e-ISSN : 2541-2019
p-ISSN : 2541-044X

AI-Based Pharyngitis Detection Expert System Using the Certainty Factor Method

Emy Susanti[1)*], Robby Cokro Buwono[2)], Dison Librado[3)], Faiz Rifki Syuhada[4)]
[1,2,4)]Faculty of Information Technology, Universitas Teknologi Digital Indonesia, Yogyakarta, Indonesia
[3)]Faculty of Business Management, Universitas Teknologi Digital Indonesia, Yogyakarta, Indonesia
[1)]emy.susanti@utdi.ac.id, [2)]robbycokro@utdi.ac.id, [3)]dison@utdi.ac.id, [4)]faiz.rifki23@students.utdi.ac.id

Submitted : Jun 24, 2026 | **Accepted** : Jul 12, 2026 | **Published** : Jul 14, 2026

Abstract: Pharyngitis, an inflammation of the pharynx commonly known as a sore throat, is one of the most frequent complaints in Indonesian primary health care, yet the uneven distribution of physicians, especially in remote regions, often delays timely triage. This study develops an Android-based expert system that provides an early-screening indication of two types of pharyngitis, acute and chronic, using the Certainty Factor (CF) method. Rather than treating the disease application itself as the main contribution, the novelty lies in the validation framework: a dual-source certainty model that keeps expert-elicited rule weights separate from user-reported symptom confidence, and a class-level accuracy evaluation, including a confusion matrix, sensitivity, specificity, precision, recall, and F1-score, that is rarely reported alongside comparable Certainty Factor systems. Knowledge was acquired from a general practitioner and clinical references and encoded as thirteen symptoms, two disease classes, and a rule base of IF-THEN rules with expert certainty weights. User certainty is captured through six linguistic terms and combined with the expert CF values using the single-evidence formula CF[H,E] = CFuser x CFexpert and the parallel combination formula. The application was built with Android Studio and a local SQLite knowledge base so that it operates entirely offline. A worked example involving five symptoms of acute pharyngitis produced a combined certainty of 0.9890, or 98.90%, illustrating a high-confidence screening output. The system's screening conclusions were compared with a general practitioner's diagnosis on 30 patient cases, agreeing in 28 cases for an overall agreement accuracy of 93.33% (class-level sensitivity of 94.12% for acute and 92.31% for chronic pharyngitis); because this figure reflects agreement with a single practitioner rather than a laboratory-confirmed reference standard, it is reported as an agreement accuracy rather than a universal clinical accuracy. Black-box testing confirmed that all functional features operated as intended. The results indicate that the Certainty Factor method can quantify screening uncertainty effectively and that the resulting offline mobile application can serve as an accessible early-screening aid for the public and a decision-support tool for paramedics, complementing rather than replacing a professional medical examination.

Keywords: Android; Certainty Factor; Expert System; Forward Chaining; Pharyngitis

INTRODUCTION

Pharyngitis, an inflammation of the pharynx located behind the mouth and nasal cavity, is one of the most common reasons people seek primary health care in Indonesia and a frequent component of acute upper respiratory tract infections, which consistently rank among the leading outpatient diagnoses at community health centres (Tabroni et al., 2022). Clinically, pharyngitis is commonly grouped into an acute form, which usually follows a viral or bacterial infection and presents with fever together with a swollen and painful throat, and a chronic form, which arises from prolonged irritation such as repeated infection, smoking, or exposure to dust and pollutants, and presents with a recurring sore throat and persistent throat discomfort (Kementerian Kesehatan Republik Indonesia, 2020). Although the two forms differ in cause and management, their early symptoms overlap considerably, so distinguishing between them through self-observation alone is difficult for the general public.

The public-health importance of correctly recognising pharyngitis extends beyond the discomfort it causes. Because the condition is so common and its symptoms are familiar, many patients treat themselves with over-the-

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

counter or leftover antibiotics without a proper diagnosis. Inappropriate antibiotic use of this kind does not help the large proportion of cases that are viral, and it contributes to antimicrobial resistance, which is widely regarded as one of the most serious threats to global health. An untreated bacterial throat infection, on the other hand, can occasionally progress to more serious complications. An accurate and timely indication of the likely type of pharyngitis therefore supports both individual patient outcomes and the broader effort to use antibiotics responsibly.

Access to professional diagnosis in Indonesia, however, remains uneven. The number and geographical distribution of physicians are still concentrated in urban centres, while many remote districts and small islands are served by only a handful of paramedics and have limited access to specialists. As a result, patients often delay seeking help, self-medicate inappropriately, or reach a health facility only after their condition has worsened (Ruhsyahadati et al., 2023). Smartphone ownership, by contrast, has spread rapidly across almost all regions of the country, and the Android platform in particular reaches users even where clinical services are scarce. This contrast highlights the opportunity for an accessible mobile tool that can provide an early, preliminary indication of a patient's condition and encourage timely consultation with a health professional.

One way to deliver such guidance is through an expert system, a branch of artificial intelligence that transfers the knowledge of a human expert into a computer program so that the program can reason about a problem and offer advice in a specific domain (Nasyuha et al., 2023). In medical diagnosis, expert systems are valuable precisely because they capture the knowledge of scarce specialists and make it available anywhere and at any time, allowing non-experts to obtain a structured and repeatable assessment that would otherwise require direct access to a clinician.

Medical reasoning, nevertheless, is rarely absolute. A single symptom may be consistent with several diseases to differing degrees, and patients themselves are often uncertain about how strongly they experience a given symptom. A diagnostic system therefore has to represent and reason with uncertainty rather than rely on strict yes-or-no rules. The Certainty Factor (CF) method addresses this need by attaching a numeric measure of belief to each rule, so that the system can express how confident it is in a conclusion and rank competing diagnoses by their degree of certainty (Sumiati et al., 2021).

The Certainty Factor method was introduced in the MYCIN system and has since been applied widely to diagnostic problems because it is computationally simple, requires only modest input from an expert, and yields results that are easy to interpret as a percentage of certainty (Hutasuhut & Nasyuha, 2021). Unlike probabilistic or machine-learning approaches, which generally require large and well-labelled datasets to estimate prior probabilities or train a model, the CF method can operate effectively with knowledge elicited directly from a domain expert. This property makes it particularly suitable for settings such as throat-disease screening, where structured clinical data are scarce but experienced practitioners can readily express the strength of association between symptoms and diseases.

Several studies have applied expert systems and the Certainty Factor method to medical diagnosis and have implemented them on the Android platform to improve accessibility (Kridalaksana et al., 2019; Lauryn et al., 2021). Despite this progress, comparatively few studies have combined the Certainty Factor method with an offline Android application aimed specifically at distinguishing the types of pharyngitis, and this weakness is concrete rather than merely descriptive. The most directly comparable prior system, a forward-chaining and Certainty Factor expert system for pharyngitis and laryngitis, was deployed as a web application that depends on an active internet connection, reported only per-case certainty percentages for seven disease sub-types rather than a single aggregate accuracy figure, and validated its output through black-box functional testing rather than a case-by-case comparison against a physician's independent diagnosis (Muhammad Siddik Hasibuan, Triase, 2024). More broadly, many existing Certainty Factor diagnostic applications either depend on an internet connection for their knowledge base, restrict validation to functional (black-box) testing without a quantitative diagnostic-accuracy metric, or report only an overall percentage of agreement without a class-level breakdown such as sensitivity, specificity, or a confusion matrix. These limitations restrict their usefulness in exactly the low-connectivity settings where such tools are most needed and leave their diagnostic reliability difficult to compare across studies.

This study addresses that gap. Its objective is to design and build an Android-based expert system that provides an early-screening indication of two types of pharyngitis, acute and chronic, using the Certainty Factor method, operates entirely offline through a local knowledge base, and is evaluated quantitatively against the judgment of a general practitioner. Rather than presenting the pharyngitis application as the primary novelty, this study positions its contribution at the level of validation strategy and system design, since the Certainty Factor method itself has already been applied to many disease domains. The main contributions are fourfold: first, an offline Android implementation whose knowledge base is stored in an editable SQLite schema, so that an expert can revise rule weights without changing application code, addressing the connectivity dependence identified as a gap in prior work; second, a dual-source certainty model that keeps the expert-elicited rule weight and the user-reported symptom confidence as separate, independently traceable quantities rather than a single blended score, which makes it possible to audit whether a given screening output was driven by strong expert evidence, strong user

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

confidence, or both; third, a case-by-case, class-level validation protocol against a practicing physician's diagnosis that goes beyond a single aggregate accuracy figure to report a confusion matrix, sensitivity, specificity, precision, recall, and F1-score for each disease class, a level of evaluation rigor that is uncommon in comparable Certainty Factor systems; and fourth, an explicit discussion of the method's limitations and of how the CF approach compares with alternative reasoning strategies such as Dempster-Shafer theory and Bayesian inference, situating the work within, rather than beside, the broader body of uncertainty-reasoning research.

LITERATURE REVIEW

An expert system consists of four core components: a knowledge base that stores facts and rules elicited from a domain expert, an inference engine that applies those rules to the facts supplied by the user, a working memory that holds the current state of the consultation, and a user interface. Two inference strategies are commonly used. Forward chaining starts from the known facts, in this case the symptoms reported by the user, and works toward a conclusion, whereas backward chaining starts from a hypothesised conclusion and searches for the facts that would support it. Because diagnosis in this study is driven by the symptoms a patient observes, the forward chaining strategy is the natural choice, and earlier work has confirmed that it is intuitive for symptom-driven reasoning (Hamidi et al., 2017; Muhammad Siddik Hasibuan, Triase, 2024; Putri, 2018).

Several studies have demonstrated the effectiveness of the Certainty Factor method for medical and non-medical diagnosis. Research on diagnosing digestive-tract diseases reported that a CF-based expert system reached an agreement of more than 90% with specialist diagnosis, confirming that the method can model diagnostic uncertainty reliably (Hutasuhut & Nasyuha, 2021; Lauryn et al., 2021). Similar work on skin-disease identification showed that CF produced consistent certainty values even when patients reported overlapping symptoms (Panjaitan et al., 2021; Rahmi et al., 2022; Shofia et al., 2017), and a study on early detection of respiratory diseases reached comparable conclusions when the CF computation was driven by a forward chaining engine (Mandiri et al., 2024; Ningsih et al., 2023).

On the implementation side, the Android platform has become a preferred medium for health-screening applications because of its wide adoption among Indonesian users and its ability to run offline. Studies that deployed expert systems on Android have noted improved accessibility for users in areas with weak connectivity, particularly when the knowledge base is stored locally using SQLite (Chen et al., 2025; Falatehan et al., 2018; Simanjorang; & Karnadi, 2020). Comparable mobile and web implementations have been reported for tonsillitis and other ear-nose-throat conditions, indicating that the approach generalises across related domains.

Beyond the Certainty Factor, researchers have applied alternative reasoning techniques such as fuzzy logic and the Dempster-Shafer theory of evidence to the same class of diagnostic problems. Comparative analyses have found that these methods can achieve similar accuracy, but that the Certainty Factor remains attractive because it needs only a single expert weight per rule and yields a result that is straightforward to interpret as a percentage (Ginting et al., 2021; Rizky et al., 2023). This balance of simplicity and interpretability is the main reason CF was selected for the present work.

To position this study within the existing literature, Table 1 summarises representative prior works drawn from the reference list of this paper in terms of the disease domain addressed, the reasoning method used, the deployment platform, and whether a physician-validated accuracy figure was reported. Where a source reported only functional (black-box) testing or a per-case certainty value rather than an aggregate accuracy validated against an independent diagnosis, the accuracy column states this explicitly rather than assigning an approximate figure, so that the comparison does not overstate what each prior study actually measured. The comparison shows that, although the Certainty Factor method and the Android platform have each been used widely, the two are seldom combined in an offline application that targets the specific distinction between acute and chronic pharyngitis and that reports a case-by-case, physician-validated accuracy; the most disease-relevant prior system was web-based and did not report such a figure at all.

Table 1. Comparison of representative related studies

Study	Domain	Method	Platform	Accuracy
Muhammad Siddik Hasibuan & Triase (2024)	Pharyngitis & laryngitis	Forward Chaining + CF	Web	Not physician-validated
Falatehan et al. (2018)	Liver disease	Fuzzy Tsukamoto	Android	Not physician-validated
Simanjorang & Karnadi (2020)	Cat disease	Forward Chaining	Android (offline)	Not physician-validated

*name of corresponding author

DOI : https://doi.org/10.33395/sinkron.v10i3.16463
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Mandiri et al. (2024)	Respiratory infection (ISPA)	Certainty Factor	Not specified	Not physician-validated
Ginting et al. (2021)	Autism in children	CF vs. Bayes theorem	Not specified	Comparative study, no single figure
Rizky et al. (2023)	Porphyria	CF vs. Dempster-Shafer	Not specified	Comparative study, no single figure
This study	Acute & chronic pharyngitis	Certainty Factor + forward chaining	Android (offline)	93.33% (agreement)

Comparison with Alternative Reasoning Methods. Certainty Factor is not the only technique used to reason under uncertainty in expert-system diagnosis; Dempster-Shafer theory, Bayesian (Bayes' theorem) methods, fuzzy logic, and, more recently, supervised machine-learning classifiers have all been applied to comparable diagnostic problems. Studies that compared the Certainty Factor method directly against the Bayes theorem for diagnosing a single condition in children reported that the two methods were broadly comparable in their conclusions, but that the Bayes theorem required a larger and more complete set of prior and conditional probabilities, which is harder to elicit reliably from a single expert than the one certainty weight per rule that the Certainty Factor method needs (Ginting et al., 2021). A similar comparison between the Certainty Factor method and Dempster-Shafer theory for a different diagnostic problem found that Dempster-Shafer's explicit handling of combined and conflicting evidence can be advantageous when multiple experts disagree, at the cost of a more complex belief-combination procedure than the Certainty Factor's single combination rule (Rizky et al., 2023). Fuzzy logic, by contrast, is best suited to symptoms that are inherently continuous, such as a measured body temperature, whereas most of the symptoms used in this study are reported as present or absent with a linguistic confidence term, which fits the Certainty Factor's evidence structure more directly than a fuzzy membership function would. Machine-learning classifiers such as Naïve Bayes or decision trees can, in principle, learn symptom-disease associations from data rather than from an expert interview, but they require a labelled dataset that is far larger than the 30 validation cases available here, and they would not by themselves produce the interpretable, rule-traceable weights that let a clinician audit why the system reached a particular conclusion. Taken together, these comparisons support the Certainty Factor method as an appropriate choice for the present setting, where an expert is available but a large labelled dataset is not, and where a transparent, auditable weight per symptom is preferable to a black-box classification score; the trade-off is that scaling this study's approach to a machine-learning-based classifier remains a natural direction for future work once a larger case dataset has been accumulated.

METHOD

This research adopted a prototyping software-development model combined with the standard stages of expert-system construction. The work proceeded through five stages: knowledge acquisition, knowledge representation and knowledge-base construction, inference design using the Certainty Factor method, implementation of the application on the Android platform, and testing. A prototyping model was chosen because the knowledge base is refined iteratively: an initial set of rules is built, the resulting prototype is reviewed with the expert, and the rules and certainty weights are adjusted until the behaviour is satisfactory. The overall research and inference flow is shown in Fig. 1.

Knowledge acquisition was carried out through structured interviews with a general practitioner and supported by clinical references on upper-respiratory infections. The expert was asked to enumerate the symptoms associated with acute and chronic pharyngitis, to indicate which symptoms belong to each disease, and to assign a degree of confidence to every symptom-disease relationship. The acquired knowledge was then represented as symptoms, diseases, and IF-THEN production rules, each carrying an expert certainty weight, and was encoded into the knowledge base.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16463
e-ISSN : 2541-2019
p-ISSN : 2541-044X

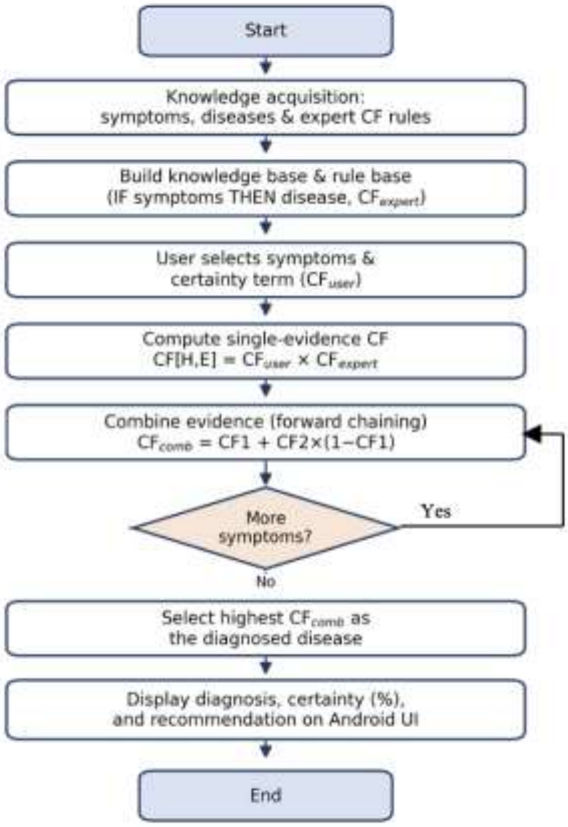

Fig. 1 Research and inference flow of the pharyngitis detection expert system

Symptoms and Diseases

Based on knowledge acquisition, thirteen symptoms (G01-G13) and two pharyngitis types (P1-P2) were defined. The symptoms are listed in Table 2, and the diseases are listed in Table 3.

Table 2. List of pharyngitis symptoms

Code	Symptom	Code	Symptom
G01	Anorexia (loss of appetite)	G08	Vomiting
G02	Fever	G09	Purplish pharyngeal wall (on exam)
G03	Palate deeper than in a normal person	G10	Throat base often feels itchy
G04	Dysphagia (difficulty swallowing)	G11	Headache
G05	Pharyngeal edema or swelling	G12	Frequent sneezing
G06	Itchy tongue (urge to scratch)	G13	Recurring sore throat
G07	Nausea		

Table 3. List of pharyngitis types

Code	Disease
P1	Acute pharyngitis
P2	Chronic pharyngitis

Certainty Factor Method

The Certainty Factor is a measure introduced by Shortliffe and Buchanan in the MYCIN system to quantify a degree of belief in a hypothesis H given a piece of evidence E. It is defined as the difference between the measure of belief (MB), the increase in belief in H caused by E, and the measure of disbelief (MD), the increase in disbelief in H caused by E, as shown in Equation (1). The resulting value ranges from −1, denoting absolute disbelief, to +1, denoting absolute certainty, with 0 meaning that the evidence neither supports nor contradicts the hypothesis (Ginting et al., 2021; Rizky et al., 2023).

*name of corresponding author

$$CF[H,E] = MB[H,E] - MD[H,E] \quad (1)$$

In this system, each rule stores an expert certainty value (*CFexpert*), while the user provides a certainty value (*CFuser*) through a linguistic term when selecting a symptom. The certainty of a single symptom-disease relationship is obtained by multiplying the two values, as in Equation (2).

$$CF[H,E] = CFuser \times CFexpert \quad (2)$$

The linguistic terms offered to the user and their corresponding numeric weights are given in Table 4. These weights let users who are unsure express partial confidence rather than a simple yes or no answer, which is essential for capturing the gradations of real symptom reports.

Table 4. Linguistic certainty terms and weights

Certainty term	CFuser weight
No / not present	0
Maybe	0.2
A little sure	0.4
Quite sure	0.6
Sure	0.8
Very sure	1.0

When more than one symptom supports the same disease, the individual CF values must be combined. The Certainty Factor method defines three combination rules depending on the signs of the two values being merged. When both values are positive, Equation (3) is used; when both are negative, Equation (4) is used; and when the two values have opposite signs, Equation (5) is used, where the denominator uses the smaller of the two absolute values.

$$CFcombine = CF1 + CF2 \times (1 - CF1) \quad (3)$$
$$CFcombine = CF1 + CF2 \times (1 + CF1) \quad (4)$$
$$CFcombine = (CF1 + CF2) / (1 - \min(|CF1|, |CF2|)) \quad (5)$$

In the present diagnostic task all evidence is supporting evidence, so the user weights and expert weights are positive and Equation (3) is applied repeatedly: the result of each combination becomes the old certainty (CF1) that is merged with the next single-evidence value (CF2). Finally, the combined certainty is converted to a percentage to make the result easier to read, as in Equation (6). The disease with the highest percentage is presented as the system's conclusion.

$$Certainty\ (\%) = CFcombine \times 100\% \quad (6)$$

Knowledge Base

The rule base links each disease to its supporting symptoms together with the expert certainty weight. Table 5 presents the complete rule base used in the system. For example, rule R02 states that if the patient has anorexia, fever, a palate deeper than normal, dysphagia, an itchy tongue, a purplish pharyngeal wall on examination, an itchy throat base, frequent sneezing, and a recurring sore throat, then the patient has chronic pharyngitis with the listed expert certainties. The two examination-based symptoms, the deeper palate (G03) and the purplish pharyngeal wall (G09), carry high expert weights because they are characteristic of the chronic form and rarely appear in the acute form.

Table 5. Knowledge base (rule base) with expert certainty values

Disease (Rule)	Symptom	CFexpert
P1 (R01)	G01	0.4
	G02	0.8
	G04	0.7
	G05	0.9
	G07	0.5
	G08	0.5
	G11	0.6
P2 (R02)	G01	0.3
	G02	0.4
	G03	0.8

*name of corresponding author

G04	0.6
G06	0.7
G09	0.9
G10	0.7
G12	0.6
G13	0.8

System Architecture

The application is organised into three layers on top of the underlying knowledge acquisition process, as illustrated in Fig. 2. The presentation layer is the Android user interface through which the patient selects symptoms, chooses a certainty term, and reads the result. The inference engine implements the forward chaining procedure and the Certainty Factor computation. The knowledge base, stored in a local SQLite database, holds the symptoms, diseases, and rules together with their expert certainty weights. Because the knowledge base resides on the device, the application performs the entire diagnosis offline.

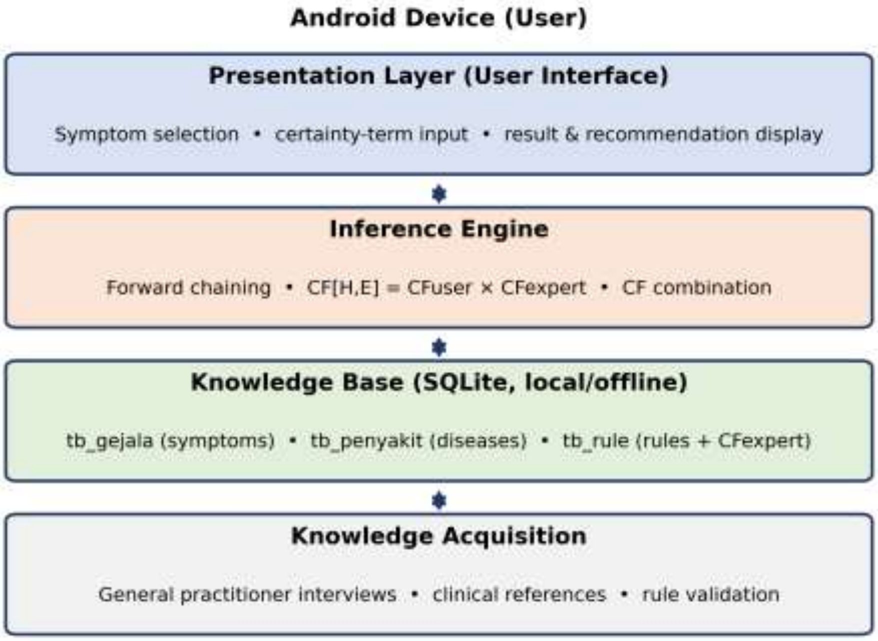

Fig. 2 Layered architecture of the expert system

Database Design

The knowledge base is implemented as three related tables in SQLite, summarised in Table 6. The symptom table stores the symptom code and description, the disease table stores the disease code and name, and the rule table links a disease to each supporting symptom and records the expert certainty weight for that pairing. Storing the certainty weight as a column of the rule table makes it easy to update the knowledge base when the expert revises a weight, without changing the application code.

Table 6. Knowledge-base database design

Table	Field	Description
tb_gejala	kode_gejala	Symptom code (primary key, e.g. G01)
	nama_gejala	Symptom description
tb_penyakit	kode_penyakit	Disease code (primary key, e.g. P1)
	nama_penyakit	Disease name
tb_rule	kode_penyakit	Disease code (foreign key)
	kode_gejala	Symptom code (foreign key)
	cf_pakar	Expert certainty weight (0-1)

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16463
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Inference Process

During a consultation the inference engine follows the forward chaining procedure shown in Fig. 1. First, the user selects the symptoms that apply and assigns a certainty term to each. For every selected symptom that appears in a rule, the engine computes the single-evidence certainty with Equation (2). The single-evidence values that belong to the same disease are then combined one by one with Equation (3), so that each disease accumulates a single combined certainty. After all symptoms have been processed, the engine converts the combined certainty of each candidate disease to a percentage with Equation (6) and selects the disease with the highest value as the screening conclusion, which is then displayed together with its certainty percentage and a recommendation.

System Implementation

The application was developed using Android Studio with the Java programming language. The knowledge base, comprising symptoms, diseases, and rules, was stored in a local SQLite database so that diagnosis can be performed entirely offline. The user interface presents the list of symptoms, lets the user choose a certainty term for each selected symptom, and then executes the Certainty Factor computation through a forward chaining procedure. The result screen displays the most likely pharyngitis type, the certainty percentage, and a short recommendation advising the user to consult a health professional.

RESULT

This section presents the inference computation and the testing of the implemented system. Two worked examples illustrate the calculation, one for each disease, followed by the results of functional and validation testing and a description of the application interface.

Worked Example 1: Acute Pharyngitis

Consider a patient who reports five symptoms associated with acute pharyngitis (P1) and assigns a certainty term to each. The selected symptoms, the user weights, the corresponding expert weights, and the single-evidence CF values are shown in Table 7.

Table 7. Single-evidence CF for worked example 1 (P1)

Symptom	CFuser (term)	CFexpert	CF[H,E]
G02	Very sure (1.0)	0.8	0.80
G04	Sure (0.8)	0.7	0.56
G05	Sure (0.8)	0.9	0.72
G07	Quite sure (0.6)	0.5	0.30
G11	Quite sure (0.6)	0.6	0.36

Each single-evidence value in Table 7 is obtained from Equation (2). For example, for symptom G02, CF[H,E] = 1.0 × 0.8 = 0.80. The five CF values are then combined sequentially with Equation (3):

CF1 = 0.80 ; CF2 = 0.56 ; CF3 = 0.72 ; CF4 = 0.30 ; CF5 = 0.36

Step 1: CFc1 = 0.80 + 0.56 × (1 − 0.80) = 0.80 + 0.112 = 0.912

Step 2: CFc2 = 0.912 + 0.72 × (1 − 0.912) = 0.912 + 0.06336 = 0.97536

Step 3: CFc3 = 0.97536 + 0.30 × (1 − 0.97536) = 0.97536 + 0.007392 = 0.982752

Step 4: CFc4 = 0.982752 + 0.36 × (1 − 0.982752) = 0.982752 + 0.006209 = 0.988961

Applying Equation (6), the certainty percentage is 0.988961 × 100% = 98.90%. The system therefore concludes that the patient most likely has acute pharyngitis with a certainty of 98.90%, which is a high-confidence result and would be accompanied by a recommendation to seek a physician for confirmation and appropriate treatment.

Worked Example 2: Chronic Pharyngitis

As a second illustration, consider a patient who reports five symptoms associated with chronic pharyngitis (P2). The symptoms include the two examination-based findings that are characteristic of the chronic form. The single-evidence CF values are shown in Table 8.

Table 8. Single-evidence CF for worked example 2 (P2)

Symptom	CFuser (term)	CFexpert	CF[H,E]
G03	Sure (0.8)	0.8	0.64
G06	Quite sure (0.6)	0.7	0.42
G09	Very sure (1.0)	0.9	0.90
G10	Sure (0.8)	0.7	0.56

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16463 p-ISSN : 2541-044X

G13	Sure (0.8)	0.8	0.64

Combining the five values with Equation (3) gives:
CF1 = 0.64 ; CF2 = 0.42 ; CF3 = 0.90 ; CF4 = 0.56 ; CF5 = 0.64
Step 1: CFc1 = 0.64 + 0.42 × (1 − 0.64) = 0.64 + 0.1512 = 0.7912
Step 2: CFc2 = 0.7912 + 0.90 × (1 − 0.7912) = 0.7912 + 0.18792 = 0.97912
Step 3: CFc3 = 0.97912 + 0.56 × (1 − 0.97912) = 0.97912 + 0.011693 = 0.990813
Step 4: CFc4 = 0.990813 + 0.64 × (1 − 0.990813) = 0.990813 + 0.005880 = 0.996693
Applying Equation (6), the certainty is 0.996693 × 100% = 99.67%. The system concludes chronic pharyngitis with a certainty of 99.67%. The high value is driven mainly by the purplish pharyngeal wall (G09), whose strong expert weight and very-sure user input make it the most decisive piece of evidence for the chronic form.

System Testing

Two kinds of testing were performed: black-box testing to verify functionality and validation testing to measure the system's screening accuracy against a general practitioner's diagnosis. The full validation set comprised all 30 patient cases, and the accuracy of 93.33% reported below is computed over that complete set; to keep the presentation manageable, Table 9 lists only ten representative cases out of the thirty as a transparent illustration of how individual cases were scored, covering both matches and the two mismatches, while the complete tally over all 30 cases is summarised in Table 9.

Table 9. Representative validation cases

No	Reported symptoms	System (CF%)	Physician	Result
1	G02, G04, G05, G07, G11	P1 (98.90%)	P1	Match
2	G03, G06, G09, G10, G13	P2 (99.67%)	P2	Match
3	G01, G02, G05, G08	P1 (95.86%)	P1	Match
4	G02, G04, G05, G11	P1 (97.66%)	P1	Match
5	G03, G09, G12, G13	P2 (98.94%)	P2	Match
6	G01, G02, G04	P1 (78.40%)	P2	Mismatch
7	G06, G10, G13	P2 (96.18%)	P2	Match
8	G02, G05, G07, G08	P1 (96.40%)	P1	Match
9	G01, G04, G13	P2 (72.16%)	P1	Mismatch
10	G03, G06, G09, G10	P2 (99.21%)	P2	Match

The black-box testing results, summarized in Table 10, show that every functional feature behaved as expected.

Table 10. Black-box testing results

Feature tested	Expected result	Status
Select symptoms and certainty term	Inputs recorded correctly	Valid
Run CF computation	Combined CF produced	Valid
Display diagnosis and percentage	Result shown correctly	Valid
Show recommendation	Recommendation displayed	Valid
Operate without internet	Runs fully offline	Valid

Validation testing compared the system's conclusion with the diagnosis of a general practitioner for all 30 patient cases. The system agreed with the practitioner in 28 cases and disagreed in 2 cases, where overlapping symptoms between acute and chronic pharyngitis caused the system to favour a different type. A summary by disease is presented in Table 11.

Table 11. Validation against a practitioner's diagnosis

Disease	Cases	Correct	Accuracy (%)
P1 Acute	17	16	94.12
P2 Chronic	13	12	92.31
Total	30	28	93.33

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

The overall accuracy is computed as the number of matching diagnoses divided by the total number of cases: (28 / 30) × 100% = 93.33%. This indicates that the system reproduces a practitioner's judgment in the large majority of cases.

Confusion Matrix and Class-Level Metrics

Because an aggregate agreement figure can mask systematic bias toward one disease class, the 30-case validation set was additionally re-expressed as a confusion matrix, using the two documented mismatches (cases 6 and 9 in Table 9) to determine, for each disease, how many cases were correctly identified, missed, or over-called. Table 12 presents the resulting confusion matrix, and Table 13 reports the derived class-level sensitivity, specificity, precision, and F1-score, together with their macro-averages.

Table 12. Confusion matrix for the 30-case validation set

	Predicted P1 (Acute)	**Predicted P2 (Chronic)**
Actual P1 (Acute), n=17	16 (TP)	1 (FN)
Actual P2 (Chronic), n=13	1 (FP)	12 (TP)

Table 13. Class-level sensitivity, specificity, precision, and F1-score

Disease class	**Sensitivity (Recall)**	**Specificity**	**Precision**	**F1-score**
P1 Acute pharyngitis	94.12%	92.31%	94.12%	94.12%
P2 Chronic pharyngitis	92.31%	94.12%	92.31%	92.31%
Macro average	93.22%	93.22%	93.22%	93.22%

The two metrics that most directly answer the reviewer concern about overall reliability are sensitivity (the proportion of true acute or chronic cases the system correctly flags) and precision (the proportion of the system's positive calls for a class that are correct); both are close to the aggregate accuracy for each class, at 94.12% and 92.31% respectively, and neither class is systematically favoured over the other, since the two misclassified cases fall one in each direction. The macro-averaged F1-score of 93.22% is close to the overall agreement accuracy of 93.33%, confirming that the aggregate figure is not being inflated by good performance on one disease at the expense of the other.

Application Interface

The application guides the user through three screens, shown in Fig. 3. On the first screen the user ticks the symptoms that apply; on the second screen the user selects a certainty term for each chosen symptom; and on the third screen the system shows the most likely pharyngitis type, the certainty percentage, and a recommendation to consult a physician. The layout was kept deliberately simple so that it can be used on low-specification devices and by people with limited experience of mobile applications.

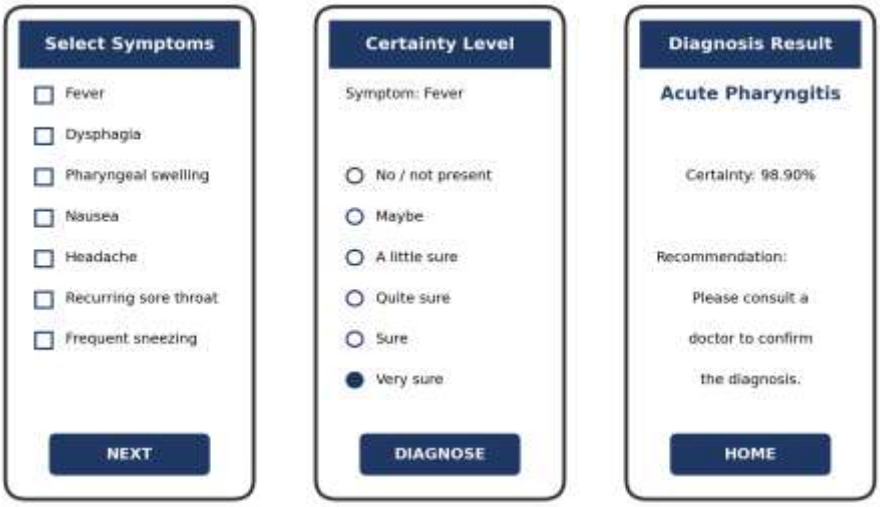

Fig. 3 Main screens of the Android application: symptom selection, certainty input, and result

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16463
e-ISSN : 2541-2019
p-ISSN : 2541-044X

DISCUSSIONS

The two worked examples demonstrate an important property of the Certainty Factor method: combining several pieces of supporting evidence raises the overall certainty without ever exceeding 1, which keeps the percentage interpretable. In the first example the certainty rose from 0.80 after the first symptom to 0.9890 after five symptoms, and in the second from 0.64 to 0.9967. This monotonic but progressively smaller increase reflects how each additional consistent symptom reinforces the conclusion while contributing less than the previous one, a behaviour that matches intuitive clinical reasoning.

The second example also shows why the choice of symptoms in the rule base matters. The purplish pharyngeal wall (G09) and the deeper palate (G03) are findings that a clinician observes on examination and that are characteristic of the chronic form. Giving them high expert weights allowed the system to reach a confident chronic-pharyngitis screening conclusion even though some of the patient's other symptoms, such as anorexia and dysphagia, also occur in the acute form. Symptoms that are unique to one disease therefore act as strong discriminators, whereas shared symptoms contribute little to telling the two forms apart.

The measured agreement accuracy of 93.33% compares favourably with the studies summarised in Table 1, several of which either did not report a physician-validated accuracy at all or validated their system only through functional (black-box) testing. The most disease-relevant prior system, a web-based forward-chaining and Certainty Factor application for pharyngitis and laryngitis, reported per-case certainty values ranging from 78.81% to 98.15% across seven disease sub-types but no aggregate accuracy against an independent diagnosis (Muhammad Siddik Hasibuan, Triase, 2024), which makes a direct numerical comparison impossible and is itself evidence for the research gap this study addresses. Where the present result can be compared, at the level of method choice rather than a single headline number, studies that pitted the Certainty Factor method against the Bayes theorem or Dempster-Shafer theory for other diagnostic problems reported broadly comparable classification outcomes between the methods, with the Certainty Factor method requiring less elicited knowledge per rule (Ginting et al., 2021; Rizky et al., 2023). This suggests that the accuracy achieved here is not an artefact of an unusually favourable method choice, but a plausible outcome of applying a well-established reasoning technique to a rule base built from careful knowledge acquisition and a modest, well-separated set of examination-based discriminator symptoms. As Table 1 indicates, most prior systems addressed either a different disease domain or used a single platform, whereas the present work combines an offline Android implementation with an explicit, case-by-case, class-level validation against a physician. The representative cases in Table 9 make the two errors easy to interpret: in case 6 the patient reported only the three symptoms shared by both forms, and in case 9 a recurring sore throat dominated an otherwise acute picture. In both, the absence of an examination-based symptom left the system without a strong discriminator, which explains the misclassification and is consistent with the confusion matrix in Table 12, where each of the two disease classes contributed exactly one error rather than one class dominating the mistakes.

From a practical standpoint, storing the knowledge base locally with SQLite allowed the application to run entirely offline, which directly addresses the accessibility problem in remote areas identified in the introduction, and the simple three-screen interface keeps the tool usable on low-specification devices. Beyond individual self-screening, the same offline architecture could plausibly support primary health care in three related ways. First, at the community health centre (Puskesmas) level, paramedics and community health workers without direct physician supervision could use the certainty output as a structured triage aid to decide whether a patient's presentation warrants referral to a physician, rather than as a stand-alone diagnostic tool. Second, in a telemedicine setting, a patient's self-reported symptoms and the resulting certainty percentage could be transmitted alongside a remote consultation request, giving the consulting physician a structured starting point before the call rather than an unstructured verbal description. Third, in regions with very limited medical staffing, periodic batch review of the locally logged screening outputs could help local health authorities notice a cluster of high-certainty cases that may warrant closer attention, provided such use is accompanied by clear communication that the tool's output is a screening indication and never a substitute for clinical judgment.

Limitations.

Several limitations should be kept in mind when interpreting these results. The disease scope is narrow: the rule base covers only two pharyngitis types out of the broader set of throat and upper-respiratory conditions that present with overlapping symptoms, so the system cannot rule out other causes of sore throat, such as tonsillitis or laryngitis, that a physician would routinely consider. The knowledge base reflects the judgment of a single general practitioner; a different expert, or a panel of experts, might assign different certainty weights to the same symptom-disease relationships, and no inter-rater reliability check between multiple experts was performed. The validation sample of 30 cases, while sufficient to illustrate the evaluation protocol and to compute a class-level confusion matrix, is small relative to the number of cases typically required for a clinically robust accuracy estimate, and the two disease classes are not evenly represented (17 acute versus 13 chronic cases). The reference standard against which the system was compared was a general practitioner's clinical diagnosis rather than a laboratory-confirmed

e-ISSN : 2541-2019
p-ISSN : 2541-044X

diagnosis such as a throat culture or rapid antigen test, so the reported accuracy reflects agreement with clinical judgment rather than agreement with a definitive biological reference. Finally, the system has so far been evaluated retrospectively against recorded cases rather than piloted with real patients in a live clinic or community setting, so its usability, acceptability, and impact on actual referral behaviour remain untested. For these reasons the application is positioned as an early-screening aid, and its recommendation always directs the user to a health professional, so the tool supports rather than replaces clinical examination.

CONCLUSION

This study designed and built an Android-based expert system that provides an early-screening indication of two types of pharyngitis, acute and chronic, using the Certainty Factor method with a forward chaining inference procedure and a local SQLite knowledge base that operates offline. Rather than the disease application alone, the study's contribution is the combination of a dual-source certainty model, an editable offline knowledge base, and a class-level validation protocol that is more rigorous than is typical for comparable Certainty Factor systems. A worked example for acute pharyngitis produced a combined certainty of 0.9890, or 98.90%, illustrating how the method aggregates multiple symptoms into a single interpretable confidence value. Comparison with a general practitioner's diagnosis on 30 cases yielded an agreement accuracy of 93.33% (class-level sensitivity of 94.12% for acute and 92.31% for chronic pharyngitis, with a macro-averaged F1-score of 93.22%), and black-box testing confirmed that all features functioned correctly; because this figure reflects agreement with a single practitioner rather than a laboratory-confirmed reference standard, it should be read as an agreement accuracy rather than a universal clinical accuracy. These findings indicate that the Certainty Factor method is suitable for modelling screening uncertainty in this domain and that the resulting application can serve as a practical early-screening aid. The main limitations, detailed above, are the narrow disease scope, the reliance on a single expert for the certainty weights, the modest and class-imbalanced validation sample, and the absence of a live clinical pilot. Future work should pursue a more visionary roadmap in addition to expanding the validation dataset and involving multiple physicians in weighting the rules: extending the rule base toward a hybrid expert system that combines Certainty Factor reasoning with a machine-learning classifier trained once enough labelled cases have accumulated, so that the interpretable rule base and a data-driven model can cross-check one another; migrating the knowledge base to a cloud-synchronised store that still allows offline operation but lets multiple experts collaboratively update rule weights over time; and exploring integration with electronic medical records (EMR) so that a patient's screening result and eventual physician diagnosis can be linked for continuous, real-world accuracy monitoring rather than a one-time retrospective validation.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the **Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kemdiktisaintek)** for the support and funding provided through the **Beginner Lecturer Research Grant (Hibah Penelitian Dosen Pemula) 2026**. This research would not have been possible without the valuable assistance, encouragement, and financial support provided by this grant. The authors also extend their appreciation to all parties who contributed to the successful implementation of this research.

REFERENCES

Chen, Z., Wang, T., Zhang, K., Guo, Y., & Shao, Z. (2025). A Q -Learning-Based Display Energy Optimization Scheme for Android Systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *44*(4), 1262–1275. https://doi.org/10.1109/TCAD.2024.3486244

Falatehan, A. I., Hidayat, N., & Brata, K. C. (2018). Sistem Pakar Diagnosis Penyakit Hati Menggunakan Metode Fuzzy Tsukamoto Berbasis Android. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, *2*(8), 2373–2381.

Ginting, R., Zarlis, M., & Rosnelly, R. (2021). Analisis Perbandingan Metode Certainty Factor dan Teorema Bayes untuk Mendiagnosa Penyakit Autis Pada Anak. *Jurnal Media Informatika Budidarma*, *5*(2), 583. https://doi.org/10.30865/mib.v5i2.2930

Hamidi, R., Anra, H., & Pratiwi, H. S. (2017). Analisis Perbandingan Sistem Pakar Dengan Metode Certainty Factor dan Metode Dempster-Shafer Pada Penyakit Kelinci. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, *5*(2), 142–147.

Hutasuhut, M., & Nasyuha, A. H. (2021). *Analisis Aritmia (Gangguan Irama Jantung) Menerapkan Metode Certainty Factor*. *5*, 1386–1393. https://doi.org/10.30865/mib.v5i4.3289

Kridalaksana, A. H., Hidayat, A., Mulawarman, U., & Cahyadi, D. (2019). Sistem Pakar Diagnosa Kerusakan Kamera DSLR Menggunakan Metode Certainty Factor Sequential. *Prosiding Seminar Nasional Ilmu Kridalaksana, A. H., Hidayat, A., Mulawarman, U., & Cahyadi, D. (2019). Sistem Pakar Diagnosa*

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Kerusakan Kamera DSLR Menggunakan Metode Certainty Factor Sequential. Prosiding Seminar Nasional Ilmu Komputer Dan Teknologi Info, *4*(1), 1–5.

Lauryn, M. S., Akhmad Saparudin, & Muhamad Ibrohim. (2021). SISTEM PAKAR DIAGNOSA PENYAKIT HEWAN TERNAK KAMBING DENGAN METODE CERTAINTY FACTOR (CF). *JSiI (Jurnal Sistem Informasi)*, *8*(1), 18–23. https://doi.org/10.30656/jsii.v8i1.2947

Mandiri, P. D., Hartanti, D., Sari, A. A., Komputer, F. I., Informatika, T., Duta, U., Surakarta, B., Pakar, S., & Factor, C. (2024). *Prototipe Sistem Pakar untuk Diagnosis Penyakit Infeksi Saluran Pernapasan (ISPA) Menggunakan Metode Certainty Factor*. *7*(September 2023), 180–191.

Muhammad Siddik Hasibuan, Triase, D. W. H. H. (2024). SISTEM PAKAR DIAGNOSIS AWAL PENYAKIT FARINGITIS DAN LARINGITIS MENGGUNAKAN METODE FORWARD CHAINING DAN CERTAINTY FACTOR. *Journal of Science and Social Research*, *3*(August), 1137–1146.

Nasyuha, A. H., Syahra, Y., Perangin-Angin, M. I., Habibie, D. R., & Subagyo, A. A. (2023). Sistem Pakar Dalam Mendiagnosis Penyakit Leishmaniasis Menerapkan Metode Case-Based Reasoning (CBR). *Jurnal Media Informatika Budidarma*, *7*(2), 747. https://doi.org/10.30865/mib.v7i2.6057

Ningsih, W., Hasibuan, N. A., & Hatmi, E. (2023). Analisa Perbandingan Metode Certainty Factor dan Teorema Bayes Untuk Mendiagnosa Penyakit Asam Urat. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, *6*(1), 234–241.

Panjaitan, Z., Hafizah, H., Ginting, R. I., & Amrullah, A. (2021). Perbandingan Metode Certainty Factor dan Theorema Bayes dalam Mendiagnosa Penyakit Kandidiasis pada Manusia Menggunakan Metode Perbandingan Eksponensial. *Jurnal Media Informatika Budidarma*, *5*(3), 1097. https://doi.org/10.30865/mib.v5i3.3078

Putri, N. A. (2018). Sistem Pakar untuk Mengidentifikasi Kepribadian Siswa Menggunakan Metode Certainty Factor dalam Mendukung Pendekatan Guru. *INTECOMS: Journal of Information Technology and Computer Science*, *1*(1), 78–90. https://doi.org/10.31539/intecoms.v1i1.164

Rahmi, R., Buaton, R., & Simanjuntak, M. (2022). *Sistem Pakar Mendiagnosa Penyakit Kambing Menggunakan Metode Certainty Factor (Studi Kasus Dinas Pertanian Dan Ketahanan Pangan Kabupaten Langkat, Sumatera Utara)*. *6*(3).

Rizky, F., Zulham, Z., Nasyuha, A. H., Elyas, A. H., & Kartadie, R. (2023). Analisis Perbandingan Certainty Factor dan Dempster Shafer Dalam Diagnosis Penyakit Porfiria Menggunakan Metode Perbandingan Eksponensial. *Building of Informatics, Technology and Science (BITS)*, *5*(1), 171–180. https://doi.org/10.47065/bits.v5i1.3611

Ruhsyahadati, R., Eldrian, F., Liana, N., & Triyana, R. (2023). *Edukasi Penyakit Faringitis Akut terhadap Masyarakat di Kecamatan Bungus Teluk Kabung Kota Padang Education of Acute Pharyngitis for Communities in Bungus Teluk Kabung District , Padang City*. *8*(4), 959–964.

Shofia, E. N., Putri, R. R. M., & Arwan, A. (2017). Sistem Pakar Diagnosis Penyakit Demam : DBD , Malaria dan Tifoid Menggunakan Metode K-Nearest Neighbor – Certainty Factor. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *1*(5), 426–435.

Simanjorang;, S. A., & Karnadi, V. (2020). SISTEM PAKAR DIAGNOSIS PENYAKIT KUCING DENGAN METODE FORWARD CHAINING BERBASIS ANDROID. *Comasie*, *3*(3), 21–30.

Sumiati, S., Akip Suhendar, & Fahriyanti. (2021). CERTAINTY FACTOR UNTUK KLASIFIKASI KELAINAN JANTUNG BERDASARKAN REKAM MEDIS ELEKTROKARDIOGRAM. *JSiI (Jurnal Sistem Informasi)*, *8*(1), 49–52. https://doi.org/10.30656/jsii.v8i1.3074

Tabroni, S. B., Wijayanta, S., & Fahyudi, A. (2022). PEMETAAN SEBARAN PENYAKIT FARINGITIS AKUT BERBASIS SISTEM INFORMASI GEOGRAFIS DI PUSKESMAS CANDILAMA TAHUN 2023. *Jurnal Rekam Medis Dan Informasi Kesehatan Indonesia (Jurmiki)*, *01*(02), 22–29. https://doi.org/https://doi.org/10.62951/jurmiki.v1i2.15

*name of corresponding author