Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal **Vol. 6, No. 2, April 2024, pp: 382-396; eISSN: 2656-8632**

RESEARCH ARTICLE **OPEN ACCESS**

Manuscript received May 3, 2024; revised May 25, 2024; accepted May 27, 2024; date of publication October 8, 2024
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v6i4.448

How to cite: Setia Mukti Azizah, Wiharto, and Brilyan Hendrasuryawan, "Expert System for Pregnancy Risk Diagnosis Using Decision Tree and Dempster-Shafer Method", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 6, no. 3, pp. 382-396, October 2024.

Expert System for Pregnancy Risk Diagnosis Using Decision Tree and Dempster-Shafer Method

Setia Mukti Azizah, Wiharto, and Brilyan Hendrasuryawan
Department of Informatics, Universitas Sebelas Maret, Surakarta, Indonesia

Corresponding author: Wiharto (e-mail: wiharto@staff.uns.ac.id).

ABSTRACT The high Maternal Mortality Rate (MMR) remains a severe concern in maternal healthcare. One of the reasons is the delay in recognizing early danger signs during pregnancy. Therefore, in order to address this issue, there is a proposed solution in the form of developing an expert system which aims to diagnose pregnancy risks in pregnant women quickly and efficiently by using the Decision Tree and Dempster Shafer methods. The Decision Tree method was used for symptom classification while Dempster Shafer provided confidence values for existing facts. This study collected data from the dataset, the Poedji Rochjati Score Card (KSPR), and qualitative data through expert interviews. From the collected data, knowledge acquisition processes were then conducted to extract knowledge by using the ID3 Decision Tree and combine all symptoms from the gathered data. Furthermore, the processed data was represented as a decision tree and assigned confidence values. The development of this expert system utilized the Laravel framework with PHP language and MySQL database. System validation involved patients as participants and midwives as experts and testers. Testing was conducted on March 13th and 16th 2024, involving 16 patients at the Gatak Community Health Center. The system evaluation results show an accuracy rate of 93.75%. This value shows that the system can operate effectively. Thus, it can be recommended for use in diagnosing pregnancy risks.

INDEX TERMS Expert System, Pregnancy Risk Diagnosis, Decision Tree, Dempster Shafer

I. INTRODUCTION

The Maternal Mortality Ratio (MMR) over the past 10 years remains above 300 per 100,000. Currently, it even stands at 350 per 100,000. It shows that the target reduction of MMR, which should reach 183 per 100,000, is still far from being achieved. One of the reasons for this high figure is the delay in recognizing early warning signs during pregnancy [1]. Therefore, in order to address this matter, pregnant women need to have early knowledge about pregnancy risks [2]. Pregnancy risk refers to conditions in pregnant women which leads to death due to complications [3]. Based on the KSPR (Poedji Rochjati Score Card) there are three groups of pregnancy risks that are Low-Risk Pregnancy (KRR), High-Risk Pregnancy (KRT), and Very High-Risk Pregnancy (KRST) [4]. These pregnancy risk groups are determined based on scores obtained from pregnancy risk factors [4]. Meanwhile, each group of pregnancy risk factors includes factors; such as, age, height, and history of cesarean section, categorized under Risk Factor 1; twin pregnancy and breech presentation, categorized under Risk Factor 2; and severe preeclampsia, categorized under Risk Factor 3 [5].

Pregnancy risk can only be determined through examination by healthcare professionals [2]. Furthermore, according to the Ministry of Health Data 2023, 650 community health centers are without doctors, 5,354 are without complete sets of nine types of healthcare workers, and 170 district/city public hospitals lack seven specialist doctor services. Among the seven types of specialist doctor services still lacking, obstetrician-gynecologists rank first with a shortage of 3,941 doctors. Therefore, from this data, it can be concluded that there is a shortage of healthcare providers for pregnant patients [6]. An expert system is needed to address the shortage of healthcare providers by detecting pregnancy risks in pregnant women more quickly and efficiently [2], [1], [6].

Several studies have implemented expert systems in pregnancy cases. One such system is for emergency pregnancy referrals by using Forward Chaining, with an accuracy of 78.4% [7]. Previously, an expert system for diagnosing pregnancy complaints was developed by using Forward Chaining and Bayesian methods [8] with an accuracy of 70%. Therefore, based on the two studies above, it shows that the resulting accuracy is considered to the good category, but it has not yet reached 80%. Forward Chaining is good for fact-finding with forward tracking [9], [10], [11], but it is difficult to recognize one important fact compared to others [6]. The Decision Tree method, which can help simplify existing knowledge [12], is used to overcome the weakness of Forward Chaining so that it constructs the symptom-searching process.

Decision Tree is a decision support method that employs a tree-like hierarchical structure in order to classify categories through a sequence of questions [13]. This method is effective for classification [14], [15] and it can optimize efficiency by reducing redundancy in the knowledge base and improving information integrity [16] so that it can increase accuracy. This statement is supported by research on using Decision Tree methods in sleep stage scoring systems with an accuracy of 80.70% [17] and research protocol ethics reviews with an accuracy reaching 92% [18]. Although Decision Tree is quite good at classification [19], it lacks the weight of confidence scores from an expert.

There are several studies on weighting methods in expert systems; such as, Certainty Factor, Bayesian Network, and Dempster-Shafer [20]. Dempster-Shafer is a mathematical method used in order to integrate independent evidence in decision-making involving uncertainty, which also known as evidence theory, highly renowned and versatile across various applications [21]. This method achieves the highest accuracy compared to the Certainty Factor [22], [23] and Bayesian Network [24]. In addition to higher accuracy levels, the Dempster-Shafer method can implement non-monotonic reasoning, reasoning which can change original rules because of new facts so that it produces valid diagnoses by minimizing uncertainty [25], while the Bayesian Network only relies on existing facts [26]. The Dempster-Shafer method, suitable for pregnancy diagnosis cases, which contains facts or symptoms experienced by pregnant women as considerations in order to determine the level of pregnancy risk; besides, it is considered superior to the Certainty Factor since it only has confidence values and it does not have uncertainty values [27], [28]. The accuracy of the Dempster-Shafer method can be demonstrated through research on vegetable disease diagnosis, which yielded an accuracy of 90% [29].

The combination of Decision Tree and Dempster-Shafer methods can strengthen the diagnosis performed since it is based on rules and it has accurate confidence values [19]. Decision Trees function as rules or support in determining diseases based on symptoms [30]. Meanwhile, Dempster-Shafer determines density values for each known fact based on expert knowledge [31]. The weight values of each fact are then combined, and the final results are the facts with the highest weight. The combination of these two methods has been conducted in cases of lung disease diagnosis, resulting in a matching rate of 83.08% [19].

Based on the problems outlined:

a. A proposed solution is to create an expert system using Decision Tree and Dempster-Shafer methods based on KSPR.
b. This system aims to facilitate a structured approach to analyzing pregnancy risk factors comprehensively.
c. The development of such an expert system is expected to enhance maternal health outcomes by providing pregnant women with accessible tools for independent pregnancy risk diagnosis.
d. Additionally, it can support healthcare professionals in improving the accuracy and efficiency of pregnancy risk detection, potentially reducing Maternal Mortality Ratio (MMR) through proactive intervention and timely medical care.

II. MATERIALS AND METHODS

A. DECISION TREE

The Decision Tree is a decision support method which utilizes a hierarchical structure in the form of a tree in order to classify classes by using a series of questions [13]. One of the algorithms in the Decision Tree is ID3 (Iterative Dichotomizer 3). The ID3 Algorithm is a classification method using a decision tree based on available data [32]. The ID3 algorithm builds decision trees top-down by first examining relevant attributes to be placed at the root, then evaluating all attributes based on statistical calculations; such as, information gain to determine how effective an attribute is in classifying data [33]. Therefore, the ID3 algorithm has the ability to select the most informative attribute based on entropy, enabling it to generate decision trees for effective data classification [34]. The working procedures of the ID3 Decision Tree algorithm are as follow:

1. Calculating the total entropy using Eq. (1)
$$Entropi\ (S) = \sum_{j=1}^{k} -p_j\ \log_2 p_j \quad (1)$$
where S is a set of cases, K is a number of partitions in S, and p_j is probability obtained from the sum of (Yes) divided by Total Cases.
2. Calculating the entropy for each attribute value in order to measure how well a node performs using Eq. (2)
$$Entropi\ (Si) = \sum_{j=1}^{k} -p_j \log_2 p_j, \quad (2)$$
where $Entropy\ (S_i)$ is entropy for samples that have the value, K is a number of partitions in S, and p_j is probability obtained from the sum of (Yes) divided by Total Cases.
3. Calculating the Information Gain of each attribute in order to measure the effectiveness of the attribute in classifying data using Eq. (3)
$$Gain(A) = Entropy(S) - \sum_{i=1}^{k} \frac{|S_i|}{|S|} xEntropy(S_i) (3)$$
where S is sample space in the training process, A is an attribute, $|S_i|$ represents the number of samples for the

value, $|S|$ is the total number of data samples, and $Entropy\ (S_i)$ is entropy for samples that have the value
4. The attribute with the highest gain is used as the root node.
5. Determining the leaf nodes from the root.
6. Repeat steps 1-4. The selected attribute is not included again. Repeat until all data has been used.

B. DEMPSTER SHAFER

The Dempster-Shafer method employs the Belief approach in order to measure the strength of evidence supporting a particular set of propositions. A value of 0 indicates that there is no existing evidence while a value of 1 indicates the presence of definite evidence [35]. Moreover, the Dempster-Shafer method is capable of integrating beliefs from various sources by using different operators in order to generate new beliefs while considering all available evidence, making it suitable for information fusion [36]. In this method, the belief function can be formulated as Eq. (4)

$$Bel\ (X) = \sum_{Y \subseteq X} m(Y), \quad (4)$$

where $Bel\ (X)$ is belief(X) and $m(Y)$ is mass function(Y). Meanwhile, the Plausibility (Pls) value can be formulated as Eq. (5)

$$Pls\ (X) = 1 - Bel\ (X') = 1 - \sum_{Y \subseteq X} m(X'), \quad (5)$$

where $Bel\ (X)$ is belief(X), $Pls\ (X)$ is plausibility(X), $m(X)$ is mass function(X), and $m(Y)$ is mass function(Y).

Plausibility has a range of values between 0 and 1. If there is belief in X', then the value of Belief(X') will become 1, causing the Plausibility(X) value to become 0 based on Eq. (5)

In the Dempster-Shafer theory, the concept of a frame of discernment is symbolized by θ. The frame of discernment represents the entire hypothesis space under discussion in an environment [37]. This concept of the frame of discernment can be formulated as Eq. (6)

$$\theta = \{\theta 1, \theta 2, \dots, \theta n), \quad (6)$$

where θ is FOD or environment and θ 1 ... θ n is elements or components within the environment.

The environment encompasses units which lead to various possible answers, with only one correct answer based on the need. In Dempster-Shafer, these possibilities are known as the power set and are denoted by P(Θ). Each unit in this power set is located within the range between 0 and 1, which can be formulated as Eq. (7) [38]

$$m = P(\Theta) \rightarrow [0,1], \quad (7)$$

where m is a mass function and P(Θ) is the power set. Next, it can be formulated as follows in Eq. (8)

$$\sum_{x \in P(\theta)} m(X) = 1 \approx \sum_{X \in P(\theta)} m(X) = 1, \quad (8)$$

where P(Θ) is the power set and $m(X)$ is a mass function of (X).

In the Dempster-Shafer theory, the mass function (m) represents the level of belief in evidence or the measure of evidence and it is denoted by (m). The formula for calculating the mass function can be seen in the Eq. (9) [39]

$$m1 \oplus m2\ (Z) = \sum_{X \cap Y = Z} m1(X) m2(Y) \quad (9)$$

where $m1 \oplus m2\ (Z)$ is a mass function on evidence (Z), $m1(X)$ is a mass function on evidence (X), $m2(Y)$ is a mass function on evidence (Y) and $\oplus$ is a direct sum operator.

Dempster-Shafer uses Dempster's Rule of Combination in order to combine various pieces of evidence in decision-making, which can then be formulated in Eq. (10) [39]

$$m1 \oplus m2\ (Z) = \frac{\sum_{X \cap Y = Z} m1(X)\ m2(Y)}{1-k} \quad (10)$$

where $m1 \oplus m2\ (Z)$ is a mass function on evidence (Z), k is the number of evidential conflicts. The value of evidential conflict (k) is formulated in the Eq. (11)

$$m1 \oplus m2\ (Z) = \sum_{X \cap Y = \theta} m1(X) m2(Y) \quad (11)$$

where $m1 \oplus m2\ (Z)$ is a mass function on evidence (Z), $m1(X)$ is a mass function on evidence (X), $m2(Y)$ is a mass function on evidence (Y) and $\oplus$ is a direct sum operator. Next, Eq. (11) is substituted into Eq. (10), resulting in Eq. (12) [39]

$$m1 \oplus m2\ (Z) = \frac{\sum_{X \cap Y = Z} m1(X) m2(Y)}{1 - \sum_{X \cap Y = \theta} m1(X) m2(Y)} \quad (12)$$

where $m1 \oplus m2\ (Z)$ is a mass function of evidence (Z), $m1(X)$ is a mass function of evidence (X), $m2(Y)$ is a mass function of evidence Y, and k is a number of evidential conflicts.

C. DESIGN PROCESS

The development of an expert system for diagnosing pregnancy risks is a multi-faceted process that involves several crucial stages, as illustrated in FIGURE 1. This process aims to harness the power of artificial intelligence to provide accurate and reliable diagnoses, ultimately improving maternal health outcomes. The system leverages decision tree algorithms for knowledge representation and the Dempster-Shafer method for managing uncertainty in diagnosis.

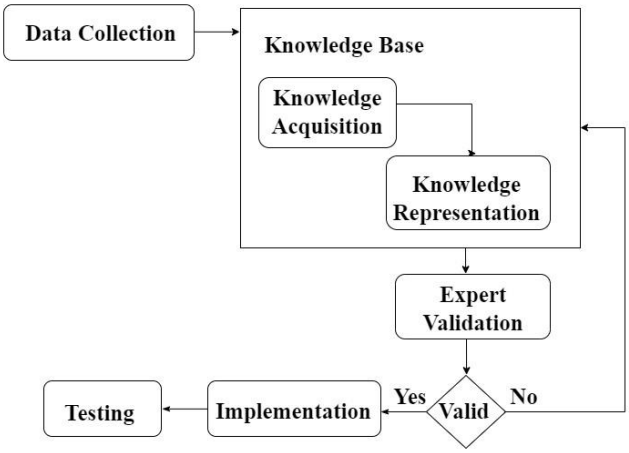

FIGURE 1. Expert System Development Process

The development of an expert system begins with data collection, which is essential for forming the knowledge base. This involves processes of knowledge acquisition and knowledge representation. Once the knowledge base is established, it undergoes expert validation. If the knowledge base is validated by experts, the process moves to the implementation and testing stages. However, if the knowledge base is not validated, the process cycles back to the knowledge base formation until it is deemed valid by the experts. For

details of each stage, it can be seen as follows. Data Collection aims to acquire knowledge from journals, books, experts, and/or datasets which support the research topic [31], [40]. In this study, there are several ways to collect data. Dataset Collection is conducted by searching various datasets relevant to the topic online. Relevance means the dataset has columns suitable for use as parameters in determining the diagnosis of pregnancy risk levels with adequate data. From this dataset collection process, a relevant dataset was found for the research topic to be taken, namely Maternal Health Risk Data sourced from Kaggle with the following link [41]: https://www.kaggle.com/datasets/csafrit2/maternal-health-risk-data

This dataset was obtained from data collection at several hospitals, clinic communities, and maternal health care through an IoT-based risk monitoring system with 1014 data. A literature review collects materials and information relevant to the research topic to be studied. The literature review is conducted at this stage by gathering and reading references; such as, articles, journals, and books related to understanding KSPR [42]. The field study in this study involves conducting interviews with three experts from midwife to obtain data and information. By directly interviewing these experts, the insights gained are expected to be more comprehensive and valid [43], [44], [45]. The insights referred to here are critical aspects related to pregnancy diagnosis. The experts will contribute in the form of practical knowledge and clinical experience. Knowledge acquisition is obtaining and incorporating human knowledge into a system [46], [47]. In this study, the first step in knowledge acquisition is processing the obtained dataset by using the ID3 Decision Tree algorithm in Jupyter Notebook with the Python programming language in order to generate a decision tree containing symptoms and rules for diagnosing pregnancy risks.

The second step is to process and combine the information obtained from experts and books regarding KSPR with the decision tree which has been created. It is conducted to improve the quality of the decision tree since the symptoms and rules become more comprehensive. Knowledge Representation is the representation of knowledge in computer systems or models which enable computers to understand, store, organize, and manipulate knowledge [9], [47].

Based on FIGURE 2, the processed data would be represented as a decision tree containing rules for determining the level of pregnancy risk. Experts would then assign the created rules weights as part of implementing the Dempster-Shafer method in order to measure the confidence level of a diagnosis. Weighting is conducted by taking the mode value from the three experts. Mode is the most frequent data or the data with the highest occurrence within a dataset. There are various types of mode:

a. Single mode, which occurs when there is only one data point with the highest frequency, referred to as the "capital union".
b. Double mode, which occurs when there are two data points with high frequencies, known as "bi capital".
c. Multi-mode, identifies within a dataset where there are more than two modes, termed as "multimodal".

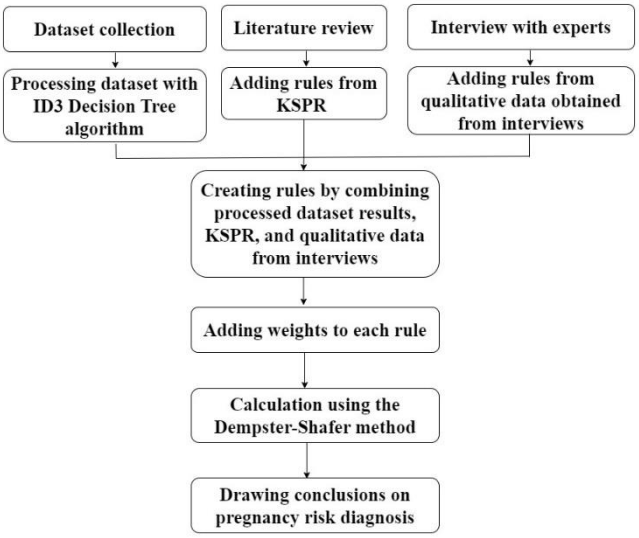

FIGURE 2. Knowledge Representation Diagram

When dealing with grouped data, it is expressed as Eq. (13) [48]

$$M_0 = Tb + \left(\frac{d_1}{d_1+d_2}\right).i \quad (13)$$

where M_0 is mode, Tb is bottom edge of mode class, d_1 is difference in mode class frequency with the previous frequency, d_2 is difference in mode class frequency with the frequency afterward, and i is the length of interval class. In this stage, experts validate that the rules and weights that have been constructed are conducted. Experts would review whether the rules and weights result in appropriate pregnancy risk diagnoses. In the implementation process, the Decision Tree method is used as the knowledge representation method acquired, consisting of a series of rules, namely pregnancy risk symptoms, in order to determine the level of pregnancy risk. Subsequently, experts would assign these symptom weights as confidence values.

When the system receives pregnancy risk symptoms experienced by the user, the system will calculate the confidence value by using the Dempster-Shafer method and provide output in the form of a prediction of the level of pregnancy risk experienced by the user. The development of the expert system would be implemented in the form of a website by using the Laravel framework with the PHP programming language and MySQL database storage. The developed expert system has two roles: admin and user. The user role has access rights to conduct consultations, view and download bookmarks (consultation history), and view guidelines and available articles. Meanwhile, the admin role has all the user's access rights and several other access rights; such as, CRUD (Create, Read, Update, Delete) features for articles, symptoms, diagnoses, and knowledge bases, read user consultation history feature and edit user role feature.

The expert system which had been developed would be tested on 16 patients offline on March 13th and 16th 2024, at Gatak Public Health Center, Sukoharjo. The testing process aims to evaluate the performance and reliability of the system [6], [49] in providing pregnancy diagnoses.

III. RESULT

There will be a detailed presentation regarding the design of the expert system for pregnancy risk diagnosis from the data collection stage, data processing (knowledge acquisition), representing the results of data processing into a knowledge base, which will then be implemented and tested in order to determine the expert system performance which has been developed.

A. DATA COLLECTION

The three sources of data collected by the authors are as follows;

1) KAGGLE DATASET

From the Kaggle dataset available at [41], various pieces of information were extracted to serve as benchmarks for determining pregnancy risk levels. The dataset columns include essential parameters that contribute to the comprehensive analysis of maternal health risks. These parameters are classified based on internet sources and expert validation, as detailed in APPENDIX A.

Age groups are categorized into three brackets: less than or equal to 16 years, 17-34 years, and greater than or equal to 35 years, each associated with unique health considerations. Teenage pregnancies, due to physiological immaturity and social factors, pose heightened risks, while advanced maternal age increases the likelihood of complications like gestational diabetes and hypertension. Body temperature, categorized as less than 36.0°C, 36.0-37.5°C, and greater than or equal to 37.5°C, serves as a vital indicator for detecting infections and inflammation affecting maternal health during pregnancy. Elevated temperatures often signify underlying infections necessitating timely medical intervention to prevent adverse outcomes.

Blood glucose levels, categorized as less than 140 mg/dL, 140-200 mg/dL, and greater than 200 mg/dL, are crucial in managing gestational diabetes mellitus, a common complication. Elevated glucose levels can lead to conditions like macrosomia and preterm birth, emphasizing the importance of proper management through diet, exercise, or medication. Diastolic blood pressure, classified as less than 80 mmHg, 80-90 mmHg, and greater than 90 mmHg, poses risks such as preeclampsia and eclampsia when elevated, underscoring the need for vigilant monitoring and management to safeguard maternal and fetal well-being.

Similarly, systolic blood pressure, categorized into less than 130 mmHg, 130-150 mmHg, and greater than 150 mmHg, indicates hypertension, associated with increased risks of cardiovascular complications and adverse pregnancy outcomes. Regular monitoring and appropriate medical interventions are essential to manage hypertension effectively and optimize maternal health outcomes. Overall, the Kaggle dataset supports evidence-based decision-making in maternal healthcare by integrating expert-validated classifications, thereby enhancing understanding and prediction of pregnancy risks based on quantitative data. The Kaggle dataset provides a valuable resource for understanding and predicting pregnancy risks based on quantitative data. It facilitates evidence-based decision-making in maternal healthcare by integrating comprehensive datasets with expert-validated classifications.

2) POEDJI ROCHJATI SCORE CARD

Here are some statements presented in the Poedji Rochjati Score Card, which serves as one of the benchmarks in determining pregnancy risk level, as seen in APPENDIX B.

Teenage pregnancy poses higher risks due to the physical and psychological immaturity necessary to manage pregnancy and childbirth effectively. Marrying at an advanced age (≥ 35 years old) with a marriage duration of ≥ 4 years can increase the risk of pregnancy complications such as hormonal disorders and decreased fertility. Advanced maternal age (≥ 35 years old) for first-time pregnancies elevates the risks of complications like preeclampsia and other pregnancy-related issues.

Too close consecutive pregnancies spaced ≤ 2 years apart increase the risk of premature delivery and maternal health complications. Having ≥ 4 children heightens the risk of multiple pregnancies, placental disorders, and other complications during pregnancy and childbirth. Advanced maternal age (≥ 35 years old) is often associated with higher risks of complications such as gestational diabetes and other maternal health issues. Women with shorter stature (≤ 145 cm) face additional risks of complications such as intrauterine growth restriction and difficulties during childbirth. Previous miscarriages increase the risk of recurrent miscarriages during subsequent pregnancies.

Delivering with a breech presentation or previous delivery assisted with forceps/vacuum can increase the risk of difficult labor and potential injuries to both mother and baby. Conditions such as manually removing the placenta or using infusions/transfusions during childbirth increase the risk of infections and blood loss. A history of cesarean section surgeries can increase the risk of complications like placenta accreta in subsequent pregnancies. Concurrent chronic diseases during pregnancy, such as diabetes or heart disease, increase the risk of complications for both mother and baby. Excessive oedema and high blood pressure during pregnancy indicate preeclampsia, requiring strict medical monitoring.

Twin pregnancies increase the risk of complications such as miscarriage, premature birth, and maternal health problems. Hydramnios (excessive amniotic fluid) and fetal demise within the uterus are critical conditions requiring immediate medical attention. Pregnancy beyond the estimated due date, breech or transverse fetal positions, and bleeding during pregnancy are signs of serious complications necessitating prompt medical intervention. Preeclampsia and eclampsia are

severe conditions affecting maternal blood pressure during pregnancy, posing dangerous complications for both mother and baby.

APPENDIX B provides a comprehensive collection of symptomatic data crucial for identifying and evaluating pregnancy risks. This information aids in evidence-based clinical decision-making to ensure optimal maternal health during pregnancy and childbirth.

3) QUALITATIVE DATA

Qualitative data was obtained from interviews with experts regarding pregnancy risk symptoms beyond the Kaggle dataset and Poedji Rochjati Score Card statements. Here is the list of symptoms obtained, as seen in APPENDIX C.

Having sexually transmitted diseases during pregnancy can lead to complications such as preterm birth and neonatal infections. Positive urine levels may indicate conditions like urinary tract infections, which, if untreated, can escalate to more severe health issues. Conversely, negative urine levels signify the absence of certain urinary conditions or infections that could affect pregnancy. A smaller upper arm circumference less than 23.5 cm may indicate malnutrition or inadequate maternal health, influencing fetal development and birth outcomes. Low hemoglobin (HB) levels below 8 suggest anemia, which poses risks of maternal fatigue, low birth weight, and other complications. Moderate HB levels between 8 and 10 still indicate the risk of anemia, necessitating monitoring and potential intervention to prevent complications. HB levels around 11 are generally considered normal, indicating healthy oxygen transport capacity for both mother and baby.

Different levels of preeclampsia severity, including low, moderate, and severe stages, indicate varying risks of hypertension and organ damage during pregnancy. Asthma or other respiratory issues can complicate pregnancy, requiring specialized care to manage symptoms and ensure maternal and fetal well-being. Stages of pre-eclampsia, from not experiencing it to mild and severe stages (including eclampsia), impact maternal blood pressure regulation and can lead to eclamptic seizures. Viral infections during pregnancy, such as Rubella, Toxoplasma, or Chickenpox, can affect fetal development and health, necessitating careful management and monitoring.

HIV or other sexually transmitted infections (STIs) pose significant risks to both maternal health and vertical transmission to the fetus. Bleeding during pregnancy can indicate various complications, from minor issues to serious conditions requiring medical intervention. Oedema, or swelling, during pregnancy can be a sign of preeclampsia or other underlying health issues requiring monitoring and management. Previous obstetric history, including delivery with breech presentation, forceps/vacuum, miscarriages, cesarean sections, hydramnios, or twin pregnancies, influences current pregnancy risks, necessitating tailored care and monitoring.

Managing asthma during pregnancy is crucial to prevent exacerbations and ensure adequate oxygen supply to both mother and fetus. Pre-existing maternal health conditions, such as anemia, malaria, tuberculosis, heart disease, and diabetes, can exacerbate during pregnancy, requiring coordinated care and management. Prolonged pregnancy duration beyond the normal gestational period may indicate post-term pregnancy, necessitating careful monitoring to prevent complications.

B. KNOWLEDGE ACQUISITION

Data processing was conducted by combining all the symptoms which were collected. The entire symptom data can be seen in detail in APPENDIX D. Starting with age, the categorization includes three groups: G01 represents pregnant women aged 16 years or younger, G02 denotes those aged 17-34 years, and G03 encompasses women aged 35 years or older. This classification acknowledges the varying implications of age on maternal health and pregnancy outcomes.

Marriage age is divided into G04 for women married before the age of 4 years, and G05 for those married at or after 4 years of age. Early marriage can impact maternal health due to increased physical and emotional immaturity, while later marriage often correlates with greater readiness for childbearing. The number of children a woman has previously borne is categorized into G06 for those without children, G07 for women with 1-3 children, and G08 for those with 4 or more children. This categorization recognizes the potential influence of parity on maternal health, including risks associated with multiple pregnancies. Pregnancy spacing, categorized as G09 (spacing less than or equal to 2 years) and G10 (more than 2 years), addresses the health implications of short versus optimal intervals between pregnancies on maternal and fetal well-being.

Physical characteristics such as height and upper arm circumference are categorized into G11/G12 and G13/G14 respectively, considering their impact on nutritional status and overall health during pregnancy. Weight gain during pregnancy, categorized into G15 (7-12 kg), G16 (12-15 kg), and G17 (more than 15 kg), reflects variations in nutritional adequacy and metabolic health during gestation. Body temperature (G18-G20), hemoglobin levels (G21-G23), and blood pressure readings (G24-G29) provide crucial indicators of maternal physiological status during pregnancy, informing risk assessment and management.

Laboratory findings, such as glucose levels (G30-G32) and urine protein levels (G33-G34), help screen for conditions like gestational diabetes and preeclampsia, respectively. Fetal presentation (G35-G37), maternal complications (G38-G44), obstetric history (G45-G50), and specific health conditions (G51-G55) further delineate the comprehensive profile used to assess pregnancy risk and guide clinical decision-making. Each category in APPENDIX D represents a key aspect of maternal health and pregnancy outcomes, integrating both quantitative data and clinical observations to enhance the accuracy of risk assessment and optimize maternal and fetal health outcomes.

C. KNOWLEDGE REPRESENTATION

1) PREGNANCY RISK LEVEL DIAGNOSIS RULE BASE

The following is the rule base visualized as a Decision Tree to determine the pregnancy risk diagnosis based on the symptoms experienced. This visualization can be seen in FIGURE 3.

2) DEMPSTER SHAFER WEIGHTING

At this stage, weights are assigned to each pregnancy risk symptom using the Dempster-Shafer method. The weight values are determined by the three experts by using the mode value. Meanwhile, for symptoms which do not have a mode value, validation will be conducted again with the experts, considering the consistency between the rules created and the diagnosis results from the system. Here are the results of filling in the weights of pregnancy risk symptoms, as shown in TABLE 1.

TABLE 1
Weight Values of Pregnancy Risk Symptoms

No	Symptom	Weight		
		KRR	KRT	KRST
1	Age of pregnant women less than or equal to 16 years		0.7	
2	Age of pregnant women 17-34 years	0.6	0.2	0.2
3	Age of pregnant women more than or equal to 35 years	0.2	0.4	0.4
4	Marriage age less than 4 years	0.6	0.2	0.2
5	Marriage age more than or equal to 4 years		0.6	0.2
6	Having no children	0.4	0.4	0.2
7	Having 1-3 children	0.2	0.2	0.2
8	Having 4 or more children		0.8	0.2
9	Pregnancy spacing less than or equal to 2 years		0.7	0.3
10	Pregnancy spacing of more than 2 years	0.6	0.2	0.2
11	Height less than 145 cm		0.7	
12	Height more than or equal to 145 cm	0.6		
13	Upper arm circumference less than 23.5 cm		0.7	0.4
14	Upper arm circumference more than or equal to 23.5 cm	0.6		
15	Weight gain 7-12 kg	0.6	0.2	0.2
16	Weight gain 12-15 kg		0.6	0.7
17	Weight gain more than 15 kg		0.6	0.8
18	Body temperature less than 36 °C	0.2	0.6	0.2
19	Body temperature 36-37.5°C	0.6	0.2	0.2
20	Body temperature more than 37.5°C		0.4	0.6
21	Hemoglobin (HB) value less than 8 grams/dL		0.6	0.4
22	HB value 8-10 grams/dL		0.6	0.4
23	HB value 11 grams/dL	0.6		
24	Systolic blood pressure less than 130 mmHg	0.6	0.2	0.2
25	Systolic blood pressure 130-150 mmHg	0.6	0.2	0.2
26	Systolic blood pressure more than 150 mmHg		0.6	0.4
27	Diastolic blood pressure less than 80 mmHg		0.5	
28	Diastolic blood pressure 80-90 mmHg		0.4	
29	Diastolic blood pressure more than 90 mmHg			0.9
30	Glucose level less than 140 mg/dL	0.6		
31	Glucose level 140-200 mg/dL	0.6	0.4	
32	Glucose level more than 200 mg/dL		0.6	0.4
33	Positive urine protein (+)		0.4	1
34	Negative urine protein (-)	0.6	0.4	
35	Breech baby position		0.7	0.9
36	Transverse baby position		0.7	0.9
37	Normal baby position	0.6	0.2	0.2
38	Non-pre-eclampsia	0.2	0.6	0.2
39	Mild pre-eclampsia		0.6	
40	Severe pre-eclampsia (eclampsia)			1
41	Experiencing intrauterine fetal death		0.6	0.4
42	Infected with viruses (Rubella, Toxoplasma, Chickenpox, etc.)		0.4	0.6
43	Manual placenta		0.6	0.9
44	Experiencing HIV/STIs (Sexually Transmitted Infections)		0.7	0.9
45	Experiencing bleeding during pregnancy		0.7	0.9
46	Experiencing oedema		0.7	0.9
47	History of delivery with breech presentation		0.7	0.9
48	History of delivery with forceps/vacuum		0.7	0.9
49	Miscarriage history		0.7	
50	Cesarean section surgery history		0.7	0.8
51	Hydramnios		0.7	
52	Twin pregnancy		0.7	0.9
53	Having asthma		0.8	0.7
54	Having pregnancy-related diseases (anemia, malaria, TB, heart disease, diabetes)		0.7	0.8
55	Pregnancy duration more than 9 months		0.6	0.4

D. IMPLEMENTATION

The implementation phase will provide a detailed explanation of comparing the calculation results conducted by using Dempster-Shafer systematically and manually then observe whether the diagnosis results are consistent with the decision tree which had been created. Here is the implementation using

manual calculation in pregnant woman case with non-pre-eclampsia symptoms, glucose levels of 140-200 mg/dL, and systolic blood pressure exceeding 150 mmHg.

Symptom 38: Non-pre-eclampsia

m1{KRR, KRT, KRST} = 0,6

⊖ = 1-0,6

= 0,4

Symptom 31: Glucose levels of 140-200 mg/dL

m2{KRR, KRT} = 0,6

⊖ = 1-0,6

= 0,4

TABLE 2
Combination Rule M3

m1		m2			
		{KRR, KRT}	{0,6}	⊖	(0,4)
{KRR, KRT, KRST}	(0,6)	{KRR, KRT}	(0,36)	{KRR, KRT, KRST}	(0,24)
⊖	(0,4)	{KRR, KRT}	(0,24)	⊖	(0,16)

With the occurrence of two symptoms that are not experiencing pre-eclampsia and glucose levels of 140-200 mg/dL, the next step is calculating a new density for several combinations (m3). A table is used to simplify the calculation process by placing the formed subset into it. The first column refers to the first symptom (m1) while the first row refers to the second symptom (m2). In this way, we can find the value of m3 as a result of the combination between m1 and m2, which is then shown in TABLE 2.

The calculated density values of the m3 combination show that the {KRR, KRT} combination has a higher level than the other symptoms, with a density of 0,6. Furthermore, when an additional symptom exists, namely systolic blood pressure exceeding 150 mmHg (m4 {KRR, KRT, KRST}), the next step is to calculate a new density, m5, which can be seen in TABLE 3.

m4{KRT, KRST} =0,6

TABLE 3
Combination Rule M5

m3		m4			
		{KRT, KRST}	(0,6)	⊖	(0,4)
{KRR, KRT}	(0,6)	{KRT}	(0,36)	{KRR, KRT}	(0,24)
{KRR, KRT, KRST}	(0,24)	{KRT, KRST}	(0,144)	{KRR, KRT, KRST}	(0,096)
⊖	(0,16)	{KRT, KRST}	(0,096)	⊖	(0,064)

From the calculation results using the Dempster-Shafer method, the highest density value is 0.36; it can be concluded that the diagnosis experienced by the patient is KRT. Furthermore, the system implementation results of the patient's case can be seen in FIGURE 4 and FIGURE 5.

FIGURE 4. Patient Consultation Result Page

FIGURE 5. Patient Consultation Page

Based on FIGURE 5, it shows that there is a consistency between the calculation results from the system and the manual calculation, which resulted in a confidence value of 36% with the diagnosis result of KRT. Then, it can be checked on the Decision Tree shown in FIGURE 3; patients with the symptoms mentioned earlier also have a diagnosis of KRT, which is in line with the results presented by the system.

E. TESTING

The testing phase is necessary in order to evaluate the system's performance in providing diagnostic results regarding the level of pregnancy risk experienced by patients. From the testing process conducted on 16 patients, 15 data were validated correctly, and 1 data was not. Furthermore, based on the test results, the accuracy value can be obtained as follows [50]:

$$Accuracy = \frac{The\ accurate\ data\ count}{The\ total\ data\ count} \times 100\% \quad (1)$$

Therefore, based on the testing process conducted on the expert system by implementing the Decision Tree and Dempster Shafer methods, a fairly good accuracy value of 93.75% is obtained.

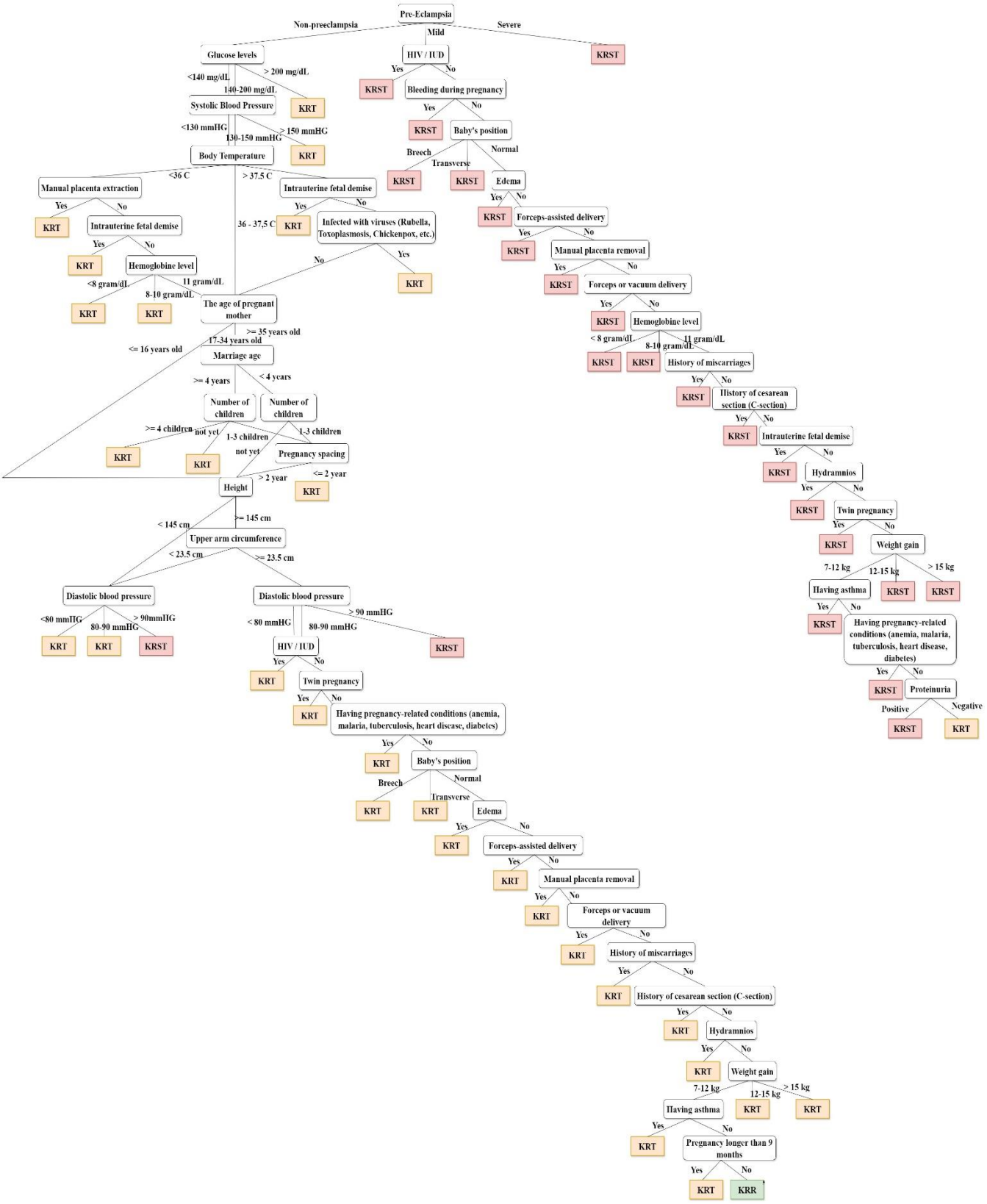

FIGURE 3. **Knowledge representation in a Decision Tree**

IV. DISCUSSION

Based on the results of the study, the expert system which has been developed is capable of diagnosing pregnancy risks effectively, with an accuracy rate reaching 93.75%. Out of a total of 16 test cases, only 1 test case is not correctly validated. The identification error occurred due to the patient's input of less specific symptoms so that it affects the weighting calculation in Dempster-Shafer overall. When the selected symptoms are not relevant or do not reflect important aspects of a particular diagnosis, the obtained results become less valid.

Based on the Decision Tree in FIGURE 3, when a symptom shows a history of cesarean section, there are 2 symptom inputs from the patient outside the symptom node leading to the diagnosis of KRT: HB value 11 grams/dL and negative urine protein (-) symptom. Both of these inputs impact the Dempster-Shafer computation process in generating a diagnosis so that the expected diagnosis of KRT appeared as KRR. If all testing data is tested by using only the Decision Tree which has been created, the results obtained are 100% valid based on the expert's decision.

Furthermore, compared to previous research [19], this study combines data collection methods from 3 sources: dataset, literature review, and field study. Additionally, this study includes a weighting process from the three experts, which takes the mode value. For symptoms that do not have a mode value, further validation will be conducted by consulting with the expert to ensure compatibility between the rules that have been created. Thus, Dempster-Shafer weighting is indirectly influenced by the created decision tree, making it more effective.

This is an innovation since, in previous studies, the weighting in Dempster-Shafer was purely derived from experts. For instance, in [27], the expert system developed relied solely on expert opinions for assigning weights to the evidence. Similarly, in [28], the system's diagnosis process was based entirely on the expertise of plant pathologists. In another study [31], the expert system to identify malaria types also relied exclusively on doctors for its weighting mechanism. In contrast, the current study advances the methodology by integrating diverse data sources and employing a robust validation process. This comprehensive approach enhances the effectiveness and reliability of the system. Furthermore, in this study, the system implementation is limited to Community Health Centers, with the system being utilized solely as an adjunct while examinations remain the responsibility of midwives.

Comparison with related studies underscores the system's advancements and limitations. While achieving high accuracy, weaknesses persist, primarily in symptom variability and expert validation protocols. Addressing these limitations is crucial for enhancing diagnostic precision. Moreover, implications of this study are significant. It underscores the potential for further research into refining Dempster-Shafer and Decision Tree methodologies across various medical contexts. Additionally, there is an opportunity to explore alternative expert system approaches to align outputs more closely with clinical decisions.

Furthermore, this study opens avenues for future research to explore and refine the Dempster-Shafer method and Decision Trees in broader applications. It also encourages the development of expert systems that can accurately mirror expert decisions in diagnostic settings. This research contributes to advancing both theoretical understanding and practical applications in maternal health diagnostics.

V. CONCLUSION

Based on the results of the study, it can be concluded that an expert system for pregnancy risk diagnosis by using Decision Tree and Dempster Shafer methods has been successfully implemented with an accuracy rate of 93.75%. This value is obtained from testing 16 test cases, where 15 are correctly verified, while 1 is not verified correctly. It shows that the expert system is not yet working perfectly, but it can be recommended for conducting pregnancy risk diagnosis well, considering the high generated accuracy value.

Furthermore, this study opens up opportunities for further research in order to continue exploring the Dempster-Shafer method and Decision Trees in other cases. In addition, there are opportunities to further develop other expert system methods so that the system outputs accurately reflect experts' decisions. It is expected that there will be the development of expert systems based on Android/iOS platforms for easier and broader access for users and enable integration with additional features which can enhance the effectiveness and accessibility of healthcare services.

REFERENCE

[1] E. F. Wati, E. Sunita Perangin-Angin, and B. Sudrajat, "Pregnancy Disease Diagnostic Expert System With Certainty Factor Method," *International Journal of Information System & Technology Akreditasi*, vol. 6, no. 158, pp. 754–760, 2023, doi: https://doi.org/10.30645/.v6i6.292.

[2] E. F. Wati and B. Sudrajat, "Application of Naive Bayes Method For Diagnosis of Pregnancy Disease," *International Journal of Information System & Technology Akreditasi*, vol. 6, no. 158, pp. 93–100, 2022, doi: https://doi.org/10.30645/ijistech.v6i1.216.

[3] R. Ilbert and D. Marfuah, "Depressive Symptoms in High-Risk Pregnancies During the COVID-19 Pandemic," *The International Virtual Conference on Nursing (IVCN)*, Feb. 2022, doi: 10.18502/kls.v7i2.10339.

[4] F. T. Anggraeny, F. Muttaqin, and M. Syahrul Munir, "Modeled early detection of pregnancy risk based on Poedji Rochjati score card using relief and neural network," *Proceeding International Joint Conference on Science and Technology (IJST)*, 2017, Accessed: Apr. 13, 2024. [Online]. Available: https://ojs.pnb.ac.id/index.php/Proceedings/article/view/920

[5] I. Putu, K. Aryadhi, I. Made Darmayasa, M. Bagus, and D. Aryana, "Characteristics of Risk Factor for Obstetric Emergency Cases Referred to Sanglah General Hospital," *International Journal of Science and Research*, 2018, doi: 10.21275/SR201006200628.

[6] G. Aditia, A. Ziqri, A. T. Maulana, and F. D. Adhinata, "Expert System for Identifying Pregnant Using Forward Chaining," *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, vol. 5, no. 2, pp. 136–143, May 2023, doi: 10.20895/inista.v5i2.494.

[7] S. Puspitaningsih, Suryono, and Farikhin, "Design and Implementation of an Emergency Pregnancy Referral System Using Rule-Based Expert System Forward Chaining Method," in *2021 6th*

International Conference on Informatics and Computing, ICIC 2021, 2021. doi: 10.1109/ICIC54025.2021.9632953.

[8] E. F. Wati and A. Puspitasari, "Expert System for Diagnosing Pregnancy Complaints by Forward Chaining," *SinkrOn*, vol. 5, no. 1, p. 7, Oct. 2020, doi: 10.33395/sinkron.v5i1.10593.

[9] N. Kapoor and N. Bahl, "Comparative Study Of Forward And Backward Chaining In Artificial Intelligence," *International Journal Of Engineering And Computer Science*, Apr. 2016, doi: 10.18535/Ijecs/v5i4.32.

[10] M. R. Anwar, "Analysis of Expert System Implementation in Computer Damage Diagnosis with Forward Chaining Method," *International Transactions on Artificial Intelligence (ITALIC)*, vol. 1, no. 2, pp. 139–152, 2023.

[11] M. Indriani, F. Anggraini, D. Hartama, and M. Ridwan Lubis, "Expert System Diagnosing Damage to Canon Ir5000 Copier With Forward Chaining Method," *Journal of Artificial Intelligence and Engineering Applications*, vol. 1, no. 2, 2022, [Online]. Available: https://ioinformatic.org/

[12] A. D. S. Nicolau, J. P. D. S. C. Augusto, and R. Schirru, "Accident diagnosis system based on real-time decision tree expert system," *AIP Conf Proc*, vol. 1836, Jun. 2017, doi: 10.1063/1.4981957.

[13] I. M. K. Karo, M. Y. Fajari, N. U. Fadhilah, and W. Y. Wardani, "Benchmarking Naïve Bayes and ID3 Algorithm for Prediction Student Scholarship," *IOP Conf Ser Mater Sci Eng*, vol. 1232, no. 1, p. 012002, Mar. 2022, doi: 10.1088/1757-899x/1232/1/012002.

[14] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

[15] Priyanka and D. Kumar, "Decision Tree Classifier: A Detailed Survey," *International Journal of Information and Decision Sciences*, vol. 12, no. 3, pp. 246–269, Apr. 2020.

[16] A. Li and X. Zhang, "Fault Diagnosis Knowledge Base Optimization Based on Decision Tree," *2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, pp. 324–327, 2018, doi: 10.1109/IHMSC.2018.00081.

[17] K. M. Gunnarsdottir, C. E. Gamaldo, R. M. E. Salas, J. B. Ewen, R. P. Allen, and S. V. Sarma, "A Novel Sleep Stage Scoring System: Combining Expert-Based Rules with a Decision Tree Classifier," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Institute of Electrical and Electronics Engineers Inc., Oct. 2018, pp. 3240–3243. doi: 10.1109/EMBC.2018.8513039.

[18] R. N. E. Anggraini, N. F. Ariyani, A. F. Septiyanto, R. Sarno, Z. D. Meilani, and T. Soendoro, "A Decision Tree Knowledge-based System for Reviewing Research Ethics Protocol," in *Proceeding - IEEE International Conference on Communication, Networks and Satellite, COMNETSAT 2022*, 2022, pp. 50–55. doi: 10.1109/COMNETSAT56033.2022.9994414.

[19] A. M. Alfatah, R. Arifudin, and M. A. Muslim, "Implementation of Decision Tree and Dempster Shafer on Expert System for Lung Disease Diagnosis," *Scientific Journal of Informatics*, vol. 5, no. 1, p. 57, May 2018, doi: 10.15294/sji.v5i1.13440.

[20] N. Nurdin, E. Susanti, H. A.-K. Aidilof, and D. Priyanto, "Comparison of Naive Bayes and Dempster Shafer Methods in Expert System for Early Diagnosis of COVID-19," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 1, pp. 215–228, Nov. 2022, doi: 10.30812/matrik.v22i1.2280.

[21] A. Anggrawan, H. Hairani, C. Satria, and A. D. Dayani, "Diagnosing Learning Disorders in Children: A Comparison of Certainty Factor and Dempster-Shafer Methods," *International Journal of Information and Education Technology*, vol. 13, no. 9, pp. 1422–1429, Sep. 2023, doi: 10.18178/ijiet.2023.13.9.1945.

[22] E. K. Panggabean, "Comparative Analysis Of Dempster Shafer Method With Certainty Factor Method For Diagnose Stroke Diseases," *International Journal of Artificial Intelligence Research*, vol. 2, no. 1, p. 32, Mar. 2018, doi: 10.29099/ijair.v2i1.53.

[23] N. Astuti Triana, E. Maya Safitri, A. Wulansari, and S. Fitri Ana Wati, "Accuracy of Kidney Disease Expert System Based On Certainty Factor and Dempster Shafer Algorithm," *IJCONSIST JOURNAL*, vol. 3, no. 2, pp. 19–24, 2022, [Online]. Available: https://doi.org/10/33005/ijconsist.v3i2.66

[24] K. Adhianto, R. Andrian, and Y. Nuhricha Sari, "Comparative Analysis of Cow Disease Diagnosis Expert System using Bayesian Network and Dempster-Shafer Method," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 10, no. 4, 2019, doi: 10.14569/IJACSA.2019.0100427.

[25] Septiana, N. Yusliani, and K. J. Miraswan, "Expert System To Diagnose Disease In Toddlers Using Dempster Shafer Method," *Sriwijaya Journal of Informatic and Applications*, vol. 2, no. 2, pp. 75–80, Jul. 2021.

[26] M. S. A. Sawa, A. Naba, and H. S. Dahlan, "Bayesian Network Expert System for Early Diagnosis of Heart Diseases," *EECCIS*, vol. 7, no. 2, 2014, Accessed: Apr. 13, 2024. [Online]. Available: https://jurnaleeccis.ub.ac.id/index.php/eeccis/article/download/222/196

[27] M. Nugraheni, R. Nuraini, M. Tonggiroh, and S. Nurhayati, "Expert System for Diagnosing Learning Disorders in Children Using the Dempster-Shafer Theory Approach," *sinkron*, vol. 8, no. 4, pp. 2513–2525, Oct. 2023, doi: 10.33395/sinkron.v8i4.12960.

[28] W. Ilham, S. E. Putra, D. Aldo, Y. S. R. Nur, and A. Utami, "Expert System for Disease Identification in Palawija Plants with the Dempster Shafer Method," *Jurnal Ecotipe (Electronic, Control, Telecommunication, Information, and Power Engineering)*, vol. 10, no. 2, pp. 258–267, Oct. 2023, doi: 10.33019/jurnalecotipe.v10i2.4418.

[29] I. Parapat Hotel -Medan, E. Astuti, N. E. Saragih, N. Sribina, and R. Ramadhani, "Dempster-Shafer Method for Diagnose Diseases on Vegetable," *The 6 th International Conference on Cyber and IT Service Management (CITSM 2018)*, 2018, doi: http://dx.doi.org/10.1109/CITSM.2018.8674377.

[30] D. Chrimes, "Using Decision Trees as an Expert System for Clinical Decision Support for COVID-19," *Interact J Med Res*, vol. 12, p. e42540, Jan. 2023, doi: 10.2196/42540.

[31] R. M. Marusaha, D. Noviandri, and A. H. Lubis, "The Expert System of Determining the Type of Malaria by using Dempster-Shafer Method," *Bulletin of Information Technology (BIT)*, vol. 4, no. 2, pp. 377–383, 2023, doi: 10.47065/bit.v3i1.

[32] Y. Li, W. Xu, W. Li, A. Li, and Z. Liu, "Research on hybrid intrusion detection method based on the ADASYN and ID3 algorithms," *Mathematical Biosciences and Engineering*, vol. 19, no. 2, pp. 2030–2042, 2021, doi: 10.3934/MBE.2022095.

[33] B. Mahbooba, M. Timilsina, R. Sahal, and M. Serrano, "Explainable Artificial Intelligence (XAI) to Enhance Trust Management in Intrusion Detection Systems Using Decision Tree Model," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6634811.

[34] F. Hafidh, M. Arsyad Al Banjari Banjarmasin, and S. Kalimantan Indonesia, "Enhancing Special Needs Identification for Children: A Comparative Study on Classification Methods Using ID3 Algorithm and Alternative Approaches," *Journal of Engineering, Electrical and Informatics*, vol. 3, no. 2, 2023, doi: 10.55606/jeei.v3i1.1468.

[35] Veronika, S. Ramadhani, and I. M. Khamidah, "Expert System for Diagnosis Papaya Plant Disease with Backward Chaining and Dempster Shafer," *TEPIAN*, vol. 3, no. 2, pp. 92–100, Jun. 2022, doi: 10.51967/tepian.v3i2.767.

[36] M. Huang and C. Zhang, "A Novel Multi-Source Domain Adaptation Method with Dempster–Shafer Evidence Theory for Cross-Domain Classification," *Mathematics*, vol. 10, no. 15, Aug. 2022, doi: 10.3390/math10152797.

[37] A. Setiawan, S. N. Fauzia, K. Kusmaya, K. S. Haryana, I. Abadi, and E. Yulianto, "Expert System for Diagnosing Disease Symptoms of Rice Pests Using the Dempster Shafer Algorithm and Fuzzy Tsukamoto Algorithm," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 11, no. 3, pp. 407–414, Dec. 2022, doi: 10.32736/sisfokom.v11i3.1425.

[38] S. I. Sezer, G. Elidolu, E. Akyuz, and O. Arslan, "An integrated risk assessment modelling for cargo manifold process on tanker ships under FMECA extended Dempster–Shafer theory and rule-based Bayesian network approach," *Process Safety and Environmental Protection*, vol. 174, pp. 340–352, Jun. 2023, doi: 10.1016/j.psep.2023.04.024.

[39] N. E. I. Hamda, A. Hadjali, and M. Lagha, "Multisensor Data Fusion in IoT Environments in Dempster–Shafer Theory Setting: An Improved Evidence Distance-Based Approach," *Sensors*, vol. 23, no. 11, Jun. 2023, doi: 10.3390/s23115141.

[40] R. G. Tanjung, "Expert System for Diagnosing Cataract Using Forward Chaining Method," *Journal of Computer Scine and*

Information Technology, pp. 13–17, Jan. 2023, doi: 10.35134/jcsitech.v9i1.56.
[41] Marzia Ahmed, "Maternal Health Risk Data," UCI Machine Learning Repository. [Online]. Available: https://www.kaggle.com/datasets/csafrit2/maternal-health-risk-data
[42] E. Fitri Rahayu, "Comparative Analysis Expert System Of Certainty Factor Method With Dempster Shafer Method By Using Open Decision Maker Tool For Toddler At Vaccine House," *UG JURNAL*, vol. 14, 2020.
[43] M. Iwan Shofyan *et al.*, "Expert System Design for Web-Based Lecturer Scheduling," *J Phys Conf Ser*, vol. 1175, no. 1, Jun. 2019, doi: 10.1088/1742-6596/1175/1/012119.
[44] S. Elmie Dansy *et al.*, "Expert System for Diagnosis Coronavirus Disease," *International Journal of Artificial Intelligence*, vol. 1, pp. 39–44, 2023, doi: 10.36079/lamintang.ijai-1001.537.
[45] Y. B. Arya, "Flash Damage Expert System Using the Dempster Shafer Method," *MECOMARE (International Journal of Mechanical Computational and Manufacturing Research)*, vol. 12, no. 1, pp. 10–15, 2023, doi: 10.35335/computational.v12i1.92.
[46] S. Y. Choi and S. H. Kim, "Knowledge acquisition and representation for high-performance building design: A review for defining requirements for developing a design expert system," *Sustainability (Switzerland)*, vol. 13, no. 9, May 2021, doi: 10.3390/su13094640.
[47] J. Fajaryanti and C. Author Abstrak, "Expert System Development To Identify Employee Personality Types Using Dempster Shafer Theory," *Jurnal Riset Informatika*, vol. 4, no. 3, pp. 227–232, 2022, doi: https://doi.org/10.34288/jri.v4i3.389.
[48] S. R. D. Putranti, "Relationships between Mean, Median and Mode," *Aloha International Journal of Multidisciplinary Advancement (AIJMU)*, vol. 3, no. 4, p. 87, Apr. 2021, doi: 10.33846/aijmu30403.
[49] H. Henderi, F. Al Khudhorie, G. Maulani, S. Millah, and V. T. Devana, "A Proposed Model Expert System for Disease Diagnosis in Children to Make Decisions in First Aid," *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 6, no. 2, pp. 139–149, Aug. 2022, doi: 10.29407/intensif.v6i2.16912.
[50] M. Abdul Aziz Sujudi and dan Deny Hidayatullah, "Expert System Certainty Factor Method in Website-Based Pest Diagnosis of Eggplant Plants," *Jurnal Mantik*, vol. 3, no. 36, pp. 428–438, 2020, [Online]. Available: https://iocscience.org/ejournal/index.php/mantik/index

AUTHORS BIOGRAPHY

SETIA MUKTI AZIZAH is from Pati, Central Java. Since 2020, she has been studying as a student in the Informatics Study Program at Universitas Sebelas Maret. Her current research focuses on data science and expert systems. She chose these fields since she has a deep fascination with them and intends to expand her knowledge further. In addition, she uses Decision Tree and Dempster Shafer methods for her final project in order to develop an Expert System for diagnosing pregnancy risks.

WIHARTO received obtained a bachelor's degree in electrical engineering (B.E.) from Universitas Telkom, Indonesia, in 1999. He obtained a master's degree in computer science from Universitas Gadjah Mada, Indonesia, in 2004 and a Doctoral degree from the same University, in 2017. Currently he works as a lecturer in the Informatics Department, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia. His experience and areas of interest focus on artificial intelligence, clinical decision support systems, machine learning, and expert systems.

BRILYAN HENDRASURYAWAN is from Boyolali, Central Java. He finished his master's at Universitas Indonesia. Now, he is a lecturer and a researcher at the Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Central Java, Indonesia. His work interest is Computer Sciences, especially in software engineering and artificial inteligence. His work area includes all aspects of it.

APPENDIX

APPENDIX I
Symptom Data from Dataset

No	Column	Symptom
1	Age	Age of pregnant women less than or equal to 16 years
		Age of pregnant women 17-34 years
		Age of pregnant women more than or equal to 35 years
2	BodyTemp	Body temperature less than 36.0 degrees Celsius
		Body temperature 36.0-37.5 degrees Celsius
		Body temperature greater than or equal to 37.5 degrees Celsius
3	BS	Blood glucose level less than 140 mg/dL
		Blood glucose level 140 - 200 mg/dL
		Blood glucose level greater than 200 mg/dL
4	DiastolicBP	Diastolic blood pressure less than 80 mmHg
		Diastolic blood pressure 80-90 mmHg
		Diastolic blood pressure greater than 90 mmHg
5	SistolicBP	Systolic blood pressure less than 130 mmHg
		Systolic blood pressure 130-150 mmHg
		Systolic blood pressure greater than 150 mmHg

APPENDIX II
Symptom Data from KSPR

No	Symptom
1	Too young to get pregnant (≤ 16 years old)
2	Marry at the late age, where the duration of marriage has been ≥ 4 years
3	Too old for first pregnancy (≥ 35 years old)
4	Pregnancy spacing ≤ 2 years
5	Having ≥ 4 children
6	Too old to get pregnant (≥ 35 years old)
7	Short stature (≤ 145 cm)
8	Miscarriage History
9	Delivery with a breech presentation
10	Previous delivery with forceps/vacuum
11	Manual placenta, infusion/transfusion
12	Cesarean section surgery history
13	Diseases during pregnancy (anemia, malaria, TB, heart disease, diabetes, PMS)
14	Oedema during pregnancy (face and legs) and high blood pressure
15	Twin pregnancy
16	Hydramnios
17	Fatal intrauterine death
18	Overdue date or overdue pregnancy

19	Breech position
20	Transverse position
21	Bleeding during pregnancy
22	Preeclampsia/eclampsia

APPENDIX III
Qualitative Data

No	Symptom
1	Having sexually transmitted diseases
2	Positive urine levels
3	Negative urine levels
4	Upper arm circumference less than 23.5 cm
5	Hemoglobin (HB) value below 8
6	HB value between 8 and 10
7	HB value 11
8	Low preeclampsia
9	Moderate preeclampsia
10	Severe preeclampsia
11	Having asthma/breathing difficulties
12	Not experiencing pre-eclampsia
13	Mild pre-eclampsia
14	Severe pre-eclampsia (eclampsia)
15	Infected with viruses (Rubella, Toxoplasma, Chickenpox, etc.)
16	Contracting HIV/STIs (Sexually Transmitted Infections)
17	Experiencing bleeding during pregnancy
18	Experiencing oedema
19	History of delivery with breech presentation
20	History of delivery with forceps/vacuum
21	Miscarriage History
22	Cesarean section surgery History
23	Hydramnios
24	Twin pregnancy
25	Having asthma
26	Having pregnancy-related diseases (anemia, malaria, TB, heart disease, diabetes)
27	Pregnancy duration of more than 9 months

APPENDIX IV
Combination of Data

Code	Symptom
G01	Age of pregnant women less than or equal to 16 years
G02	Age of pregnant women 17-34 years
G03	Age of pregnant women more than or equal to 35 years
G04	Marriage age less than 4 years
G05	Marriage age more than or equal to 4 years
G06	Having no children
G07	Having 1-3 children
G08	Having 4 or more children
G09	Pregnancy spacing less than or equal to 2 years
G10	Pregnancy spacing of more than 2 years
G11	Height less than 145 cm
G12	Height more than or equal to 145 cm
G13	Upper arm circumference less than 23.5 cm
G14	Upper arm circumference more than or equal to 23.5 cm
G15	Weight gain 7-12 kg
G16	Weight gain 12-15 kg
G17	Weight gain of more than 15 kg
G18	Body temperature less than 36 °C
G19	Body temperature 36-37.5°C
G20	Body temperature more than 37.5°C
G21	Hemoglobin (HB) value less than 8 grams/dL
G22	HB value 8-10 grams/dL
G23	HB value 11 grams/dL
G24	Systolic blood pressure less than 130 mmHg
G25	Systolic blood pressure 130-150 mmHg
G26	Systolic blood pressure more than 150 mmHg
G27	Diastolic blood pressure less than 80 mmHg
G28	Diastolic blood pressure 80-90 mmHg
G29	Diastolic blood pressure more than 90 mmHg
G30	Glucose level less than 140 mg/dL
G31	Glucose level 140-200 mg/dL
G32	Glucose level more than 200 mg/dL
G33	Positive urine protein (+)
G34	Negative urine protein (-)
G35	Breech baby position
G36	Transverse baby position
G37	Normal baby position
G38	Not experiencing pre-eclampsia
G39	Mild pre-eclampsia
G40	Severe pre-eclampsia (eclampsia)
G41	Experiencing intrauterine fetal death
G42	Infected with viruses (Rubella, Toxoplasma, Chickenpox, etc.)
G43	Manual placenta
G44	Experiencing HIV/STIs (Sexually Transmitted Infections)
G45	Experiencing bleeding during pregnancy
G46	Experiencing oedema
G47	History of delivery with breech presentation
G48	History of delivery with forceps/vacuum
G49	Miscarriage history
G50	Cesarean section surgery history
G51	Hydramnios
G52	Twin pregnancy
G53	Having asthma
G54	Having pregnancy-related diseases (anemia, malaria, TB, heart disease, diabetes)
G55	Pregnancy duration more than 9 months