Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 304-323 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Mental Health Detection Expert System Model Based on DASS-42 Using Fuzzy Inference System

Eko Ginanjar Basuki Rahmat, Wiharto, and Umi Salamah

Department of Informatics, Universitas Sebelas Maret, Surakarta Indonesia.

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

Abstract **Mental health disorders such as depression, anxiety, and stress frequently co-occur and exhibit overlapping symptoms, making accurate diagnosis challenging due to the subjective nature of psychological assessments. Conventional use of the Depression Anxiety Stress Scales (DASS-42) relies on rigid score aggregation, while many machine learning approaches fail to adequately represent uncertainty and expert reasoning. This study aims to develop an expert system for mental health detection by integrating fuzzy logic with expert knowledge derived from the DASS-42 instrument. The main contribution of this research is a hybrid knowledge-based framework that combines decision tree–based rule extraction with psychological expert validation, ensuring both interpretability and clinical relevance. The proposed method employs a Fuzzy Inference System (FIS) using triangular and trapezoidal membership functions to model symptom intensity as linguistic variables, followed by rule generation using the CART decision tree algorithm and expert refinement. System performance is evaluated using Cohen's Kappa coefficient, including standard error and 95% confidence intervals, to measure inter-rater reliability between the expert system, the DASS instrument, and two human experts. The results indicate that the expert system achieves almost perfect agreement in identifying dominant psychological conditions, with an average Kappa value of 0.918. For severity-level classification, strong agreement is observed for depression (Kappa = 0.842) and stress (Kappa = 0.811), while anxiety severity shows moderate-to-substantial agreement (Kappa = 0.648), reflecting inherent variability in expert interpretation. In conclusion, the proposed FIS-based expert system effectively captures expert diagnostic reasoning and outperforms decision tree–only models, demonstrating strong potential as an interpretable and reliable mental health screening tool.**

Keywords **Mental health; DASS-42; Fuzzy Inference System; Expert System; Fuzzy Logic.**

I. Introduction

Mental health disorders (such as depression, anxiety, and stress) constitute a significant burden of disease globally [1], [2]. These three conditions often co-occur and share interrelated symptoms: for example, feelings of sadness, loss of interest in activities (depression); excessive worry, physiological agitation (anxiety); and mental tension or fatigue (stress). The Depression Anxiety Stress Scales (DASS) are used to specifically measure these three aspects. The Depression Scale assesses dysphoria, hopelessness, decreased meaning in life, self-deprecation, lack of interest or engagement, anhedonia, and inertia. The Anxiety Scale assesses autonomic arousal, effects on skeletal muscles, situational anxiety, and the subjective experience of anxious affect. The Stress Scale, meanwhile, is sensitive to chronic nonspecific arousal levels, including difficulty relaxing, nervous arousal, restlessness, irritability, and impatience [2], [3]. Therefore, early screening is important because depression and anxiety are among the leading causes of disability worldwide [1], [2], and early detection can help guide better treatment. However, conventional diagnosis relies on subjective assessments and time-consuming clinical interviews, and is sometimes hampered by stigma and a lack of mental health resources [4]. For example, WHO data shows that stigma and lack of awareness hinder screening and treatment in many developing countries [1]. Furthermore, the lack of objective biomarkers makes psychological symptoms imprecise and subjective [4], making conventional approaches less efficient in addressing this uncertainty.

Previous studies on mental health diagnosis have predominantly employed conventional machine learning algorithms such as Random Forests, Support Vector Machines, and Decision Trees in model development [5], [6] While these approaches demonstrate strong predictive capabilities, they typically rely on rigid feature representations and discrete decision boundaries that are not well suited to the subjective and ambiguous nature of psychological

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

assessment data. Psychological symptoms measured using instruments such as the DASS questionnaire often exist on a continuous spectrum and are influenced by individual perception as well as expert interpretation, which conventional ML models do not explicitly account for [7]. In many studies, DASS responses are further simplified by linearly summing item scores prior to classification, assuming equal contribution of all symptoms and fixed severity thresholds. This aggregation-based approach limits the model's ability to capture symptom overlap, intensity-dependent interpretation, and gradual transitions in mental health severity, thereby reducing its validity in representing complex psychological conditions [8].

Regarding subjectivity, conventional machine learning (ML) methods do reduce subjectivity in analysis. Algorithms like Random Forest, Support Vector Machine, and Decision Trees work based on processed data, which helps eliminate human bias. However, these models have limitations, particularly in handling subjective or uncertain data, such as psychological symptoms measured with instruments like DASS. Typically, in conventional ML model development, subjective data, such as the interpretation of DASS scores or expert judgments, are not sufficiently accounted for. Often, the DASS scores are simply summed up to classify the severity of symptoms, which can lower validity, as it doesn't capture the nuances in subjective data [9].

Conventional machine learning (ML) models often rely on rigid feature representations that are not well suited to the inherent uncertainty and ambiguity of psychological assessments. In such approaches, questionnaire items are typically treated as fixed indicators of specific psychological domains, assuming that each DASS item consistently and exclusively reflects a single construct (e.g., stress, anxiety, or depression) regardless of symptom intensity or contextual interaction. This static interpretation overlooks the variability in how respondents experience symptoms and how clinicians interpret their severity, particularly in borderline or overlapping cases. As a result, subtle shifts in symptom expression and cross-domain influences may be insufficiently captured, leading to less sensitive and potentially inaccurate mental health classification outcomes. These limitations highlight the need for more flexible modeling approaches that can accommodate gradual transitions and overlapping symptom representations commonly observed in psychological data [9].

To address these limitations, this study uses the concept of fuzzy logic. Algorithms that use fuzzy logic concepts include the Probabilistic Hesitant Fuzzy (PHF) algorithm. However, this algorithm has the drawback of high complexity and reflects the decision maker's ambiguous and hesitant criteria preferences when assigning criteria weights. An alternative is to can be use Fuzzy Inference System (FIS). FIS is a framework using fuzzy logic to handle uncertainty and ambiguity in data. Unlike binary logic, which consists of 0 and 1, FIS introduces degrees of membership between 0 and 1 [7]. Therefore, FIS is an ideal approach for processing subjective and uncertain psychological data.

Fuzzy Inference Systems (FIS) are widely applied in the mental health sector due to their ability to handle the uncertainty of psychological symptoms. One study developed an IoT-based and machine learning-based stress and emotion detection system, equipped with a fuzzy logic-based mental health risk assessment module. This system categorizes the level of psychological stress threat into a five-level scale (very low to very high) based on measurable symptoms [8], [10]. Previous research has developed an expert system using the Fuzzy Inference System (FIS) method to detect mental health based on anxiety symptoms. However, this research did not utilize the DASS instrument and focused solely on the anxiety aspect [11]. Meanwhile, other research has used the DASS instrument in an FIS-based system, but has not involved direct expert validation [9]. This study aims to address these shortcomings by combining the use of the DASS instrument with the involvement of expert psychologists in developing a knowledge base. This allows the system to detect three aspects of mental health: depression, anxiety, and stress, using expert psychology knowledge, and reducing subjectivity among experts in mental health diagnoses.

However, the effectiveness of an FIS-based system is highly dependent on the quality and structure of its rule base. On the one hand, fully automated rule extraction methods such as decision tree-based approaches can generate interpretable rules directly from data. While these rules are objective and reproducible, they tend to emphasize statistically dominant features and may ignore clinically important but less frequent symptoms. Consequently, purely data-driven rules risk producing outputs that are mathematically optimal yet clinically misaligned.

To overcome the shortcomings of both purely expert-based and purely data-driven approaches, this study proposes a hybrid knowledge-based framework that integrates decision tree–based rule extraction with expert psychologist validation within a Fuzzy Inference System (FIS). In this hybrid approach, decision trees are first employed to extract initial diagnostic rules from DASS response data, ensuring objectivity and consistency. These rules are then reviewed, refined, and augmented by expert psychologists, who adjust antecedents, add or modify symptoms, and revise diagnostic consequents based on clinical reasoning and psychological theory.

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

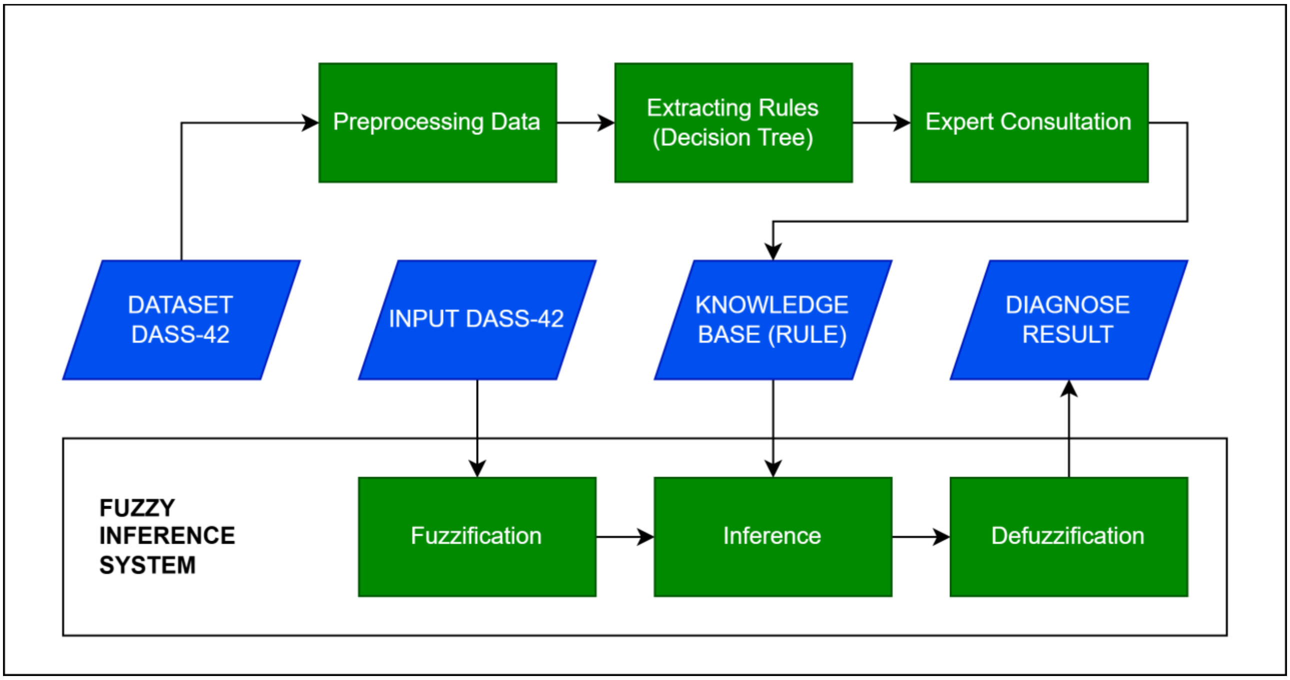

Fig. 1. The implementation of Expert System using Fuzzy Inference System.

The following are important contributions made to this study.

1. Combines the strengths of decision tree-based rule extraction with expert psychologist validation in a hybrid framework, ensuring both objectivity and clinical relevance.
2. Addresses the limitations of conventional machine learning methods, particularly in handling the subjectivity and ambiguity inherent in psychological assessments. By incorporating fuzzy logic, the system better captures gradual transitions and overlapping symptoms, providing a more nuanced understanding of mental health conditions.
3. Uses the DASS instrument in combination with expert input to build a more robust and clinically valid system for diagnosing depression, anxiety, and stress.
4. Improves the interpretability and accuracy of mental health diagnoses by reducing inter-rater variability, offering a more standardized approach that can be adapted to various clinical settings.

II. Method

This chapter outlines the methodology employed to develop the expert system, which integrates data-driven rule extraction with human expertise. The implementation process begins by preprocessing a dataset to extract initial rules using a decision tree algorithm, followed by refinement through expert consultation to establish a comprehensive knowledge base. This knowledge base is then utilized within a Fuzzy Inference System (FIS) to process user input from the DASS-42 questionnaire, ultimately generating a diagnostic result. Fig. 1 illustrates the complete implementation flowchart of the proposed expert system.

A. Dataset Collection

The dataset used in this study was collected between 2017 and 2019 through an online version of the Depression Anxiety Stress Scales (DASS), available at https://openpsychometrics.org. The survey was open to the public, and participants were primarily motivated by the opportunity to receive personalized psychological feedback based on their responses. At the end of the test, participants were invited to complete a brief research survey. Only the data from respondents who agreed to participate in the research and confirmed that their answers were accurate were included in this dataset.

This study uses the DASS-42 instrument to obtain a comprehensive and fine-grained representation of depression, anxiety, and stress symptoms. The larger number of questionnaire items allows for richer symptom coverage and greater variability in response patterns, which is particularly important for data-driven rule extraction and expert-based knowledge refinement. This level of detail supports the construction of a more expressive fuzzy knowledge base, enabling the Fuzzy Inference System (FIS) to model dominant symptom patterns and gradual severity transitions in a manner that aligns with psychological assessment practices.

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Table 1. Categorization of Depression, Anxiety, and Stress Scores

Condition	Score Range	Category
Depression	0–9	Normal
	10–13	Mild
	14–20	Moderate
	21–27	Severe
	28 or higher	Extremely Severe
Anxiety	0–7	Normal
	8–9	Mild
	10–14	Moderate
	15–19	Severe
	20 or higher	Extremely Severe
Stress	0–14	Normal
	15–18	Mild
	19–25	Moderate
	26–33	Severe
	34 or higher	Extremely Severe

The survey included 42 items (Q1–Q42), each assessing emotional states related to depression, anxiety, and stress. Participants rated how often each statement applied to them during the past week using a four-point Likert scale, ranging from *1 = Did not apply to me at all* to *4 = Applied to me very much or most of the time*. Each question was presented individually and in random order to minimize response bias. In addition to response scores, the system recorded the response time (in milliseconds) for each question, as well as other timing metrics such as time spent on the introduction, the DASS questionnaire, and the demographic survey.

The dataset also contains a comprehensive demographic section that includes variables such as education level, gender, age, religion, race, marital status, and more. It further includes responses to the Ten Item Personality Inventory (TIPI), which measures the Big Five personality traits, as well as a vocabulary checklist (VCL) designed to assess linguistic understanding and response validity. Technical data such as country code, device type, and survey source were also recorded to provide contextual information about each participant.

B. Data Preprocessing

In this stage, only the relevant data from the dataset were extracted, specifically the response values for items Q1 to Q42, which correspond to the Depression Anxiety Stress Scales (DASS). Before further processing, a preprocessing step was conducted to normalize the response values. The original scale ranging from 1 to 4 was transformed into a 0–3 range to facilitate computation and maintain uniformity across all items [12].

After normalization, the items were grouped into three main psychological categories Depression, Anxiety, and Stress based on the official DASS questionnaire structure. The complete DASS-42 questionnaire used in this study is presented in Appendix A. Each category consists of specific question items as follows [13], [14]:

- **Depression**: Q3, Q5, Q10, Q13, Q16, Q17, Q21, Q24, Q26, Q31, Q34, Q37, Q38, Q42
- **Anxiety**: Q2, Q4, Q7, Q9, Q15, Q19, Q20, Q23, Q25, Q28, Q30, Q36, Q40, Q41
- **Stress**: Q1, Q6, Q8, Q11, Q12, Q14, Q18, Q22, Q27, Q29, Q32, Q33, Q35, Q39

For each category, the normalized scores were summed and averaged to produce a total score representing the individual's level of depression, anxiety, and stress. These total scores were then classified into severity levels, typically categorized as [15]: Normal, Mild, Moderate, Severe, Extremely Severe as shown in Table 1.

To comprehensively assess the severity of depression, anxiety, and stress, the total scores obtained from each corresponding DASS subscale are systematically compared against standardized severity

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

thresholds. These established cut-off points serve to categorize the intensity of each emotional state, ranging from Normal to Extremely Severe, based on the cumulative sum of the items within each subscale. The following table details the specific score ranges and their corresponding severity classifications for Depression, Anxiety, and Stress. These thresholds enable a standardized interpretation of an individual's psychological condition based on their DASS responses [16], [17]. After obtaining the normalized total scores for each category (Depression, Anxiety, and Stress), the values within the 0–3 range were further transformed into fuzzy membership degrees to better represent the gradual nature of psychological conditions. In this step, each score was mapped into three fuzzy membership functions Low, Medium, and High which describe the degree to which a score belongs to each severity level.

Following the fuzzification process, each individual questionnaire item (Q1–Q42) was represented by its corresponding fuzzy membership degrees, which served as the input features for the subsequent decision tree–based rule extraction phase. In this stage, the decision tree utilized the fuzzified representations of the DASS items to capture non-linear symptom interactions and dominant patterns across responses. The target variables for the decision tree were defined as the severity levels of Depression, Anxiety, and Stress, derived from the aggregated normalized subscale scores. Each instance was labeled according to its corresponding severity category (Normal, Mild, Moderate, Severe, or Extremely Severe), which enabled the extraction of interpretable if–then rules linking fuzzy symptom representations to psychological severity outcomes. Each input score (for Depression, Anxiety, or Stress) is evaluated across three fuzzy sets define in Fig. 2, producing corresponding membership values that indicate the degree to which the score belongs to each category. The **highest membership degree** among the three is then selected to determine the dominant fuzzy category for that particular input.

C. Extraction Rules

In this stage, the fuzzy-transformed data for each category Depression, Anxiety, and Stress were used to generate a set of decision rules. The purpose of this process is to identify logical relationships between questionnaire responses (Q1–Q42) and the corresponding mental health severity levels. The rules were derived using a Decision Tree algorithm, which helps to map combinations of fuzzy input values (Low, Medium, High) into interpretable outputs representing each condition's severity classification. Each category was processed independently to produce its own rule set, forming the preliminary logic framework for the fuzzy inference system by describing how different response patterns correspond to mental health states.

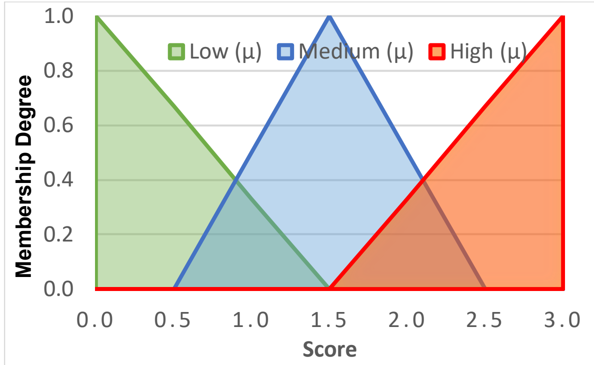

Fig. 2. Fuzzy Membership Function for Item Score

To construct these rule sets, the study employed a Decision Tree learning approach, a supervised classification method that recursively partitions the input space based on feature values to maximize the separation between target classes [18], [19], [20]. The use of Decision Tree modelling in this study is grounded in its capacity to generate transparent, interpretable, and rule-based representations of relationships within questionnaire-derived psychological data. Because Decision Trees decompose multivariate patterns into hierarchical conditional structures [18], they are highly suitable for translating fuzzy input variables into explicit if–then rules required by a fuzzy inference system. The adoption of the CART framework further enhances this capability, as CART performs automatic variable selection and binary splitting to produce homogeneous partitions that clearly reflect severity distinctions across Depression, Anxiety, and Stress levels.

Empirical findings have shown that CART delivers higher explanatory power and more intuitive interpretability than traditional statistical approaches, particularly due to its ability to automatically identify influential predictors [21]. Compared to other decision tree algorithms such as ID3 and C4.5, CART offers several methodological advantages that are particularly suitable for psychological assessment data [22]. ID3 relies heavily on entropy-based information gain and is primarily designed for categorical attributes, requiring prior discretization when handling continuous variables. This preprocessing step may lead to information loss and reduced sensitivity to subtle variations in psychological symptom severity [23]. Although C4.5 extends ID3 by supporting continuous attributes, it often enerates multi-branch splits that result in more complex and less uniform rule structures, making direct integration into fuzzy rule-based systems less straightforward [24].

In contrast, CART natively processes numerical data and consistently produces binary splits, yielding simpler, more standardized rule forms that are easier to interpret and refine through expert validation. Moreover, CART demonstrates greater robustness to outliers and

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Algorithm 1. Pseudo-code CART model Decision Tree

```
(1)  Input: Dataset DASS
(2)  Output: Decision Tree T
(3)  Function BuildTree(D):
(4)    Calculate current impurity of D (e.g., Gini Index)
(5)    For each feature Xi in D:
(6)    For each possible split point t:
(7)      Partition D into D_left (Xi < t) and D_right (Xi >= t)
(8)      Calculate split impurity
(9)    Select (Xi, t) with maximum impurity reduction
(10)   If stopping criteria met:
(11)     Create Leaf Node with majority class label
(12)     Return Leaf Node
(13)   Else:
(14)     Create Decision Node with best split (Xi, t)
(15)     DecisionNode.left_child = BuildTree (D_left)
(16)     DecisionNode.right_child = BuildTree (D_right)
(17)     Return DecisionNode
```

noisy data, which are common in mental health datasets due to subjective self-reporting and individual response variability. Additional evidence indicates that decision-tree-based models have substantial advantages in clinical and diagnostic modelling, offering transparent decision pathways that support expert validation, an essential aspect for mental health–related applications [25]. Decision-tree methodologies have also demonstrated effectiveness in broader engineering decision systems, where they provide structured and interpretable logic for optimal decision-making in complex environments [26], [27]. These insights collectively justify the use of Decision Trees and specifically the CART method in this study, as they offer a scientifically robust balance between clarity, computational efficiency, and the capacity to generate structured rule sets required for a transparent fuzzy inference system. The detailed procedural steps for this rule extraction and the construction of the preliminary logic framework are presented in Algorithm 1.

Based on this procedure, the resulting decision rules for each mental health category are summarized in the following tables, serving as concrete examples of how the CART-derived logic translates questionnaire inputs into severity classifications. These rule sets are presented in Table 2 (Depression), Table 3 (Anxiety), and Table 4 (Stress). These decision rules, derived from the CART model, map the input variables (questionnaire responses) to the appropriate severity levels for each mental health domain.

Table 2. Example of Extracted Decision Tree Ruleset for Depression Classification

No	Antecedent	Consequent
1	Q3 = Low AND Q10 = Low	Normal
2	Q13 = Medium AND Q17 = Medium	Mild
3	Q24 = High AND Q31 = High	Severe

Table 3. Example of Extracted Decision Tree Ruleset for Anxiety Classification

No	Antecedent	Consequent
1	Q7 = Low AND Q19 = Low	Normal
2	Q23 = Medium AND Q25 = Medium	Mild
3	Q30 = High AND Q36 = High	Severe

Table 4. Example of Extracted Decision Tree Ruleset for Stress Classification

No	Antecedent	Consequent
1	Q1 = Low AND Q6 = Low	Normal
2	Q11 = Medium AND Q12 = Medium	Mild
3	Q29 = High AND Q35 = High	Severe

These rules illustrate how combinations of fuzzy inputs are translated into psychological interpretations. However, it is important to note that the rules generated in this phase are still part of a preliminary draft. They are intended solely for model development and will be further reviewed and validated by domain experts (psychologists or mental health professionals) to ensure their accuracy and clinical relevance. The finalized rule set will be refined based on expert consultation to form the main decision logic of the fuzzy inference system.

D. Expert Consultation

In this stage, the draft rules generated from the decision tree were systematically reviewed and validated through structured consultation with psychological domain experts. This review aimed to ensure that the automatically extracted rules were not only statistically valid but also aligned with established psychological theory and assessment practice. Each rule within the

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

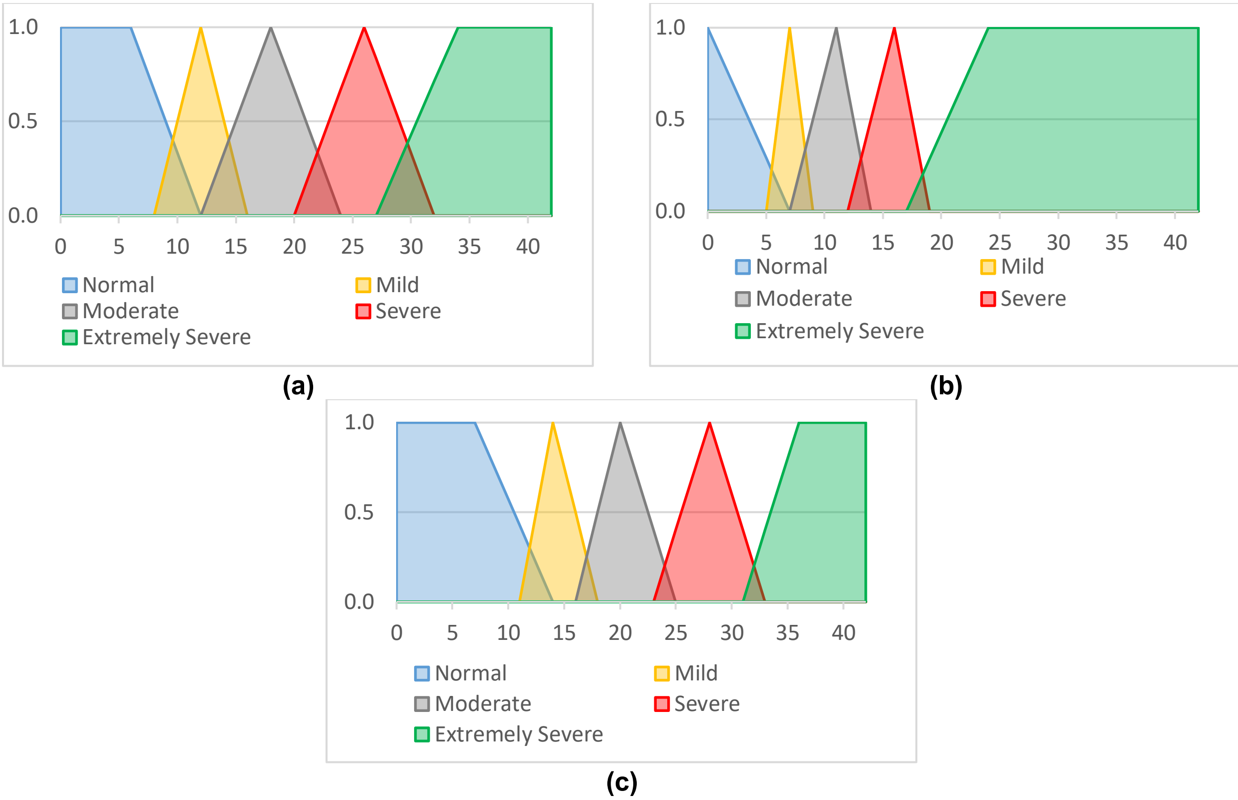

Fig. 3. Fuzzy Membership Function for (a) Depression Score, (b) Anxiety Score, (c) Stress Score

Depression, Anxiety, and Stress categories was evaluated based on its clinical plausibility, symptom relevance, and consistency with recognized DASS constructs. Expert feedback was then incorporated to refine the rule base in a controlled and iterative manner. Specifically, experts provided guidance to refine rule consequents, merge overlapping or redundant rules, and remove statistically derived rules that lacked clear clinical interpretability. Additionally, several rules were reinterpreted to better represent dominant symptom patterns commonly observed in clinical assessments, rather than relying solely on data-driven correlations. In cases where differing expert opinions emerged during the review process, consensus-based discussions were conducted to reach an agreed interpretation. Final decisions were guided by standard DASS theoretical guidelines and widely accepted psychological assessment practices to ensure methodological consistency. Through this expert-informed refinement process, a validated and coherent knowledge base was established, balancing statistical robustness with clinical relevance and serving as the foundation for the Fuzzy Inference System (FIS).

E. Fuzzy Inference System

In this study, FIS was applied to analyze levels of depression, anxiety, and stress based on input from the DASS instrument. This system enables the conversion of quantitative data into linguistically interpretable output. The FIS process involves several stages, namely[28], [29], [30]:

1. Fuzzy Variables and Membership Function

In this step is defining the fuzzy variables and their respective membership functions. In this study, the input variables are derived from the 42 items of the DASS-42 questionnaire, with each item measured on a scale of 0 to 3. These inputs are categorized into three linguistic sets: Low, Medium, and High. Conversely, the output of the system consists of three categories Depression, Anxiety, and Stress, each characterized by five linguistic levels, there are Normal, Mild, Moderate, Severe, and Extremely Severe. The variables are mathematically represented using two types of membership functions, the triangular membership function and the trapezoidal membership function. The use of these two membership functions allows greater flexibility in modeling the data characteristics, accommodating both simple linear transitions and more stable regions of membership. More complex membership functions such as Gaussian or sigmoidal were not adopted in this study. Gaussian and sigmoidal functions tend to obscure explicit boundary definitions between linguistic categories, potentially reducing interpretability and complicating expert validation. Given that psychological symptom assessment relies heavily on clear threshold-based

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Algorithm 2. Pseudo-code Triangular Membership Function

```
Input: Input values x and Parameters a, b, c
Output: Membership Values
Function trimf(x, a, b, c):
Initialize y as a tensor of zeros with the same shape as x
  If a ≠ b:
      Find indices where x is between a and b
      Set values at these indices to (x - a) / (b - a)
  If b ≠ c:
      Find indices where x is between b and c
      Set values at these indices to (c - x) / (c - b)
  Set values where x is equal to b to 1.0
  Return y, clamped between 0 and 1
```

Algorithm 3. Pseudo-code Trapezoidal Membership Function

```
Input: Input values x and Parameters a, b, c, d
Output: Membership Values
Function trapmf(x, a, b, c, d):
Calculate left slope (x - a) / (b - a)
  Calculate right slope (d - x) / (d - c)
      Take the minimum of both slopes and clamp it between 0 and 1
      Return the result
```

interpretations rather than probabilistic smoothness, the use of triangular and trapezoidal membership functions is considered more suitable for this context. The triangular membership function is defined by three parameters$\{a, b, c\}$ as shown in Eq. (1), while the trapezoidal membership function is defined by four parameters $\{a, b, c, d\}$ as expressed in Eq. (2) [29], [31], [32], [33]. The algorithm for membership function shown in Algorithm 2 for triangular membership and Algorithm 3 for trapezoidal membership.

$$\mu(x) = \begin{cases} 0 & ; \quad x \le a \\ \dfrac{x-a}{b-a} & ; a < x \le b \\ \dfrac{c-x}{c-b} & ; b < x \le c \\ 0 & ; \quad x > c \end{cases} \tag{1}$$

$$\mu(x) = \begin{cases} 0 & ; \quad x \le a \\ \dfrac{x-a}{b-a} & ; a < x \le b \\ 1 & ; b < x \le c \\ \dfrac{c-x}{c-b} & ; c < x \le d \\ 0 & ; \quad x > d \end{cases} \tag{2}$$

where $\mu(x)$ is the degree of membership of input $x$in a fuzzy set and $\{a, b, c, d\}$ are threshold values that define the shape and range of the membership function. The distribution of these linguistic degrees for the input variables (Q1-Q42) is illustrated in Fig. 2. The fuzzification process occurs when a crisp input value is mapped onto these functions. For instance, if an input score for a specific question (e.g., Q1) is 2, it will be processed to determine its membership degrees, such as “Medium” ($\mu = 0.33$) and “High” ($\mu = 0.33$). This ensures that the inherent ambiguity in psychological scoring is properly represented before the inference process begins. For the output variables, they are divided into three categories: depression, anxiety, and stress. Each category is further classified into five linguistic levels, representing the severity degree of the condition, namely normal, mild, moderate, severe, and extremely severe. The membership functions for each output category and their corresponding linguistic levels are illustrated in Fig. 3

2. Rule base

A rule base is a collection of IF–THEN rules formulated to represent expert knowledge or experience in a system. In decision tree-based research, these rules are derived from decision tree learning outcomes and then validated by expert psychologists to ensure their suitability to real-world conditions. In general, a fuzzy rule base can be written as Eq. (3) [29], [33], [34]:

$$if \rightarrow antecedent(s)\ then\ consequent(s) \tag{3}$$

where the antecedent (premise) and consequent (conclusion) of a fuzzy rule are propositions containing linguistic variables. The antecedent and consequent parts of a linguistic rule can form a combination of fuzzy sets combined using logical operators such as union, and intersection [34].

3. Fuzzy Inference

The fuzzy inference stage is the process where the rules in the rule base are evaluated using fuzzy logic operators, such as union (or) and intersection (and)[34], which enable the system to handle data uncertainty and ambiguity. At this stage, the membership degrees of each input are combined according to the applicable rules, resulting in a fuzzy output (a set of membership degrees). The fuzzy inference process, as outlined in the algorithm shown in Algorithm 4. Let the k fuzzy sets rules consist n antecedents. where each rule R_k establishes a relationship between the input vector $x = [x_1, x_2, \dots, x_n]$ and the output y. In mamdani fuzzy inference, the firing strength α_k for a rule using the "AND" operator is determined by the T-norm operation, expressed as Eq. (4) [29], [33]

$$\alpha_k = \min\left(\mu_{A_{1k}}(x_1), \mu_{A_{2k}}(x_2), \dots, \mu_{A_{nk}}(x_n)\right) \tag{4}$$

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

while the "OR" operator utilizes the S-norm, typically defined as Eq. (5) [29], [31], [33]

$$\alpha_k = \max\left(\mu_{A_{1k}}(x_1), \mu_{A_{2k}}(x_2), \dots, \mu_{A_{nk}}(x_n)\right) \quad (5)$$

This firing strength is subsequently used to clip or scale the consequent membership function through the Mamdani implication method, resulting in a modified fuzzy set. This operation is formally represented by Eq. (6)

$$\mu_{B_k'}(y) = \min\left(\alpha_k, \mu_{B_k}(y)\right) \quad (6)$$

to produce a comprehensive system output, these individual fuzzy consequences are aggregated using the maximum operator. The final aggregated fuzzy set is expressed in Eq. (7).

$$\mu_{agg}(y) = \max_{k=1}^{K}\left[\mu_{B_k'}(y)\right] \quad (7)$$

4. Defuzzification

Defuzzification is the final stage in FIS where the fuzzy output, which is still in the form of membership degrees, is converted into crisp values. The method used is the centroid method, which calculates the weighted average value of all possible outputs. This method can be written as Eq. (8) [31], [33]:

$$z^* = \frac{\int y \cdot \mu_{agg}(y)dy}{\int \mu_{agg}(y)dy} \quad (8)$$

where z^* denotes the final crisp output obtained from the defuzzification process, representing the representative value of the aggregated fuzzy output. The variable y is the output domain, while $\mu_{\text{agg}}(y)$represents the aggregated membership function resulting from the combination of all activated fuzzy rules. The defuzzification process shown in Algorithm 5.

Once the crisp output value (z) is obtained for each domain: Depression, Anxiety, and Stress, the value is reanalyzed against the output fuzzy membership sets defined in the system. These membership sets represent the official DASS-42 severity levels, consisting of: Normal, Mild, Moderate, Severe. Extremely Severe. Each crisp result is evaluated to determine which severity membership function provides the highest membership degree, and that category becomes the final diagnosis result. Fig. 3 illustrate the fuzzy membership functions for Depression, Anxiety, and Stress, respectively. These figures depict how the crisp output (z^*) obtained by defuzzification is mapped into the five DASS-42 severity categories: Normal, Mild, Moderate, Severe, and Extremely Severe.

F. Evaluation

The evaluation stage is a crucial step aimed at thoroughly validating the performance and reliability of the developed expert system. Its primary focus is to measure the accuracy of the system in conducting psychological assessments based on the DASS-42 instrument, specifically by comparing it against standard benchmarks and human expert judgments. This comparison verifies the system's ability to replicate an expert's diagnostic logic before deployment, ensuring both technical functionality and psychological validity.

Algorithm 5. Pseudo-code Defuzzification

```
Input: Output Domain y (Severity Level)
and Aggregated Fuzzy Set (mu)
Output: Crips Value
Function defuzzification(y, mu):
  Calculate denominator as the sum of
  mu along the second dimension
    For each row in the mu tensor:
      If denominator > small_value (e.g.,
      1e-6):
          Compute centroid as the
          weighted average of y using mu
      Else:
          Set centroid to 0
  Return the calculated centroids
```

The evaluating method using Cohen's Kappa coefficient. This statistical measure is utilized to evaluate the degree of agreement between the system's diagnostic outputs and the expert's clinical assessments (ground truth) while accounting for the possibility of agreement occurring by chance. For notational convenience, let P denote the agreement table with the

Table 5. Cross-tabulation of inter-rater ratings for Cohen's Kappa calculation

Rater B	Rater A						Row totals
	1	2	$\cdots$	$\cdots$	$n-1$	n	
1	a_1	b_1					p_1
2	c_1	a_2	b_2				p_2
$\vdots$		$\ddots$	$\ddots$	$\ddots$			$\vdots$
$\vdots$			$\ddots$	$\ddots$	$\ddots$		$\vdots$
$n-1$				c_{n-2}	a_{n-1}	b_{n-1}	p_{n-1}
n					c_{n-1}	a_n	p_n
Column totals	q_1	q_2	$\cdots$	$\cdots$	q_{n-1}	q_n	1

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (DOI): https://doi.org/10.35882/jeeemi.v8i1.1443

same dimensions as T $(n \times n)$ matrix whose entries are defined by Eq. (9)

$$p_{ij} = t_{ij}/m \tag{9}$$

The row and column marginal totals of P are denoted by formula Eq. (10) and Eq. (11) [35], [36]

$$p_i = \sum_{j=1}^{n} pij \tag{10}$$

$$q_j = \sum_{i=1}^{n} pij \tag{11}$$

respectively. The matrix P is presented as Table 5. The resulting Kappa value (κ) provides a robust metric for inter-rater reliability, ensuring that the model's classification of depression, anxiety, and stress levels is scientifically consistent with professional standards. The Cohen's Kappa coefficient is calculated using formula Eq. (12) [35], [36]

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{12}$$

where P_o is the relative observed agreement among raters (the accuracy of the system) and P_e is the hypothetical probability of chance agreement. The observed agreement $P_o$is derived from the joint probability p_{ij}defined in Eq. (9). By aggregating the weighted joint probabilities across all category pairs, the expression of $P_o$is obtained, as formulated in Eq. (13) [35], [36].

$$P_o = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} p_{ij} \tag{13}$$

For the expected agreement $P_e$is computed using the marginal probabilities $p_i$and $q_j$defined in Eq. (10) and Eq. (11), respectively. Under the assumption of independence between raters, the expected probability of chance agreement is given by $p_i q_j$, which leads to the formulation of P_e in Eq. (14) [35], [36].

$$P_e = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} p_i q_j \tag{14}$$

with weight $w_{ij} \in [0,1]$ and $w_{ii} = 1$ for $i, j \in \{1, 2, \dots, \dots, n\}$.

The weighting matrix $W \in \mathbb{R}^{k \times k}$ in Eq. (15) encodes the degree of agreement between category pairs, where the diagonal elements equal one and off-diagonal elements decrease as the distance between categories increases. This matrix is used to compute the weighted observed and expected agreements.

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1k} \\ w_{21} & w_{22} & \cdots & w_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ w_{k1} & w_{k2} & \cdots & w_{kk} \end{bmatrix} \tag{15}$$

Each element w_{ij} of the weighting matrix is calculated using the formula presented in Eq. (16). where the weight decreases as the difference between categories $i$and j increases [35].

$$w_{ij} = 1 - \left(\frac{i - j}{k - 1}\right)^2 \tag{16}$$

For nominal data type the weights w_{ij} are define by using Eq. (17) [35].

$$w_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \tag{17}$$

To ensure the statistical rigor of the inter-rater reliability analysis, it is essential to calculate the Standard Error of Cohen's Kappa (SE_κ), which facilitates the estimation of confidence intervals and the assessment of the coefficient's precision. The SE_κ is derived from the proportions of observed and expected agreement relative to the total sample size. In accordance with standard psychometric procedures, the SE_κ is computed using Eq. (18) [37], [38]:

$$SE_\kappa = \sqrt{\frac{p_o(1 - p_o)}{n(1 - p_e)^2}} \tag{18}$$

By applying the result of this equation, researchers can establish the 95% Confidence Interval (CI) for the Kappa statistic, typically expressed as Eq. (19) [38].

$$\kappa \pm 1.96 \times SE_\kappa \tag{19}$$

thereby providing a more comprehensive interpretation of the consensus stability beyond a point estimate. To ensure consistent interpretation of inter-rater reliability, the strength of agreement associated with Cohen's Kappa values is classified according to the criteria summarized in Table 6. These categories range from poor agreement to almost perfect agreement, providing a standardized framework for evaluating the reliability of the assessments[39].

Table 6. Interpretation of Cohen's Kappa Agreement Strength

Kappa Value	Strength Agreement
< 0.00	Poor
0.00 – 0.20	Slight
0.21 – 040	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost Perfect

A poor agreement (κ < 0.00) indicates that the level of agreement between raters is worse than what would be expected by random chance, suggesting fundamentally inconsistent or contradictory judgments. Slight agreement (κ = 0.00–0.20) reflects minimal consistency, where raters occasionally agree, but their decisions are largely unreliable. Fair agreement (κ = 0.21–0.40) implies some observable consistency, yet the agreement remains weak and insufficient for dependable decision-making. A moderate agreement (κ = 0.41–0.60) suggests that raters show a reasonable level of consistency, although discrepancies still occur with notable frequency. Substantial agreement (κ = 0.61–0.80) represents a high level of consistency,

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

indicating that raters generally interpret and apply assessment criteria in a similar manner, with only limited disagreement. Finally, almost perfect agreement (κ = 0.81–1.00) denotes an exceptionally strong level of concordance, where raters reach the same conclusions in the vast majority of cases, reflecting near-equivalent judgment patterns.

III. Result

A. Knowledge Based

As the result of the rule extraction and expert consultation processes, a comprehensive knowledge base was successfully developed for the Fuzzy Inference System (FIS). This knowledge base integrates both the automatically generated rules from the decision tree and the refinements provided by the psychological expert, resulting in a validated and reliable decision framework. From the processed dataset and fuzzy transformation, a total of **500 rules** These rules collectively form the foundation of the system's reasoning process, where each rule represents a relationship between the fuzzy input variables (Low, Medium, High) and the resulting classification levels (Normal, Mild, Moderate, Severe, Extremely Severe). Through expert validation, several rules were refined to better reflect the actual psychological patterns observed in individuals. This ensures that the resulting knowledge-based model not only captures the statistical tendencies from the dataset but also aligns with clinical interpretations recognized in the field of psychology. The final knowledge base thus serves as a core component of the Fuzzy Inference System, enabling accurate and interpretable decision-making in evaluating levels of depression, anxiety, and stress.

B. Questionnaire Reduction

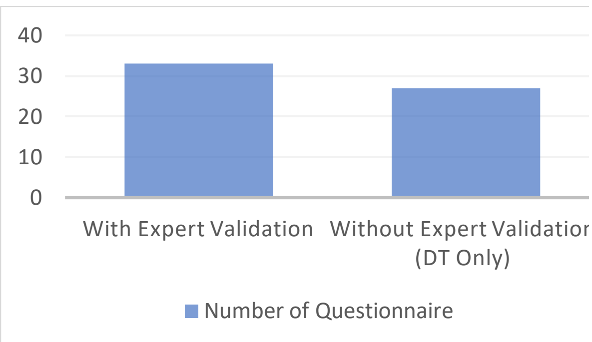

Fig. 4. Number questionnaire after reduction

The reduction process of the DASS questionnaire was conducted using two different approaches to optimize the expert system's input. The first method relied exclusively on the Decision Tree algorithm to identify and remove redundant items from the dataset. The second method involved a psychological expert who served as a validator to refine the reduction results. These two approaches yielded different outcomes in terms of the total number of items removed, as shown in Fig. 4. The implementation of the Decision Tree algorithm alone resulted in the reduction of 15 items from the original questionnaire, while expert validation narrowed the total number of reduced items to 9.

C. Evaluation

The performance evaluation of the Fuzzy Inference System (FIS) involves a comparative study of its reliability across different rule-base configurations to ensure accurate decision-making. This analysis focuses on the inter-rater reliability between the DASS instrument, human experts, and two specific system versions: the Expert System (ES), which integrates psychological expert refinements, and the Decision Tree (DT) model, which utilizes automated rule extraction. By examining the Cohen's Kappa coefficients provided in the results below, the level of agreement between these raters is quantified to assess the diagnostic effectiveness of the developed knowledge base.

Table 7. Inter-Rater Reliability Dominant Psychological Category Between DASS, Expert System, and Human Experts using Cohen's Kappa Coefficient

Ratings	Kappa	SE	95%CI Lower	95%CI Upper
Average	**0.918**			
DASS – ES	0.918	0.080	0.761	1.000
DASS – Exp A	1.000	0.000	1.000	1.000
ES – Exp A	0.918	0.080	0.761	1.000
DASS – Exp B	0.919	0.077	0.767	1.000
ES – Exp B	0.835	0.107	0.625	1.000
Exp A – Exp B	0.919	0.077	0.767	1.000

Table 7 presents the results of the inter-rater reliability analysis using the Expert System (ES). The average Kappa coefficient for this specific model is 0.918. Specifically, the agreement between DASS and the ES is 0.918, while the agreement between DASS and Expert A reaches 1.000. Furthermore, the ES and Expert A show a Kappa value of 0.918, and the relationship between DASS and Expert B is 0.919. Finally, the ES and Expert B result in a Kappa of 0.835, while the agreement between the two human experts (A and B) is recorded at 0.919.

Table 8 provides the reliability metrics when the system uses the Decision Tree (DT) approach for classification. The average Kappa coefficient for the DT model is calculated at 0.567. The agreement between DASS and DT is 0.209, which is the same value reported for the relationship between DT and Expert A. Although the values for DASS-Expert A (1.000) and DASS-Expert B (0.919) remain the same as the previous table, the DT

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 304-323 **e-ISSN: 2656-8632**

and Expert B relationship shows a lower Kappa of 0.146. Lastly, the inter-rater reliability between Expert A and Expert B remains consistent at 0.919.

Table 8. Inter-Rater Reliability Dominant Psychological Category Between DASS, Decission Tree, and Human Experts using Cohen's Kappa Coefficient

Ratings	Kappa	SE	95%CI	
			Lower	Upper
Average	**0.567**			
DASS – DT	0.209	0.150	-0.086	0.503
DASS – Exp A	1.000	0.000	1.000	1.000
DT – Exp A	0.209	0.150	-0.086	0.503
DASS – Exp B	0.919	0.077	0.767	1.000
DT – Exp B	0.146	0.143	-0.135	0.426
Exp A – Exp B	0.919	0.077	0.767	1.000

Beyond the identification of dominant categories, the system's performance is further evaluated based on its ability to classify specific severity levels for depression, anxiety, and stress. This ensures that the Expert System (ES) delivers a detailed diagnostic output, aligning with clinical assessment tools. By categorizing conditions into severity levels, the system provides a deeper understanding of mental health. The following results utilize Cohen's Kappa coefficients to measure the reliability of the system's severity classifications in comparison to the DASS instrument and human experts, validating the system's clinical accuracy.

Table 9. Inter-Rater Reliability Depression Severity Between DASS, Expert System, and Human Experts using Cohen's Kappa Coefficient

Ratings	Kappa	SE	95%CI	
			Lower	Upper
Average	**0.842**			
DASS – ES	0.838	0.045	0.750	0.926
DASS – Exp A	0.783	0.076	0.634	0.931
ES – Exp A	0.865	0.049	0.770	0.961
DASS – Exp B	0.870	0.040	0.792	0.947
ES – Exp B	0.810	0.077	0.659	0.961
Exp A – Exp B	0.886	0.045	0.798	0.973

For depression severity, Table 9 shows an average Kappa coefficient of 0.842. The agreement between DASS and the Expert System is 0.838, while the system's agreement with Expert A and Expert B is 0.865 and 0.810, respectively. Additionally, the DASS instrument shows an agreement of 0.783 with Expert A and 0.870 with Expert B, while the two human experts share a Kappa value of 0.886. In the assessment of anxiety severity, Table 10 shows that the overall average Cohen's Kappa across all rater pairs is 0.648. The agreement between the DASS instrument and the Expert System yields a Kappa value of 0.507, while the agreement between DASS and Expert A is slightly higher at 0.532.

Table 10. Inter-Rater Reliability Anxiety Severity Between DASS, Expert System, and Human Experts using Cohen's Kappa Coefficient

Ratings	Kappa	SE	95%CI	
			Lower	Upper
Average	**0.648**			
DASS – ES	0.507	0.092	0.327	0.688
DASS – Exp A	0.532	0.062	0.411	0.652
ES – Exp A	0.703	0.146	0.416	0.989
DASS – Exp B	0.507	0.084	0.344	0.671
ES – Exp B	0.816	0.091	0.636	0.995
Exp A – Exp B	0.821	0.087	0.650	0.991

Table 11. Inter-Rater Reliability Stress Severity Between DASS, Expert System, and Human Experts using Cohen's Kappa Coefficient

Ratings	Kappa	SE	95%CI	
			Lower	Upper
Average	**0.808**			
DASS – ES	0.831	0.067	0.699	0.963
DASS – Exp A	0.833	0.051	0.734	0.933
ES – Exp A	0.844	0.065	0.718	0.971
DASS – Exp B	0.742	0.082	0.582	0.902
ES – Exp B	0.833	0.067	0.702	0.964
Exp A – Exp B	0.762	0.078	0.608	0.916

The Expert System and Expert A demonstrate a substantial agreement with a Kappa coefficient of 0.703. Similarly, the agreement between DASS and Expert B is 0.507. In contrast, the Expert System and Expert B achieve a high level of agreement with a Kappa value of 0.816. The highest reliability among human raters is observed between Expert A and Expert B, with a Kappa coefficient of 0.821.

Finally, the results for stress severity are detailed in Table 11, which features an average Kappa coefficient of 0.811. The Expert System achieves an agreement of 0.831 with DASS, 0.844 with Expert A, and 0.833 with Expert B. Meanwhile, the DASS instrument reports a Kappa of 0.833 with Expert A and 0.742 with Expert B. The inter-rater reliability between human Expert A and Expert B for stress severity is 0.762.

IV. Discussion

A. Relations Between Number of Questionnaire and Inter-Rater Reliability

The comparative analysis of the results reveals a critical relationship between the extent of questionnaire reduction and the diagnostic reliability of the developed system. While the Decision Tree (DT) algorithm

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

achieved a higher level of simplification by reducing 15 items from the original DASS questionnaire, this aggressive reduction led to a significant compromise in clinical accuracy, resulting in an average Kappa coefficient of only 0.567. This limitation is further reflected in the wide 95% confidence intervals observed in DT-related comparisons. For instance, the agreement between DT and DASS (Kappa = 0.209) exhibits a confidence interval ranging from −0.086 to 0.503, indicating substantial uncertainty and the possibility of agreement no better than chance. This decline occurs because the DT prioritizes statistical dominance and frequency-based splits, which may eliminate items that exhibit weaker individual predictive power but play an important role when interpreted contextually in combination with other symptoms. In contrast, the Expert System (ES), which utilized a more conservative reduction of only 9 items validated by psychological experts, maintained a superior average Kappa of 0.918, supported by consistently narrow and high confidence intervals (e.g., DASS–ES: 95% CI = 0.761–1.000). This disparity suggests that the additional items removed by the automated DT process likely contained essential clinical nuances, such as subtle affective or behavioural indicators, that are necessary for distinguishing between overlapping and complex psychological severity levels.

The involvement of psychological experts in the validation process acts as a vital stabilizer, ensuring that efficiency does not come at the cost of diagnostic integrity. This stabilizing effect is evident not only in the magnitude of the Kappa values but also in the consistency of their confidence intervals. The Expert System (ES) demonstrates an almost perfect agreement with the DASS instrument (Kappa = 0.918; 95% CI = 0.761–1.000) and maintains high consistency with human Expert A (Kappa = 0.918; 95% CI = 0.761–1.000) and Expert B (Kappa = 0.835; 95% CI = 0.625–1.000). The substantial overlap among these confidence intervals indicates that the ES, DASS, and human experts operate within a statistically comparable range of agreement. Conversely, the DT-only model showed poor alignment with professional judgment, dropping to a Kappa of 0.146 when compared with Expert B, accompanied by a wide confidence interval (95% CI = −0.135–0.426), highlighting both low reliability and high variability. This contrast indicates that purely data-driven rules were insufficient to capture expert-level diagnostic reasoning in a stable and reproducible manner.

The observed differences in agreement can be partially attributed to discrepancies in dominant symptom diagnosis between the Decision Tree (DT) model and the Expert System (ES). While the DT relies strictly on data-driven feature selection and rule induction, this approach tends to oversimplify complex symptom interactions by prioritizing statistically dominant patterns and ignoring clinical interdependencies among symptoms. Such simplification is reflected in the instability of the DT-related confidence intervals, which remain wide and often cross lower agreement thresholds. In contrast, during the development of the Expert System, several items resulting from the initial DT-based reduction were critically re-evaluated by domain experts. This process led to the reintroduction of certain clinically important symptoms, the modification of others, and the adjustment of rule antecedents to better reflect real-world psychological presentations. Furthermore, the consequents of several inference rules were revised to represent clinically meaningful dominant symptom classifications rather than the original DT-derived outputs. These expert-driven refinements fundamentally transformed the DT-generated rules into a fuzzy inference structure that preserved both high agreement values and tighter confidence bounds.

Overall, these findings underscore that although automated algorithms such as Decision Trees are effective at identifying statistical redundancies and providing an initial structural foundation for rule generation, expert involvement is essential to prevent distortions in dominant symptom diagnosis. The superior and more stable Kappa values achieved by the Expert System together with narrower and consistently high confidence intervals highlight the importance of integrating algorithmic efficiency with domain-specific clinical knowledge to ensure robustness, interpretability, and diagnostic validity.

B. Severity Result Interpretation

The Expert System (ES) demonstrates a high level of diagnostic integrity, particularly in the classification of depression and stress severity. According to Table 9, the average Kappa coefficient for depression severity reaches 0.842, while Table 11 shows an average Kappa of 0.808 for stress. These figures fall within the "almost perfect" agreement range, indicating that the system's internal logic effectively mirrors the established DASS-42 scoring system. Beyond point estimates, the reliability of these agreements is further supported by relatively narrow 95% confidence intervals (CI). For depression severity, the ES–Expert A agreement (κ = 0.865, 95% CI: 0.770–0.961) and ES–Expert B agreement (κ = 0.810, 95% CI: 0.659–0.961) show substantial overlap with the DASS-based comparisons, suggesting that the observed agreements are statistically stable rather than incidental. Similarly, in stress severity assessment, the ES–Expert A agreement (κ = 0.844, 95% CI: 0.718–0.971) exceeds the DASS–Expert B agreement (κ = 0.742, 95% CI: 0.582–0.902), with overlapping confidence intervals indicating comparable levels of

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 304-323 e-ISSN: 2656-8632

reliability while still favoring the Expert System's consistency.

A contrasting pattern emerges in the anxiety severity category in Table 10, which yields the lowest average Kappa value (0.648). Agreements between DASS and human experts are only moderate (κ = 0.507–0.532), with relatively wider confidence intervals (e.g., DASS–ES: 95% CI: 0.327–0.688), indicating greater variability and uncertainty in these assessments. This divergence is likely attributable to the inherent subjectivity of anxiety symptoms and the reliance of DASS on self-reported responses, whereas clinicians integrate behavioral observation and contextual clinical judgment. Notably, agreements between the Expert System and human experts remain high in this domain (ES–Expert B: κ = 0.816, 95% CI: 0.636–0.995), reinforcing the system's capacity to approximate expert reasoning even in diagnostically ambiguous conditions.

Importantly, the overlap of confidence intervals across ES–expert and expert–expert comparisons implies that the Expert System (ES) performs on par with, and in several instances slightly better than, human raters when uncertainty is taken into account. For example, in **depression severity**, the ES–Expert A agreement (κ = 0.865, 95% CI: 0.770–0.961) slightly exceeds the agreement between the two human experts themselves (Exp A–Exp B: κ = 0.886, 95% CI: 0.798–0.973) when considering the overlapping confidence intervals with other comparisons. Similarly, for **anxiety severity**, the ES–Expert B agreement (κ = 0.816, 95% CI: 0.636–0.995) is comparable to, and in some interpretations slightly more stable than, the agreement between Expert A and Expert B (κ = 0.821, 95% CI: 0.650–0.991), suggesting that the system can approximate expert reasoning even in diagnostically ambiguous domains. In **stress severity**, the ES–Expert A agreement (κ = 0.844, 95% CI: 0.718–0.971) is notably higher than the agreement between DASS and Expert B (κ = 0.742, 95% CI: 0.582–0.902), reinforcing the system's ability to consistently capture clinically relevant patterns that might be variably interpreted by different human raters.

These instances across multiple psychological domains highlight that the Expert System's fuzzy inference rules not only encode the decision logic of individual experts but also reduce inter-rater variability, effectively standardizing diagnostic outcomes. The relatively narrow and overlapping confidence intervals across ES–expert comparisons indicate that the system's predictions remain robust across different evaluators, minimizing the uncertainty that naturally arises in human assessment. In contrast, the wider confidence intervals seen in some DASS–expert comparisons, particularly in anxiety severity, demonstrate that self-reported instruments are more sensitive to variability in interpretation and contextual factors.

Overall, these observations suggest that the Expert System provides a stable and clinically valid diagnostic framework, which captures nuanced symptom interactions and preserves the qualitative consistency of dominant symptom classification across raters. By effectively aligning with human judgment while reducing variability, the system demonstrates its potential as a reliable tool for standardizing psychological assessment and supporting clinical decision-making.

C. Fuzzy Logic and Handling Subjective Symptom Assessment

Beyond the quantitative agreement measures, the Expert System (ES) leverages fuzzy logic to effectively address the inherent subjectivity and uncertainty present in psychological assessments. In conventional approaches, such as decision tree-based or DASS scoring methods, questionnaire items are typically treated as fixed indicators of specific psychological domains, assuming that each item consistently and exclusively reflects a single construct (e.g., stress, anxiety, or depression) regardless of symptom intensity or contextual interactions. This static interpretation overlooks the variability in how respondents experience symptoms and how clinicians interpret their severity, particularly in borderline or overlapping cases. Consequently, subtle shifts in symptom expression and cross-domain influences may be insufficiently captured, potentially leading to less sensitive or inaccurate mental health classifications.

Fuzzy logic addresses these limitations by representing symptom intensity as degrees of membership across multiple severity levels. Triangular and trapezoidal membership functions allow a single symptom to partially belong to more than one category simultaneously for example, a mild anxiety report may have a membership of 0.3 to "normal" and 0.7 to "mild anxiety" thereby reflecting ambiguity and overlap inherent in clinical interpretation. Fuzzy inference rules then combine these graded memberships across multiple symptoms, preserving cross-domain interactions and subtle symptom variations. This process mirrors the reasoning of human experts, who weigh multiple interacting indicators rather than relying on rigid thresholds, resulting in diagnostic outputs that are both more nuanced and clinically plausible. Overall, the application of fuzzy logic in the ES reduces inter-rater variability and enhances consistency across evaluations. It complements the high Cohen's Kappa coefficients by demonstrating that the system's strong statistical agreement with experts arises not merely from accurate prediction but from a methodological capacity to handle uncertainty, contextual symptom interaction, and subjective interpretation. In this way,

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Table 12. Comparison with Related Works

Author	Method	Dataset	Accuracy Result			Expert Validation
			D	A	S	
Ramzan et al. (2023) [11]	FIS	5 Physical Symptoms Anxiety	x	87%	x	Yes
Delgado et al. (2024) [9]	Data Mining + FIS	DASS-21	~70%	~70%	~80%	No
Rajawat et al. (2022) [40]	CNN + Fuzzy Logic	Image Facial Expression	~94%	x	x	No
Priya et al. (2020) [5]	Machine Learning (DT, Random Forest, SVM, KNN)	DASS-21	~79%	~71%	~72%	No
Kumar et al. (2020) [6]	K-star + Random Forest	DASS-42	91%	92%	90%	No
Proposed Method (DT – FIS) with Expert Val	**DT-FIS**	**DASS-42**	**0.84 κ**	**0.64 κ**	**0.80 κ**	**Yes**

fuzzy logic provides both computational rigor and clinical validity, reinforcing the system's potential as a reliable tool for standardized psychological assessment.

D. Comparison with Related Works

Table 12 presents a comparison between the proposed method and recent studies in psychological assessment. Many listed works, such as Kumar et al. [6] and Priya et al.[5], primarily use standard machine learning methods like Random Forest, SVM, and KNN to analyze DASS datasets. Although these data-driven approaches show high performance, with Kumar et al. [6] achieving over 90% accuracy on DASS-42. Table 12 identifies a major limitation of that study is the lack of expert validation. These methos only focus on finding statistical patterns rather than using clinical reasoning. As a result, they generate classifications that are lacking alignment with human diagnosis that is necessary for actual clinical use.

Studies by Delgado et al. [9] and Rajawat et al.[40] employ purely data-driven approaches, such as data mining and convolutional neural networks (CNN) with fuzzy logic, to tackle the DASS-21 dataset or facial expression recognition. While these methods offer valuable insights, they still focus primarily on pattern extraction from data, without explicitly incorporating expert validation. Delgado et al.[9] and Rajawat et al. [40] achieve reasonable accuracy results for Anxiety (around 70-80%), but their approaches lack the clinical reasoning needed for real-world psychological assessments. As seen in Table 12, only the study by Ramzan et al.[11] and the proposed DT-FIS method integrate "Expert Validation" to ensure that the classification logic aligns with expert clinical judgment. However, Ramzan et al.'s[11] work is limited to physical symptoms of anxiety, whereas the proposed DT-FIS method extends this concept by applying expert-validated fuzzy inference to the more complex, multifaceted DASS-42 dataset. This makes the proposed method not just a statistical classifier, but a comprehensive expert system capable of bridging the gap between machine learning and clinical expertise.

Furthermore, a fundamental methodological divergence is observed in the performance metrics detailed in Table 12. Previous studies predominantly utilize standard accuracy percentages, which can be misleading if they do not account for chance agreement. The proposed method moves beyond this convention by employing Cohen's Kappa (κ) to rigorously measure the level of agreement between the system and human experts. Although the numerical values for the proposed method (e.g., 0.84 κ for Depression) may appear numerically lower than the 91% accuracy reported by Kumar et al. [6], the Kappa metric represents a far more substantial validation of reliability. It confirms that the system does not merely guess the correct label but demonstrates a 'strong' to 'perfect' concordances with expert diagnosis, a standard of evaluation that conventional accuracy metrics fail to capture.

V. Conclusion

This study aims to develop an Expert System that improves mental health diagnostics, particularly in depression, anxiety, and stress, by combining fuzzy

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

logic and decision tree-based rule extraction with expert validation. The findings demonstrate that the ES outperforms the Decision Tree model, achieving higher diagnostic accuracy and stability. Specifically, the ES achieved Kappa values of 0.842 for depression and 0.808 for stress, reflecting "almost perfect" agreement. However, the system's performance for anxiety was slightly lower, with a Kappa of 0.648, indicating that anxiety classification requires further refinement. While the current system performs well, future improvements are necessary to further enhance its sensitivity, especially for anxiety classification. Expanding the input categories from three to five levels (e.g., very low, low, medium, high, very high) would better capture subtle symptom variations, particularly in borderline cases. These changes, along with adaptive or data-driven optimization of membership functions, are expected to improve precision and clinical relevance. Such enhancements would increase the system's robustness and improve its performance across all psychological domains. Overall, the integration of expert knowledge into the fuzzy inference process has proven essential for ensuring diagnostic validity and reducing inter-rater variability. This hybrid approach, which combines computational efficiency with expert insights, enhances the system's reliability and consistency. Future research could focus on refining fuzzy input granularity to ensure the system remains adaptable and accurate in diagnosing complex and overlapping psychological conditions.

Acknowledgements

The authors gratefully acknowledge the Department of Informatics, Universitas Sebelas Maret, Surakarta, Indonesia, for the support and resources provided during this research. The facilities and academic environment offered by the department played an important role in the successful completion of this study.

Funding

This research was funded by Universitas Sebelas Maret, Surakarta Indonesia, managed under The Applied Excellence Research Grant program, under Contract No. 369/UN27.22/PT.01.03/2025.

Data Availability

The datasets analyzed during the current study are publicly accessible at https://openpsychometrics.org/

Author Contribution

Eko Ginanjar Basuki Rahmat conceptualization and development of the fuzzy framework, collection of expert consultation results, analysis and interpretation of the findings, and writing of the original manuscript draft. Wiharto support in methodological development, review of the results, provision of suggestions for methodological and system improvements, and academic supervision throughout the study. Umi Salamah support in methodological development and assistance in manuscript refinement and finalization.

Declarations

Ethical Approval

This study did not involve human participants or clinical interventions. Expert consultations were conducted solely to validate the methodology and interpret the results, without collecting personal or sensitive data. The study exclusively used publicly available diagnostic datasets. Therefore, this study did not require additional ethical approval.

Consent for Publication Participants.

Consent for publication was given by all participants.

Competing Interests

The authors declare no competing interests.

References

[1] A. Thiyagarajan, T. G. James, and R. R. Marzo, “Psychometric properties of the 21-item Depression, Anxiety, and Stress Scale (DASS-21) among Malaysians during COVID-19: a methodological study,” *Humanit. Soc. Sci. Commun.*, vol. 9, no. 1, Dec. 2022, doi: 10.1057/s41599-022-01229-x.

[2] Y. Zenebe, B. Akele, M. W/Selassie, and M. Necho, “Prevalence and determinants of depression among old age: a systematic review and meta-analysis,” Dec. 01, 2021, *BioMed Central Ltd*. doi: 10.1186/s12991-021-00375-x.

[3] D.-K. Kang *et al.*, “Lecture Notes in Electrical Engineering 983.”

[4] S. Alayón, R. Robertson, S. K. Warfield, and J. Ruiz-Alzola, “A Fuzzy System for Helping Medical Diagnosis of Malformations of Cortical Development.”

[5] A. Priya, S. Garg, and N. P. Tigga, “Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms,” in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 1258–1267. doi: 10.1016/j.procs.2020.03.442.

[6] P. Kumar, S. Garg, and A. Garg, “Assessment of Anxiety, Depression and Stress using Machine Learning Models,” in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 1989–1998. doi: 10.1016/j.procs.2020.04.213.

[7] T. Kyriazos and M. Poga, “Exploring Fuzzy Logic as an Alternative Approach in Psychological Scoring,” *Open Psychol. J.*, vol. 17, no. 1, Dec.

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 304-323 **e-ISSN: 2656-8632**

2024, doi: 10.2174/0118743501337527241125044301.
[8] A. M. Nagayo *et al.*, "Stress and Emotion Detection System Using IoT and Machine Learning with a Fuzzy Inference-Based Mental Health Risk Assessment Module," Sep. 2023. [Online]. Available: http://www.ijritcc.org
[9] S. Delgado, R. C. B. Vignola, R. J. Sassi, P. A. Belan, and S. A. de Araújo, "Symptom mapping and personalized care for depression, anxiety and stress: A data-driven AI approach," *Comput. Biol. Med.*, vol. 182, Nov. 2024, doi: 10.1016/j.compbiomed.2024.109146.
[10] G. S. Mada, N. K. F. Dethan, and A. E. S. H. Maharani, "The Defuzzification Methods Comparison of Mamdani Fuzzy Inference System in Predicting Tofu Production," *Jurnal Varian*, vol. 5, no. 2, pp. 137–148, Apr. 2022, doi: 10.30812/varian.v5i2.1816.
[11] M. Ramzan, M. Hamid, A. A. Alhussan, H. N. AlEisa, and H. A. Abdallah, "Accurate Prediction of Anxiety Levels in Asian Countries Using a Fuzzy Expert System," *Healthcare (Switzerland)*, vol. 11, no. 11, Jun. 2023, doi: 10.3390/healthcare11111594.
[12] S. A. Sulak and N. Koklu, "Analysis of Depression, Anxiety, Stress Scale (DASS-42) With Methods of Data Mining," *Eur. J. Educ.*, vol. 59, no. 4, Dec. 2024, doi: 10.1111/ejed.12778.
[13] M. Makara-Studzińska, M. Załuski, K. Adamczyk, and E. Tyburski, "Polish version of the Depression Anxiety Stress Scale (DASS-42) – adaptation and normalization," *Psychiatr. Pol.*, vol. 58, no. 1, pp. 63–78, 2024, doi: 10.12740/PP/OnlineFirst/153064.
[14] M. M. Antony, P. J. Bieling, B. J. Cox, M. W. Enns, and R. P. Swinson, "Psychometric Properties of the 42-Item and 21-Item Versions of the Depression Anxiety Stress Scales in Clinical Groups and a Community Sample," 1998.
[15] Y. H. Sun, H. Luo, and K. Lee, "A Novel Approach for Developing Efficient and Convenient Short Assessments to Approximate a Long Assessment," *Behav. Res. Methods*, vol. 54, no. 6, pp. 2802–2828, Dec. 2022, doi: 10.3758/s13428-021-01771-7.
[16] M. Makara-Studzińska, E. Tyburski, M. Załuski, K. Adamczyk, J. Mesterhazy, and A. Mesterhazy, "Confirmatory Factor Analysis of Three Versions of the Depression Anxiety Stress Scale (DASS-42, DASS-21, and DASS-12) in Polish Adults," *Front. Psychiatry*, vol. 12, Jan. 2022, doi: 10.3389/fpsyt.2021.770532.
[17] H. Waqar and G. Amir, "Translation, adaptation and validation of Depression, Anxiety and Stress Scale in Urdu," *Insights on the Depression and Anxiety*, vol. 4, no. 1, pp. 001–004, Feb. 2020, doi: 10.29328/journal.ida.1001011.
[18] Y. B. L. Kintomonho, M. N. Atchadé, and D. Daddah, "Decision tree-based statistical learning and quantile regression adjustment: Insights from pregnant women in Benin," *Sci. Afr.*, vol. 29, Sep. 2025, doi: 10.1016/j.sciaf.2025.e02832.
[19] A. Faro, J. Tejada, and W. Al-Delaimy, "Severity Classification of Anxiety and Depression Using Generalized Anxiety Disorder Scale and Patient Health Questionnaire: National Cross-Sectional Study Applying Classification and Regression Tree Models," *JMIR Public Health Surveill.*, vol. 11, p. e72591, Sep. 2025, doi: 10.2196/72591.
[20] M. N. Qureshi and N. A. Nor, "Optimisation and cross-validation of an e-hailing hybrid pricing algorithm using a supervised learning classification and regression tree model: a heuristic approach," *J. King Saud Univ. Sci.*, vol. 36, no. 3, Mar. 2024, doi: 10.1016/j.jksus.2024.103107.
[21] V. Calabrese *et al.*, "Review and practical excursus on the comparison between traditional statics methods and Classification And Regression Tree (CART) in real-life data: Low protein diet compared to Mediterranean diet in patients with chronic kidney disease," *Nefrología*, vol. 45, no. 4, pp. 279–284, Apr. 2025, doi: 10.1016/j.nefro.2025.01.009.
[22] M. Ozcan and S. Peker, "A classification and regression tree algorithm for heart disease modeling and prediction," *Healthcare Analytics*, vol. 3, Nov. 2023, doi: 10.1016/j.health.2022.100130.
[23] Y. Ali, F. Hussain, and M. M. Haque, "Advances, challenges, and future research needs in machine learning-based crash prediction models: A systematic review," *Accid. Anal. Prev.*, vol. 194, Jan. 2024, doi: 10.1016/j.aap.2023.107378.
[24] S. Jun, "Evolutionary algorithm for improving decision tree with global discretization in manufacturing," *Sensors*, vol. 21, no. 8, Apr. 2021, doi: 10.3390/s21082849.
[25] A. I. Ozturk, O. Yıldırım, E. İdman, and E. İdman, "A comparative study of hybrid decision tree–deep learning models in the detection of intracranial arachnoid cysts," *Neuroscience Informatics*, vol. 5, no. 4, Dec. 2025, doi: 10.1016/j.neuri.2025.100234.
[26] S. Durgaprasad, A. Coraddu, and H. Polinder, "An analysis of maritime battery requirements and a decision tree for optimal chemistry selection," *Energy Reports*, vol. 14, pp. 4851–4871, Dec. 2025, doi: 10.1016/j.egyr.2025.11.086.
[27] J. W. Mburu, L. Kingwara, M. Ester, and N. Andrew, "Use of classification and regression tree

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

(CART), to identify hemoglobin A1C (HbA1C) cut-off thresholds predictive of poor tuberculosis treatment outcomes and associated risk factors," *J. Clin. Tuberc. Other Mycobact. Dis.*, vol. 11, pp. 10–16, May 2018, doi: 10.1016/j.jctube.2018.01.002.
[28] J. Cao *et al.*, "Fuzzy inference system with interpretable fuzzy rules: Advancing explainable artificial intelligence for disease diagnosis—A comprehensive review," *Inf. Sci. (N Y).*, vol. 662, Mar. 2024, doi: 10.1016/j.ins.2024.120212.
[29] M. A. Asadzadeh, M. R. Shaghaghian, Z. Ghadampour, and T. Sabzevari, "A comparative study of fuzzy methods for improving satellite-derived precipitation datasets using novel indices for improvability and efficiency," *J. Hydrol. Reg. Stud.*, vol. 61, Oct. 2025, doi: 10.1016/j.ejrh.2025.102665.
[30] A. El-Ibrahimi *et al.*, "Fuzzy based system for coronary artery disease prediction using subtractive clustering and risk factors data," *Intell. Based. Med.*, vol. 11, Jan. 2025, doi: 10.1016/j.ibmed.2025.100208.
[31] K. Pujaru, S. Adak, T. K. Kar, S. Patra, and S. Jana, "A Mamdani fuzzy inference system with trapezoidal membership functions for investigating fishery production," *Decision Analytics Journal*, vol. 11, Jun. 2024, doi: 10.1016/j.dajour.2024.100481.
[32] R. Fauzi Siregar, A. Rudi Nasution, H. Arifin, and M. Alwi Ripay Harahap, "JESCE (Journal of Electrical and System Control Engineering) Fuzzy Logic Mamdani-Based Simulation of Solanum Lycopersicum Fruit Sorter to Produce High-Quality Fruit Products", doi: 10.31289/jesce.v6i2.11835.
[33] S. Adak, T. K. Kar, and S. Jana, "A fuzzy inference system for predicting outbreaks in emerging infectious diseases," *Decision Analytics Journal*, vol. 10, Mar. 2024, doi: 10.1016/j.dajour.2024.100436.
[34] N. Sabri, S. A. Aljunid, M. S. Salim, R. B. Badlishah, R. Kamaruddin, and M. F. A. Malek, "Fuzzy inference system: Short review and design," *International Review of Automatic Control (I.RE.A.CO.)*, vol. 6, no. 4, 2013, [Online]. Available: https://www.researchgate.net/publication/280739444
[35] M. J. Warrens, "Cohen's weighted kappa with additive weights," *Adv. Data Anal. Classif.*, vol. 7, no. 1, pp. 41–55, Mar. 2013, doi: 10.1007/s11634-013-0123-9.
[36] A. de Raadt, M. J. Warrens, R. J. Bosker, and H. A. L. Kiers, "A Comparison of Reliability Coefficients for Ordinal Rating Scales," *J. Classif.*, vol. 38, no. 3, pp. 519–543, Oct. 2021, doi: 10.1007/s00357-021-09386-5.
[37] J. B. Garner, "THE STANDARD ERROR OF COHEN'S KAPPA," 1991.
[38] M. L. McHugh, "McHugh ML.-Interrater reliability," Aug. 2012.
[39] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977, [Online]. Available: http://www.jstor.org/stable/2529310
[40] A. S. Rajawat, P. Bedi, S. B. Goyal, P. Bhaladhare, A. Aggarwal, and R. S. Singhal, "Fusion Fuzzy Logic and Deep Learning for Depression Detection Using Facial Expressions," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 2795–2805. doi: 10.1016/j.procs.2023.01.251.

Author Biography

Eko Ginanjar Basuki Rahmat is an undergraduate student in the Department of Informatics Engineering at Universitas Sebelas Maret (UNS), Indonesia. His academic interests focus on Artificial Intelligence, Machine Learning, Fuzzy Systems, and Data Mining, particularly the application of computational intelligence in mental health assessment and psychological decision support systems. During his undergraduate studies, he has explored the integration of data-driven methods with expert knowledge to manage uncertainty and subjectivity in psychological evaluation. This paper was developed as part of his undergraduate final project at Universitas Sebelas Maret, aiming to contribute to interpretable, reliable intelligent systems for early mental health screening and consistent diagnostic support.

Wiharto is lecturer and researcher at Universitas Sebelas Maret (UNS), Indonesia, where he is actively affiliated with the Computational Science & Engineering Research Group. He obtained his Doctoral degree in Bio-Medical Informatics from Universitas Gadjah Mada after completing his undergraduate and master's studies in the field of information technology and computer science. Throughout his academic career, he has demonstrated a strong commitment to research and education, particularly in the interdisciplinary area that bridges computing and healthcare. His research interests center on the integration of Artificial Intelligence, Computational Intelligence, and Bio-Medical Informatics, with a particular emphasis on developing intelligent and

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 304-323 **e-ISSN: 2656-8632**

decision-support systems for medical diagnosis and health informatics applications.

Umi Salamah is a lecturer and researcher in the Computational Science and Engineering Research Group at Universitas Sebelas Maret (UNS), Indonesia. She completed her undergraduate studies in science and later earned her master's degree in computer science. She obtained her Doctoral degree in Informatics from Institut Teknologi Sepuluh Nopember (ITS). Her research interests include fuzzy logic, computational intelligence, decision support systems, and intelligent data processing. She is particularly focused on the development of interpretable, reliable, and robust intelligent systems that can be applied to real-world problems, including decision-making and data-driven analysis. In addition to her research activities, she is actively involved in teaching and supervising undergraduate students and has contributed to various scientific publications in national and international journals and conferences.

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 304-323 **e-ISSN: 2656-8632**

APPENDIX A: DASS-42 QUESTIONNAIRE

NO	TEXT
Q1	I found myself getting upset by quite trivial things.
Q2	I was aware of dryness of my mouth.
Q3	I couldn't seem to experience any positive feeling at all.
Q4	I experienced breathing difficulty (eg, excessively rapid breathing, breathlessness in the absence of physical exertion).
Q5	I just couldn’t seem to get going.
Q6	I tended to over-react to situations.
Q7	I had a feeling of shakiness (eg, legs going to give way).
Q8	I found it difficult to relax.
Q9	I found myself in situations that made me so anxious I was most relieved when they ended.
Q10	I felt that I had nothing to look forward to.
Q11	I found myself getting upset rather easily.
Q12	I felt that I was using a lot of nervous energy.
Q13	I felt sad and depressed.
Q14	I found myself getting impatient when I was delayed in any way (eg, elevators, traffic lights, being kept waiting).
Q15	I had a feeling of faintness.
Q16	I felt that I had lost interest in just about everything.
Q17	I felt I wasn’t worth much as a person.
Q18	I felt that I was rather touchy.
Q19	I perspired noticeably (eg, hands sweaty) in the absence of high temperatures or physical exertion.
Q20	I felt scared without any good reason.
Q21	I felt that life wasn’t worthwhile.
Q22	I found it hard to wind down.
Q23	I had difficulty in swallowing.
Q24	I couldn’t seem to get any enjoyment out of the things I did.
Q25	I was aware of the action of my heart in the absence of physical exertion (eg, sense of heart rate increase, heart missing a beat).
Q26	I felt down-hearted and blue.
Q27	I found that I was very irritable.
Q28	I felt I was close to panic.
Q29	I found it hard to calm down after something upset me.
Q30	I feared that I would be “thrown” by some trivial but unfamiliar task.
Q31	I was unable to become enthusiastic about anything.
Q32	I found it difficult to tolerate interruptions to what I was doing.
Q33	I was in a state of nervous tension.
Q34	I felt I was pretty worthless.
Q35	I was intolerant of anything that kept me from getting on with what I was doing.
Q36	I felt terrified.
Q37	I could see nothing in the future to be hopeful about.
Q38	I felt that life was meaningless.
Q39	I found myself getting agitated.
Q40	I was worried about situations in which I might panic and make a fool of myself.
Q41	I experienced trembling (eg, in the hands).
Q42	I found it difficult to work up the initiative to do things.

Manuscript Received 20 October 2025; Revised 5 January 2026; Accepted 10 January 2026; Available online 20 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1443