INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

A Non-Intrusive Fuzzy Logic Framework for Real-Time Participant Engagement Assessment Using Multimodal Parameters

Mohd Najib Ali Mokhtar[1], Mahasan Mat Ali[1] & Mohd Nazrin Muhammad[2]

[1]Centre of Smart System and Innovative Design (COSSID), Fakulti Teknologi dan Kejuruteraan Industri dan Pembuatan, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

[2]Mechate Sdn Bhd, Kulim Hi-Tech Park, Kedah, Malaysia.

DOI: https://doi.org/10.47772/IJRISS.2026.100700780

Received: 28 July 2026; Accepted: 03 August 2026; Published: 12 August 2026

ABSTRACT

Classroom and training session engagement is difficult to measure objectively at scale, since traditional methods are either subjective or impractical for large groups. This study addresses that gap by developing a non-intrusive framework that includes participant engagement in real time from ordinary camera video. Three vision-derived behavioral indicators head-pose deviation, facial expression, and posture. Parameters are fuzzified using triangular and trapezoidal membership functions and combined through a Mamdani-type fuzzy inference system. The rule base and membership functions were validated through manual centroid calculation and cross-checked against a converged Python simulation; for a representative multi-rule case, the manual and simulated scores agreed to within 0.002%, and convergence analysis across discretization grid resolutions confirmed the numerical stability of the implementation. By combining lightweight camera-based feature extraction with an interpretable fuzzy inference layer, the proposed framework avoids the scalability limitations of sensor-based or manual assessment while explicitly modeling the uncertainty inherent in behavioral interpretation. The result is a practical, real-time engagement analytics tool that can support instructors and adaptive learning systems in identifying disengagement and personalizing intervention using contactless sensor.

Key Words: Fuzzy Logic, Engagement Detection, Computer Vision, Mamdani Inference System, Multimodal Behavioral Analysis

INTRODUCTION

Participant engagement is important in classroom teaching or training environment. However, it remains difficult to measure in real classroom settings. Traditional approaches use questionnaires [1], teacher observation checklists, or physiological sensors such as EEG, eye-trackers and wearables sensors [2-3]. The methods have certain limitations. It is subjective and difficult to measure classroom activity across many students simultaneously. This creates a practical gap between the pedagogical needs to detect students' disengagement.

Computer vision offers a good alternative. Using a single classroom-facing or student-facing camera, it is possible to extract behavioral proxies for engagement such as head pose and gaze direction, facial expression, and body posture without requiring physical contact or specialized hardware beyond a standard camera [4-5]. However, mapping these raw visual cues onto a meaningful, human-interpretable engagement cannot be translated directly. Engagement is not a binary state, and the boundaries between attentive and distracted are inherently vague. A certain degree gaze deviation is hardly classified as student or participant disengagement. Moreover, can a neutral facial expression classified as bored. Crisp threshold-based rules tend to produce brittle, discontinuous classifications that misrepresent the gradational nature of attention and are highly sensitive to sensor noise, lighting conditions, and individual variation in expressiveness.

Fuzzy logic is well suited to this problem because it explicitly models the imprecision inherent in behavioral interpretation [6-8]. Rather than forcing each vision-derived feature into a hard category, fuzzy membership

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

functions allow a given gaze angle, expression, or posture reading to partially belong to multiple linguistic categories with varying degrees of membership. A fuzzy inference system (FIS) then combines these partial memberships across modalities using a rule base to produce a smooth, continuous engagement score. This approach mirrors how human observers such as experienced teachers. It assesses holistically rather than applying rigid cutoffs.

This paper proposes a fuzzy logic–based framework that fuses three vision-derived indicators which is head pose, facial expression, and posture into a single interpretable engagement score. The framework is designed to operate passively using standard camera input, requiring no physical sensors attached to students, and to produce outputs that are directly actionable for instructors and adaptive learning systems.

METHODOLOGY

The proposed framework follows an eight-stage pipeline that transforms raw camera input into a stable, interpretable engagement score for each student. Table 1 summarizes the pipeline stages; each is described in detail in the subsections that follow.

Table 1. Stages and Functions of the Fuzzy Logic-Based Classroom Engagement Detection System

Stage	**Function**
1. Acquisition	Standard RGB camera captures a continuous video stream of the student(s) in the classroom.
2. Detection & tracking	Face detection and landmark tracking locate each student's face and body region frame-by-frame.
3. Feature extraction	Three parallel modules estimate (a) head pose or gaze deviation angle, (b) facial expression class and confidence, and (c) posture or activity state from pose key points.
4. Fuzzification	Each crisp feature value is mapped to membership degrees (0-1) in the linguistic sets Low / Medium / High using triangular or trapezoidal membership functions.
5. Rule evaluation	A Mamdani-type fuzzy inference engine evaluates the IF-THEN rule base, combining input memberships via AND (minimum) or product T-norm.
6. Aggregation	Outputs of all fired rules are aggregated into a single fuzzy output set for Engagement Level.
7. Defuzzification	The aggregated fuzzy set is converted to a crisp engagement score (0-100) using the centroid (center of gravity) method.
8. Temporal smoothing	A moving average smoothing filter over a sliding time window reduces frame-to-frame jitter and produces a stable per-student engagement trend.

The framework as shown in Figure 1 converts gaze, facial expression, and posture into fuzzy values, processes them through Mamdani fuzzy rules, defuzzifies the result into a single crisp engagement score, then visualizes each student's score as a classroom-wide heatmap for real-time monitoring.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

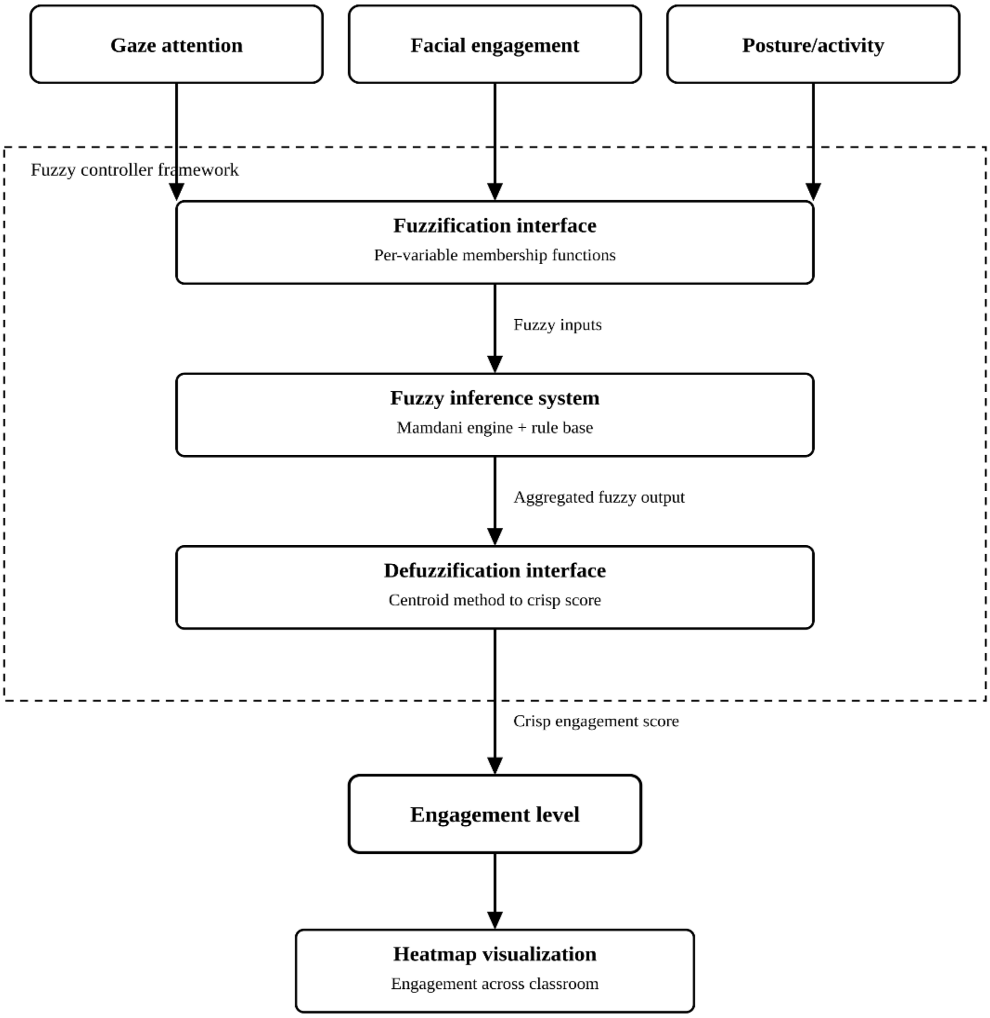

Figure 1. System Architecture of the Fuzzy Logic-Based Classroom Engagement Detection Framework.

2.1 Input and Output Variable Definitions

Table 2 formally defines the input and output variables of the fuzzy inference system, including their type, a descriptor of what each variable represents, its crisp numerical range prior to fuzzification (or after defuzzification, for the output), and the linguistic terms used to partition that range.

Table 2. Input and Output Variable Definitions

Variable type	**Variable name**	**Descriptor**	**Range**	**Linguistic terms**
Input	Gaze attention (head pose deviation)	Angular deviation of head/eye gaze from the teacher/screen reference point, derived from head-pose estimation	0°–90°	Low, Medium, High
Input	Facial engagement	Confidence-weighted classification of facial expression, indicating attentiveness or emotional valence	0–1 (normalized)	Low, Medium, High

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

Variable type	Variable name	Descriptor	Range	Linguistic terms
Input	Posture	Composite score of torso inclination, shoulder symmetry, and motion magnitude, indicating physical engagement	0–1 (normalized)	Low, High
Output	Overall engagement level	Defuzzified composite score representing the student's holistic engagement state	0–100 (crisp score)	Very Low, Low, Medium, High, Very High

Note. Normalized variables are scaled to a 0-1 range prior to fuzzification. The crisp score represents the defuzzified output of the fuzzy inference system on a 0-100 scale.

2.2 Feature Extraction

Three parallel vision modules process each tracked student region per frame (or per short temporal window):

- Head Pose: A head-pose estimator yields yaw or pitch angles relative to the camera or a designated target such as teacher or screen position. The output is a deviation angle in degrees.
- Facial Expression: A facial emotion recognition model classifies expression into engagement-relevant categories (e.g., neutral/attentive, positive/interested, bored, confused) and outputs a confidence score per category.
- Posture : Body key point estimation derives features such as torso inclination, shoulder height symmetry, and motion magnitude to distinguish upright/active participation from slouching or fidgeting.

These three raw, crisp measurements form the input vector to the fuzzy inference system at each evaluation interval.

2.3 Fuzzification

Each crisp input is converted into linguistic membership degrees using overlapping triangular or trapezoidal membership functions, consistent with the sets described earlier (e.g., Gaze Attention: Low for deviation >30°, Medium for moderate deviation, High for <10°). Overlap between adjacent sets ensures a smooth, continuous transition rather than an abrupt category change as the underlying measurement crosses a threshold.

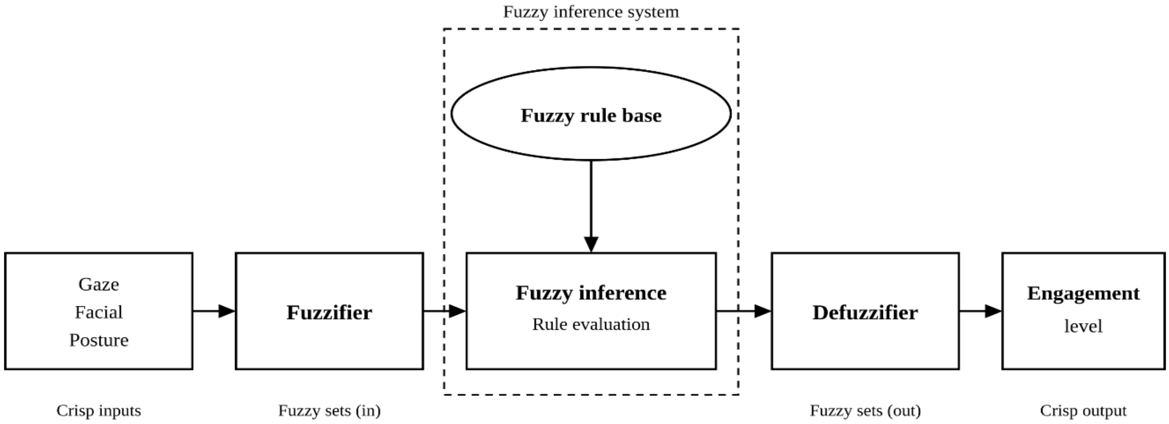

Figure 2. Block Diagram of the Fuzzy Inference System (FIS) Architecture

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

All linguistic terms in this framework study are represented using piecewise-linear membership functions of either triangular or trapezoidal shape [9]. Each shape is defined over a crisp input x by a small set of breakpoints.

The trapezoidal membership function [10] is parameterized by four breakpoints a ≤ b ≤ c ≤ d and is mathematically defined as:

$$\mu(x;a,b,c,d)=\begin{cases}0, & x\le a \text{ or } x\ge d\\ \dfrac{x-a}{b-a}, & a<x\le b\\ 1, & b\le x\le c\\ \dfrac{d-x}{d-c}, & c<x<d\end{cases} \quad (1)$$

where,

X is the input variable evaluated by the membership function.
M is the degree of membership, ranging continuously between 0 and 1.
a is the lower support limit, representing the point below which the membership degree is zero.
b is the lower core limit, representing the onset point where the membership degree reaches its maximum value of one.
c is the upper core limit, representing the endpoint where the membership degree begins to decrease from one.
D is the upper support limit, representing the point above which the membership degree returns to zero.

The triangular membership function is represented by equation (2).

$$\mu(x;a,b,d)=\begin{cases}0, & x\le a \text{ or } x\ge d\\ \dfrac{x-a}{b-a}, & a<x\le b\\ \dfrac{d-x}{d-b}, & b<x<d\end{cases} \quad (2).$$

where,

x: The input variable evaluated by the membership function.
μ: The degree of membership, ranging continuously between 0 and 1.
a: The lower support limit, representing the point below which the membership degree is zero.
b: The peak value, representing the exact point where the membership degree reaches its absolute maximum value of one.
d: The upper support limit, representing the point above which the membership degree returns to zero.

The parameters of membership function represented by equation (1) and equation (2) are listed in Table 4.

Table 3. Membership Function Parameters for All Input and Output Linguistic Terms

Variable	**Linguistic term**	**Shape**	**Parameters (a, b, c, d)**
Gaze attention	High (focused)	Trapezoidal	(0, 0, 10, 25)
Gaze attention	Medium	Triangular	(10, 22.5, 45)
Gaze attention	Low (distracted)	Trapezoidal	(25, 40, 90, 90)
Facial engagement	Low	Trapezoidal	(0, 0, 0.2, 0.4)
Facial engagement	Medium	Triangular	(0.25, 0.5, 0.75)

Variable	Linguistic term	Shape	Parameters (a, b, c, d)
Facial engagement	High	Trapezoidal	(0.6, 0.8, 1, 1)
Posture	Low	Trapezoidal	(0, 0, 0.3, 0.6)
Posture	High	Trapezoidal	(0.4, 0.7, 1, 1)
Engagement level (output)	Very low	Trapezoidal	(0, 0, 10, 30)
Engagement level (output)	Low	Triangular	(10, 30, 50)
Engagement level (output)	Medium	Triangular	(30, 50, 70)
Engagement level (output)	High	Triangular	(50, 70, 90)
Engagement level (output)	Very high	Trapezoidal	(70, 90, 100, 100)

Gaze Attention membership function as shown in Figure 3 defined over the universe of discourse [0°, 90°], where 0° represents a gaze perfectly aligned with the teacher/screen reference point. “High” is a right shoulder (saturating at 1 for $x \leq 10°$), “Medium” is a triangular set peaking at 22.5°, and “Low” is a left shoulder (saturating at 1 for $x \geq 40°$).

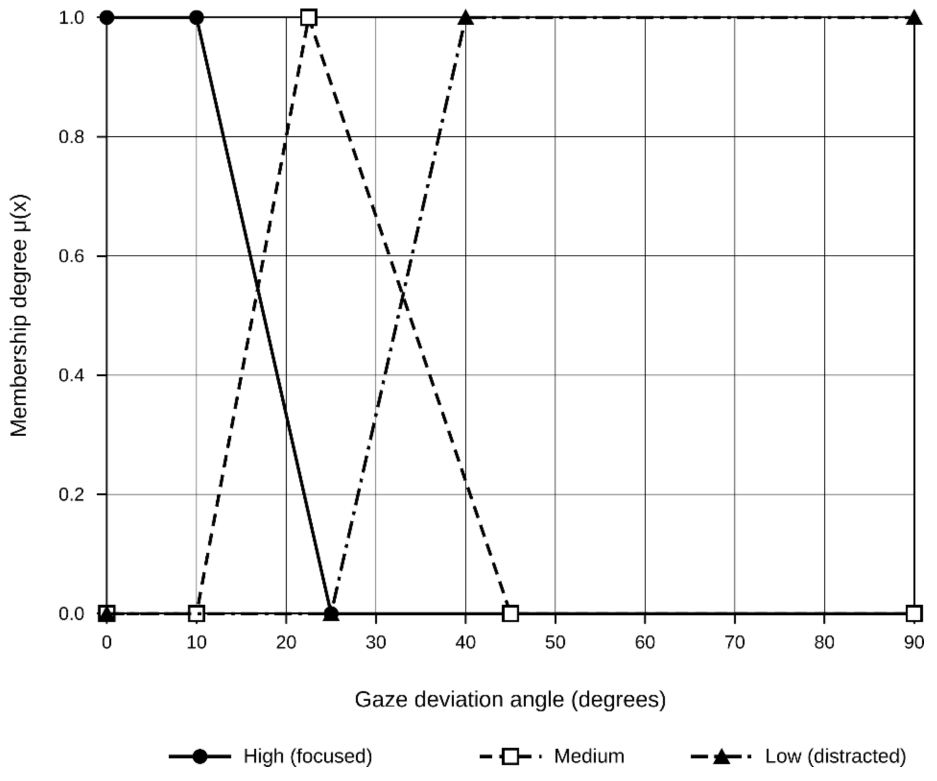

Figure 3. Membership functions for the Gaze Attention input variable.

Facial Engagement membership function as shown in Figure 4 is defined over the normalized confidence range [0, 1] produced by the emotion recognition module. The three sets are distributed symmetrically across the range, with “Medium” (neutral expression) centered at 0.5.

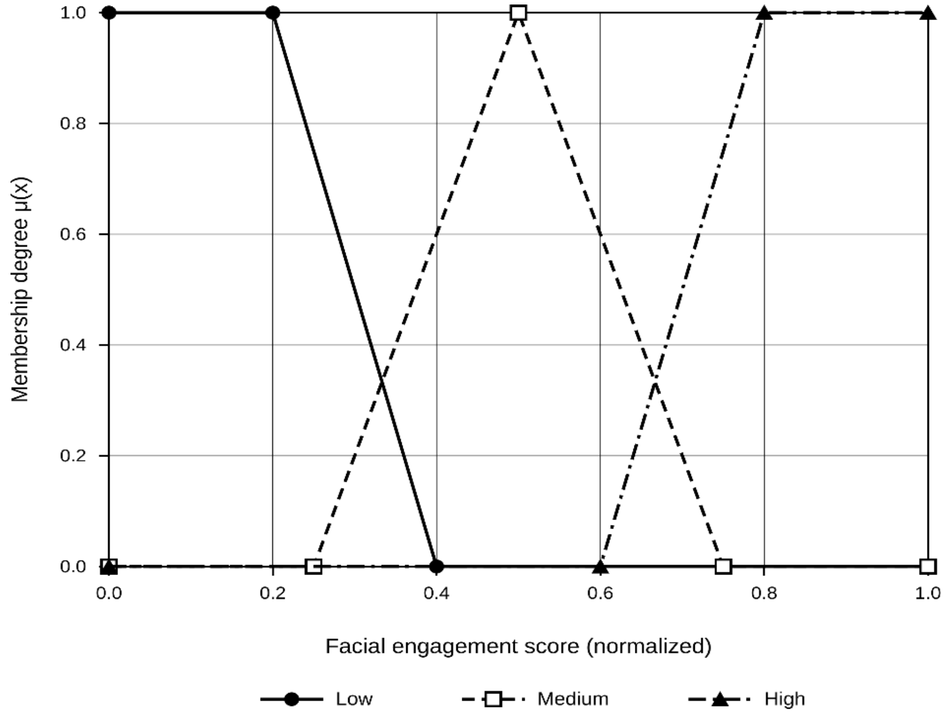

Figure 4. Membership functions for the Facial Engagement input variable.

Posture membership function as shown in Figure 5 is defined over the normalized range [0, 1] derived from torso inclination, shoulder symmetry, and motion magnitude. Consistent with Table 4 in Section 3.1, only two linguistic terms are used, with the crossover point ($\mu = 0.5$ for both sets) at approximately $x = 0.5$.

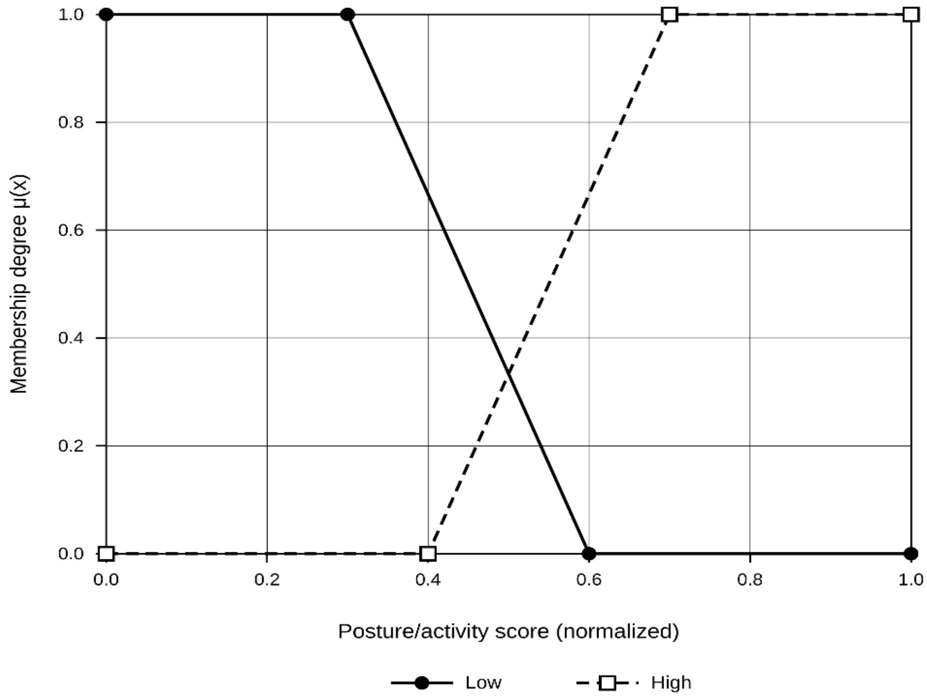

Figure 5. Membership functions for the Posture input variable

The output membership function is depicted in Figure 6. The output covers the normalized engagement score [0, 100]. Five evenly spaced, symmetrically overlapping sets (Very Low, Low, Medium, High, Very High) partition this range, which the defuzzification stage uses to compute the final crisp engagement score via the centroid method.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

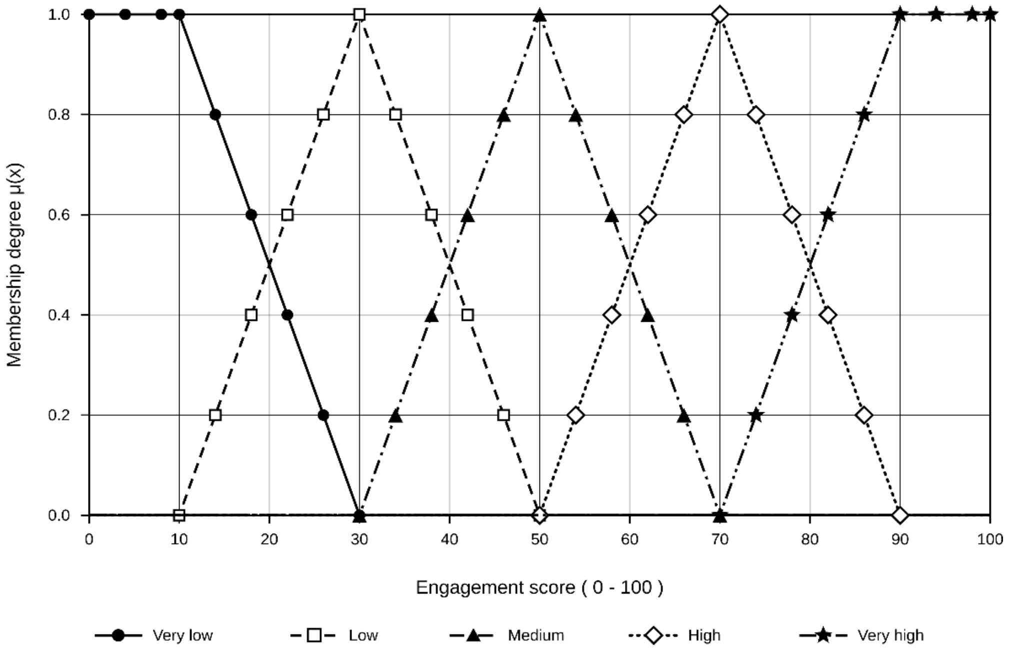

Figure 6. Membership functions for the Engagement Level output variable.

A Mamdani-type fuzzy inference system combines the three fuzzified inputs using a bank of expert-defined IF-THEN rules. A representative of the rule base is shown in Table 3. It lists all 27 combinations of the three three-level input sets (with reduced sets for edge cases).

Table 3. Representative Fuzzy Rule Base (Mamdani-Type)

Gaze attention	**Facial engagement**	**Posture/activity**	**Overall engagement (output)**
High	High	High	**Very high**
High	High	Low	**High**
High	Medium	High	**High**
High	Medium	Low	**Medium**
High	Low	High	**Medium**
High	Low	Low	**Low**
Medium	High	High	**High**
Medium	Medium	Medium	**Medium**
Medium	Low	Low	**Low**

Gaze attention	Facial engagement	Posture/activity	Overall engagement (output)
Low	High	High	**Medium**
Low	Medium	Medium	**Low**
Low	Low	Low	**Very low**
Low	Low	High	**Low**

For each rule, the firing strength is computed as the minimum (or product) of the input membership degrees, following standard Mamdani AND-composition. Rules with non-zero firing strength contribute a clipped or scaled version of their consequent membership function to the output aggregation stage.

RESULT

3. 1 Result Validation

A manual calculation approach is conducted to test the model. Sample input as shown in Table 5 is used.

Table 5. Sample Crisp Input Values for a Single Classroom Observation

Input variable	Crisp value
Gaze attention (deviation angle)	18°
Facial engagement (normalized score)	0.33
Posture/activity (normalized score)	0.45

The results comparison of the developed model and manual calculation are shown in Table 6.

Table 6. Validation and Simulation Comparison with Manual Calculation

Quantity	Value
Manual defuzzified score (centroid)	≈ 45.63 / 100
Python simulation (converged, reference)	45.6287 / 100
Absolute / percent difference	≈ 0.001 (≈ 0.002%)
Final engagement classification	Medium

The manual defuzzified score was computed using the centroid method on discretized membership values. The fuzzy model are solved using the Python simulation as depicted in Figure 7.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

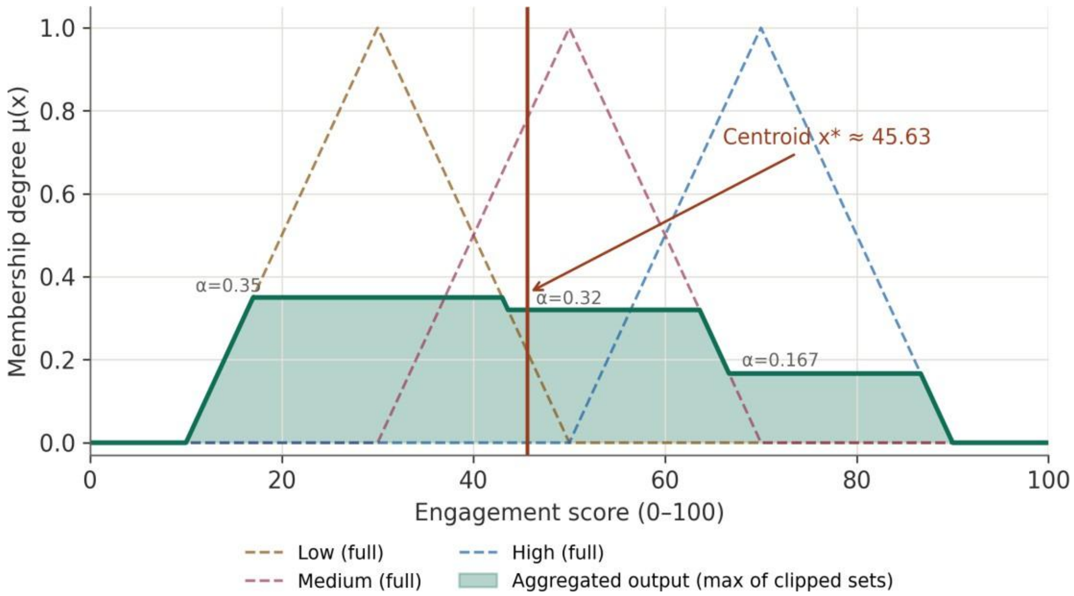

Figure 7. Fuzzy inference process for engagement evaluation using Phyton simulation with the result centroid marked at x* ≈ 45.63.

The plots in Figure 8 show membership functions and crisp values for Gaze Attention (18.0°), Facial Engagement (0.33), and Posture (0.45). The final subplot displays the aggregated output area and the defuzzified engagement score ($x^{*} = 45.6$, Medium). The results show that the manual calculation and simulation agree, and both yield a final classification of Medium engagement.

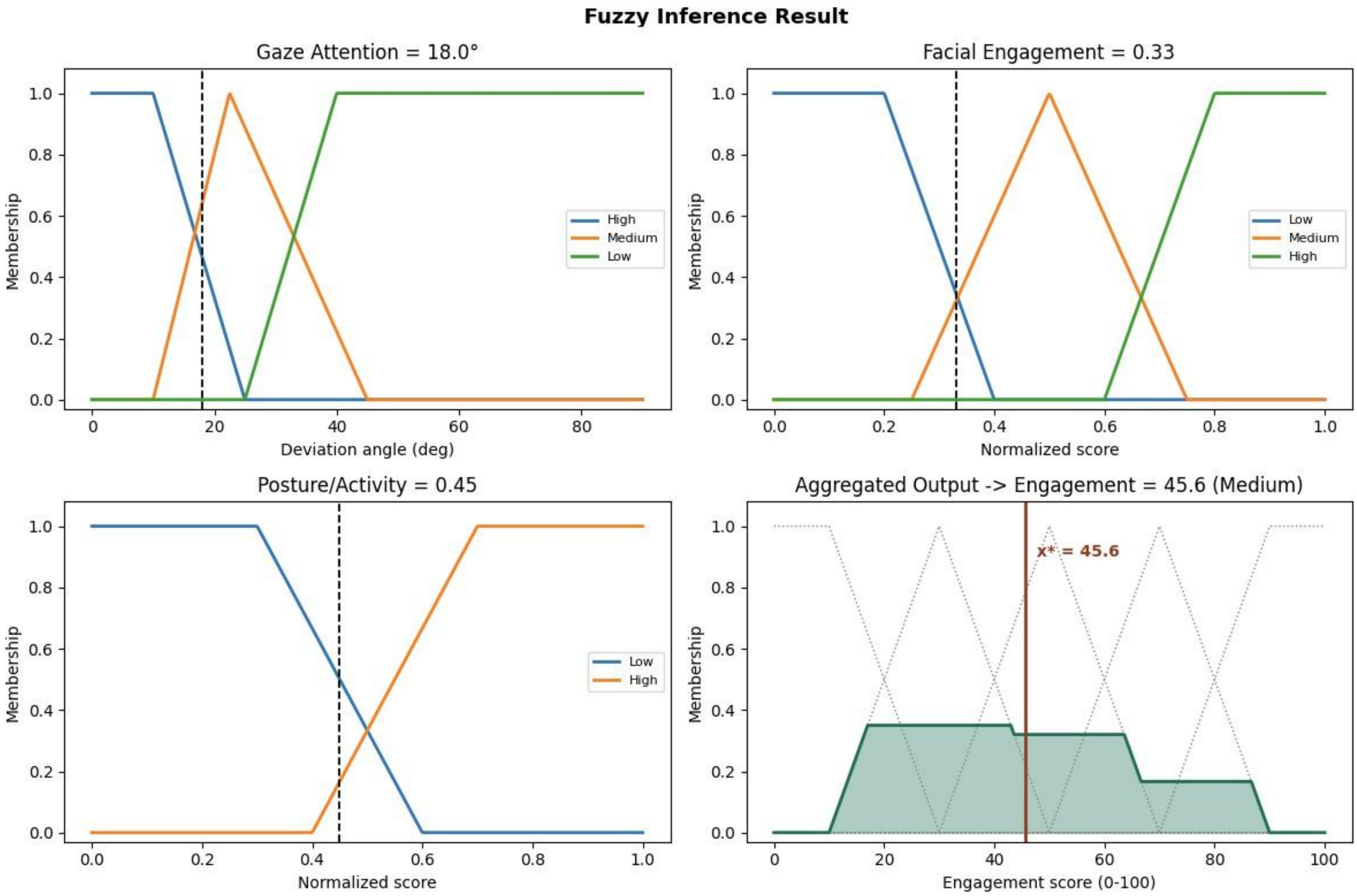

Figure 8: The input membership functions and crisp values for Gaze Attention (18.0°), Facial Engagement (0.33), and Posture/Activity (0.45) and output membership function.

The small residual difference from the Python simulation (45.63 vs. 45.6287, about 0.002%) is attributable entirely to decimal rounding in the hand arithmetic, not to any methodological discrepancy.

To test behavior beyond what a manual calculation can reasonably verify, the Python simulation is repeated using different number of grid points to increase resolution. The convergence plot is depicted in Figure 9 and the deviation magnitude is listed in Table 7. The phyton 3 code executed with Intel(R) Xeon(R) CPU with clock speed of 2.20GHz. The system RAM used to hold the active data, variables, and libraries of the Python 3 code while executing is 1.2 GB out of total 12.7 GB available.

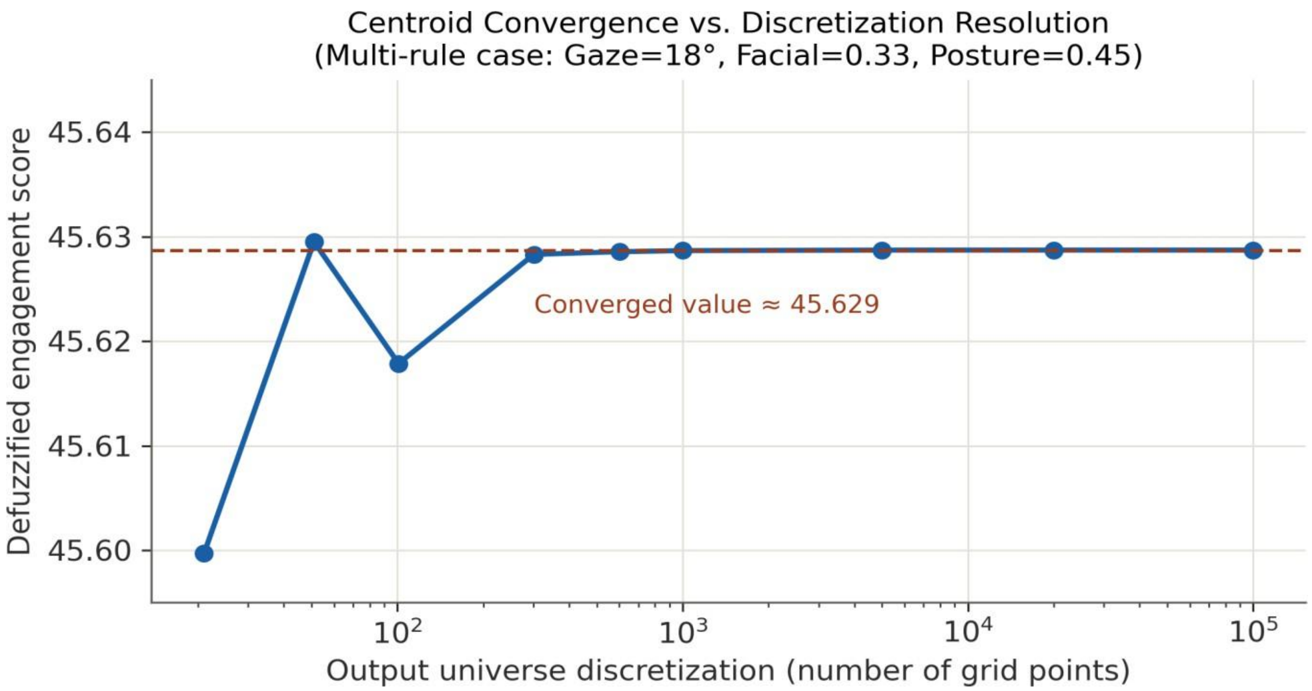

Figure 9. Defuzzified engagement score for the 8-rule multi-firing case as a function of output-universe discretization (grid resolution), converging to ≈ 45.629.

Table 7: Deviation error based on the grid resolutions.

Grid Resolution (points)	Defuzzified Score	Deviation error (%) from Converged Value
21	45.5997	2.90
51	45.6295	0.08
101	45.6179	1.08
301	45.6283	0.04
601	45.6286	0.01
1,001 (used in classroom notebook)	45.6287	0.005
5,001	45.62871	0.0001
100,001 (reference)	45.62871	0 (reference)

In general, the score converges as grid resolution increases. This confirms the simulation is numerically stable and that the default settings already operate well within the convergence region.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

3. 2 Classroom Fuzzy Engagement

The model was tested for classroom seating conditions and the resulted value were listed in Table 8.

Table 1: Classroom Fuzzy Engagement Observations and Extracted Parameter Scores

Student	**Row**	**Col**	**Gaze (deg)**	**Facial (0-1)**	**Posture (0-1)**	**Engagement Score**	**Engagement Level**
S01	1	1	69.7	0.76	0.44	39.1	Medium
S02	1	2	39.5	0.35	0.83	50.4	High
S03	1	3	77.3	0.97	0.70	70.0	High
S04	1	4	62.8	0.89	0.31	30.0	Low
S05	2	1	8.5	0.78	0.83	88.9	Very High
S06	2	2	87.8	0.19	0.80	30.0	Low
S07	2	3	68.5	0.47	0.39	30.0	Low
S08	2	4	70.7	0.04	0.29	10.8	Low
S09	3	1	11.5	0.15	0.68	66.8	High
S10	3	2	40.5	0.68	0.14	36.5	Medium
S11	3	3	33.4	0.74	0.20	39.7	Medium
S12	3	4	83.4	0.97	0.01	30.0	Low
S13	4	1	57.9	0.33	0.79	39.0	Medium
S14	4	2	74.0	0.37	0.66	45.2	Medium
S15	4	3	39.9	0.47	0.71	55.2	High
S16	4	4	20.5	0.19	0.78	56.4	High
S17	5	1	49.9	0.13	0.46	19.0	Low
S18	5	2	5.7	0.48	0.57	66.7	High
S19	5	3	74.5	0.23	0.14	11.2	Low
S20	5	4	56.8	0.67	0.11	30.0	Low
Mean			**51.6**	**0.49**	**0.49**	**42.2**	

Figure 10 shows example of the classroom engagement heatmap where the values are derived from tracking metrics across a 5x4 classroom seating matrix grid setup

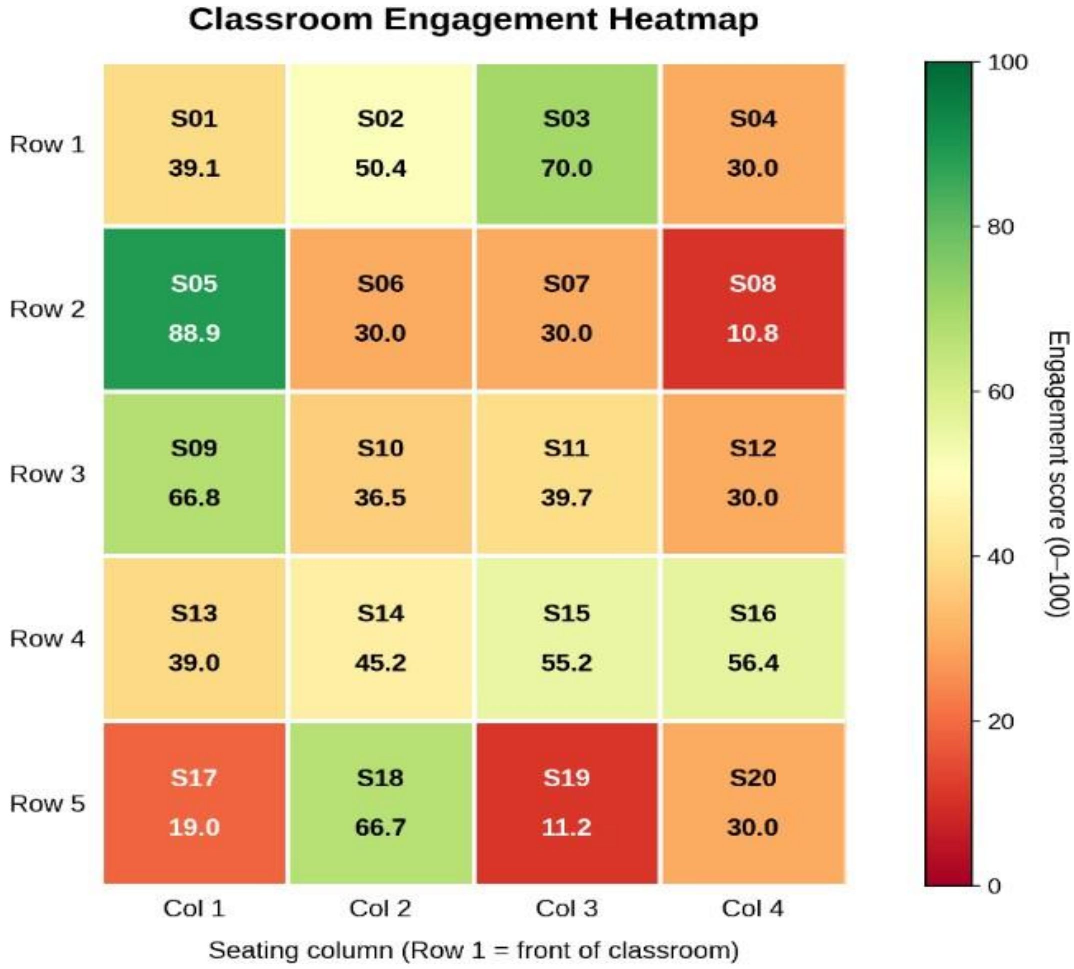

Figure 10. Engagement Heatmap

CONCLUSION

This study presented a non-intrusive, fuzzy logic-based framework that estimates classroom or training-session engagement in real time from three camera-derived behavioral indicators: gaze deviation, facial expression, and posture. By fuzzifying these inputs and combining them through a Mamdani-type inference engine, the framework maps inherently ambiguous behavioral cues into a continuous, interpretable engagement score rather than forcing them into brittle, threshold-based categories. Validation against manual centroid calculation showed agreement to within 0.002%, and convergence analysis across output-universe discretization confirmed that the implementation is numerically stable, supporting the reliability of the underlying inference pipeline. Application to a simulated 5×4 classroom seating matrix further demonstrated that the framework practicality, per-student engagement scores that can be aggregated into a real-time heatmap. As the approach relies solely on a standard camera and requires no wearable or physiological sensors, it offers a scalable and low-cost alternative to self-report surveys, observation checklists, and sensor-based monitoring. Future work should extend the validation to real classroom video across diverse lighting conditions, occlusion levels, and demographic groups, incorporate multi-rater ground truth for the fuzzy rule base, and benchmark the framework against machine learning-based engagement estimators to quantify its relative accuracy and computational cost.

ACKNOWLEDGMENT

The authors extend sincere appreciation to the Universiti Teknikal Malaysia Melaka (UTeM) and the Fakulti Teknologi dan Kejuruteraan Industri dan Pembuatan (FTKIP) for the valuable support provided, including financial assistance and access to essential resources, which facilitated the successful execution of this study. This work was conducted under project reference INDUSTRI(URMG)/G7AEROSPACE/FTKIP/2025/I00106.

REFERENCES

1. Garcia-Ceja, E., Alvarado-Uribe, J., Escamilla-Ambrosio, P. J., Lara, A., Mena-Martinez, A., Gallegos-Garcia, G., Gonzalez-Mendoza, M., Monroy, R., Martinez Luna, G., & Fernández-Cárdenas, J. M. (2026). A dataset of university students' stress and anxiety levels based on questionnaires and wearable sensors. Scientific Data, 13(1), Article 732. https://doi.org/10.1038/s41597-026-07085-7
2. Manikandan, S., & Vijayakumar, D. (2026). Secure cloud-based wearable sensor networks for teaching effectiveness evaluation in physical education through virtualized load balancer. International Journal of Computational Intelligence Systems, 19(1), Article 105. https://doi.org/10.1007/s44196-026-01230-1
3. Wang, Y., Zhou, K., Lin, Z., Shen, H., Liu, R., Su, Z., Womer, F. Y., Wang, Y., Wei, Y., Zhu, R., Zheng, J., Zhang, X., & Wang, F. (2026). Evaluation of the efficacy of group cognitive behavioral therapy on anxiety in college students based on wearable devices and mobile applications: A randomized controlled trial. BMC Psychiatry, 26(1), Article 6. https://doi.org/10.1186/s12888-025-07572-y
4. Dan, Z., Yali, Z., & Tong, Z. (2026). A multimodal AI framework for real-time student engagement detection and adaptive feedback in higher education. Journal of King Saud University - Computer and Information Sciences, 38(5), Article 243. https://doi.org/10.1007/s44443-026-00639-0
5. Bazhenova, G. A. (2026). The study of involvement in the educational process knowledge fatigue and the level of emotional intelligence of students using neural networks. Problemy Sotsial'noi Gigieny, Zdravookhraneniia i Istorii Meditsiny, 34, 371–377. https://doi.org/10.32687/0869-866X-2026-34-s1-371-377
6. Duan, D. (2026). Using a predictive model based on fuzzy logic methods to test the mental health of students. Discover Computing, 29(1), Article 168. https://doi.org/10.1007/s10791-026-09962-6
7. Deepakraj, J., Reddy, S. S., Khalil, B. T., Viswanadham, T. S., Balamurugan, M., & Mandal, R. (2025). Fuzzy logic and deep learning hybrid model for physiological stress detection in higher education. 2025 5th International Conference on Intelligent Technologies (CONIT). https://doi.org/10.1109/CONIT65521.2025.11167821
8. Irhamsyah, M., Melinda, M., Zakaria, L. Q., Yunidar, Y., Koprawi, T. Y. P., & Miftahujjannah, R. (2025). A fuzzy logic-based stress recognition system for final-year students using PSS-10 neuropsychological measures. International Journal on Advanced Science, Engineering and Information Technology, 15(5), 1663–1671. https://doi.org/10.18517/ijaseit.15.5.20941
9. Waqas, S. M., Alim, A., Talpur, K., Su'ud, M. M., Djenouri, Y., Ali, S. M., & Alam, M. M. (2026). A fuzzy membership-based enhancement to density peak clustering with comparative performance evaluation. Array, 30, Article 100785. https://doi.org/10.1016/j.array.2026.100785
10. Şenel, A. A., Adiloğulları, G. E., & Şenel, E. (2026). Modelling the effect of motivation on mental health components with fuzzy logic among elite athletes. Scientific Reports, 16(1), Article 8076. https://doi.org/10.1038/s41598-026-39718-7