Combination certainty factor method and fuzzy expert system module to determine the dose of leukemia drugs

Dwi Krisbiantoro¹, Linda Perdana Wanti², Nur Wachid Adi Prasetya²
¹Department of Computer Science, Universitas Amikom Purwokerto, Purwokerto, Indonesia
²Department of Computer and Business, Politeknik Negeri Cilacap, Cilacap, Indonesia

ABSTRACT
Leukemia is a type of blood cancer. Treatment for leukemia patients can last for years because the dose of medication given is adjusted to the patient's immune system. The aim of this research is the use of information technology through a combination of certainty factors and the development of a fuzzy expert system (FES) module to determine the therapeutic schedule for administering leukemia drugs. The urgency of this research is to help medical personnel in measuring the dose of leukemia medication to be given to patients so as to increase the cure rate for leukemia patients. The method used is certainty factor and fuzzy logic. The combination of the certainty factor method and the FES module which is carried out using input variables in the form of the severity of the leukemia suffered by the patient is to produce an appropriate therapeutic schedule for administering leukemia drugs. The result of this research is a combination of the factor certainty method and the FES module which has been tested and the accuracy level is 95.17%, the same as recommendations from experts.

Keywords: Dose, Fuzzy expert system, Leukemia, Medicine, Modular

This is an open access article under the CC BY-SA license.

Corresponding Author:
Linda Perdana Wanti
Department of Computer and Business, Politeknik Negeri Cilacap
Street Dr. Sutomo, No. 1, Cilacap Selatan, Cilacap 53212, Jawa Tengah, Indonesia
Email: linda_perdana@pnc.ac.id

1. INTRODUCTION
Leukemia is blood cancer caused by the body producing too many abnormal white blood cells. White blood cells are part of the immune system and are produced in the bone marrow [1]. When bone marrow function is disrupted, the white blood cells produced will experience changes and no longer carry out their role effectively [2]. Leukemia is often difficult to detect because its symptoms resemble those of other diseases. Early examination needs to be carried out so that this disease can be treated quickly [3]. Leukemia can be chronic or acute. In some types of chronic leukemia, the cancer develops slowly with initial symptoms that are usually mild [4]. Meanwhile, for acute leukemia, the development of cancer cells occurs very quickly and the symptoms can get worse very quickly [5]. The conclusion is that acute leukemia is more dangerous than chronic leukemia [6]. There are four types of leukemia based on the type of white blood cells that identify the type of leukemia, including the first, acute lymphoblastic leukemia (ALL), which occurs when the bone marrow produces too many white blood cells, immature lymphocytes or lymphoblasts, and the second, chronic lymphocytic leukemia (CLL), which occurs when the bone marrow produces too many abnormal lymphocytes and slowly causes cancer, the third is acute myeloblastic leukemia (AML) which occurs when the bone marrow produces too many immature myeloid cells or myeloblasts and the last is chronic myelocytic leukemia (CML) occurs when the bone marrow is no longer able to produce mature myeloid cells [7]. Data from the Indonesian Pediatric Center Registry states that in the period 2021-2022
there were 3,834 new cases of childhood cancer in Indonesia. A total of 1,373 children were still undergoing treatment as of December 2022. 833 were confirmed to have died [8]. A total of 519 child patients were recorded as having dropped out of treatment, meaning they did not continue treatment, and 148 children with cancer were confirmed to have completed cancer therapy treatment [9].

Symptoms of leukemia are often not clearly visible, new symptoms appear when the cancer cells are growing and starting to attack body cells [10]. The symptoms vary depending on the type of leukemia suffered, but in general some complaints from leukemia patients are fever accompanied by chills, fatigue that does not go away even after resting, weight loss drastically, experiencing symptoms of anemia, red spots appearing on the skin, frequent nosebleeds, the body bruises easily, sweats excessively, especially at night, easily gets infections, lumps appear in the neck due to swollen lymph nodes and the stomach often feels uncomfortable due to swollen liver and spleen, severe headaches, nausea and vomiting, loss of muscle control, pain bones, and spasms [11]. The problem that is often faced by both patients and health workers in the field is the side effects of leukemia treatment that is too long and this causes various side effects such as retention in treatment, decreased immunity, emotional and psychological impacts, large costs of treatment, and supportive care for leukemia patients such as blood transfusions, treatment for patients who experience excessive fatigue after undergoing leukemia chemotherapy treatment. This is the focus of this research, namely helping medical personnel in measuring the dose of leukemia medication that will be given to patients so as to increase the recovery rate of leukemia patients so that leukemia treatment can be more effective and efficient both in terms of time, medication given and minimizing side effects that will occur experienced by the patient.

Expert systems are able to support and sometimes replace experts in the medical world in diagnosing a disease [12]. Medical expert system applications provide direct access to doctors and patients through knowledge stored in the expert system knowledge base [13]. The development of expert systems also considers the advantages of a diagnosis and warning system approach through artificial intelligence, without forgetting the importance of good validation to assess system functionality [14]. Expert system users must follow several mutually agreed criteria, ranging from fuzzy logic to solutions that can be used together for treatment outside the clinical environment [15]. A fuzzy modular expert system uses fuzzy logic to overcome uncertainty and complexity in processing information into a decision or diagnosis result. In this research, the certainty factor method and fuzzy logic will be combined. The certainty factor method is used to produce calculations based on the level of confidence of the symptoms experienced by leukemia patients so that it can produce more accurate answers [16]. Fuzzy logic is used to help determine the dose of leukemia drugs to increase the effectiveness of the drug’s effect on the patient’s recovery [17]. The modular approach in expert system development refers to dividing a problem into smaller modules and each part is used to address certain aspects of a large problem [18]. The certainty factor and modular fuzzy expert system (FES) methods were implemented in this research to determine the module used in administering the dose of leukemia cancer medication based on the input variable, namely the severity of the leukemia suffered by the patient. Fuzzy logic is used to model knowledge or decisions that are uncertain [19].

The latest of this research with research that has already been carried out, including [17], [20]-[23] is that this research combines the certainty factor method and the FES module which was developed to determine the dosage of leukemia drugs given to patients. This dose will later be compared with the dose based on experts, namely doctors, to determine the level of accuracy of the recommended dose from the expert system and the dose from the expert. Providing the right dose will support the acceleration of patient recovery and increase the patient’s resistance to leukemia [24]. The next update is that the symptom data used in this study was taken from patients who were categorized into 4, including ALL, CLL, AML, CML. The next update is that the input variable used is the level of cancer suffered by the patient which is categorized into four, namely stage 1 (minimal), stage 2 (minor), stage 3 (moderate) and stage 4 (major) [25].

2. METHOD

This paper presents two methods, namely the certainty factor method and the fuzzy logic method. The fuzzy logic method is an implementation of FES-A which is designed manually and adjusts the rule base and database from FES. Meanwhile, the certainty factor method is implemented because it provides a certainty value regarding the symptoms of leukemia experienced by the patient. The explanation of each method is as follows:

2.1. Certainty factor method for calculating the level of confidence

Certainty factor is a method in expert systems that is used to handle uncertainty in determining conclusions based on the evidence or facts provided [26]. Each rule in the system is assigned a certainty factor, which indicates the extent to which the rule is reliable or how strong its influence is on the conclusion [27].
Combination certainty factor method and fuzzy expert system module to determine ...

(Deby Krisbiantoro)

CF ranges from -1 to 1, where a positive value indicates certainty of a positive conclusion, a negative value indicates certainty of a negative conclusion, and a value of zero indicates uncertainty [28]. Stages of the certainty factor method [29]:

i) Determine the value of CF


(1)

ii) Determine the value of CF combination determined by one premise

\[ CF[A \land B] = \text{Min}(CF[a], CF[b] \ast CF[\text{RULE}]) \]  

(2)

iii) Determine the value of CF combination determined by more than one premise

\[ CF[A \land B] = \text{Max}(CF[a], CF[b] \ast CF[\text{RULE}]) \]  

(3)

iv) Determine the CF value from the same conclusions

\[ CF\text{Combination}[CF1, CF2] = CF1 + CF2 \ast (1 - CF1) \]  

(4)

The stages in the certainty factor method are as follows: the initial assertion or proposition that needs to be evaluated. This can be a hypothesis, rule, or any statement about the problem domain. Gathering evidence or data relevant to the statement. This evidence may come from various sources such as sensors, databases, or user input. Calculation of the certainty factor based on the evidence collected, calculate the certainty factor for the statement. The certainty factor is a measure of the degree to which evidence supports or contradicts the statement. It usually ranges from -1 to 1, where -1 indicates complete contradiction, 0 indicates neutrality, and 1 indicates complete support. Combination of certainty factors: If there are several items of evidence, combine the certainty factors using the appropriate combination rules. Common combination rules include the product rule, minimum rule, and weighted sum rule. Apply a threshold to the combined certainty factors to determine the overall certainty of the statement. This threshold step helps in making decisions or drawing conclusions based on existing evidence. Based on the overall certainty obtained in the previous step, draw a conclusion or decision. If the certainty is above a certain threshold, then the statement can be considered true or acceptable. If not, further investigation or action may be required. Figure 1 shows the flowchart of the certainty factor method. Starting with determining the CF value for each premise of the rule used, then proceeding with determining the combination CF value determined by one or more premises, and ending with determining the CF value for the same conclusion, namely the diagnosis of leukemia cancer.

![Flowchart of certainty factor method](image-url)
2.2. Implementation of manually designed fuzzy expert system

Figure 2 explains the FES-A module model architecture where the input variables are the type and grade of cancer from leukemia patients which are divided into four categories, namely stage one which is symbolized by the description minimal, stage two by the description minor, stage 3 by moderate and stage four by the description major [30]. This module explains the definition of fuzzy variables in the fuzzification process, then continues with the formation of fuzzy rules in the inference engine and defuzzification to end the FES-A module development process [31]. The process continues by adjusting the percentage of the leukemia drug dose given to the patient which has implications for increasing the drug dose. The dose will be adjusted to the chemotherapy period and leukemia drug administration schedule. The final process is to develop a module that has been tested by paying attention to the level of toxicity of cancer cells and the condition of the patient’s body which is continuously monitored clinically. Based on Figure 2, the stages in the FES-A module: The first stage is defining the input and output variables in linguistic form. This involves defining fuzzy sets that represent variable values and membership functions that describe how strongly a value belongs to that set. Once the variables are defined, the input values received from the environment or user are converted into fuzzy values using predefined membership functions. This allows processing of uncertain or ambiguous data. Use of fuzzy rules: once the input values are converted into fuzzy values, fuzzy rules are used to take decisions or provide recommendations. These rules consist of a collection of “if-then” statements that express the relationship between input and output values in a given domain. Fuzzy Inference: This stage involves evaluating fuzzy rules to produce intermediate fuzzy values that represent the membership level of each relevant output set. Aggregation: Intermediate fuzzy values resulting from the inference stage are combined to produce overall fuzzy values that represent the output of the system. The commonly used method for this aggregation is to use aggregation operators such as maximum or minimum. Defuzzification: The final stage is to change the overall fuzzy values into concrete or crisp values. This is done using defuzzification methods such as centroid, mean of maximum, or other methods described previously.

![Figure 2. FES-A Modular based model architecture](image)

Based on expert opinions, user expectations, software requirements, records of existing field data from a previous release or similar system, and metrics thresholds, a linguistic variable can be determined [32]. Note that the membership function distributions and rule-base of the fuzzy logic system are developed with an expert assistance, a literature review, and numerous trial and error processes [33]. In other words, in a fuzzy classification approach, a collection of n data objects \((x^1, x^2, ..., x^n)\) is represented by a set of m attributes \(x_1^{(n)}, x_2^{(n)}, ..., x_m^{(n)}\). Each attribute in \(x^{(n)}\) shows a set of t discrete linguistic variables \(LV(x_m^{(n)})=\{LV(x_{m1}), LV(x_{m2}), ..., LV(x_{mt})\}\). Each input vector is classified into p different fuzzy sets, which are shown by FESA, FESB, ..., FESN. The membership function \(\mu_{FESi}(x^{(i)})\) shows the degree to which \(x^{(i)}\) belongs to FESi. In addition, the membership value \(\mu_{LV}(x_{i}^{(j)})\) depicts the degree to which attribute j of input vector i belongs to linguistic variable l (please refer to (5) and (6)) [34].

\[
IF \left(x_1^{(i)} \text{ is } LV_1^{(i)} \right) \text{ AND } \left(x_2^{(i)} \text{ is } LV_2^{(i)} \right) \text{ AND } ... \text{AND} \left(x_p^{(i)} \text{ is } LV_p^{(i)} \right) \text{ THEN } (FS \text{ is } FS_k) \quad (5)
\]

\[
IF \left(x_1^{(i)} \text{ is } LV_1^{(i)} \right) \text{ AND } \left(x_2^{(i)} \text{ is } LV_2^{(i)} \right) \text{ AND } ... \text{AND} \left(x_p^{(i)} \text{ is } LV_p^{(i)} \right) \text{ THEN } y = f(x) \quad (6)
\]
2.3. Adjusting the rule-base and data-base of the FES

In this section, the manually developed FES-A contains a database and rule base section that is optimized with an inference engine which is one of the processes in the expert system [35]. Rule data and symptom data are processed in the inference engine to produce a disease recommendation that refers to four categories of leukemia, namely ALL, CLL, AML, and CML. Data explaining the upper and lower limits of each membership function of the input variables used in this research are explained in Table 1. This data will later be processed using python programming language to visualize the results of calculating the degree of membership for each input variable involved.

Figure 3 explains the trapezoidal curve of the input variable membership function, namely the severity level of leukemia cancer patients, which is divided into four levels, namely minimal, minor, moderate and the most severe, namely major. Each input variable has a center point/centroid. In this research, the centroid method was used in the defuzzification process. The center point or centroid of a fuzzy set indicates the center of mass or "middle value" of the set. This can be used to formulate inference rules in fuzzy systems. Based on Table 1, the lower and upper limits have been determined for each input variable for the severity of the cancer suffered by the patient. The center point for minimum variables is 2, for minor variables is 4.5, for moderate variables is 8 and the center point for major variables is 10.5.

![Trapezoidal curve of severity level input variable membership function](image)

<table>
<thead>
<tr>
<th>Table 1. Range of input variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Minimal</td>
</tr>
<tr>
<td>Minor</td>
</tr>
</tbody>
</table>

3. RESULTS AND DISCUSSION

3.1. Certainty factor method calculation

The certainty factor method calculation with symptoms for each category of leukemia is as follows:

a) Determine the value of CF

Using (1) to find the user CF and expert CF values for each of the factor that cause of leukemia symptoms. Table 2 shows the user CF and expert CF values for each factor causing leukemia.

![Graph showing CF values](image)

<table>
<thead>
<tr>
<th>Table 2. An expert and an user interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms code</td>
</tr>
<tr>
<td>S1</td>
</tr>
<tr>
<td>S2</td>
</tr>
<tr>
<td>S3</td>
</tr>
<tr>
<td>S4</td>
</tr>
<tr>
<td>S5</td>
</tr>
<tr>
<td>S6</td>
</tr>
<tr>
<td>S7</td>
</tr>
<tr>
<td>S8</td>
</tr>
<tr>
<td>S9</td>
</tr>
<tr>
<td>S10</td>
</tr>
<tr>
<td>S11</td>
</tr>
<tr>
<td>S12</td>
</tr>
<tr>
<td>S13</td>
</tr>
<tr>
<td>S14</td>
</tr>
<tr>
<td>S15</td>
</tr>
<tr>
<td>S16</td>
</tr>
<tr>
<td>S17</td>
</tr>
</tbody>
</table>
b) Determine the value of CF combination determined by more than one premise
After knowing the user CF value and CF expert value, proceed with determining the CF Combine value, which is determined by more than one premise using (3). Table 3 shows the results of the CF Combination values of each rule used to determine the type of leukemia cancer suffered by a patient who experiences symptoms with the following codes: S1 (severe headache), S3 (muscles lose control), S4 (bone pain), S5 (dazed), S7 (fever and chills) and S8 (fatigue that doesn't go away even after resting).

<table>
<thead>
<tr>
<th>Rule code</th>
<th>CF1</th>
<th>CF2</th>
<th>Rule code</th>
<th>CF1</th>
<th>CF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>R01</td>
<td>0.614</td>
<td>0.732</td>
<td>R09</td>
<td>0.624</td>
<td>0.742</td>
</tr>
<tr>
<td>R02</td>
<td>0.789</td>
<td>0.807</td>
<td>R10</td>
<td>0.728</td>
<td>0.807</td>
</tr>
<tr>
<td>R03</td>
<td>0.637</td>
<td>0.566</td>
<td>R11</td>
<td>0.637</td>
<td>0.516</td>
</tr>
<tr>
<td>R04</td>
<td>0.762</td>
<td>0.455</td>
<td>R12</td>
<td>0.752</td>
<td>0.455</td>
</tr>
<tr>
<td>R05</td>
<td>0.637</td>
<td>0.807</td>
<td>R13</td>
<td>0.637</td>
<td>0.807</td>
</tr>
<tr>
<td>R06</td>
<td>0.596</td>
<td>0.752</td>
<td>R14</td>
<td>0.576</td>
<td>0.724</td>
</tr>
<tr>
<td>R07</td>
<td>0.634</td>
<td>0.807</td>
<td>R15</td>
<td>0.642</td>
<td>0.803</td>
</tr>
<tr>
<td>R08</td>
<td>0.614</td>
<td>0.890</td>
<td>R16</td>
<td>0.614</td>
<td>0.817</td>
</tr>
</tbody>
</table>

Table 3. The value of CF combination

From the picture above, it can be concluded that the level of accuracy of expert diagnosis by calculating using the certainty factor method found. In the second iteration the level of accuracy obtained was 95.17% because the CF calculation results for each type of leukemia cancer were 0.69 for ALL, 0.93 for CLL, 0.85 for AML and 0.69 for CML. While the expert diagnosis results for each disease were 0.67 for ALL, 0.92 for CLL, 0.84 for AML and 0.68 for CML.

3.2. Results of the combination of certainty factor and fuzzy logic methods
From the results of the certainty factor method calculation above, it is stated that patients who experience symptoms with codes S1, S3, S4, S5, S7, and S8 are more likely to suffer from cancer with the type CLL. The recommended dosing schedule using the fuzzy logic method with a modular concept is explained as follows: the input variables used are as in Table 1, namely the level of severity suffered by the patient in four categories. So the first thing to do is separate the expert system into different modules according to their functions and tasks, namely modules that are developed based on input from the results of certainty factor method calculations [36]. The FES module developed to recommend a leukemia drug administration schedule accommodates all expert knowledge in the form of fuzzy rules. The diagram obtained by data processing is shown in Figure 4. Figure 5 shows the processing of patient data with a sample of symptoms as previously explained above is that Figure 5(a) explains the low dose of leukemia medication given with blue shading and the administration schedule according to the recommended time while Figure 5(b) explains that the dose of medication given is high with green shading or the recommended time range. Figure 5(c) shows the combination when a low dose of leukemia drug and a high dose of leukemia drug are combined, you can see that there is a shaded area with a combination of blue and green.
Combination certainty factor method and fuzzy expert system module to determine ...

The final step is the defuzzification process. In this research, the defuzzification process carried out is that the fuzzy inference machine provides output in the form of a fuzzy set which is converted to crisp or firm values to recommend a schedule for administering leukemia drugs. With the input variable, the severity of the cancer suffered by patients with CLL and the severity of the cancer suffered is minor according to the fuzzy set conversion [37]. The method used for defuzzification in this research is the centroid method where the center of mass of the fuzzy set is used as the defuzzification value [38]. The center of mass is the point where the area under the fuzzy set curve is equal to its area. Figure 6 shows the results of the defuzzification process using the centroid method to determine the center of mass when determining the schedule for administering leukemia drugs. Figure 6(a) shows the center of mass for the leukemia drug administration schedule for low doses and Figure 6(b) shows the center of mass for the leukemia drug administration schedule for high doses.

4. CONCLUSION

The proposed system presents optimal schedule recommendations by combining the certainty factor method to determine the confidence value of the type of cancer suffered by a patient based on the symptoms experienced and the fuzzy logic method to recommend a schedule for administering leukemia drugs that is tailored to the physical condition of the patient. The diagnosis results using the certainty factor method compared with the doctor's diagnosis resulted in an accuracy rate of 95.17%. The resulting FES module shows that the research carried out is by expert recommendations in providing drug dosage and adjusted to the leukemia drug administration schedule while still paying attention to the patient's physical condition and the side effects of chemotherapy carried out according to the recommended schedule. Recommendations for the dosage schedule for leukemia drugs are generated from the FES module for low doses and high doses using the centroid method which shows the central value based on the inference rules in the FES. All doses and schedules given are by the monitoring of the doctor treating the patient to minimize unwanted events. In

Figure 5. Trapezoidal curve of severity level input variable membership function: (a) trapezoidal curve low; (b) trapezoidal curve high; and (c) combination trapezoidal curve low and high

Figure 6. Trapezoidal curve for defuzzification with centroid method (a) low dose and (b) high dose
future research, it is recommended to use more patient data so that the resulting level of accuracy is also higher and can also be analyzed again regarding the methods used to increase the accuracy rate of success of leukemia treatment in a shorter time and minimize the impact of treatment on patients.

REFERENCES


Combination certainty factor method and fuzzy expert system module to determine ... (Dwi Krisbiantoro)