

Implementation of Forward Chaining and Certainty Factor Method on Android-Based Expert System of Tomato Diseases Identification

Kurnia Muludi¹

Department of Computer Science
Lampung University
Lampung, Indonesia

Radix Suharjo²

Department of Plant Protection
Lampung University
Lampung, Indonesia

Admi Syarif³, Fitria Ramadhani⁴

Department of Computer Science
Lampung University
Lampung, Indonesia

Abstract—Plant disease is one of the reasons that cause the destruction of plant. It affects plant productivity and quality. Most of the farmers made mistake in cope with this problem because of the lack of knowledge. Expert system is a solution that has been widely used for identifying disease. This paper presents an Android-based expert system to help identifying tomato diseases. Data used in this expert system consist of 16 data of tomato diseases, 53 data of symptoms, and 20 variety of rules. This paper implements forward chaining and certainty factor method. Forward chaining is used as a reasoning method to get the result of disease identification. Certainty factor is used as a calculation method to obtain accuracy degree of identification results. Testing has been done through two stages, internal and external. The result from internal testing shows that tomato expert system works properly and fit perfectly in various android devices. External testing is done by giving questionnaire to 44 respondents. The result of questionnaires shows that tomato expert system is categorized as “good” by them.

Keywords—Expert system; forward chaining; certainty factor; tomato diseases; android

I. INTRODUCTION

Tomato is a commodity plant that is being cultivated by many farmers in Indonesia due to its benefit. Tomato can be categorized as vegetable and fruit. Tomato is also used as food ingredient and medicine [1]. Based on statistical data of vegetable productivity in Indonesia in 2016, tomato becomes big seven along with onion, garlic, and chili with 883.233 ton^a. Tomato becomes one of horticulture commodities that has high economic value yet still needs more concern, especially in production and quality enhancement [2]. This matter is showed by statistical data of tomato in Lampung province between 2014 until 2016 in Table I^a. Crop areas increase twice in 2016 yet is not followed by production result. So, the productivity decreases treble with only 4,54 Ton/Ha.

Plant disease is one of the most important reasons that lead to the destruction of plants and crops [3]. Detecting diseases at early stage enable people to overcome and treat them appropriately. This process requires an expert but the number of experts are still far from ideal condition. Moreover, it needs high cost to do consultation. Therefore, expert system helps a great deal in identifying diseases and describing methods of treatment to be carried out.

TABLE I. TABLE OF TOMATO STATISTICAL DATA IN LAMPUNG PROVINCE BETWEEN 2012 - 2016

Year	Crop area (Ha)	Production (Ton)	Productivity (Ton/Ha)
2014	2.366	23.776	10,05
2015	2.143	24.490	11,43
2016	5.211	23.638	4,54

This paper develops an expert system for identifying tomato diseases implementing forward chaining and certainty factor method. These methods are used to get identification results with accuracy degree. This expert system is designed to be android based, so that it can be used by many people anywhere and anytime.

II. LITERATURE REVIEW

A. Artificial Intelligence

Artificial Intelligence (AI) is the area of computer science that focus on creating expert machines that can engage on behaviors that humans consider intelligent [4]. It has many areas in computer science. There are 12 fields in AI, such as Natural Language Processing (NLP), Speech (voice) Understanding, Robotics and Sensory Systems, and Expert Systems [5].

B. Expert System

Expert system is a field in artificial intelligence intended to serve as consultant for decision making. It uses a collection of facts, rules of thumb, and other knowledge about a limited domain to help making inferences in the domain. It is called expert systems because it addresses problems normally thought to require human specialists for solution [6]. Expert systems can be viewed as having two environments: the development environment and the consultation environment. The development environment is used by an expert system builder to build the components and put knowledge into the knowledge base. The consultation environment is used by a non-expert to obtain expert knowledge and advice. The structure of expert system is showed in Fig 1 [5].

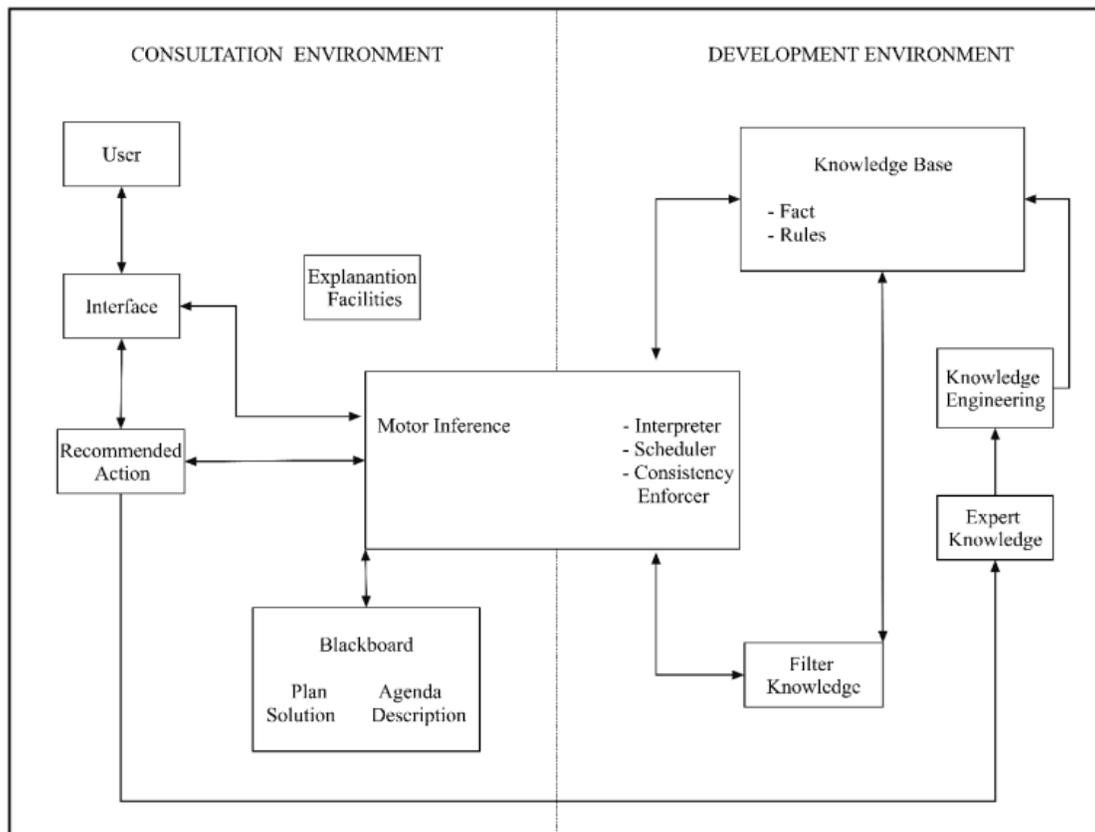


Fig. 1. The Structure of Expert System.

C. Inference Engine

Expert knowledge often can not be represented in single rule. Instead, rules may be chained together dynamically to cover numerous conditions. The process of chaining multiple rules together is called inference. The component conducting inference in an expert system is called the inference engine. Inference engine provides a methodology for reasoning about information in the knowledge base and on the blackboard, and for formulating conclusions. This component provides directions about how to use the system's knowledge by developing the agenda that organizes and controls the steps taken to solve problems whenever consultation takes place. Therefore, inference engine is controller structure or rule interpreter, also is known as brain in expert system. There are two popular approaches for inferencing, such as forward chaining and backward chaining [5].

Forward-chaining is a bottom-up computational model. Forward chaining works with a set of known facts and applies rules to generate new facts whose premises match the known facts, and continues until it reaches a predetermined goal, or until no further facts can be derived whose premises match the known facts. Backward chaining is a reasoning model in expert system that using a goal-driven search. It is a top-down computational design and starts with a goal or hypothesis and then looks for rules to support the hypothesis. It attempts matching the variables that lead to valid facts in the data and indicates that the inference moves backward from the intended goal to determine facts that would satisfy that goal [7].

Forward chaining is appropriate for problems in which most of the data is given in the initial problem statement as how this model works. Forward chaining works forward to find conclusion from facts, starts with the initial facts. Meanwhile, backward chaining starts with hypothesis or goal to know whether the hypothesis accepted or rejected. Backward chaining is appropriate for problem in which goal or hypothesis is given in the problem statement [7]. In Conclusion, forward chaining is used as a reasoning because the system is identifying diseases based on symptoms or facts.

D. Certainty factor

To develop an expert system calculation method is needed to get accuracy degree of the identification results. There are some calculation methods used in expert system, such as heuristic method, classical probability, bayes theorem, and certainty factor.

Heuristic is a method introduced by George Polya in 1945. He defines heuristic as rules for choosing branches in a state space that are most likely to lead to an acceptable problem solution. This method has weakness because it often works based on experience or intuition, so the result maybe inaccurate. Also, it uses limited information and seldom able to predict the exact behavior of the state space farther along in the search [8].

Theory of classical probability was proposed by Pascal and Fermat in 1654. Classical probability considers ideal system that do not become worn out like dice, card, and coin.

Classical probability only considers probability that will be searched, as showed by the following formula:

$$P = W/N \quad (1)$$

Where W is the number of wins and N is the number of equally possible events [9].

Bayes theorem relates cause and effect in such a way that by understanding the effect people can learn the probability of its cause. This theory was proposed by Thomas Bayes in 1763 in his paper *Essay towards Solving a Problem in the Doctrine of Chances* [8]. Bayes theorem and classical probability requires the application of mathematical probability in uncertainty management. Bayes theorem is better than classical probability because it has prior value. Unfortunately, to obtain past data is becoming problem for expert. As a result, if prior data is lacking then it will affect the conclusion, even can give wrong conclusion [10].

One of the most known expert systems attempts to develop representations of uncertainty as an extension of deterministic rule-based expert systems is MYCIN. MYCIN was designed to assist physicians in the diagnosis and treatment of bacterial infections. Shorliffe Buchanan in 1975 introduced certainty factor (CF) method through MYCIN. When a CF-based system recommends a decision (for example, when MYCIN suggests treatment for a suspected infection), it makes use of the CFs assigned to the competing diseases to assess the amount of evidence for each [11]. Here is the formula of CF [12]:

$$CF(H,E) = MB(H,E) - MD(H,E) \quad (2)$$

Description :

$CF(H,E)$: Certainty factor from hypothesis H which is affected by symptom (evidence) E

$MB(H,E)$: Measurement of increased belief on hypothesis H affected by symptom E

$MD(H,E)$: Measurement of increased disbelief on hypothesis H affected by symptom E

1) *Parallel CF*: Parallel CF is a CF obtained by some premises from rule. The value of parallel CF is affected by CF User for each premise and operator used in the premise. Here is the formulas for parallel CF [12]:

$$CF(x \text{ AND } y) = \min(CF(x), CF(y)) \quad (3)$$

$$CF(x \text{ OR } y) = \max(CF(x), CF(y)) \quad (4)$$

$$CF(\text{NOT } x) = -CF(x) \quad (5)$$

2) *Sequential CF*: Sequential CF is obtained from calculation result of parallel CF of all premises in a rule with CF from expert. Sequential CF is defined as follows [12]:

$$CF(x,y) = CF(x) \times CF(y) \quad (6)$$

Description :

$CF(x,y)$: Sequential CF

$CF(x)$: Parallel CF from all premises in a rule

$CF(y)$: CF from expert

3) *Combined CF*: Combined CF is final CF from conclusion candidate. This CF is affected by all parallel CF from a rule that will produce conclusion. Combined CF is used when a conclusion obtained from several rules at once. Here is the formula for Combine CF [12]:

$$CF(x,y) = \begin{cases} CF(x)+CF(y)-(CF(x) \times CF(y)), & \text{where } CF(x)>0 \\ & \text{and } CF(y)>0 \end{cases} \quad (7)$$

$$\begin{cases} CF(x)+CF(y)/(1-(\min(|CF(x)|,|CF(y)|))), & \text{where} \\ & \text{one of } (CF(x),CF(y))<0 \end{cases} \quad (8)$$

$$\begin{cases} CF(x)+CF(y) \times (1+CF(x)), & \text{where } CF(x)<0 \text{ and} \\ & CF(y)<0 \end{cases} \quad (9)$$

E. Expert Systems in Agriculture

Expert system has been implemented in many areas, such as medical, finance, automotive, and agriculture. During this five years, many expert systems in agriculture were developed, like tobacco expert system, sugarcane expert system, and red chili peppers expert system. CaneDES is a web-based expert system for disorder diagnosis in sugarcane. This expert system diagnoses disorders in sugarcane which is caused by biotic and abiotic and implements a new classified knowledge representation technique [13]. Another expert system that has been developed is an expert system for tobacco disease management that is applied in web. This tobacco expert system uses two methods for disease diagnosis, such as rule based inference and ontology [14]. Expert system for diagnosing diseases of red chili uses forward chaining and certainty factor method. Based on testing results, this expert system can diagnose red chili plant disease with high accuracy [15].

F. Android

Nowadays web-based systems which can be accessed on the PC or laptop is slowly getting unpopular. People move to palm tools gadget like smartphone which makes activity more comfortable and efficient. So that, every business is developing applications for the mobile devices. One of the most popular operating systems used in mobile device is android [16]. Android is a software stack for mobile device that includes an operating system, middle-ware and key applications. Android is mainly based on Linux operating system which uses java-like languages for running applications [17]. The platform was officially announced in October 2008 [18].

G. Diseases of Tomato

In this research, tomato diseases are divided into two categories, biotic and abiotic. Biotic disease is caused by organism such as fungus, virus, and bacteria. While, abiotic disease is caused by environmental factors like pesticide, less of calcium, and sunshine intensity. There are 12 diseases in biotic category, such as Late Blight (*Phytophthora infestans*), early blight (*Alternaria solani*), leaf mold (*Fulvia fulva*), Fusarium wilt (*Fusarium oxysporum f.sp. lycopersici*), Bacteria wilt (*Ralstonia solanacearum*), Tobacco Mosaic (Tobacco Mosaic Virus/TMV), Cucumber Mosaic (Cucumber

Mosaic Virus/CMV), crinkle (Tobacco Crinkle Virus), Tomato Yellow Leaf Curl (Tomato Yellow Leaf Curl Virus/TYCLV), gray mold (*Cercospora sp.*), powdery mildews (*Oidium sp.*) [19]. There are 4 diseases in abiotic category, such as pesticide poisoning, lack of calcium, high intensity of sunshine, and low intensity of sunshine [20].

III. SYSTEM DESIGN

A. Use Case Diagram

Use case diagram describes interaction between user and system, like what action user can do in this system. This diagram shows that tomato expert system has 5 functions as shown in Fig 2, such as do diagnose, see disease's detail information, access help menu, access about menu, and access instruction menu.

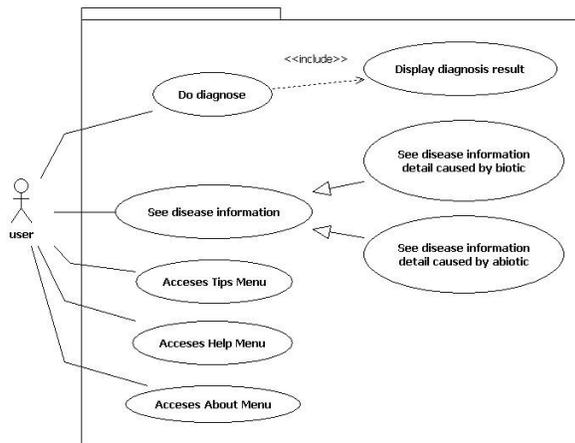


Fig. 2. Use Case Diagram of Tomato Expert System.

B. Calculation of Certainty Factor Method

Calculation method implemented in this expert system is certainty factor. In developing expert system, rule-based is made into two types operator, OR and AND. Operator AND is used for main symptoms. Meanwhile, symptoms using operator OR is displayed in list form. For example, user done a consultation on the tomato leaf with the following symptoms.

- There's spot with diameter 0,5 cm (G22)
- There's gray-brownish ring on the spot (G23)
- There's brown-reddish mold on the spot (G24)

Expert gives CF value 0.79 for the selected symptoms. The rule: **“IF** (There's spot with diameter 0,5 cm (G22) **OR** There's gray-brownish ring on the spot (G23)) **AND** There's brown-reddish mold on the spot (G24) **THEN** *Corynespora* Spot Leaf (B12)”.

Then the calculation is:

$$\begin{aligned}
 \text{Parallel CF} &= CF (G24) \text{ AND } CF (G22 \text{ OR } G23) \\
 &= \min (G24, \max (G22, G23)) \\
 &= \min (0.83, \max (0.79, 0.81)) \\
 &= \min (0.83, 0.81) \\
 &= 0.81
 \end{aligned}$$

$$\begin{aligned}
 \text{Sequential CF} &= CF \text{ from Expert} \times \text{Parallel CF} \\
 &= 0.79 \times 0.81 = 0.6399
 \end{aligned}$$

Percentage of accuracy = 63,99 %

Based on the calculation result, the conclusion for following symptoms chosen by user is *Corynespora* Spot Leaf with the accuracy 63,99%.

IV. TESTING AND RESULTS

The system was developed by using Android Studio, XAMPP on TOSHIBA Satellite L840 and android smartphone OPPO F1s A1601. Two stages of testing, internal and external, is done to see the effectiveness of the system. Internal testing is done by testing system's compatibility and accuracy degree of identification results. To see whether the system is compatible or not, the system is installed in several android devices with different version and screen size. As a result, the system is compatible with Android 4.4 (Kitkat), 5.0 (Lollipop), 6.0 (Marshmallow), and 7.0 (Nougat) version and with android 4,5 inch, 5 inch, and 5,5 inch. The system's function works properly and the interface fits perfectly as shown in Fig 3.

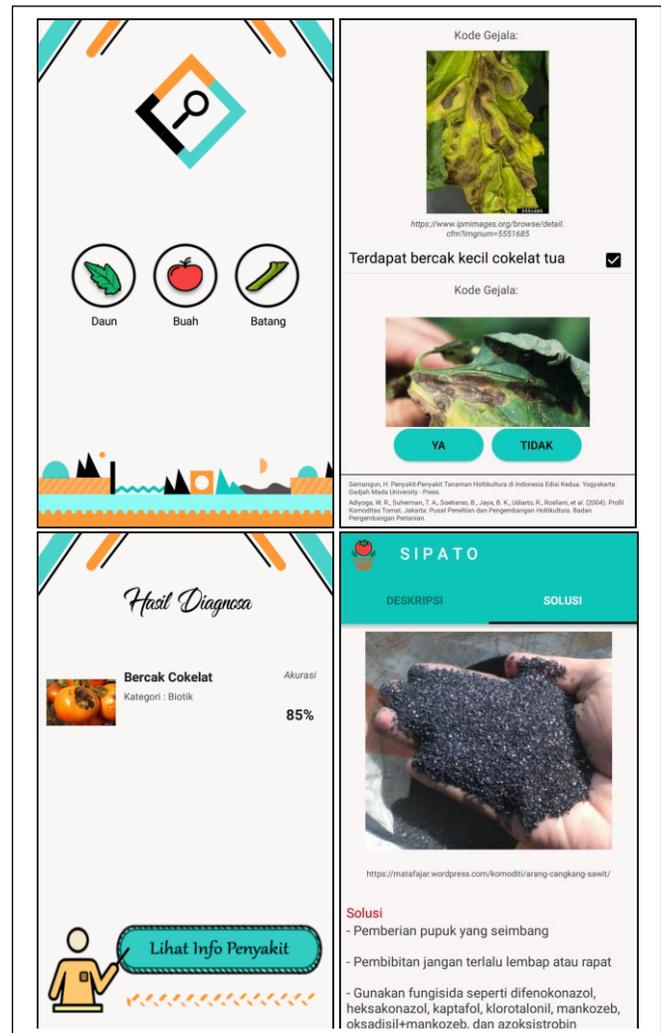


Fig. 3. Interface of Tomato Expert System.

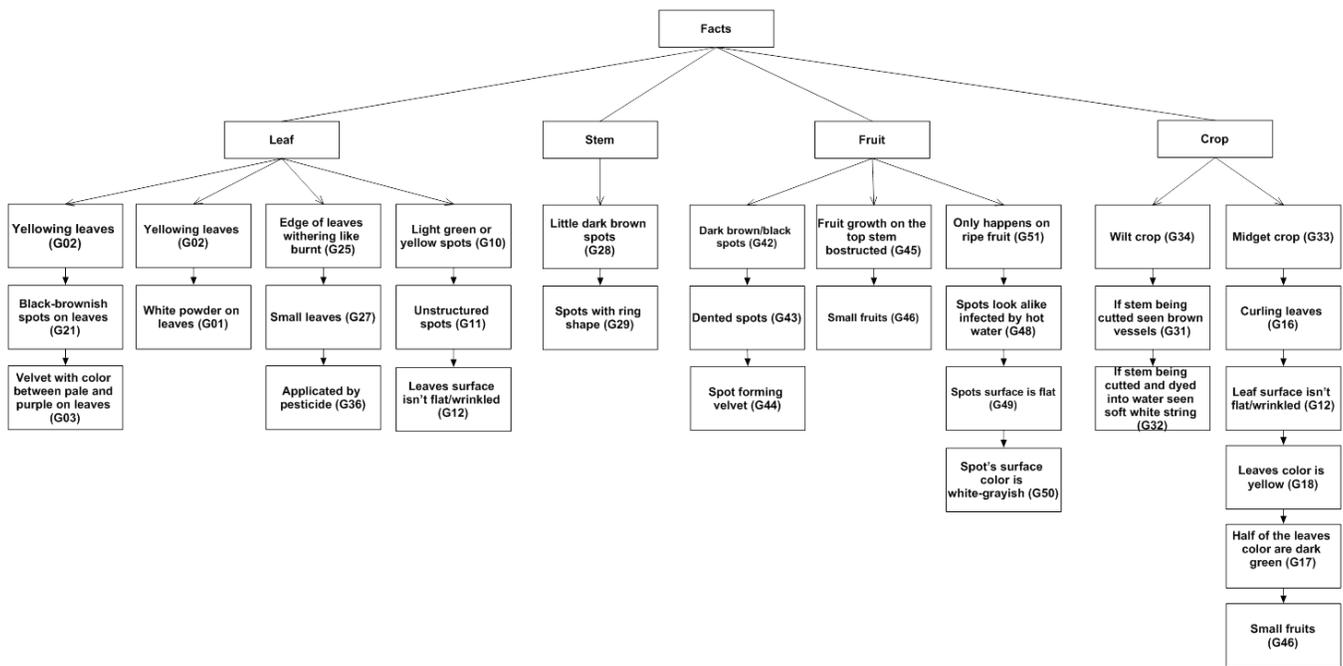


Fig. 4. Facts of Symptom.

This paper provides some facts and compares identification results between the system and expert. It use 33 facts divided into 4 symptom locations, such as leaf, stem, fruit, and crop as shown in Fig 4.

Testing result shows that system gives the accurate identification results. Also, there's degree of accuracy from implementation of certainty factor method. Comparison of identification results is showed in Table II.

TABLE II. TABLE OF COMPARISON IDENTIFICATION RESULTS BETWEEN TOMATO EXPERT SYSTEM AND EXPERT

Facts	Diagnosis	
	System Calculation	Expert Judgement
G02, G21, G03	Late Blight : 66,4%	Late Blight
G02, G01	Powdery Mildews : 61,62%	Powdery Mildews
G25, G27, G36	Pesticide Poisoning : 62,37%	Pesticide Poisoning
G10, G11, G12	Tobacco Mosaic : 63,99%	Tobacco Mosaic
G28, G29	Early Blight : 65,6%	Early Blight
G42, G43, G44	Early Blight : 66,42%	Early Blight
G45, G46	Cucumber Mosaic : 57,75%	Cucumber Mosaic
G51, G48, G49, G50	High intensity of sunshine : 65,57%	High intensity of sunshine
G34, G31, G32	Bacteria Wilt : 71,4%	Bacteria Wilt
G33, G16, G12, G18, G19, G46	Yellow Leaf Curl : 65,57%	Yellow Leaf Curl

External testing is done by giving questionnaire to several respondents. The respondent is divided into 3 groups. Group 1 is experts consist of 4 lecturers from Department of Plant Protection Lampung University. Group 2 is people who are related with this domain consist of 10 farmers and 20 students of Agriculture Faculty of University of Lampung. Group 3 is common people who do not understand about tomato diseases consist of 10 students of Computer Science of University of Lampung.

The respondents were asked to access the system directly and give their assessment by completing the questionnaire. Each statement has 5 answer points with different value, such as Completely Disagree (CD) = 1, Disagree (D) = 2, Slightly Agree (SA) = 3, Agree (A) = 4, and Completely Agree (GA) = 5. Then, those questionnaires being generated into percentage with this following formula [21]:

$$P = \frac{(a \times A)}{(n \times N)} \times 100\% \tag{10}$$

P = Percentage of statement

a = Number of respondents choosing this point

A = Value of statement

n = Number of all respondents

N = Highest value of statement

Next step is deciding interval for each category using this following formula [21]:

$$I = 100\% / k \tag{11}$$

Description :

I = Interval

K = Number of categories

TABLE III. TABLE OF INTERVAL CATEGORY

Percentage	Description
$0\% \geq P \leq 20\%$	Not Good
$20\% > P \leq 40\%$	Less Good
$40\% > P \leq 60\%$	Good Enough
$60\% > P \leq 80\%$	Good
$80\% > P \leq 100\%$	Very Good

REFERENCES

This research uses 5 categories, so interval for each category is 20%. The description for each category is showed in Table III.

Based on the questionnaires result, it shows that this expert system provides useful information for people especially tomato farmers with 87,33%. It helps in identifying diseases well with 79,44%. System works properly and appropriate with 84%. System is easy to understand and to use with 82,77%. Data in the system is complete and clear with 74,44%. It makes people understand about expert system with 79,11%. System provides right solution with 76,11%. Graphic for external testing result is showed in Fig 5.

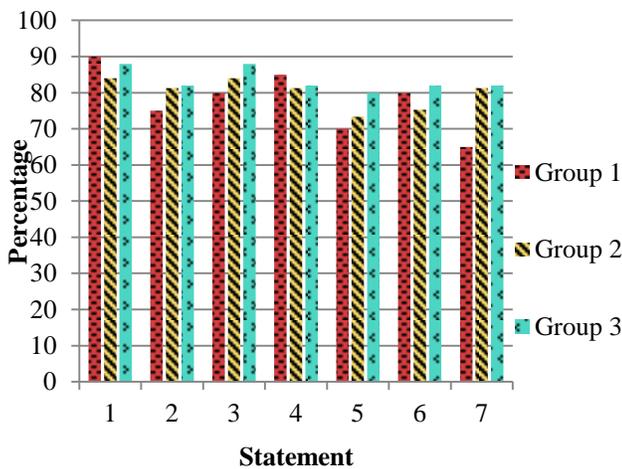


Fig. 5. Graphic of External Testing Result.

V. CONCLUSIONS

An Android-based tomato expert system implements forward chaining and certainty factor method to identifying tomato diseases has been established. The system which has been installed on various android devices shows the accurate identification result as it is expert. Moreover, all functions work properly and the interface fits perfectly. The result from 44 respondents shows that this tomato expert system is categorized as “good”.

- [1] Bambang Cahyono, *Technique of Cultivating Superior Tomato with Organic and Anorganic Method*. Depok: Pustaka Mina, 2016.
- [2] Roydatul Zikria, *Tomato Commodity Outlook 2014*. Jakarta: Agriculture Data and Information System Center, 2014.
- [3] S S Abu-Naser, K A Kashkash, and M Fayyad, "Developing an Expert System for Plant Disease Diagnosis," *Journal of Artificial Intelligence*, vol. 1, no. 2, pp. 78-85, 2008.
- [4] P Santosh Kumar Patra, Dipti Prava Sahu, and Indrajit Mandal, "An Expert System for Diagnosis of Human Diseases," *International Journal of Computer Applications*, vol. 1, no. 13, pp. 71-73, 2010.
- [5] Efraim Turban, Jay E Aronson, and Ting- Peng Liang, *Decision Support Systems and Intelligent Systems*. New Jersey: Prentice-Hall, Inc, 2005.
- [6] Herbert E Rauch, "Probability Concepts For An Expert System Used For Data Fusion," in *Seventeenth Annual Asilomar Conference on Circuits, Systems and Computers*, California, 1983, pp. 55-60.
- [7] Ajlan S Al-Ajlan, "The Comparison between Forward and Backward," *International Journal of Machine Learning and Computing*, vol. 5, no. 2, pp. 106-113, 2015.
- [8] George F Luger, *Artificial Intelligence Structures and Strategies for Complex Problem Solving*. England: Pearson Education Limited, 2005.
- [9] Joseph Giarratano and Gary Riley, *Expert Systems Principles and Programming*. U.S.A: PWS Publishing Company, 1998.
- [10] Yeow Wei Liang and Rohana Mahmud, "A Comparison Model for Uncertain Information in Expert System," *International Conference on Uncertainty Reasoning and Knowledge Engineering*, pp. 127-130, 2012.
- [11] Eric J Horvitz, John S Breese, and Max Henrion, "Decision Theory in Expert Systems and Artificial Intelligence," *International Journal of Approximate Reasoning*, vol. 2, pp. 247-302, 1988.
- [12] Kusrini, *Expert System Application Deciding User Certainty Factor with Quantification Question Method*. Yogyakarta: ANDI, 2008.
- [13] S S Hasan et al., "CaneDES: A Web-Based Expert System for Disorder Diagnosis in Sugarcane," *Sugar Tech: An International Journal of Sugar Crops and Related Industries*, vol. 17, no. 4, pp. 418-427, 2015.
- [14] H Ravisankar, K Sivaraju, D Damodar Reddy, and N Srinivasa Rao, "Web Based Expert System for Tobacco Disease," *Journal of Entomology and Zoology Studies*, vol. 6, no. 1, pp. 5-11, 2018.
- [15] Fahrul Agus, Hernandha Eka Wulandari, and Indah Fitri Astuti, "Expert System With Certainty Factor For Early Diagnosis Of Red Chili Peppers Diseases," *Journal of Applied Intelligent System*, vol. 2, no. 2, pp. 52-66, 2017.
- [16] Chetan Bulla, Sachin Bakanetti, Jayant Bhosale, Kiran Patil, and Poornima Gujanal, "My Campus Android Application," *International Journal of Engineering Science and Computing*, vol. 7, no. 6, pp. 12631-12632, 2017.
- [17] Mohamed Abd El-Latif Mowad, Ahmed Fathy, and Ahmed Hafez, "Smart Home Automated Control System Using Android Application and Microcontroller," *International Journal of Scientific & Engineering Research*, vol. V, no. 5, pp. 935-939, 2014.
- [18] Suhas Holla and Mahima M Katti, "Android Based Mobile Application Development and Its Security," *International Journal of Computer Trends and Technology*, vol. 3, no. 3, pp. 486-490, 2012.
- [19] H Semangun, *Horticulture plant diseases in Indonesia*, 2nd ed. Yogyakarta: Gadjah Mada University - Press, 2007.
- [20] W Setiawati, I Sulatrini, and N Gunaeni, *Implementation of PHT Technology on Tomato Plant*. Bandung: Plant Vegetable Research Office, 2001.
- [21] Djarwanto, *Knowing Some of Statistical Test in Research*. Yogyakarta: Liberty, 1996.