AN EXPERT SYSTEM FOR EARLY DETECTION OF MENTAL HEALTH CONDITIONS USING CERTAINTY FACTOR AND DASS-42

Indri Rahmayuni[1*], Yance Sonatha[2], Tsalsabila Jilhan Haura[3], Fazrol Rozi[4]
Department of Information Technology, Politeknik Negeri Padang, West Sumatera, Indonesia[1234]
indri@pnp.ac.id[*]

Received: 07 April 2026, Revised: 17 May 2026, Accepted: 23 May 2026
**Corresponding Author*

ABSTRACT
Mental health problems such as depression, anxiety, and stress continue to increase in many countries, while access to professional services is still limited. Many digital screening systems use fixed scoring methods and do not consider uncertainty in user responses. This study developed a web-based expert system by combining the Depression Anxiety Stress Scales (DASS-42) and the Certainty Factor (CF) method to represent uncertainty in overlapping emotional symptoms and provide more flexible screening results. The knowledge base was prepared through consultation with a licensed clinical psychologist and converted into 42 production rules based on the DASS-42 items. Each rule was assigned a confidence value according to expert judgment. The system uses forward chaining to combine active rules and calculate confidence scores for depression, anxiety, and stress at the same time. System evaluation was conducted using 50 community cases aged 18–35 years and compared with independent expert assessment. The overall accuracy reached 86% (43 of 50 cases). The accuracy for each category was 88.2% for depression, 82.3% for anxiety, and 87.5% for stress. Most classification errors occurred between anxiety and stress, which may be related to overlapping symptoms in the DASS-42 instrument. The findings indicate that the proposed system can support early mental health screening through interpretable confidence-based results. However, this study used a limited dataset and only one expert in knowledge development. The system is intended as a screening support tool and not as a replacement for clinical diagnosis.
Keywords ***:*** *Expert System, Certainty Factor, DASS-42, Mental Health Screening, Decision Support*

1. Introduction

Mental health disorders have become an important global health issue and continue to affect people in many countries. Depression, anxiety, and stress contribute to reduced quality of life, impaired daily functioning, and long-term disability. According to the World Health Organization (2022), mental health conditions remain a major contributor to the global burden of disease. Large-scale epidemiological studies also reported that emotional disorders are common across different age groups and social backgrounds (Vos et al., 2020).

In Indonesia, mental health problems also require serious attention. National survey data reported that emotional and mental disorders among people aged 15 years and above reached around 6.1% of the population (Kemenkes RI, 2018). However, the actual number may be higher because many individuals do not seek help or are not formally identified. Social stigma, limited awareness, and lack of early screening services are often reported as contributing factors (Patel et al., 2018).

Access to professional mental health services remains unequal, especially in developing countries. The number of psychologists and psychiatrists is still limited compared with population needs. In addition, long waiting times, service cost, and geographical barriers often reduce access to treatment (Saxena et al., 2007). Previous studies also noted that many people delay seeking professional support until symptoms become more serious (Thornicroft et al., 2017; Corrigan et al., 2014). As a result, the treatment gap for mental health disorders remains high in many low- and middle-income countries (Kohn et al., 2004).

Digital technology has been considered one possible approach to reduce this gap. Web-based and mobile systems can provide faster access, wider reach, and more private interaction for users who may hesitate to visit mental health services directly. Recent studies showed that digital platforms may support screening, self-monitoring, and early intervention processes (Mohr et al., 2017; Torous et al., 2020).

Among available approaches, expert systems remain relevant because they can represent professional knowledge in a structured and transparent way. Unlike some black-box prediction models, expert systems allow users and developers to trace how conclusions are produced from rules and evidence (Shortliffe & Buchanan, 1975; Nunes et al., 2015). This characteristic is important in health-related applications where interpretability and trust are necessary.

Several previous studies have applied expert systems for mental health cases. Al-Hajji et al. (2019) developed an online expert system for psychiatric diagnosis. Putra and Yuhandri (2021) implemented a certainty factor model for mental disorder analysis using Indonesian expert knowledge. Anindita et al. (2023) used forward chaining and certainty factor reasoning for depression, anxiety, and stress classification. Other studies also explored intelligent systems for depression detection and related conditions (Raharja et al., 2022; Vasgi et al., 2022). These findings indicate that expert systems still have practical potential in mental health support.

Even so, several limitations are still found in previous studies. Many systems rely mainly on fixed-threshold scoring or focus only on single-condition classification without explicitly representing uncertainty and overlapping emotional symptoms. In addition, only a limited number of studies integrate DASS-42 with interpretable certainty-based reasoning to produce simultaneous confidence scores for depression, anxiety, and stress.

Another important challenge in mental health screening is symptom overlap. Depression, anxiety, and stress often share similar symptoms such as fatigue, sleep disturbance, emotional tension, and difficulty concentrating (Lovibond & Lovibond, 1995). In many existing systems, assessment results are based mainly on fixed score thresholds. Although practical, this approach may not adequately represent uncertainty in user responses or borderline symptom patterns (Zhu et al., 2023; Sulak & Köklü, 2024).

To address this issue, the Certainty Factor (CF) method can be considered. CF was introduced to model confidence levels in expert reasoning under uncertain conditions (Shortliffe & Buchanan, 1975). Instead of producing only categorical results, CF can show the degree of belief for each possible condition based on available evidence. Previous studies also reported that CF performs well in medical and psychological expert systems (Windarsyah et al., 2017; Yuwono et al., 2019).

The Depression Anxiety Stress Scales (DASS-42) is one of the most widely used instruments for measuring negative emotional states. It contains 42 items divided into three subscales: depression, anxiety, and stress. The instrument has demonstrated good psychometric properties in many countries and populations (Antony et al., 1998; Henry & Crawford, 2005; Osman et al., 2012). Indonesian studies also support the validity of this instrument in local settings (Damanik, 2006; Nada et al., 2022). Even so, DASS-42 is commonly used as a scoring instrument rather than as part of an inference system that models uncertainty.

Based on this background, this study proposes a web-based expert system that combines DASS-42 and the Certainty Factor method for early mental health screening. Each questionnaire item is represented as a rule with expert-defined confidence values. The system generates confidence scores for depression, anxiety, and stress simultaneously.

The main contributions of this study are as follows:

1. Development of a DASS-42-based expert system using Certainty Factor reasoning.
2. Implementation of a web platform for self-screening and knowledge management.
3. Performance evaluation using 50 screening cases compared with expert assessment.

This system is designed as a preliminary screening and decision-support tool. It is not intended to replace professional psychological diagnosis.

2. Literature Review

A. Expert Systems in Mental Health Screening

Expert systems are computer-based applications that are developed to imitate the reasoning process of human experts. In general, these systems use a knowledge base and a set of rules to generate recommendations or conclusions from user input. In healthcare, expert systems have been applied in diagnosis, treatment support, and decision-making processes for many years

(Jackson, 1998; Russell & Norvig, 2020). Earlier foundational work also described the development of rule-based expert systems for medical decision support (Buchanan & Shortliffe, 1984).

One important advantage of expert systems is interpretability. The reasoning process can be traced through rules, so users and developers can understand how a conclusion is produced. This characteristic is useful in health applications, where transparency and trust are important. In contrast, some machine learning models may provide accurate predictions but are more difficult to explain clearly (Doshi-Velez & Kim, 2017).

Although machine learning approaches may achieve high predictive performance, these methods often require large labelled datasets and may produce less interpretable outputs (Mohr et al., 2017; Shamseldin et al., 2025). In contrast, rule-based expert systems can represent reasoning processes more transparently, which is important in healthcare applications where explanation and user trust are necessary (Shortliffe & Buchanan, 1975; Russell & Norvig, 2020).

In the mental health field, several studies have shown that expert systems can support early screening and classification. Al-Hajji et al. (2019) developed an online system for psychiatric diagnosis using backward chaining. Putra and Yuhandri (2021) implemented a certainty factor-based system for mental disorder analysis using Indonesian expert knowledge. Nunes et al. (2015) proposed a hybrid approach for schizophrenia diagnosis and reported acceptable performance results. Anindita et al. (2023) also combined forward chaining and certainty factor methods for depression, anxiety, and stress classification. Other studies explored expert systems for depression detection and chatbot-based mental health support (Raharja et al., 2022; Vasgi et al., 2022).

These studies indicate that rule-based systems still have practical value, especially for accessible and explainable screening tools. However, several limitations are still found in previous works. First, only a limited number of studies use DASS-42 as the main source of knowledge. Second, many studies focus only on overall accuracy and do not report more detailed evaluation results such as class-level performance or confusion matrices. Third, privacy risks, false reassurance, and ethical use of mental health screening systems are not always discussed clearly. Therefore, further development is still needed.

In addition, ethical and privacy considerations remain important because mental health screening systems process sensitive psychological information (Cavoukian, 2009). Interpretability is also necessary to reduce the risk of misunderstanding or overreliance on automated screening results (Doshi-Velez & Kim, 2017). Similar studies in Indonesia also applied certainty factor and decision-support approaches for depression, anxiety, and stress screening, particularly in student and community populations (Khawarizmi et al., 2020; Kusumadewi & Wahyuningsih, 2020).

B. Depression Anxiety Stress Scales (DASS-42)

The Depression Anxiety Stress Scales (DASS-42) was introduced by Lovibond and Lovibond (1995) as a psychological instrument to measure three related emotional conditions, namely depression, anxiety, and stress. The questionnaire consists of 42 items, divided into three subscales with 14 items for each category. Respondents answer each statement using a four-point scale based on their experience during the previous week.

The final score of each subscale is commonly interpreted into several severity levels such as normal, mild, moderate, severe, and extremely severe (Lovibond & Lovibond, 1995). Because of its structured format, DASS-42 has been widely used in research, counselling, and early screening activities.

Previous studies reported that DASS-42 has good reliability and validity in different countries and populations. Antony et al. (1998) examined the psychometric properties of the instrument in clinical and community samples. Henry and Crawford (2005) also reported strong construct validity for the shorter version of DASS. Studies in Malaysia and Indonesia supported the use of this instrument in local cultural contexts (Ramli et al., 2007; Damanik, 2006; Nada et al., 2022).

Even though DASS-42 is widely accepted, it is important to understand that this instrument is not a full diagnostic tool. The questionnaire mainly measures symptom severity during a specific period. It does not fully assess clinical history, duration of symptoms, or functional impairment as required in formal diagnosis systems such as DSM-5 (American Psychiatric Association, 2013). Because of that, DASS-42 is more appropriate for screening and early indication rather than final diagnosis.

In this study, DASS-42 is used as a structured source of symptom evidence. Each item is converted into a rule in the expert system so that user responses can be processed using certainty-based reasoning.

C. Certainty Factor Method

The Certainty Factor (CF) method was first introduced in the MYCIN expert system to support reasoning under uncertain conditions (Shortliffe & Buchanan, 1975). In many medical and psychological cases, available symptoms do not always lead to one definite conclusion. A symptom may strongly support one condition, weakly support another condition, or only provide partial evidence. Therefore, a method that can represent uncertainty is needed.

The CF method expresses the degree of confidence in a hypothesis based on available evidence. Two basic components are commonly used, namely Measure of Belief (MB), which represents how strongly the evidence supports a hypothesis, and Measure of Disbelief (MD), which represents how strongly the evidence rejects it. The basic certainty value is formulated as follows:

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

where H represents the hypothesis and E represents the observed evidence. The CF value generally ranges from -1 to 1, where higher positive values indicate stronger confidence in the hypothesis.

When more than one symptom supports the same condition, the certainty values are combined step by step using the following equation:

$$CF\_combined = CF_1 + CF_2(1 - CF_1) \quad ...(2)$$

This combination process allows the system to accumulate several pieces of evidence into one final confidence score. As additional supporting evidence is added, the certainty value increases gradually.

Several previous studies have applied CF in medical and mental health expert systems. Windarsyah et al. (2017) used CF for schizophrenia diagnosis and reported good agreement with experts. Yuwono et al. (2019) compared CF with the Dempster-Shafer method and found similar performance results. Nopi et al. (2022) also implemented CF for mental disorder diagnosis. Other health-related studies showed that CF is practical and suitable for rule-based systems with limited training data (Febrianto et al., 2023).

Compared with more complex probabilistic approaches, CF has several practical advantages. First, the method is relatively simple to implement. Second, each rule has an explicit confidence value, so the reasoning process is easier to understand. Third, CF can directly incorporate expert judgement without requiring a large historical dataset. Based on these considerations, the CF method was selected as the inference mechanism in this study.

3. Research Methodology

A. Research Design

This study used an applied research approach focused on system development and performance evaluation. The study aimed to design and evaluate a web-based expert system for early mental health screening using the DASS-42 instrument and the Certainty Factor (CF) method.

The research was conducted through several sequential stages. First, the problem background and system requirements were identified through literature review and analysis of current mental health screening needs. Second, expert knowledge was collected and transformed into decision rules. Third, the system model and application architecture were designed. Fourth, the proposed system was implemented as a web-based platform. Finally, system performance was evaluated by comparing system outputs with expert assessment results.

This development flow follows common stages found in expert system research and decision-support system design (Jackson, 1998; Turban et al., 2020). The overall research procedure is presented in Fig. 1.

Vol 4(1) 2026 : 20-31

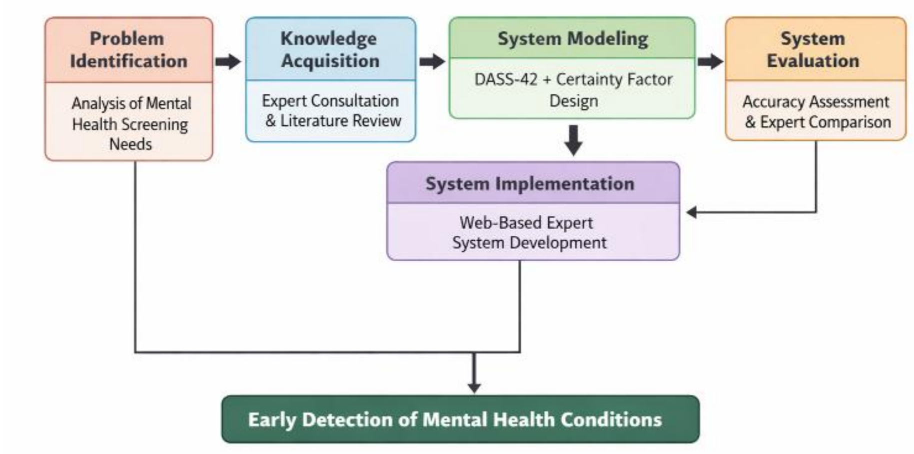

Figure 1. Research Design for Expert System Development

The proposed system was designed as a preliminary screening tool and not a replacement for professional psychological diagnosis. Therefore, all outputs were presented as confidence-based screening indicators with recommendations for further consultation when needed.

B. Knowledge Acquisition and Expert Calibration

Knowledge acquisition was conducted using two main sources. The first source was literature related to DASS-42, mental health screening, and certainty-based reasoning methods. The second source was direct consultation with a licensed clinical psychologist.

The expert involved in this study has more than ten years of professional experience and is an active member of the Indonesian Psychological Association (HIMPSI), West Sumatra branch. Three structured consultation sessions were conducted to gather domain knowledge and practical judgement. All interview sessions, discussion notes, and supporting materials were properly documented and archived to maintain transparency and research traceability. These consultation activities focused on three main objectives:

1. Identifying which DASS-42 items are strongly associated with depression, anxiety, or stress.
2. Assigning confidence values for each symptom-condition relationship.
3. Reviewing severity interpretations used in practical screening settings.

Based on the consultation process, each of the 42 DASS-42 items was linked to one dominant target category and assigned a certainty value. Symptoms considered more specific received higher weights, while symptoms with possible overlap received moderate values. The final knowledge base was reviewed again by the expert before implementation.

The acquired knowledge was then represented as production rules in the following form:

$$\text{IF } S_i \text{ THEN } H_j \ (CF_{ij}) \quad \text{...(3)}$$

where S_i represents a symptom item, H_j represents the target condition, and CF_{ij} represents the certainty value assigned by the expert. Examples of the resulting rules are shown in Table 1. This expert calibration stage was important to reduce arbitrary rule assignment and to align the system reasoning process with practical professional judgement.

Table 1. Sample Knowledge Base Rules with Expert-Assigned CF Weights

DASS-42 Item (abbreviated)	Target Condition	CF (MB)
Could not experience positive feeling	Depression	0.85
Aware of dryness of mouth	Anxiety	0.75
Could not seem to wind down	Stress	0.70
Felt down-hearted and blue	Depression	0.85
Felt close to panic	Anxiety	0.80
Felt using a lot of nervous energy	Stress	0.65
Felt life was meaningless	Depression	0.90
Experienced trembling (e.g., in hands)	Anxiety	0.75
Felt intolerant of interruptions	Stress	0.70

Because the knowledge base was calibrated by one clinical expert, inter-expert agreement analysis was not performed in this study. Different experts may assign slightly different confidence values depending on clinical experience and interpretation.

C. DASS-42 Operationalization

In common practice, DASS-42 is generally applied as a scoring instrument by summing item values in each subscale. In this study, a different approach was used. Each questionnaire item was treated as an individual symptom indicator in the expert system. This design allows the system to capture symptom patterns in more detail rather than depending only on total scores. User responses were collected using the original four-point Likert scale ranging from 0 to 3. Higher values indicate more frequent symptom experience during the previous week.

The response score was used to adjust the contribution of each rule. A response value of 3 activated the rule with full certainty weight, while lower responses contributed proportionally. In this way, partial symptoms could still be represented during the reasoning process. All 42 items were processed simultaneously. Confidence values for depression, anxiety, and stress were accumulated independently so that overlapping symptom patterns could still be identified.

D. CF Inference Implementation

The inference engine used forward chaining. The reasoning process started from user responses entered through the questionnaire page. Rules with satisfied conditions were activated and contributed to the related category. When several rules supported the same category, the Certainty Factor values were combined iteratively using the following formula:

$$CF\_running = CF_1 + CF_2(1 - CF_1) \quad ...(4)$$

The formula was applied repeatedly until all related rules had been processed. The final result of each category represented the confidence level of depression, anxiety, or stress. The category with the highest confidence score was displayed as the dominant screening result. However, the system also displayed all category scores so that users could observe the relative tendency across conditions.The final confidence values were then interpreted into severity labels such as normal, mild, moderate, severe, and extremely severe based on DASS-42 references and expert adjustment.

E. System Architecture

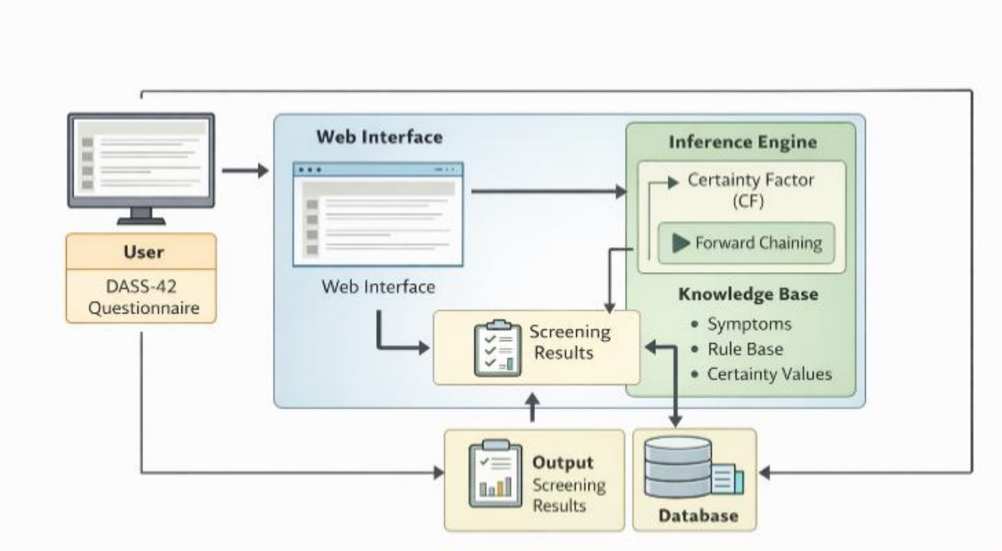

Figure 2. System Architecture of the Web-based Mental Health Expert System

The proposed application was implemented as a web-based system consisting of four main components: user interface, knowledge base, inference engine, and database. The system was developed using the Laravel framework and PostgreSQL database management system. The web interface was designed to support questionnaire processing, rule management, consultation history storage, and result visualization in a lightweight and maintainable environment. The interaction between these components is illustrated in Fig. 2.

The user interface was used for questionnaire completion and result presentation. The knowledge base stored rules, certainty weights, and severity thresholds. The inference engine processed user responses using Certainty Factor reasoning. The database stored user accounts, consultation history, and administrative records.

To support responsible use, the result page clearly stated that the output is only an early screening result and not a formal diagnosis. Users with concerning results were advised to seek professional consultation. An administrative module was also provided so experts could revise rules, certainty values, and threshold parameters without changing the application source code.

F. Evaluation Dataset and Method

System evaluation used 50 completed DASS-42 response cases collected through the web platform. Before completing the questionnaire, participants were informed about the purpose of the study and agreed voluntarily to use their anonymized responses for research evaluation purposes. Participants were adults aged between 18 and 35 years. The sample consisted of 28 female and 22 male participants. Based on independent expert review, the dominant conditions were distributed into 17 depression cases, 17 anxiety cases, and 16 stress cases. This distribution provided relatively balanced data across the three target categories. For each case, the system generated confidence scores and selected one dominant category. Separately, the same response set was reviewed independently by the clinical expert without access to the system output. The expert judgment was used as the comparison reference.

System performance was measured using overall accuracy and class-level accuracy as follows:

$$\text{Accuracy} = (N\_match / N\_total) \times 100\% \quad ...(5)$$

where N_match is the number of matching results and N_total is the total number of evaluated cases.

In addition to accuracy, a confusion matrix was used to analyze classification errors among depression, anxiety, and stress categories. This evaluation was intended to provide a more detailed view of system behaviour than overall accuracy alone.

4. Results and Discussion

A. System Implementation

The proposed system was implemented as a web-based application for early mental health screening. The application provides three main user interfaces, namely the questionnaire page, the result page, and the consultation history page. The main interfaces of the system can be seen in Figure 3.

The questionnaire page contains all 42 DASS-42 items using the standard four-point response scale. Users are required to complete all items before submitting the form, so incomplete responses do not affect the calculation process. The system automatically prevents submission when unanswered items were detected, thereby reducing missing-response problems during evaluation. After submission, the system processes the responses automatically and shows the result in a short time.

The result page displays confidence scores for depression, anxiety, and stress at the same time. The dominant category and severity level are also shown. This design allows users to see the tendency of each condition instead of receiving only one final label.

The consultation history page stores previous screening sessions so users can review earlier results. This feature may help users observe changes in their responses over time, although detailed clinical monitoring is outside the scope of this study. From the administrator side, the system also provides a management module for updating rules, certainty values, and severity thresholds. This feature supports future maintenance without changing the program code directly.

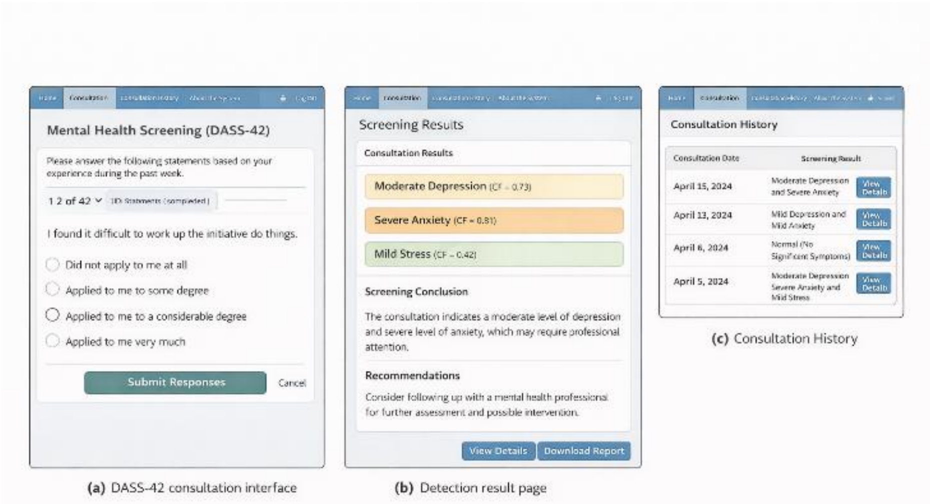

Figure 3. (a) DASS-42 questionnaire interface, (b) Detection result page showing CF values and severity levels, (c) Consultation history page

B. Certainty Factor Results

The Certainty Factor calculations produced different confidence values across the 50 evaluated cases. This result shows that the system was able to respond to different symptom patterns entered by users. Across all tested cases, the confidence scores ranged from 0.07 to 0.43, with an average value of 0.27. In general, most scores were in the low to moderate range and were not concentrated at very high values. This condition is acceptable because a screening system should be able to distinguish between stronger and weaker indications.

In several cases, anxiety scores were slightly higher than depression and stress scores. This pattern may be related to the characteristics of the tested sample rather than bias in the rule base. Stress scores also showed wider variation, which may indicate that stress-related responses were more sensitive to differences between users.

One practical advantage of the proposed method is that users receive gradual results. For example, a user may obtain a moderate confidence score for anxiety together with lower scores for depression and stress. Such information can be more useful than receiving only one category result. Unlike conventional threshold-based scoring approaches, the proposed system provides gradual confidence values that allow overlapping tendencies between depression, anxiety, and stress to be represented simultaneously. This characteristic may help reduce rigid classification outcomes in borderline cases.

C. Accuracy Evaluation

The performance of the proposed system was evaluated by comparing system outputs with independent expert judgment. A summary of the classification results is presented in Table 2.

Table 2. System Accuracy Performance

Category	Correct	Incorrect	Total Cases	Accuracy (%)
Depression	15	2	17	88.2%
Anxiety	14	3	17	82.3%
Stress	14	2	16	87.5%
Total / Average	43	7	50	86.0%

Based on the 50 evaluated cases, 43 cases produced dominant labels that matched the expert assessment. Therefore, the system achieved an overall accuracy of 86.0%. This result indicates that the proposed approach was generally able to produce outputs consistent with expert interpretation. Although the evaluation dataset was relatively limited, the class-level accuracy results showed reasonably consistent performance across the three target categories. For each category, the highest accuracy was found in depression cases (88.2%), followed by stress (87.5%) and anxiety (82.3%). Although the difference was not large, anxiety showed slightly lower performance than the other categories. This may happen because anxiety symptoms often overlap with stress symptoms.

To examine the classification pattern in more detail, a confusion matrix was also used. The results are shown in Table 3.

Table 3. Confusion Matrix of Expert vs. System Diagnosis

Expert / System	Depression	Anxiety	Stress
Depression	15	1	1
Anxiety	1	14	2
Stress	0	2	14

The confusion matrix shows that most errors occurred between anxiety and stress categories. A smaller number of cases occurred between depression and stress. No extreme contradiction pattern was found in the tested data. This result is understandable because some emotional symptoms may appear in more than one condition, especially in self-reported questionnaire data. For example, symptoms related to emotional tension, restlessness, and concentration difficulty may contribute to both anxiety and stress classifications. This overlap may explain why several cases were classified differently between the system and expert evaluation.

D. Discussion and Limitations

The results show that the proposed expert system can be used as a practical tool for early mental health screening. By combining DASS-42 and Certainty Factor reasoning, the system was able to provide interpretable outputs with relatively good agreement with expert judgment. One important strength of the system is transparency. Unlike black-box prediction models, the reasoning process in this system is based on clear rules and certainty values. This makes the decision process easier to review, revise, and explain. In health-related applications, this aspect is important for user trust and responsible use.

In addition, privacy protection should also be considered in digital mental health systems because users provide sensitive psychological information (Cavoukian, 2009). Another advantage is accessibility. Because the system was developed as a web-based platform, users can perform preliminary screening more easily without direct face-to-face consultation in the early stage. This may help people who hesitate to seek support because of stigma, time limitations, or distance barriers.

The accuracy result of 86.0% can be considered competitive when compared with several previous studies that applied expert systems in mental health fields (Nunes et al., 2015; Putra & Yuhandri, 2021; Anindita et al., 2023). Recent studies also showed that artificial intelligence methods can be used to predict depression, anxiety, and stress using psychometric data (Shamseldin et al., 2025).

However, direct comparison should be made carefully because previous studies may use different datasets, instruments, and evaluation settings. Even so, several limitations should be noted. First, the evaluation dataset consisted of only 50 cases and may not fully represent a wider population. Second, the knowledge base was developed with one clinical expert, so other experts may provide different confidence values. Third, the system is intended only for screening support and should not be interpreted as a clinical diagnosis tool.

Moreover, rule-based reasoning may not fully capture complex psychological conditions that involve dynamic behavioural, social, and environmental factors. Some emotional states may also change over time and cannot always be represented through fixed questionnaire rules alone.

Future studies may improve the system by involving more experts, increasing the dataset size, and testing the model in wider real conditions. Hybrid approaches that combine expert rules with data-driven calibration may also be explored. Overall, the findings indicate that the proposed approach provides a useful basis for developing accessible, transparent, and practical mental health screening systems.

5. Conclusion

This study developed and evaluated a web-based expert system for early mental health screening by combining the DASS-42 instrument with the Certainty Factor (CF) method. The main contribution of this study is the use of Certainty Factor reasoning to produce interpretable confidence-based screening results for overlapping emotional conditions. The system was designed to process questionnaire responses and generate confidence scores for depression, anxiety, and stress. By using rule-based reasoning, the system can provide screening results in a more flexible form than conventional score-threshold approaches.

Based on the evaluation of 50 response cases, the system achieved an overall accuracy of 86.0% when compared with independent expert judgment. The highest category accuracy was obtained for depression, followed by stress and anxiety. Most classification errors occurred between anxiety and stress categories, which may be influenced by overlapping symptoms between both conditions. The findings show that the proposed system has potential as a practical tool for preliminary mental health screening. The web-based design may improve accessibility, while the use of Certainty Factor reasoning provides more transparent results that can be easier to understand.

However, several limitations should be noted. The evaluation used a relatively small dataset and involved one clinical expert in the knowledge acquisition process. In addition, the system is intended only as a screening support tool and cannot replace professional psychological diagnosis. Future work may involve larger datasets, multiple experts, and wider field testing. Further development may also combine rule-based reasoning with data-driven approaches to improve performance and adaptability.

Overall, this study indicates that interpretable artificial intelligence can be applied to support accessible and practical early mental health screening systems when used responsibly.

References

Al-Hajji, A. A., AlSuhaibani, F. M., & Alharbi, N. (2019). An online expert system for psychiatric diagnosis. *International Journal of Artificial Intelligence & Applications*, *10*(2), 59–72. https://doi.org/10.5121/ijaia.2019.10206

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Anindita, M. T., Pristyanto, Y., Sismoro, H., Nurmasani, A., & Nugraha, A. F. (2023). Diagnose of mental illness using forward chaining and certainty factor. *Jurnal Techno Nusa Mandiri*, *20*(2), 63–70. https://doi.org/10.33480/techno.v20i2.4330

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, *10*(2), 176–181. https://doi.org/10.1037/1040-3590.10.2.176

Buchanan, B. G., & Shortliffe, E. H. (1984). *Rule-based expert programs: The MYCIN experiments of the Stanford heuristic programming project*. Addison-Wesley.

Cavoukian, A. (2009). *Privacy by design: The 7 foundational principles*. Information and Privacy Commissioner of Ontario.

Corrigan, P. W., Druss, B. G., & Perlick, D. A. (2014). The impact of mental illness stigma on seeking and participating in mental health care. *Psychological Science in the Public Interest*, *15*(2), 37–70. https://doi.org/10.1177/1529100614531398

Damanik, E. D. (2006). *Pengujian reliabilitas, validitas, analisis item dan pembuatan norma Depression Anxiety Stress Scale (DASS): Berdasarkan penelitian pada kelompok sampel Yogyakarta dan Bantul yang mengalami gempa bumi dan kelompok sampel Jakarta dan sekitarnya yang tidak mengalami gempa bumi* [Testing the reliability, validity, item analysis and norm construction of the Depression Anxiety Stress Scale (DASS): Based on research in sample groups from Yogyakarta and Bantul who experienced earthquakes and sample groups from Jakarta and its surroundings who did not experience earthquakes]. (Master's thesis). Universitas Indonesia.

Doshi-Velez, F., & Kim, B. (2017). *Towards a rigorous science of interpretable machine learning*. arXiv. https://arxiv.org/abs/1702.08608

Febrianto, K., Udayanti, E. D., Indriyono, B. V., Mahmud, W., & Zahari, I. (2023). Expert system for detection of diseases in layers using forward chaining and certainty factor methods. *Jurnal Masyarakat Informatika*, *14*(2), 80–95. https://doi.org/10.14710/jmasif.14.2.52266

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, *44*(2), 227–239. https://doi.org/10.1348/014466505X29657

Jackson, P. (1998). *Introduction to expert systems (3rd ed.)*. Addison-Wesley.

Kemenkes RI. (2018). *Hasil utama Riskesdas 2018* [Main results of Basic Health Research 2018]. Kementerian Kesehatan Republik Indonesia.

Khawarizmi, I. N., Triayudi, A., & Sholihati, I. D. (2020). Diagnosa depresi pada mahasiswa menggunakan metode certainty factor dan forward chaining [Diagnosis of depression in students using certainty factor and forward chaining methods]. *INTI Nusa Mandiri*, *14*(2), 239–246. https://doi.org/10.33480/inti.v14i2.1173

Kohn, R., Saxena, S., Levav, I., & Saraceno, B. (2004). The treatment gap in mental health care. *Bulletin of the World Health Organization*, *82*(11), 858–866.

Kusumadewi, S., & Wahyuningsih, H. (2020). Model sistem pendukung keputusan kelompok untuk penilaian gangguan depresi, kecemasan, dan stres berdasarkan DASS-42 [A group decision support system model for the assessment of depression, anxiety, and stress disorders based on DASS-42]. *Jurnal Teknologi Informasi dan Ilmu Komputer*, *7*(2), 219–228. https://doi.org/10.25126/jtiik.2020721052

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Psychology Foundation.

Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, *13*, 23–47. https://doi.org/10.1146/annurev-clinpsy-032816-044949

Nada, Q., Herdiana, I., & Andriani, F. (2022). Testing the validity and reliability of the Depression Anxiety Stress Scale (DASS-21) instrument for individuals with psychodermatology. *Psympathic: Jurnal Ilmiah Psikologi*, *7*(2), 153–168. https://doi.org/10.21580/pjpp.v7i2.11802

Nopi, N. P., Munandar, M. H., Irawan, F., & Lubis, J. R. (2022). Sistem pakar mendiagnosa gangguan mental pada diri seseorang menggunakan metode certainty factor [Expert system for diagnosing mental disorders in individuals using the certainty factor method]. *Journal of Applied Computer Science and Technology*, *3*(1), 157–164. https://doi.org/10.52158/jacost.v3i1.307

Nunes, L. C., Pinheiro, P. R., Cavalcante, T. P., & Pinheiro, M. C. (2015). Handling diagnosis of schizophrenia by a hybrid method. *Computational and Mathematical Methods in Medicine*, *2015*, Article 987298. https://doi.org/10.1155/2015/987298

Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., & Lozano, G. (2012). The Depression Anxiety Stress Scales–21 (DASS-21): Further examination of dimensions, scale reliability, and correlates. *Journal of Clinical Psychology*, *68*(12), 1322–1338. https://doi.org/10.1002/jclp.21908

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., & Unutzer, J. (2018). The Lancet Commission on global mental health and sustainable development. *The Lancet*, *392*(10157), 1553–1598. https://doi.org/10.1016/S0140-6736(18)31612-X

Putra, R. S., & Yuhandri, Y. (2021). Sistem pakar dalam menganalisis gangguan jiwa menggunakan metode certainty factor [Expert system for analyzing mental disorders using the certainty factor method]. *Jurnal Sistim Informasi dan Teknologi*, *3*(4), 227–233. https://doi.org/10.37034/jsisfotek.v3i4.70

Raharja, B., Samudera, E. B., Lay, F., & Hansun, S. (2022). Expert system for depression detection in teenagers. *System Research & Information Technologies*, *2022*(2), 143–150. https://doi.org/10.20535/SRIT.2308-8893.2022.2.12

Ramli, M., Ariff, M. F. M., & Zaini, S. (2007). Translating, adapting and validating the DASS-21 in a Malaysian context. *Malaysian Journal of Medical Sciences*, *14*(1), 22–31.

Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Saxena, S., Thornicroft, G., Knapp, M., & Whiteford, H. (2007). Resources for mental health: Scarcity, inequity, and inefficiency. *The Lancet*, *370*(9590), 878–889. https://doi.org/10.1016/S0140-6736(07)61239-2

Shamseldin, T., et al. (2025). Artificial intelligence for predicting depression, anxiety, and stress using psychometric data. *Scientific Reports*, *15*, *Article 37282*. https://doi.org/10.1038/s41598-025-21301-1

Shortliffe, E. H., & Buchanan, B. G. (1975). A model of inexact reasoning in medicine. *Mathematical Biosciences*, *23*(3–4), 351–379. https://doi.org/10.1016/0025-5564(75)90047-4

Sulak, S. A., & Köklü, N. (2024). Analysis of depression, anxiety, stress scale (DASS-42) with methods of data mining. *European Journal of Education*. https://doi.org/10.1111/ejed.12778

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., & Kessler, R. C. (2017). Undertreatment of people with major depressive disorder in 21 countries. *The British Journal of Psychiatry*, *210*(2), 119–124. https://doi.org/10.1192/bjp.bp.116.188078

Torous, J., Myrick, K. J., Rauseo-Ricupero, N., & Firth, J. (2020). Digital mental health and COVID-19: Using technology today to accelerate the curve on access and quality tomorrow. *JMIR Mental Health*, *7*(3), e18848. https://doi.org/10.2196/18848

Turban, E., Sharda, R., & Delen, D. (2020). *Analytics, Data Science, & Artificial Intelligence: Systems for Decision Support* (11th ed.). Pearson.

Vasgi, D. B. P., Urkudkar, C., Ghaisas, I., Karale, T., & Lole, P. (2022). Mental health chatbot and mental illness identification system. *International Journal for Research in Applied Science and Engineering Technology*, *10*(9), 1278–1285. https://doi.org/10.22214/ijraset.2022.46791

Vos, T., et al. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019. *The Lancet*, *396*(10258), 1204–1222. https://doi.org/10.1016/S0140-6736(20)30925-9

Windarsyah, W., Khatimi, H., & Maulana, R. (2017). Sistem pakar diagnosa jenis gangguan jiwa skizofrenia menggunakan kombinasi metode forward chaining dan certainty factor [Expert system for diagnosing types of schizophrenia using a combination of forward chaining and certainty factor methods]. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, *2*(2), 51–58. https://doi.org/10.20527/jtiulm.v2i2.20

World Health Organization. (2022). *World mental health report: Transforming mental health for all*. https://www.who.int/publications/i/item/9789240049338

Yuwono, D. T., Fadlil, A., & Sunardi, S. (2019). Comparative analysis of Dempster-Shafer method and certainty factor method on personality disorders expert systems. *Scientific Journal of Informatics*, *6*(1), 12–21. https://doi.org/10.15294/sji.v6i1.17216

Zhu, J., Li, W., & Xie, H. (2023). Beyond fixed thresholds: Challenges in automated psychological screening and the role of fuzzy reasoning. *Journal of Medical Systems*, *47*(1), 12–25. https://doi.org/10.1007/s10916-023-01912-w