0.5, 0.5, 0, 0.25, 0, 0.5625]
[0.25, 0, 0, 0, 0, 0.0625]
[3.0, 5.75, 3.0, 3.0, 5.5, 27.8125]
```

Fig. 1 Simulation Output

TABLE V. FINAL FITNESS FUNCTION TABLE AFTER SECOND ITERATION.

Patient Id	x_1	x_2	x_3	x_4	x_5	f(x)
P1	0	0.5	0.5	0.75	0.5	1.3125
P2	0	0	0.5	0.5	3	9.5
P3	1	3	1	0.5	1	**12.25**
P4	0.75	1	0.25	1	0.5	2.875
P5	0.5	0.25	0	0	0.25	0.375
P6	0	0	0	0	0.25	0.0625
P7	0	0.5	0	0	0	0.25
P8	0	0	0.75	0	0	0.5625
P9	0.5	0.5	0	0.25	0	0.5625
P10	0.25	0	0	0	0	0.0625

V. APPLY MUTATION AND CROSSOVER

Step1. Calculate S =∑ chromosome finesses.

Step2. Calculate r which is a random number within the interval (0, S)

Step3. Calculate s which is cumulative fitness of the chromosomes.

Select the chromosome for which s, cumulative fitness is greater than r.

```
Step4:  cross over implementation
    4.1for all chromosomes selected in step 1 to 3,
    4.2for two attribute regions ai and aj
    4.3if ai-aj> =0.5
    4.4interchange the value of (ai, aj)
    4.5 if ai >aj
    4.6 ai = (ai - aj) - 0.54.7if ai<aj
        4.8value (aj) = (value (aj) - value (ai)) - 0.5
        End for
            Update the support values.
        End for
```

Consider the fuzzy association rule PlsF y=>BMIz

The above algorithm is applied for patient P3 in Table 4.

TABLE VI. FUZZY VALUES OF PLSF Y AND BMI_Z IN PATIENT P3.

Patient's Id	$PlsF_y$	BMI_z	BG_y	BP_y	Bm_z
P3	1	3	1	0.5	1

As Value (PlsFy) < Value (BMIz),
value of (BMIz)=value of (BMIz)-value of (PlsFy)-0.5

=3-1-0.5=1.5.

So modified value is as follows,

TABLE VII. MODIFIED FUZZY VALUES OF PLSFY AND BMIZ IN PATIENT P3

Patient's Id	PlsF y	BMI_z	BG_y	BP_y	Bm_z
P3	1	1.5	1	0.5	1

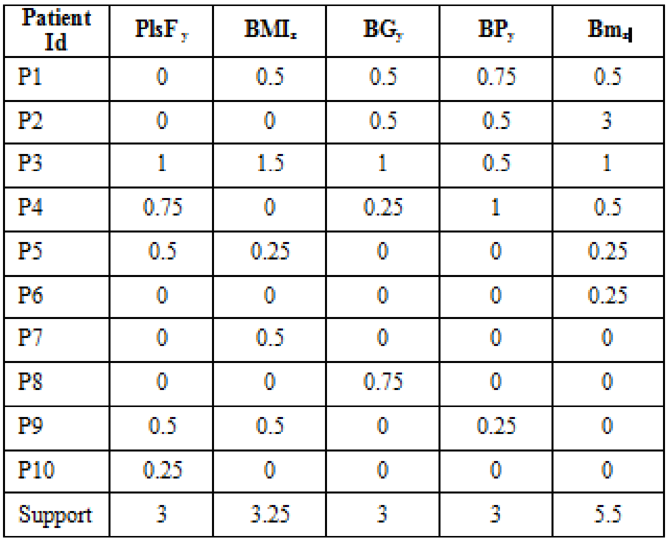
Modified Table after applying Mutation and Crossover Algorithm is provided in Table VIII.

TABLE VIII. MODIFIED TABLE AFTER APPLYING MUTATION AND CROSSOVER ALGORITHM

Patient Id	$PlsF_y$	BMI_z	BG_y	BP_y	Bm_z
P1	0	0.5	0.5	0.75	0.5
P2	0	0	0.5	0.5	3
P3	1	1.5	1	0.5	1
P4	0.75	1	0.25	1	0.5
P5	0.5	0.25	0	0	0.25
P6	0	0	0	0	0.25
P7	0	0.5	0	0	0
P8	0	0	0.75	0	0
P9	0.5	0.5	0	0.25	0
P10	0.25	0	0	0	0
Support	3	4.25	3	3	5.5

If support is greater than minimum support then if any fuzzy value is 1.0 change it to 0.0.

TABLE IX. SUPPORT CALCULATION OF MODIFIED TABLE

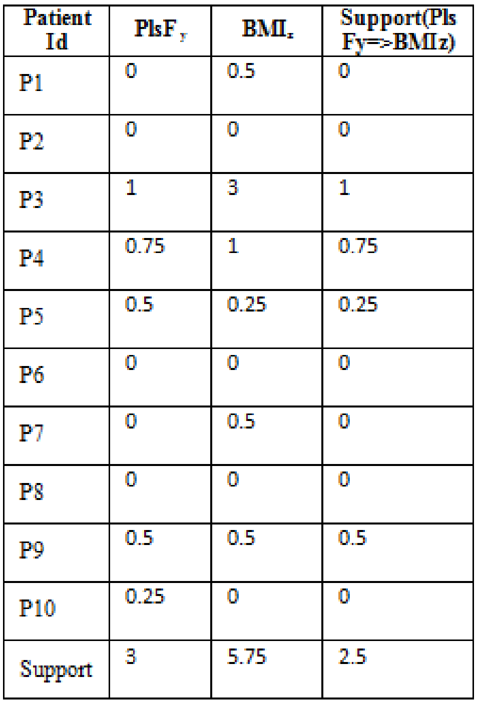


Patient Id	PlsF_y	BMI_z	BG_y	BP_y	Bm_d
P1	0	0.5	0.5	0.75	0.5
P2	0	0	0.5	0.5	3
P3	1	1.5	1	0.5	1
P4	0.75	0	0.25	1	0.5
P5	0.5	0.25	0	0	0.25
P6	0	0	0	0	0.25
P7	0	0.5	0	0	0
P8	0	0	0.75	0	0
P9	0.5	0.5	0	0.25	0
P10	0.25	0	0	0	0
Support	3	3.25	3	3	5.5

Then the child's fitness is computed. Depending on the mutation probability If fitness of child is greater than parent, then parent is replaced by child depending on the mutation probability.

Following Table shows the computation of support count of PlsFy=>BMIz

TABLE X. COMPUTATION OF SUPPORT AND SUPPORT COUNT OF PLSFY=>BMIZ

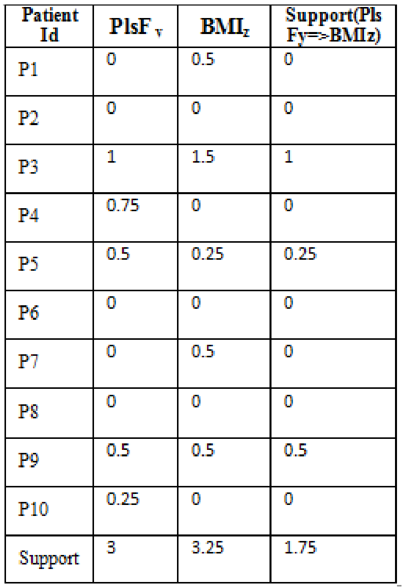


Patient Id	PlsF_y	BMI_z	Support(Pls Fy=>BMIz)
P1	0	0.5	0
P2	0	0	0
P3	1	3	1
P4	0.75	1	0.75
P5	0.5	0.25	0.25
P6	0	0	0
P7	0	0.5	0
P8	0	0	0
P9	0.5	0.5	0.5
P10	0.25	0	0
Support	3	5.75	2.5

Support Count (PlsF y=>BMIz) =2.5

Confidentiality (PlsF y=>BMIz) =Support Count (PlsF y=>BMIz)/ support count (PlsF y) =2.5/3=83%

TABLE XI. COMPUTATION OF SUPPORT COUNT AND SUPPORT OF PLSFY=>BMIZ AFTER MUTATION AND CROSSOVER.



Patient Id	PlsF_y	BMI_z	Support(Pls Fy=>BMIz)
P1	0	0.5	0
P2	0	0	0
P3	1	1.5	1
P4	0.75	0	0
P5	0.5	0.25	0.25
P6	0	0	0
P7	0	0.5	0
P8	0	0	0
P9	0.5	0.5	0.5
P10	0.25	0	0
Support	3	3.25	1.75

Support Count (PlsF y=>BMIz) =1.75
Confidentiality (PlsF y=>BMIz) =Support Count (PlsF y) =>BMIz)/ support count (PlsFy) =1.75/3.25=53.8%
After the calculation of support and count, the confidentiality reduced from 83% to 53.8%. The sensitive rule is successfully hidden.

VI. PERFORMANCE ANALYSIS

This section gives experimental analysis to survey the feature and capability of the proposed rule hiding model. From the test results, it is proved that the system was stable by varying the measuring values and for different values of MS and MC our algorithm hides more rules.

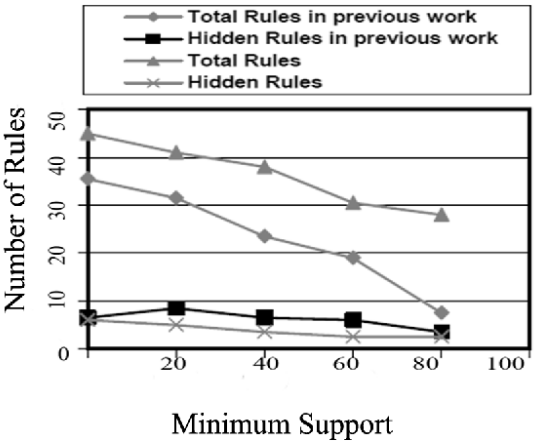


Fig. 2 Simulation Output

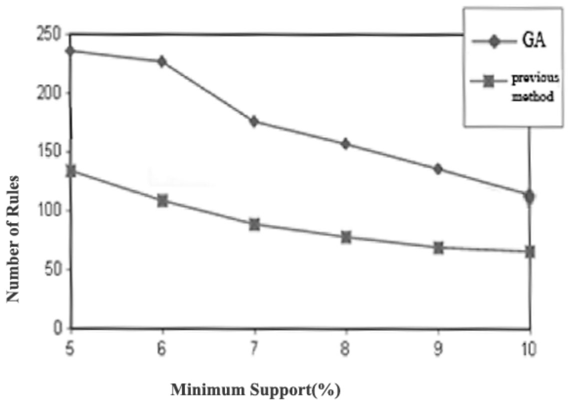


Fig. 3. No. of interesting rules and MS in Jaya based GA.

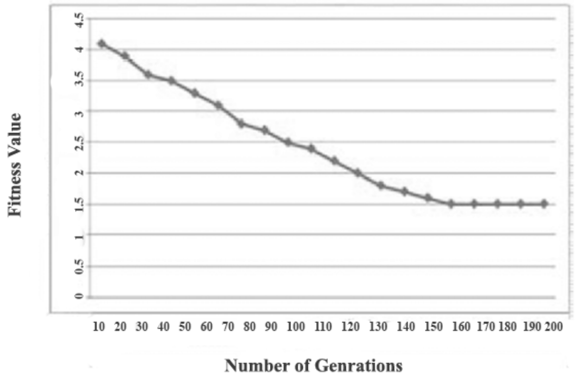


Fig 4. Fitness graph of sensitive rules.

VII. CONCLUSION

In this paper a model has been proposed to hide the sensitive rules of a database before it gets published. Also, we present a case study, in which we examine a few parameters for diabetes patient to get the patient status and consider some potential range of symptoms values. To protect the database and provide the best level of utility for the mined rules, a Jaya-based Genetic Algorithm is employed. There are many drawbacks in privacy preservation for find out the complete and efficient solution. Some of them are privacy of the database should be gain with accuracy. So, optimization of the model should be deeply researched.

REFERENCES

[1] Mary, A. G., Acharjya, D. P., & Iyengar, N. C. S. (2014). Privacy preservation in fuzzy association rules using rough computing and DSR. Cybernetics and Information Technologies, 14(1), 52-71.

[2] Machanavajjhala, A., Kifer, D., Gehrke, J., & Venkitasubramaniam, M. (2007). l-diversity: Privacy beyond k-anonymity. ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1), 3-es.

[3] Tian, H., & Zhang, W. (2009, April). Extending l-diversity for better data anonymization. In 2009 Sixth International Conference on Information Technology: New Generations (pp. 461-466). IEEE.

[4] Sweeney, L. (2001). Computational disclosure control: A primer on data privacy protection (Doctoral dissertation, Massachusetts Institute of Technology).

[5] Sweeney, L. (2002). Achieving k-anonymity privacy protection using generalization and suppression. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(05), 571-588.

[6] Li, N., Li, T., & Venkatasubramanian, S. (2006, April). t-closeness: Privacy beyond k-anonymity and l-diversity. In 2007 IEEE 23rd international conference on data engineering (pp. 106-115). IEEE.

[7] Kamakshi, P., & Babu, A. V. (2012, December). Automatic detection of sensitive attribute in PPDM. In 2012 IEEE international conference on computational intelligence and computing research (pp. 1-5). IEEE.

[8] Vatsalan, D., Christen, P., & Verykios, V. S. (2013). A taxonomy of privacy-preserving record linkage techniques. Information Systems, 38(6), 946-969.

[9] Marimuthu, V. K., & Lakshmi, C. (2021, February). Performance analysis of privacy preserving distributed data mining based on cryptographic techniques. In 2021 7th International Conference on Electrical Energy Systems (ICEES) (pp. 635-640). IEEE.

[10] Shynu, P. G., Shayan, H. M., & Chowdhary, C. L. (2020, February). A fuzzy based data perturbation technique for privacy preserved data mining. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (pp. 1-4). IEEE.

[11] Kaur, A. (2017, February). A hybrid approach of privacy preserving data mining using suppression and perturbation techniques. In 2017 international conference on innovative mechanisms for industry applications (ICIMIA) (pp. 306-311). IEEE.

[12] Malina, L., & Hajny, J. (2013, July). Efficient security solution for privacy-preserving cloud services. In 2013 36th International Conference on Telecommunications and Signal Processing (TSP) (pp. 23-27). IEEE.

[13] Aggarwal, C. C., & Yu, P. S. (2008). A general survey of privacy-preserving data mining models and algorithms. Privacy-preserving data mining: Models and algorithms, 11-52.

[14] Belwal, R. C., Varshney, J., Khan, S. A., Sharma, A., & Bhattacharya, M. (2008, October). Hiding sensitive association rules efficiently by introducing new variable hiding counter. In 2008 IEEE International Conference on Service Operations and Logistics, and Informatics (Vol. 1, pp. 130-134). IEEE.

[15] Yi, X., & Zhang, Y. (2013). Equally contributory privacy-preserving k-means clustering over vertically partitioned data. Information systems, 38(1), 97-107.

[16] Rao, R. (2016). Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. International Journal of Industrial Engineering Computations, 7(1), 19-34.

[17] Mandapati, S., Bhogapathi, R. B., & Chekka, R. B. (2013). A hybrid algorithm for privacy preserving in data mining. International Journal of Intelligent Systems and Applications, 5(8), 47.