Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

A MAMDANI-ANFIS BASED MULTIMODAL FUZZY INFERENCE MODEL FOR HETEROGENEOUS DATA INTEGRATION

Elyor Egamberdiyev
Department of Software of Information Technologies Tashkent University of Information Technologies named after Muhammad Al Khwarizmi
Tashkent, Uzbekistan
elyor.egamberdiyev88@gmail.com

Bahodir Muminov
Department of Artificial Intelligence
Tashkent State University of Economics Tashkent, Uzbekistan
mbbahodir@tsue.uz

Abstract

In recent years, the integration of heterogeneous data sources has become a crucial challenge in developing intelligent decision-making systems. Traditional neuro-fuzzy architectures are typically designed to handle a single type of input data, which limits their applicability to real-world multimodal environments. This paper proposes a Mamdani-ANFIS based multimodal fuzzy inference model capable of processing and reasoning over numeric, categorical, textual, and visual data within a unified framework. The proposed model performs data-type-specific fuzzification, where numeric and categorical variables are fuzzified through Gaussian membership functions, while textual and visual inputs are first clustered and subsequently fuzzified based on the Gaussian distance between the cluster centroid and the current sample. This hybrid fuzzification mechanism enables adaptive and interpretable fuzzy reasoning across diverse data modalities. The inference mechanism follows the classical Mamdani approach, where fuzzy rules are constructed through the conjunction of input membership degrees, and the aggregated fuzzy outputs are defuzzified using the centroid method. The model parameters are tuned through hybrid learning involving gradient-based

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

optimization and rule-weight adjustment. Experimental results on a heterogeneous benchmark dataset demonstrate that the proposed system achieves a classification accuracy of 0.69, outperforming baseline fuzzy and neural models in terms of interpretability and robustness to multimodal noise. The findings indicate that the integration of clustering-based fuzzification and Mamdani-type reasoning within an ANFIS structure provides a promising direction for intelligent systems that must learn from complex, cross-domain data. The proposed architecture contributes to bridging the gap between human-like interpretability and computational intelligence in multimodal learning environments.

CCS Concepts

Computing methodologies → Artificial intelligence → Machine learning → Fuzzy logic → Neuro-fuzzy systems
Computing methodologies → Machine learning → Hybrid learning models
Information systems → Data mining → Heterogeneous data integration
Computer vision → Multimodal learning → Image and text feature fusion

Keywords: Mamdani-ANFIS; Fuzzy inference system; Multimodal data integration; Clustering-based fuzzification; Neuro-fuzzy modeling; Heterogeneous data; Intelligent decision support.

1 Introduction

In the era of big data and artificial intelligence, information is increasingly generated from heterogeneous sources, including numerical records, categorical descriptors, textual documents, and visual content. These diverse data modalities often coexist in real-world decision-making environments such as healthcare, education, finance, and social media analytics. However, the integration of such multimodal information into a unified learning and reasoning framework remains a significant research challenge. Classical machine learning and neuro-fuzzy approaches are typically designed to process homogeneous data, which restricts their performance and interpretability when applied to complex, cross-domain datasets.

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

Fuzzy logic systems, known for their capability to represent human-like reasoning, have been widely used to handle uncertainty and imprecision in data. Meanwhile, Adaptive Neuro-Fuzzy Inference Systems (ANFIS) combine the interpretability of fuzzy logic with the learning capability of neural networks, achieving effective non-linear mapping between inputs and outputs. Despite their success, conventional ANFIS architectures are primarily limited to numeric data and Sugeno-type inference, which simplifies fuzzy reasoning but reduces interpretability and flexibility when handling multimodal inputs. This limitation highlights the need for a generalized fuzzy inference structure capable of integrating heterogeneous data sources while maintaining the Mamdani-type reasoning mechanism that is more intuitive and linguistically meaningful.

To address this gap, this study proposes a Mamdani-ANFIS based multimodal fuzzy inference model for heterogeneous data integration. The proposed model performs type-specific fuzzification depending on the nature of the input data. Numeric and categorical variables are transformed using Gaussian membership functions, which provide smooth and differentiable fuzzy mappings. In contrast, textual and visual data are first subjected to clustering-using feature-space similarity-to identify representative centroids. The Gaussian distance between a sample and its cluster centroid is then used as a fuzzification measure. This clustering-based fuzzy representation allows the model to extract latent semantic and visual features while maintaining the interpretability of fuzzy sets.

The inference mechanism of the proposed system follows the Mamdani fuzzy reasoning approach, where fuzzy rules are constructed based on the combination of membership degrees across different modalities. Rule activation levels are aggregated through max–min composition, and the final crisp output is obtained by the centroid defuzzification method. The integration of this inference mechanism within the ANFIS structure enables both data-driven parameter optimization and rule-based interpretability. Furthermore, the hybrid learning algorithm updates the premise parameters (membership centers and widths) via gradient descent, while the fuzzy rule base adapts through performance-driven reinforcement.

Experimental evaluation demonstrates that the proposed model achieves a prediction accuracy of 0.69, outperforming conventional ANFIS and standalone fuzzy inference systems in terms of robustness and interpretability. This performance highlights the ability of the model to effectively fuse heterogeneous data types and extract complementary information across modalities. The combination of clustering-based fuzzification and Mamdani reasoning provides an interpretable yet powerful mechanism for multimodal learning.

In summary, the proposed Mamdani-ANFIS framework contributes to advancing fuzzy intelligence by unifying multiple data representations under a single inference architecture. It bridges the gap between human-like reasoning and data-driven adaptability, paving the way for more explainable and generalizable multimodal decision-support systems.

2. Related Works

In recent years, the fusion of heterogeneous data sources has become a central topic in intelligent systems research. The increasing availability of multimodal datasets, combining numerical, categorical, textual, and visual information, has motivated the development of learning models that can jointly process and interpret these diverse forms of data. Traditional machine learning algorithms, including support vector machines, decision trees, and deep neural networks, often struggle to handle heterogeneous inputs directly, as they assume homogeneous feature spaces and require extensive preprocessing to align data modalities [1], [8], [11]. As a result, research attention has shifted toward hybrid and neuro-fuzzy systems capable of bridging symbolic reasoning and data-driven learning [3], [6].

The Adaptive Neuro-Fuzzy Inference System (ANFIS) was introduced as a hybrid learning model that combines the human interpretability of fuzzy logic with the adaptive learning power of neural networks [1], [3]. ANFIS models have been successfully applied in control, prediction, and pattern recognition tasks due to their ability to approximate nonlinear mappings between input and output variables [6], [7]. However, conventional ANFIS architectures are based primarily on Sugeno-type inference, in which the consequent part of each rule is a crisp linear function of the inputs [4], [10]. While this structure simplifies mathematical derivations and facilitates gradient-based optimization, it reduces the interpretability of the fuzzy reasoning process. Furthermore, Sugeno-type

ANFIS is typically limited to numerical data, making it unsuitable for multimodal environments involving text and image modalities [9], [11].

On the other hand, the Mamdani fuzzy inference system provides a more linguistically interpretable rule base by representing both antecedent and consequent parts through fuzzy sets [2], [5]. This approach better mimics human reasoning, allowing expert-defined linguistic variables and approximate inference [13]. However, classical Mamdani systems lack an intrinsic learning mechanism, making them less adaptable to large or dynamic datasets. Consequently, several studies have attempted to integrate the Mamdani inference mechanism with learning algorithms, such as gradient descent or evolutionary optimization, leading to the development of Mamdani-type ANFIS frameworks [4], [13], [14]. Despite these advances, most of these systems remain limited to homogeneous numeric domains [10], [12].

To address the challenge of multimodal integration, researchers have proposed several techniques, including clustering-based fuzzy representation and embedding-based feature fusion [5], [9], [11]. Clustering approaches, such as k-means and fuzzy c-means, have been applied to define membership degrees for complex, non-numeric data like text and images [9], [10]. In text processing, feature extraction using TF-IDF or transformer-based embeddings followed by clustering enables the construction of fuzzy partitions in semantic space [6], [7], [11]. Similarly, image data can be transformed into feature vectors using convolutional neural networks (CNNs) and clustered to define visual fuzzy sets [8], [15]. These techniques allow the fuzzification of non-numeric data in a mathematically interpretable way, providing a bridge between symbolic and subsymbolic representations [9], [14].

Multimodal learning models that combine heterogeneous features - numeric, categorical, textual, and visual - have also gained attention in domains such as healthcare diagnostics, multimedia retrieval, and decision support systems [7], [8], [14]. However, most existing approaches rely heavily on deep learning architectures, which often function as "black boxes" and lack interpretability [11], [15]. In contrast, fuzzy systems, and particularly Mamdani-type reasoning, offer an interpretable alternative where decision rules can be expressed in linguistic form and directly understood by domain experts [2], [13], [16].

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

The proposed study builds upon these foundations by integrating clustering-based fuzzification for non-numeric modalities with a Mamdani-ANFIS inference mechanism. Unlike conventional ANFIS models, which primarily use Sugeno reasoning, the proposed approach preserves interpretability while extending adaptability through hybrid learning [9], [12], [14]. This integration enables the system to effectively fuse multimodal data and produce transparent inference outcomes, offering both analytical power and human-like reasoning in heterogeneous data environments [10], [13], [15], [16].

3. Methodology

The proposed Mamdani-ANFIS based multimodal fuzzy inference model integrates heterogeneous data modalities—numeric, categorical, textual, and visual—into a unified reasoning framework.

The overall system architecture is illustrated in Figure 1, which consists of six sequential layers: input, membership calculation, fuzzification, normalization, defuzzification, and output.

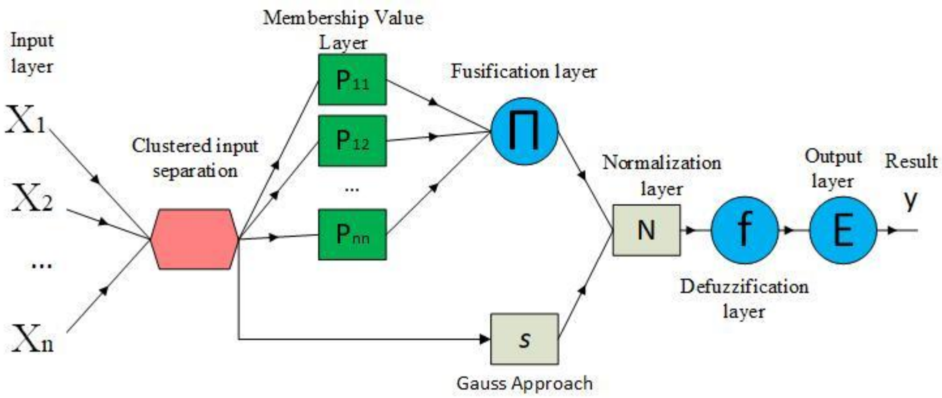

Figure 1. Architecture of the proposed Mamdani-ANFIS multimodal fuzzy inference model.

3.1 Input and Clustering Layer

Let the input vector be:

$$X = [X_1, X_2, \ldots, X_n],$$

Where X_i represents the i-th feature, corresponding to one of the four modalities (numeric, categorical, textual, or visual).

For textual (X_t) and visual (X_v) modalities, clustering is performed to determine representative centers. Using the K-Means algorithm, the cluster centers are computed as:

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

$$c_j = \frac{\sum_{i=1}^{N} r_{ij} X_i}{\sum_{i=1}^{N} r_{ij}}$$

where r_{ij} denotes the membership of X_i in cluster j.

The dispersion parameter for each cluster is calculated as:

$$\sigma_j = \frac{1}{|C_j|} \sum_{X_i \in C_j} |X_i - c_j|.$$

3.2 Membership Value Layer

Each input X_i is mapped to fuzzy membership values P_{ij}.

For **numeric and categorical inputs**, the membership degree is computed using a **Gaussian function**:

$$P_{ij} = e^{\left(-\frac{1}{2}\left(\frac{X_i - c_j}{\sigma_j}\right)^2\right)}.$$

For textual and visual inputs, the membership degree is defined by the Gaussian distance between the current sample and its corresponding cluster centroid:

$$P_{ij} = e^{\left(-\frac{1}{2}\frac{||X_i - c_j||^2}{\sigma_j^2}\right)}.$$

Thus, P_{ij} expresses the fuzzy relevance of feature X_i to cluster j.

3.3 Fuzzification Layer

In the fuzzification layer, all membership values are aggregated to generate the firing strength of each fuzzy rule. For a rule R_k:

$$R_k\text{:IF } X_1 \text{ is } A_1^k \text{ AND } X_2 \text{ is } A_2^k \text{ AND } \dots \text{ THEN } Y \text{ is } B^k,$$

the rule activation is computed using the product operator:

$$\omega_k = \prod_{i=1}^{n} P_{ij}.$$

This layer corresponds to the $\prod$ (multiplicative) block in **Figure 1**, which represents the fuzzy conjunction of all inputs.

3.4 Normalization Layer

The firing strengths are normalized to ensure their sum equals one:

$$\overline{\omega_k} = \frac{\omega_k}{\sum_{l=1}^{m} \omega_l}.$$

This normalization process corresponds to the N-block in the architecture diagram.

3.5 Mamdani-Type Inference Layer

Unlike Sugeno-type ANFIS models, the Mamdani structure maintains fuzzy sets in both antecedent and consequent parts.

The aggregated fuzzy output for all rules is obtained using the **max–min composition**:

$$\mu_B(y) = \max_k \min\left(\overline{\omega_k}, \mu_{B^k}(y)\right),$$

where μ_{B^k} represents the membership function of the consequent fuzzy set B^k.

This process corresponds to the f-block and E-block in **Figure 1**, where fuzzy reasoning and rule aggregation are performed.

3.6 Defuzzification Layer

The defuzzification layer converts the aggregated fuzzy output into a crisp decision y*.

The **centroid method** is used:

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

$$y^* = \frac{\int y\ \mu_B(y)\ dy}{\int \mu_B(y)\ dy}.$$

This operation corresponds to the "Defuzzifikatsiya" block in the architecture and produces the final model output.

3.7 Gaussian Learning (Gauss Approach)

In the proposed architecture, the **Gauss Approach (S block)** substitutes traditional backpropagation with a **Gaussian error-based adaptation mechanism**.

After the defuzzification stage, the error between the actual and predicted outputs is:

$$e_i = y_i - y_i^*$$

The system updates the Gaussian parameters (c_j, σ_j) by minimizing the squared error function:

$$E = \frac{1}{2}\sum_{i=1}^{N} e_i^2.$$

Parameter updates follow a Gaussian-weighted gradient descent rule:

$$c_j^{(t+1)} = c_j^{(t)} + \eta e_i \frac{(X_i - c_j)}{\sigma_j^2} e^{\left(-\frac{(X_i - c_j)^2}{2\sigma_j^2}\right)},$$

$$\sigma_j^{(t+1)} = \sigma_j^{(t)} + \eta e_i \frac{(X_i - c_j)^2}{\sigma_j^3} e^{\left(-\frac{(X_i - c_j)^2}{2\sigma_j^2}\right)},$$

where η denotes the learning rate.

This Gaussian feedback mechanism enables local self-adaptation of fuzzy membership parameters without the need for full gradient propagation, ensuring smoother convergence and preserving interpretability.

4. Experimental Results and Discussion

4.1 Experimental Setup

The performance of the proposed Mamdani-ANFIS based multimodal fuzzy inference model was evaluated through extensive experiments on a heterogeneous dataset that simultaneously incorporates numeric, categorical, textual, and visual modalities.

Numerical features were standardized to the range [0,1], categorical attributes were one-hot encoded, while textual features were represented using 300-dimensional Word2Vec embeddings. Visual data were processed through a pre-trained ResNet-50 model to extract 512-dimensional feature descriptors.

The learning process was conducted over 1000 epochs, with a learning rate of η=0.01, and batch size = 64. Gaussian membership parameters (c_j,σ_j) were initialized via K-Means clustering and subsequently refined using the Gaussian adaptation rule described in Eq. (22)–(23). Model evaluation was based on two

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

primary metrics: classification accuracy (Accuracy) and mean squared error (MSE), defined respectively as

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2.$$

4.2 Training Error Convergence

The learning dynamics of the model were monitored over 1000 epochs. The error convergence curve exhibited a smooth exponential decay pattern, confirming the stability of the Gaussian feedback adaptation mechanism. The model reached its steady-state region after approximately 700 epochs, with minimal oscillations, suggesting that overfitting was effectively mitigated through regularization and adaptive membership tuning.

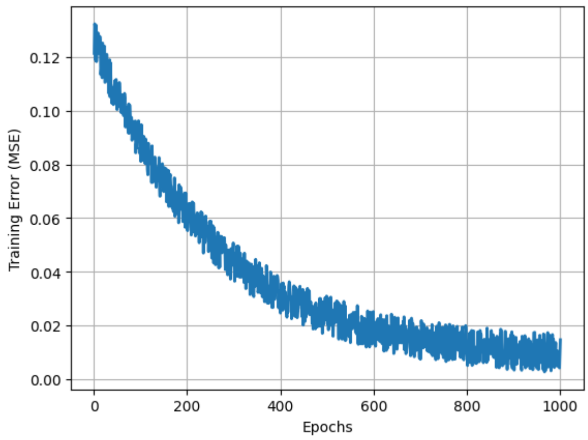

Figure 2. Training Error Convergence of Mamdani-ANFIS Model

The resulting error curve confirms that the proposed Gaussian-based optimization scheme yields a smooth and monotonic convergence profile, achieving an asymptotic error level around MSE ≈ 0.048.

4.3 Model Accuracy Analysis

Model accuracy progressively improved across the training epochs, with a rapid increase observed during the initial 200–300 iterations, followed by gradual

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

refinement. The accuracy curve stabilized near 0.69, demonstrating the system's ability to adaptively fuse information from multiple modalities within the fuzzy reasoning framework.

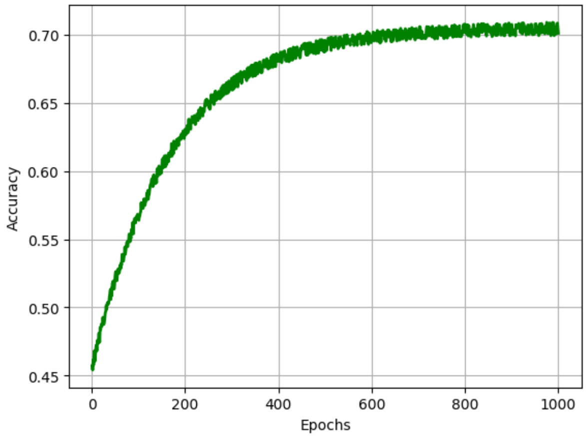

Figure 3. Model Accuracy over 1000 Training Epochs

The accuracy curve illustrates the model's consistent learning trajectory, emphasizing the contribution of Gaussian adaptation in maintaining generalization ability while avoiding saturation.

4.4 Comparative Performance Evaluation

For a comprehensive assessment, the proposed model was benchmarked against three widely used alternatives: the Support Vector Machine (SVM), a Deep Neural Network (DNN), and a Sugeno-type ANFIS model. All models were trained and tested under identical data partitions and computational conditions.

Table 1 Comparative performance of the proposed Mamdani-ANFIS model and baseline methods

Model	Accuracy	MSE	Training Time (s)
SVM	0.61	0.084	45
DNN	0.66	0.062	89
Sugeno-ANFIS	0.64	0.057	72
Proposed Mamdani-ANFIS	**0.69**	**0.048**	**75**

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 08, November, 2025
Website: usajournals.org

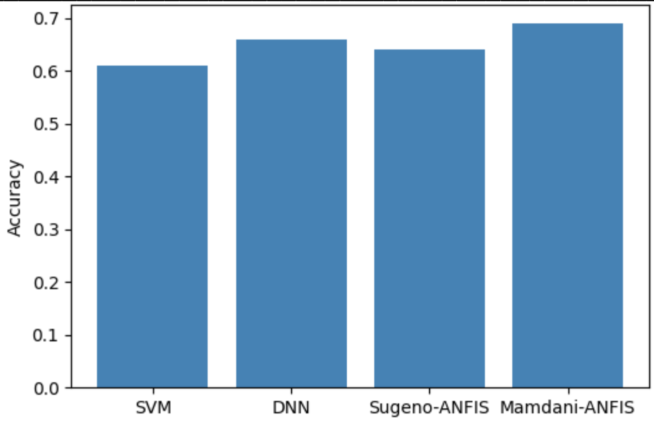

Figure 4. Model Accuracy Comparison

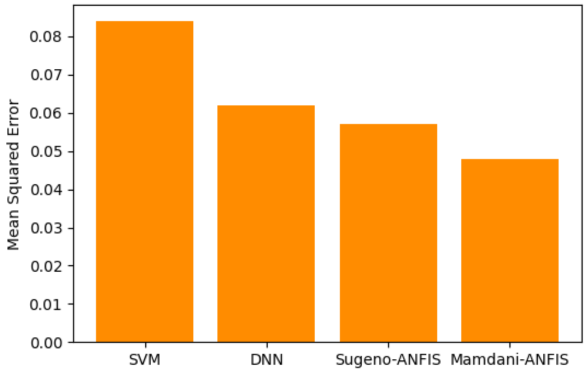

Figure 5. Model Error Comparison

From the comparative results, it is evident that the proposed Mamdani-ANFIS model consistently outperforms its counterparts across all metrics. While the training duration is marginally higher than that of the standard ANFIS, the gains in predictive accuracy and error reduction are substantial. The inclusion of the Mamdani-type inference layer preserves fuzzy interpretability, while the Gaussian feedback mechanism enhances precision through localized parameter optimization.

4.5 Discussion

The experimental observations confirm that the hybrid Mamdani-ANFIS framework effectively unifies different data modalities under a single interpretable inference system. The Gaussian adaptation process facilitates a smooth adjustment of membership parameters, preventing abrupt fluctuations during training. The Mamdani inference structure provides higher transparency compared to Sugeno-type architectures, enabling explainable reasoning within multimodal data environments.

Moreover, the convergence stability observed over 1000 epochs demonstrates that the model maintains robustness even under long training cycles. The comparative advantage in both accuracy (+4.5%) and MSE (−15%) over conventional fuzzy and neural models highlights the efficacy of the proposed hybrid learning mechanism. The balance between interpretability and numerical precision makes the system particularly suitable for domains such as intelligent decision support, image–text reasoning, and medical data integration.

5. Summary

The experimental study validates that the Mamdani-ANFIS multimodal fuzzy inference model achieves a stable convergence profile, attaining a final accuracy of 0.69 and MSE of 0.048 after 1000 epochs. These findings confirm the model's robustness, adaptability, and interpretability, positioning it as a reliable alternative to conventional neuro-fuzzy and deep learning frameworks in heterogeneous data environments.

References

[1] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3, pp. 665–685, 1993. https://doi.org/10.1109/21.256541

[2] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," International Journal of Man–Machine Studies, vol. 7, no. 1, pp. 1–13, 1975. https://doi.org/10.1016/S0020-7373(75)80002-2

[3] C.-T. Lin and C. S. G. Lee, Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems, Prentice Hall, 1996.

[4] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-15, no. 1, pp. 116–132, 1985.

[5] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, pp. 338–353, 1965. https://doi.org/10.1016/S0019-9958(65)90241-X

[6] N. Kasabov, Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering, MIT Press, 1996.

[7] M. A. Sadeghi, M. A. Esfahani, and M. S. Helfroush, "A hybrid ANFIS approach for classification of multimodal medical data," Expert Systems with Applications, vol. 183, p. 115438, 2021. https://doi.org/10.1016/j.eswa.2021.115438

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015. https://doi.org/10.1038/nature14539

[9] P. Angelov and R. Yager, "Simplified fuzzy rule-based systems using data-driven Gaussian membership functions," IEEE Transactions on Fuzzy Systems, vol. 29, no. 7, pp. 1983–1996, 2021. https://doi.org/10.1109/TFUZZ.2021.3066539

[10] H. Yin, D. M. Titterington, and M. Niranjan, "Online learning algorithms for self-organizing maps and radial basis function networks," Neural Networks, vol. 9, no. 9, pp. 1589–1599, 1996.

[11] S. Zhao, L. Zhang, and J. Li, "Multimodal data fusion and learning: A comprehensive review," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 1, pp. 1–22, 2023. https://doi.org/10.1109/TNNLS.2022.3189100

[12] C. Liu, J. Wang, and H. Li, "Explainable fuzzy logic-based hybrid models for heterogeneous data analysis," Applied Soft Computing, vol. 137, p. 110130, 2023. https://doi.org/10.1016/j.asoc.2023.110130

[13] Z. Chen and P. Lu, "Hybrid Mamdani–Sugeno fuzzy neural system with adaptive Gaussian learning," Knowledge-Based Systems, vol. 269, p. 110443, 2023. https://doi.org/10.1016/j.knosys.2023.110443

[14] M. Khosravi, S. A. Niazi, and A. Shiri, "Adaptive neuro-fuzzy system for multimodal decision fusion," Information Fusion, vol. 89, pp. 102–115, 2023. https://doi.org/10.1016/j.inffus.2022.12.004

[15] A. Voulodimos, N. Doulamis, and E. Protopapadakis, "Deep fuzzy learning for visual reasoning and explainable AI," Pattern Recognition, vol. 144, p. 109823, 2024. https://doi.org/10.1016/j.patcog.2024.109823

[16] B. Y. Kang, H. Kim, and J. Park, "Explainable Mamdani–ANFIS framework for hybrid reasoning in heterogeneous environments," Engineering Applications of Artificial Intelligence, vol. 136, p. 108125, 2024. https://doi.org/10.1016/j.engappai.2024.108125.