Optimization of Dempster-Shafer's Believe Value Using Genetic Algorithm fo Identification of Plant Diseases Jatropha Curcas

Triando Hamonangan Saragih[1], Wayan Firdaus Mahmudy[2], Yusuf Priyo Anggodo[3]
[1,2]Faculty of Computer Science, Brawijaya University, Malang 65145, Indonesia
[3]Data Analyst, Ilmuone Data, Jakarta 12190, Indonesia

Article Info	ABSTRACT
	Jatropha curcas is a plant that can be used as a substitute for diesel fuel. Lack of knowledge of farmers and the limited number of experts and extension agents to deal with the disease of the plant will result lower quality of Jatropha curcas. Dempster-Shafer method can be a solution for decision making based on previous research. The difference in beliefs of every expert in seeing Jatropha diseases may reduce the accuracy of the method. A set of numerical experiment prove that optimization of belief values using genetic algorithms can improve the accuracy Dempster-Shafer.

Corresponding Author:

Wayan Firdaus Mahmudy,
Faculty of Computer Science,
Brawijaya University, Malang 65145, Indonesia.
Email: wayanfm@ub.ac.id

1. INTRODUCTION

Jatropha curcas is a shrub that can live in dry conditions and in an area that has low rainfall (1). Jatropha can be found in Southeast Asia, southern Africa and Central and South India (2). This plant can be used as a substitute for diesel fuel (3).

The many types of diseases that attack Jatropha curcas can degrade the quality of the resulting Jatropha curcas (1). The lack of experts and farmers' knowledge about Jatropha curcas give adverse effects to Jatropha curcas. Issues that are not completed as soon as possible negative impact on the quality of Jatropha curcas. This problem can be helped using an expert system. An expert system is a system that adopts expert knowledge is then fed into a computer and then the computer can provide solutions to problems like an expert (4).

This problem can be solved by various methods, such as previous studies using Dempster-Shafer (5) method which is still one family in the methods along with Certainty Factor (6). Other studies prove the merger of two different methods can resolve these issues, such as the use of Neural Network to the implementation Backprogation structure using Genetic Algorithms (7). Other studies prove that using other method such as fuzzy neural network (8) can resolve this problem and get better result with neuron optimized with Simulated Anealing (9).

Dempster-Shafer, a method of representation, as well as the combination of propogation uncertainty. This method has the characteristics are instutitif in common with the way of thinking of an expert, but has strong mathematical basis (10).

Dempster-Shafer method uses the value of belief to make a decision. Values obtained from the belief of experts through random numbers 0-1 estimate the influence of a symptom of the disease (11). This is equivalent to changing the expert knowledge gained into a number, whereas the value obtained from

an expert can be different with other experts in the same field. This issue is never discussed in a study that questioned the credibility of expert knowledge is processed into a probability parameter. Patrick Hester then suggested their credibility measurement belief values obtained from experts, but there is still no further research to show the validity of the results of this study (12). Researcher using other existing methods to solve the problems of the value of belief, namely genetic algorithms.

Genetic Algorithm is a simple but powerful computational theory in search of improvement (13). In optimization problems, genetic algorithms are often used as a settlement, such as research classification of breast cancer using Neural Network to the implementation of the Genetic Algorithm in the Backpropogation structure provide results that this method of Neural Network using a genetic algorithm as optimization parameters generate an average value better accuracy than methods Naïve Bayes and Neural Network methods with Asociation Rules (7).

Another study conducted Wijayaningrum and Mahmudy prove that optimization for scheduling ships' route using Genetic Algorithms can generate nearly optimal solution (14). The authors intend to use genetic algorithms to optimize the value of belief in the method of Dempster Shafer.

Based on exposures that has been described authors conducted a study titled Value Belief Optimization Implementation Jatropha Curcas Plant Disease Detection. This system can identify Jatropha Curcas plant disease based on symptoms, as well as providing better results when using genetic algorithms to generate value belief.

2.   GENETIC ALGORITHM IN DEMPSTER-SHAFER

Dempster-Shafer method is a method that has a model frame of discernment which is denoted by θ (theta). Frame of discernment is the universe of discourse of a set of hypotheses to associate trust elements θ because not all evidence directly supports each element. For that we need the probability density (m), which will look for the largest density value as a result of the decision (11).

Genetic algorithm is designed to mimic of the natural system necessary for evolution, in particular the theory of evolution Charles Darwin, the survival of fitness (15). Terms used in genetic algorithm is also adopted from the science of genetics such as chromosomes, genes, crossover, mutation, and others. In addition to the terms, the process of crossover, mutation, and selection also adopted from genetic science applied in this algorithm (16). The working process of Genetic Algorithm with Dempster-Shafer is as follows:

Genetic Algorithm
1. Initialization parameter.
2. Generate random first generation
3. Evaluate the fitness value of each chromosome in the population.
4. Generate a new population using the following process:
    a. Selection: Take two parent chromosomes from the existing population
    b. Crossover: Do crossover against two parent chromosomes to produce new offspring
    c. Mutation: Offspring formed from the existing parent mutations
5. Obtain a new population in the next generation.
6. Repeat the process again from the beginning to find the desired needs.

Dempster-Shafer
7. Take a belief value of each criterion selected.
8. Determine the highest belief value of each criterion selected.
9. Determine the plausibility value of each criterion selected.
10. Doing a subset of the criteria with other criteria gradually.
11. Getting density values based on the calculation subset.
12. Make decisions based on the highest density value.

3.   RESEARCH METHOD

The datas are used as many as 30 criteria for symptoms of the disease and 9 types of illness. Symptoms are taken from several parts of Jatropha as fruits, leaves, stems and roots.

3.1. Chromosome Representation

Representation of the chromosome were used that using integer representation. There are 270 genes in one chromosome. Each gene has a value of 0-100 representing their respective belief value of jatropha curcas plant diseases. Figure 1 shows an example of chromosome representation.

| G1 | G2 | G3 | … | G100 | G101 | G102 | G103 | … | G201 | G202 | G203 | … | G270 |

Figure 1. Chromosome representation

3.2. Fitness

In the selection process using the fitness value derived from the value of the accuracy of the calculation based on the Dempster-Shafer belief contained in each chromosome. There are 50 examples of cases that are used for the calculation of fitness value using Equation (1).

$$\text{fitness} = \frac{\text{the number of cases that is true}}{\text{total number of cases}} \tag{1}$$

3.3. Reproduction

In this stage, to produce offspring. The method used is crossover and mutation. This process relies on the crossover rate and mutation rate are included. In this paper, crossover method used one-cut point and mutation method used random mutation (16). A one-cut point crossover process is done by selecting two individuals and select one point to randomly take the left from the first individual or *P1* and the right of the second individual or *P2* to form a new individual, as shown in Figure 2.

35	34	25	26	36	47	86	13	13	57	86	45	P1
54	31	25	78	76	87	57	90	18	80	23	67	P2
35	34	25	26	36	47	57	90	18	80	23	67	C1

Figure 2. One-cut point crossover

While a random mutation process is done by selecting one individuals to randomly from all individuals and then select two point to randomly, exchange to form a new individual, as shown in Figure 3.

| 35 | 34 | 25 | 26 | 36 | 47 | 86 | 13 | 13 | 57 | 86 | 45 | P1 |
| 35 | 34 | 25 | 57 | 36 | 47 | 86 | 13 | 13 | 26 | 86 | 45 | C1 |

Figure 3. Random mutation

3.4. Selection

Selection is the stage at which the selection to get the best fitness value. Selection were used that using the Selection elitism which took the best individuals based on all the existing population.

3.5. Accuracy Testing

In the process accuracy testing used the value of belief that has been optimized. Accuracy testing of data uses 31 test cases. If the system is issuing more than one decision and worth valued properly, the properly value were used that one divided by the number of decisions issued by the system as shown in Equation (2).

$$\text{accuracy} = \frac{\text{the number of cases that is true}}{\text{total number of cases}} \tag{2}$$

4. RESULT AND DISCUSSION

There are several tests performed, namely the population testing, based on a combination of cr and mr testing and iteration testing. This test aims to determine the optimal parameters to produce the best generation in the optimization.

In testing conducted using population population every multiple of 5 starting from the number 10. Rated cr and mr were used that 0.5 and the number of iterations as many as 30. The results of these tests can be seen in Figure 4.
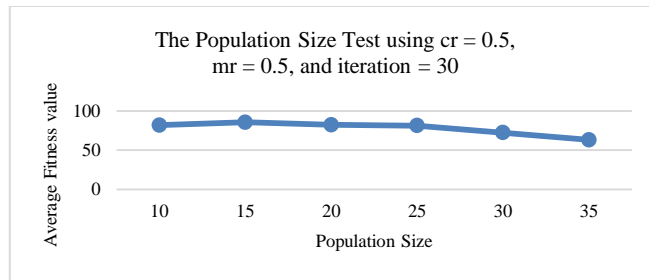

Figure 4. The results of the population size test

The results of the population testing in Figure 4 indicates that the most optimal results possessed a population of 15 with a value of 85.48% accuracy. The increasing number of the population are increasingly making the value of the accuracy of the system is declining.

In the test based on the value of cr and mr used to determine the value of cr and mr optimal as the best solution in this optimization. Population values used are 10, 15, 20 and 25 because it has an accuracy above 80%. The number of iterations used as many as 30. The results are shown in Figure 5.

In the testing based on the value of cr and mr for a total population of 10, said that a value of cr is 0.6 and mr is 0.4 had the highest accuracy of 86.56%. In the next testing the value of cr and mr performed with a total population of 15. The results are shown in Figure 6.
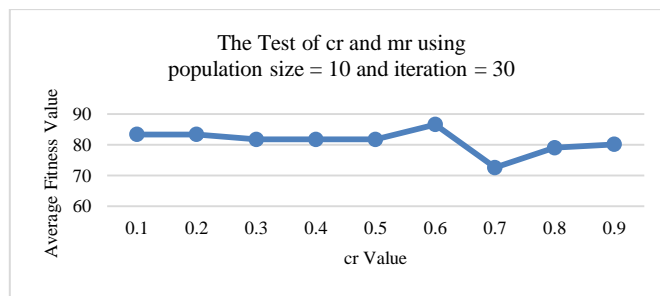

Figure 5. Testing cr and mr with popsize 10

Figure 6. Testing cr and mr with popsize 15

In the testing based on the value of cr and mr for a total population of 15, said that a value of cr is 0.5 and mr is 0.5 had the highest accuracy of 85.48%. In the next testing the value of cr and mr performed with a total population of 20. The results are shown in Figure 7.
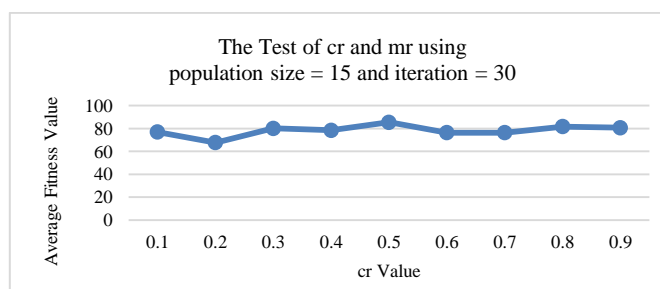

Figure 7. Testing cr and mr with popsize 20

In the testing based on the value of cr and mr for a total population of 20, said that a value of cr is 0.3 and mr is 0.7 had the highest accuracy of 87.1%. In the next testing the value of cr and mr performed with a total population of 25. The results are shown in Figure 8.
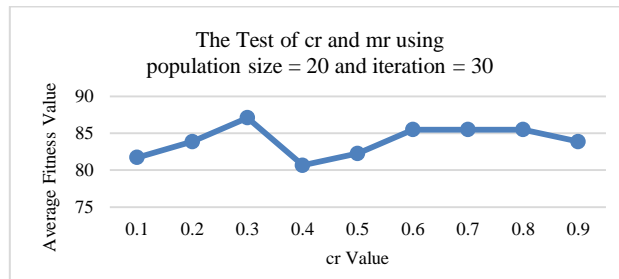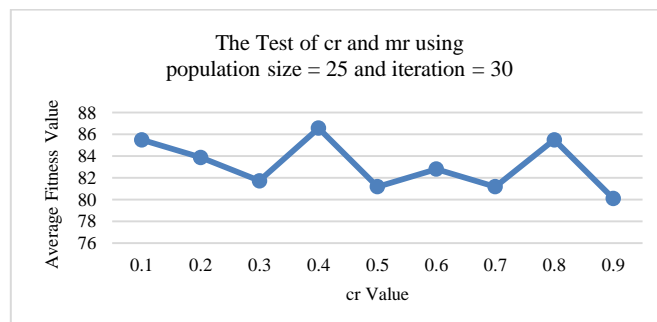

Figure 8. Testing cr and mr with popsize 25

In the testing based on the value of cr and mr for a total population of 20, said that a value of cr is 0.3 and mr is 0.7 had the highest accuracy of 86.56%. Based on result test of cr and mr value with 4 total population of different grades showed that the optimal population size is 20 and the optimum value of cr is 0.3 and mr is 0.7.

Iteration testing aims to find value in the number generation has optimal results in this optimization. Iteration testing used multiple value 5 starts at a value of 10 to 100. The results of the testing iterations can be seen in Figure 9.
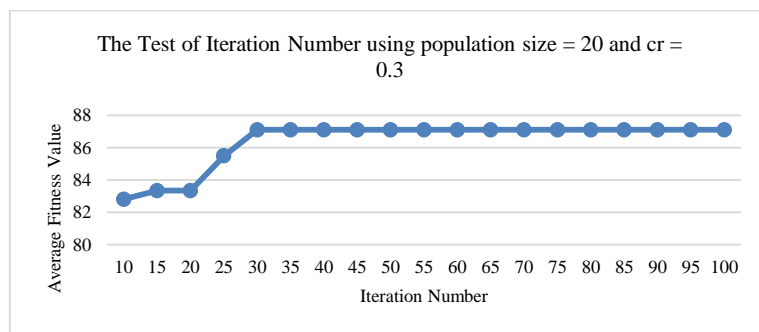

Figure 9. Iteration number testing

Based on Figure 9 for the result test obtained iteration on the optimal value generation 30. At iteration of grades 10 to 30, an increase accuracy value, while the value of 35 to 100 indicates the value of accuracy is stable and equal to the value of accuracy in the 30th generation. This causes an early convergent. Increasing number of iterations provides a long time in computing and does not always give better accuracy. Table 1 shows the result of Dempster-Shafer decision making with belief value from optimization using genetic algorithm with the best parameter.

Table 1. Result of Dempster-Shafer Decision Making

Case	Criteria	Expert Result	System Result	Accuracy
1.	• G06 • G07 • G08 • G09 • G15 • G26 • G28	Bacterial Wilt	Bacterial Wilt	1
2.	• G06 • G07 • G25 • G30	Fusarium Wilt	Fusarium Wilt	1
3.	• G06 • G09 • G10	Charcoal Rot	Charcoal Rot, Fusarium Wilt	0.5
4.	• G06 • G09 • G10 • G25	Charcoal Rot	Charcoal Rot	1
5.	• G18 • G19 • G20	Powdery Mildew	Powdery Mildew	1
6.	• G06 • G07 • G08 • G09	Bacterial Wilt	Charcoal Rot, Bacterial Wilt	0.5
7.	• G08 • G18 • G19 • G20 • G21	Powdery Mildew	Powdery Mildew	1
8.	• G05 • G09	Altenaria Leaf Blight	Altenaria Leaf Blight	1
9.	• G09 • G25 • G29 • G30	Fusarium Wilt	Fusarium Wilt	1
10.	• G23 • G24	Bacterial Blight	Bacterial Blight	1
11.	• G01 • G08	Anthracnose	Anthracnose	1
12.	• G08 • G15 • G19	Dieback	Dieback, Powdery Mildew	0.5
13.	• G06 • G07 • G08 • G09 • G10	Charcoal Rot	Charcoal Rot	1
14.	• G08 • G15 • G19 • G20	Dieback	Dieback, Powdery Mildew	0.5
15.	• G07 • G23 • G24	Bacterial Blight	Bacterial Blight	1

Case	Criteria	Expert Result	System Result	Accuracy
16.	• G06 • G07 • G09 • G10 • G25 • G26	Charcoal Rot	Charcoal Rot	1
17.	• G19 • G20 • G25	Dieback	Dieback	1
18.	• G07 • G08 • G15	Bacterial Wilt	Bacterial Wilt	1
19.	• G07 • G08 • G09 • G26	Bacterial Wilt	Bacterial Wilt	1
20.	• G08 • G19 • G20 • G25	Dieback	Dieback	1
21.	• G26 • G28	Bacterial Wilt	Bacterial Wilt	1
22.	• G09 • G25 • G28	Charcoal Rot	Charcoal Rot, Fusarium Wilt	0.5
23.	• G01 • G02 • G03	Anthracnose	Anthracnose	1
24.	• G08 • G14 • G17	Dieback	Dieback	1
25.	• G06 • G07 • G09 • G10 • G25	Charcoal Rot	Charcoal Rot	1
26.	• G15 • G19 • G20	Dieback	Dieback, Powdery Mildew	0.5
27.	• G15 • G26 • G28	Bacterial Wilt	Bacterial Wilt	1
28.	• G02 • G03 • G23	Bacterial Blight	Bacterial Blight	1
29.	• G08 • G15 • G20	Dieback	Dieback, Powdery Mildew	0.5
30.	• G08 • G15 • G26	Bacterial Wilt	Bacterial Wilt	1
31	• G09 • G25	Charcoal Rot	Charcoal Rot, Fusarium Wilt	0.5
	Total of Accuracy			27

Based on result of Dempster-Shafer decision making using Equation 1 obtained accuracy of 87.096 %. The accuracy with Genetic Algorthm Optimization is better than without optimization that just only gave accuracy 82.3%(5). It proves that with optimization of believe value can increase the accuracy of system.

5. CONCLUSION

        Based on the testing that has been done can be concluded that genetic algorithms can be used to optimize the value of belief in the Dempster-Shafer. Optimization using a genetic algorithm can improve the accuracy of the value system that takes decisions using Dempster-Shafer. Nearly optimal parameters which popsize by 20, the value of cr 0.3 and mr 0.7 and the number of iterations of 30. With these parameter values obtained an accuracy of 87.1% compared with no optimization using genetic algorithms by 82.23%.

In subsequent studies, the optimization of the value of belief in the Dempster-Shafer's case of Jatropha Curcas disease identification can be done with other methods to further enhance the value of the accuracy of the system. Particle Swarm Optimization (PSO) and hybrid genetic algorithm could form the proper and efficient solutions for the optimization.

REFERENCES

[1]   Yulianti T, Hidayah N. Jatropha Curcas Disease. Malang: Balai Penelitian Tanaman Pemanis dan Serat; 2015. 217-232 p.

[2]   Rodrigues J, Miranda I, Furquim L, Gominho J, Vasconcelos M, Barradas G, et al. Storage stability of Jatropha curcas L. oil naturally rich in gamma-tocopherol. Ind Crops Prod. 2015;64:188–93.

[3]   Fernández CM, Fiori L, Ramos MJ, Pérez Á, Rodríguez JF. Supercritical extraction and fractionation of Jatropha curcas L. oil for biodiesel production. J Supercrit Fluids [Internet]. 2015;97:100–6. Available from: http://dx.doi.org/10.1016/j.supflu.2014.11.010

[4]   Dymova L, Sevastjanov P, Kaczmarek K. A Forex trading expert system based on a new approach to the rule-base evidential reasoning. Expert Syst Appl. 2016;51:1–13.

[5]   Saragih TH, Soebroto AA, Yulianti T. Expert System For Jatropha Curcas Disease Identification Using Dempster-Shafer Method. FILKOM Doro. 2016;7(4).

[6]   Setyarini E, Putra D, Purnawan A. The Analysis of Comparison of Expert System of Diagnosing Dog Disease by Certainty Factor Method and Dempster-Shafer Method. Int J Comput Sci Issues. 2013;10(1):576–84.

[7]   Zamani AM, Amaliah B, Munif A. Implementation of Genetic Algorithm on Backpropagation Neural Network for Breast Cancer Classification. J Tek POMITS. 2012;1(1):1–6.

[8]   Saragih TH, Fajri DMN, Hamdianah A, Mahmudy WF, Anggodo YP. Jatropha Curcas Disease Identification Using Fuzzy Neural Network. In: International Conference on Sustainable Information Engineering and Technology (SIET), Batu, Indonesia, 25-25 November. 2017.

[9]   Fajri DMN, Saragih TH, Hamdianah A, Mahmudy WF, Anggodo YP. Optimized Fuzzy Neural Network for Jatropha Curcas Plant Disease Identification. In: International Conference on Sustainable Information Engineering and Technology (SIET), Batu, Indonesia, 25-25 November. 2017.

[10]   Wu D. Supplier selection in a fuzzy group setting: A method using grey related analysis and Dempster-Shafer theory. Expert Syst Appl [Internet]. 2009;36(5):8892–9. Available from: http://dx.doi.org/10.1016/j.eswa.2008.11.010

[11]   Masaleno A, Hasan MM. Skin Diseases Expert System using Dempster-Shafer Theory. Int J Intell Syst Appl. 2012;5:38–44.

[12]   Hester PT, Meyers TJ. Multi-Criteria Performance Measurement for Public and Private Sector Enterprises. Appl Manag Sci. 2012;15:183–206.

[13]   Azim I, Rahman F. Genetic Algorithm Based Reactive Power Management by SVC. Int J Electr Comput Eng. 2014;4(2).

[14]   Wijayaningrum VN, Mahmudy WF. Optimization of Ship ' s Route Scheduling Using Genetic Algorithm. Indones J Electr Eng Comput Sci. 2016;2(1):180–6.

[15]   Wibowo IS, Saragih H, Darjono W, Utomo T, Roestam R. Design Simulation Program of Runway Capacity Using Genetic Algorithm at Soekarno-Hatta. Int J Electr Comput Eng. 2011;1(2):202–12.

[16]   Mahmudy WF, Marian RM, Luong LHS. Real Coded Genetic Algorithms for Solving Flexible Job-Shop Scheduling Problem - Part II: Optimization. Adv Mater Res [Internet]. 2013 May;701:364–9. Available from: http://www.scientific.net/AMR.701.364