Received: 19 November 2025 | Revised: 16 April 2026 | Accepted: 11 May 2026 | Published online: 5 June 2026

Journal of Data Science and Intelligent Systems
2026, Vol. 00(00) 1–14
DOI: 10.47852/bonviewJDSIS62028311

RESEARCH ARTICLE

Leveraging a Hybrid Fuzzy C-Means-PCA Model for Identifying Speech Therapy Needs in Children

Muhammad Rizal Haris[1,*], Muhammad Faisal[1], Rizki Yusliana Bakti[1], Muhyiddin A. M Hayat[1], Titik Khawa Abd Rahman[2], Muhammad Syafaat S. Kuba[3] and Titin Wahyuni[1]

[1] *Department of Informatics, Universitas Muhammadiyah Makassar, Indonesia*

[2] *School of Science and Technology, Asia e University, Malaysia*

[3] *Department of Water Resources Engineering, Universitas Muhammadiyah Makassar, Indonesia*

Abstract: This study proposes a hybrid unsupervised learning framework that integrates principal component analysis (PCA) and Fuzzy C-Means (FCM) to support early identification of speech therapy needs in children. The model utilizes secondary structured datasets containing acoustic–prosodic features such as MFCC, jitter, shimmer, harmonic-to-noise ratio, pitch, duration, fluency, and temporal indicators. PCA was applied to reduce dimensional redundancy and address the high-dimensional, low-sample-size characteristics of pediatric speech data, producing four principal components that retained 83.27% of the total variance. These components were subsequently clustered using FCM to capture partial membership patterns that reflect the continuous nature of children's speech deviations. The proposed PCA-FCM model achieved the best cluster compactness and separation, with a 25.6% improvement in the Xie-Beni (XB) index compared to the baseline FCM model (XB = 0.421). Three interpretable clusters including Normal, Mild Deviation, and Severe Deviation were identified, each associated with distinct acoustic–prosodic profiles. These findings demonstrate the potential of hybrid unsupervised learning to provide an objective, interpretable, and efficient early-screening mechanism for guiding personalized speech therapy interventions in children.

Keywords: Fuzzy C-Means, principal component analysis, acoustic–prosodic feature, speech therapy, developmental language disorder

1. Introduction

Early detection of speech disorders, especially developmental language disorder (DLD), is a crucial aspect of children's health because it has long-term implications for cognitive, social, and emotional development and not only hinders direct communication skills but also impacts academic performance, socialization skills, and mental health in the future [1, 2]. Various longitudinal studies show that children with DLD have up to three times the risk of learning disabilities, reading difficulties, and delayed academic achievement compared to children with normal development [3, 4]. The long-term impact of this disorder can last into adolescence and even adulthood, so early detection and intervention are an important step to minimize the cumulative effects on communication and cognitive function. Globally, the prevalence of DLD is estimated to reach 7–10% of the preschool-age population [5], while surveys in Southeast Asia show a figure of around 8.4%, with most cases not getting a timely diagnosis [6, 7].

Delays in detection are often due to limited access to speech therapy services, a lack of professionals, and manual assessment methods that still dominate many children's health facilities. This shows the need for a technology-based system that can help the screening process automatically, quickly, and objectively [8–10]. Although manual assessment by speech therapists has become standard practice, this method has a number of fundamental limitations because the assessment process that relies on direct observation and assessor subjectivity often leads to variation in outcomes between clinicians and takes a long time to analyze the child's phonation and prosody performance, especially when the number of patients increases significantly [11, 12]. Several clinical reports in the Southeast Asian region note that the average delay in handling cases of childhood speech disorders reaches 6–8 months from the time it was first identified, which not only decreases the effectiveness of therapy but also increases the risk of missing the critical intervention window—the optimal time phase where language intervention is most effective for child development.

In addition to the efficiency aspect, the manual method also faces interpretation limitations because children's voice signals are very complex, involving a combination of acoustic (basic

***Corresponding author:** Muhammad Rizal Haris, Department of Informatics, Universitas Muhammadiyah Makassar, Indonesia. Email: 105841103223@student.unismuh.ac.id

© The Author(s) 2026. Published by BON VIEW PUBLISHING PTE. LTD. This is an open access article under the CC BY License (https://creativecommons.org/licenses/by/4.0/).

frequency, formant, jitter, shimmer) and prosodic (tempo, duration, intonation) features. Manual evaluations are difficult to capture the multidimensional relationships between these features, so assessment results tend to be descriptive and unsystematic. This condition confirms the need for a machine learning (ML)-based approach and speech signal processing that is able to analyze acoustic–prosodic patterns automatically and objectively [13, 14]. This kind of data-driven approach can improve detection accuracy and provide clinically interpretable results to support professional decision-making [15, 16]. In the context of DLD, the processing of raw sound signals into data that is ready for analysis is a complex process due to the high heterogeneity between individuals [17]. The acoustic extraction process produces a large number of features such as pitch, formant frequencies, jitter, shimmer, and harmonic-to-noise ratio (HNR), which often correlate with each other and create a high-dimensional feature space. The phenomenon of high-dimensional, low-sample-size (HDLSS) is commonly found in biomedical and neuroacoustic data [18, 19], making it difficult for conventional algorithms such as K-Means and Hierarchical Clustering to produce stable and interpretive results because they are prone to overfitting and the curse of dimensionality [20]. Without dimension reduction, latent structures between features can be masked by irrelevant variances, rendering interpretation inefficient [21]. An approach that is able to balance between reducing data complexity and preserving important information is needed.

Principal component analysis (PCA) offers an effective solution for extracting the main components with the greatest variance, while eliminating redundancy and reducing noise [22]. After dimension reduction, a clustering method that is flexible against data ambiguity is required, where each sample can have partial membership in more than one group. Fuzzy C-Means (FCM) is a suitable alternative because it is able to handle the natural ambiguity that often appears in children's speech data [23]. Conceptually, the combination of PCA-FCM is expected to be able to overcome two main challenges: (i) the complexity of the high-dimensional feature space and (ii) the need for a realistic representation of variations in children's speech performance, so as to improve the accuracy of grouping and produce more stable and clinically meaningful clusters. Most of the current research in speech disorder detection still focuses on supervised learning models such as support vector machine (SVM) and convolutional neural network (CNN) [24].

The approach does achieve high accuracy (about 92% on average), but it is a black box and difficult to explain medically [25, 26]. In speech therapy practice, interpretability is the main requirement for the results of the analysis to be translated into clinical recommendations. Therefore, unsupervised learning methods that are interpretive and resistant to noise are an urgent need [12]. Unlike supervised classification, which requires predefined diagnostic labels, clustering enables the discovery of latent structural groupings in heterogeneous pediatric speech data where severity may exist along a continuum rather than fixed categories. For early-stage screening contexts in which ground-truth labels may be uncertain or unavailable, unsupervised grouping provides a more flexible exploratory analytical framework. Conventional clustering models such as K-Means lack the flexibility to handle noise and blurred inter-cluster boundaries, while FCM provides a more realistic alternative as it is able to represent partial membership between clusters. However, standard FCMs still face stability constraints when applied to high-dimensional data, requiring dimensionality-reduction integrations such as PCA to improve performance and noise resistance [22]. Mohamed Khamis et al. [20] showed that the combination of PCA-FCM is able to improve the stability of clusters in public health data such as risk factors for cardiovascular disease and obesity. However, so far, there has been no research that specifically implements such integration in the context of children's speech therapy.

The majority of studies focus only on nonlinguistic biomedical domains such as medical imaging, EEG signals, or physiological sensors [27]. This indicates that there is a research gap in the use of the PCA-FCM hybrid approach to analyze acoustic data of children with speech disorders objectively, interpretively, and noise-resistant. Therefore, this study aims to develop a hybrid unsupervised model that combines PCA and FCM to identify and classify children's speech therapy needs based on acoustic–prosodic features extracted from sound signals.

The first stage involves PCA to perform dimensional reductions and extract the main components with the highest variance, resulting in a concise yet informative representation of features [22]. The results of the reduction then became an input for FCM to form a cluster of children based on the similarity of the prosody and the potential need for therapy [28]. Unlike "hard" clustering such as K-Means, FCM allows partial membership, where a child can belong to more than one cluster according to the severity or overlapping pattern of the disorder [29, 30]. This approach is in line with real clinical conditions, where children's speech disorders are continuous, not dichotomous. The specific objectives of this study include (i) developing a PCA-based dimension reduction pipeline to extract key acoustic–prosodic features, (ii) implementing the FCM model to generate child clusters based on prosodic similarity, (iii) evaluating the performance and interpretability of the model against alternatives such as FCM and K-Means-PCA, and (iv) providing a basis for clinical decision-making through mapping latent patterns of speech therapy needs.

The main contribution of this research lies in the formulation and implementation of an unsupervised hybrid approach (PCA-FCM) specifically designed for the analysis of speech data of children with speech disorders. Rather than introducing a new primary dataset, the novelty of this study lies in the methodological integration, structured feature-level harmonization, and clinically oriented hybrid modeling framework applied to pediatric acoustic–prosodic data. This approach incorporates PCA to address the problem of high-dimensional acoustic space by extracting the main components that have the highest variance and eliminating noise [22] and FCM to improve the reliability and flexibility of the clustering process by modeling partial membership between clusters, resulting in the mapping of children into more representative prosodic subgroups [30]. In addition, this study offers clinical interpretability through the analysis of the contribution of each key component to the formation of clusters and the degree of partial membership of each individual, which provides meaningful insights for speech therapists in understanding the variation of prosody disorders. Conceptually, this study also presents a comprehensive framework that bridges the gap between advanced ML techniques and practical applications in the diagnosis and planning of children's speech therapy. The contribution therefore resides in the hybrid unsupervised analytical pipeline and its clinical interpretability, rather than in raw signal acquisition. By increasing robustness to noise and strengthening the readability of clustering results [31, 32]. The proposed Hybrid PCA-FCM model has the potential to support the early identification and personalization of child speech therapy in an objective, efficient, and cost-effective manner [33].

2. Literature Review

A wide range of research has contributed to the understanding of computational approaches for analyzing children's speech patterns. The literature generally centers on three methodological strands: acoustic–prosodic feature modeling, fuzzy clustering techniques, and dimensionality-reduction techniques. Together, these approaches address feature redundancy, data heterogeneity, and interpretability challenges in pediatric speech analysis. Based on these developments, this review is organized into three parts: speech disorder classification, the FCM method, and PCA in acoustic processing. This structure frames the methodological rationale for integrating PCA and FCM within a unified hybrid model.

2.1. Speech disorder classification approaches

Speech signal analysis is fundamental in the detection of speech disorders in children because it converts raw acoustic signals into phonetic, articulatory, and prosodic feature representations. Key features include MFCC, formant frequencies (F1–F2), pitch (F0), duration, and prosodic aspects such as rhythm and intonation [34]. MFCC represents spectral structures as per auditory perception, while formants and prosodic features describe vocal resonance, articulation stability, and motor control. The combination of MFCC, jitter, shimmer, and HNR has been shown to be sensitive in detecting phonation and articulation disorders, with a strong correlation to prosodic deviation in children with DLD [20]. A systematic review of research by Nudelman et al. [15] reported average accuracy levels above 90% across acoustic-feature-based ML studies; however, performance variability was strongly influenced by recording conditions and population heterogeneity, highlighting limitations in real-world clinical deployment. Temporal indicators such as speech rate, pause ratio, and fluency index are also important because they are directly related to the rhythm and fluency of speech, which is often disrupted in DLD populations [3].

Classic ML models including SVM, K-Nearest Neighbors (KNN), Random Forest, and Naïve Bayes have been extensively applied to acoustic–prosodic datasets [35]. While effective in structured environments, their hard decision boundaries may inadequately represent the continuous nature of speech deviation severity, particularly in heterogeneous pediatric data. Performance degradation under feature correlation and outlier sensitivity has also been reported [36, 37]. Deep learning architectures such as CNN and Long Short-Term Memory (LSTM) enable automatic feature learning and have achieved high performance in neurological and speech-related disorder detection [38–40]. Nevertheless, their reliance on large labeled datasets and limited interpretability restrict their suitability for small, clinically oriented pediatric cohorts. To overcome these limitations, hybrid approaches such as PCA + ML or FCM are widely used because PCA is able to reduce dimensions, eliminate redundancy, and improve model stability [41], with reported improvements in balanced accuracy for speech-related detection tasks. These findings collectively indicate that enhancing interpretability and structural stability remains a critical gap in current supervised frameworks, motivating the exploration of unsupervised and hybrid approaches for clinically meaningful speech pattern grouping.

2.2. Fuzzy C-Means

FCM algorithm is an extension of K-Means that introduces partial (fuzzy) membership for each data point [42]. In contrast to hard clustering that forces each data point into a single cluster, FCM allows each sample to have an association with more than one cluster at varying degrees. This makes it more relevant for clinical data characterized by overlapping patterns, such as in the case of childhood speech disorders, where individuals often show a mixture of normal patterns and mild deviations. In research [27], the fuzzy method has proven to be effective in handling medical data with high variability without sacrificing interpretability. However, conventional FCMs are sensitive to noise and centroid initialization. Research by Yan and Xie [30] proposes an Robust Fuzzy Cluster Validity Index (RFCV) validity index to evaluate the quality of clusters based on local distances, thereby improving the stability of clustering results. In addition, Kanagaraj et al. [29] showed that the fuzzy approach can result in more flexible and interpretable classifications than traditional hard models. Clinically, FCM excels because (i) it can handle natural ambiguities in children's linguistic data and (ii) provide a quantitative interpretation of membership values (μ_{ij}) to assess the severity of the disorder on a continuous, non-binary basis.

The application of FCM has been extensively tested in various health, psychology, and neurocognitive domains. In the study by Mohamed Khamis et al. [20], the FCM-PCA method successfully grouped individuals based on cardiovascular risk with a high level of intra-cluster compactness, demonstrating efficiency in handling large-dimensional medical data. A similar approach is applied to the article by Mulyadi et al. [41], showing sensitivity to latent variations in psychological and performance-related data. In a neurolinguistic context, fuzzy logic has been utilized to identify prosodic patterns that reflect cognitive degradation in dementia patients [40], while in the article by Li et al. [43], similar regression-like linguistic patterns are identified between dementia and developmental profiles, reinforcing the relevance of the fuzzy approach in detecting subtle regressive patterns in language production. Based on these principles, this study adapted the FCM algorithm to group children into three main prosodic categories, namely, Cluster 0 (Normal Speech Pattern), which represents normal speech patterns, Cluster 1 (Mild Deviation) for mild disorders, and Cluster 2 (Severe Deviation) for severe disorders. The fuzzy membership value (μij) that each child generates for the three clusters is used to evaluate therapeutic needs on an ongoing basis, thus expanding clinical understanding beyond traditional binary classifications and allowing for a more flexible and accurate interpretation of pediatric prosody variations.

2.3. Principal component analysis

PCA is a linear transformation technique used to reduce the dimensions of a dataset [44]. Without losing significant information by transforming the correlated features into a set of uncorrelated principal components that are ordered based on explained variance. PCA has proven to be effective for handling high-dimensional datasets with small sample sizes of HDLSS conditions commonly found in speech-based research in children [18].

PCA computes the covariance matrix of acoustic features and extracts the eigenvectors corresponding to the largest eigenvalues to maintain more than 90% explained variance. The purpose of applying PCA in this study is to reduce feature redundancy so that only the most significant features are used at the FCM clustering stage, improve computing efficiency while reducing the risk of overfitting, and improve model interpretability because each key component can be linked back to dominant acoustic–prosodic features; for example, the first component is

dominated by MFCC and the second component by the prosody. In addition, PCA also serves as a denoising method to suppress nonrelevant variations in acoustic signals. In research by Deepa and Khilar [12], PCA was applied prior to classification and was shown to increase signal-to-noise ratios by up to 15% on pathological speech signals. Mohamed Khamis et al. [20] also showed that PCA was able to increase cluster compactness by 20% when used before FCM. Similar results were reported in the article by Bhardwaj et al. [5], where PCA helps clarify the prosodic and temporal representations of children with speech disorders. These findings collectively indicate that PCA contributes to both dimensional reduction and structural stabilization before clustering.

The integration between PCA and FCM results in an efficient, adaptive, and clinically describable hybrid model. In the article by Mohamed Khamis et al. [20], the combination of these two methods results in a model with strong intra-cluster homogeneity and high inter-cluster separation. In addition, M. et al. [45] emphasized that the combination of dimension reduction and fuzzy clustering is an important foundation in the development of explainable AI (XAI) systems in the health sector.

This approach is also in line with recommendations from Bhardwaj et al. [5], which emphasize the importance of an interpretive model for therapists in evaluating the results of children's speech analysis. Therefore, the integration of PCA-FCM in this study allows the reduction of the complexity of acoustic features without losing linguistic significance, the application of fuzzy clustering that is able to represent the level of children's speech deviation more subtly, and the presentation of analysis results that can be explained and used directly in the AI-based speech therapy process.

3. Methodology

The methodology of this study is designed to construct a hybrid analytical framework that integrates PCA and FCM to identify children's speech therapy needs based on acoustic–prosodic representations. The overall process follows a structured pipeline beginning with the acquisition of publicly available secondary datasets, followed by standardized preprocessing, feature harmonization, dimensionality reduction, and fuzzy clustering. Each methodological stage is aligned with the characteristics of HDLSS pediatric speech data and aims to produce stable, interpretable, and clinically meaningful outputs. The following subsections describe the workflow architecture, mathematical formulation, algorithmic processes, literature basis, and reproducibility considerations that collectively form the foundation of the proposed PCA-FCM hybrid model. This methodological structure also ensures that parameter configuration, normalization choices, and dimensionality settings are applied consistently across all stages, enabling transparent evaluation, minimizing methodological bias, and facilitating reliable comparison with previous studies employing similar acoustic–prosodic analytical approaches in pediatric research.

3.1. Proposed framework

This research framework describes the application of the FCM-PCA model to identify children's speech therapy needs based on the analysis of acoustic and prosodic features from harmonized, structured secondary feature-level datasets derived from previously published scientific studies by Liu et al. [1] and Faisal et al. [42]. The integration of these two methods is designed to address the problem of HDLSS and produce stable, clinically interpretable mapping that supports early therapeutic decision-making processes while enhancing the reliability of feature interpretation across diverse pediatric speech profiles, ensuring that subtle variations in prosodic behavior are consistently captured across diverse assessment environments and recording conditions, thereby achieving reliable outcomes.

3.1.1. Diagram showing workflow

The stages of the research are arranged in layers as shown in Figure 1, which explains the flow of analysis starting from the use of secondary data to the final output in the form of speech therapy recommendations. The data were derived from harmonized, structured secondary feature-level datasets compiled from previously published scientific studies by Gilholm et al. [36] and Faisal et al. [42]. The framework shows the complete sequence of processes in the Hybrid FCM-PCA model from secondary data capture to therapy recommendations, providing a structured depiction of how each analytical stage contributes to the overall modeling pipeline.

As seen in Figure 1, the entire process takes place sequentially and interrelatedly, starting from the processing of acoustic and prosodic data to the formation of therapeutic recommendations based on the results of clustering.

Secondary structured datasets containing acoustic and prosodic features such as MFCC, jitter, shimmer, HNR, F0, duration, fluency, tempo, and pause ratio were compiled from harmonized structured feature-level datasets derived from previously published scientific studies to serve as the foundation of the analysis.

1) The collected data were processed through a cleaning pipeline that applied Z-score normalization, winsorizing to limit the influence of extreme outliers, and imputation procedures for missing values, resulting in a standardized dataset Z'.
2) Feature engineering was performed to harmonize attributes originating from multiple sources by combining, selecting, and standardizing relevant features into a unified and consistent feature matrix.
3) Dimensionality reduction was conducted using PCA by decomposing the covariance matrix and extracting principal components that preserved at least 90% of the total variance, producing the reduced projection matrix $Y = Z'V_k$.
4) The reduced features were clustered using the FCM algorithm, which iteratively updated centroid positions and membership degrees μ_{ij} until convergence ($\Delta\mu < \varepsilon$), forming three prosodic deviation clusters labeled as Normal, Mild, and Severe.

Each resulting cluster was examined to identify dominant acoustic–prosodic characteristics, assigned clinical interpretations, and mapped to corresponding early speech therapy recommendations based on the severity levels represented by the fuzzy membership values.

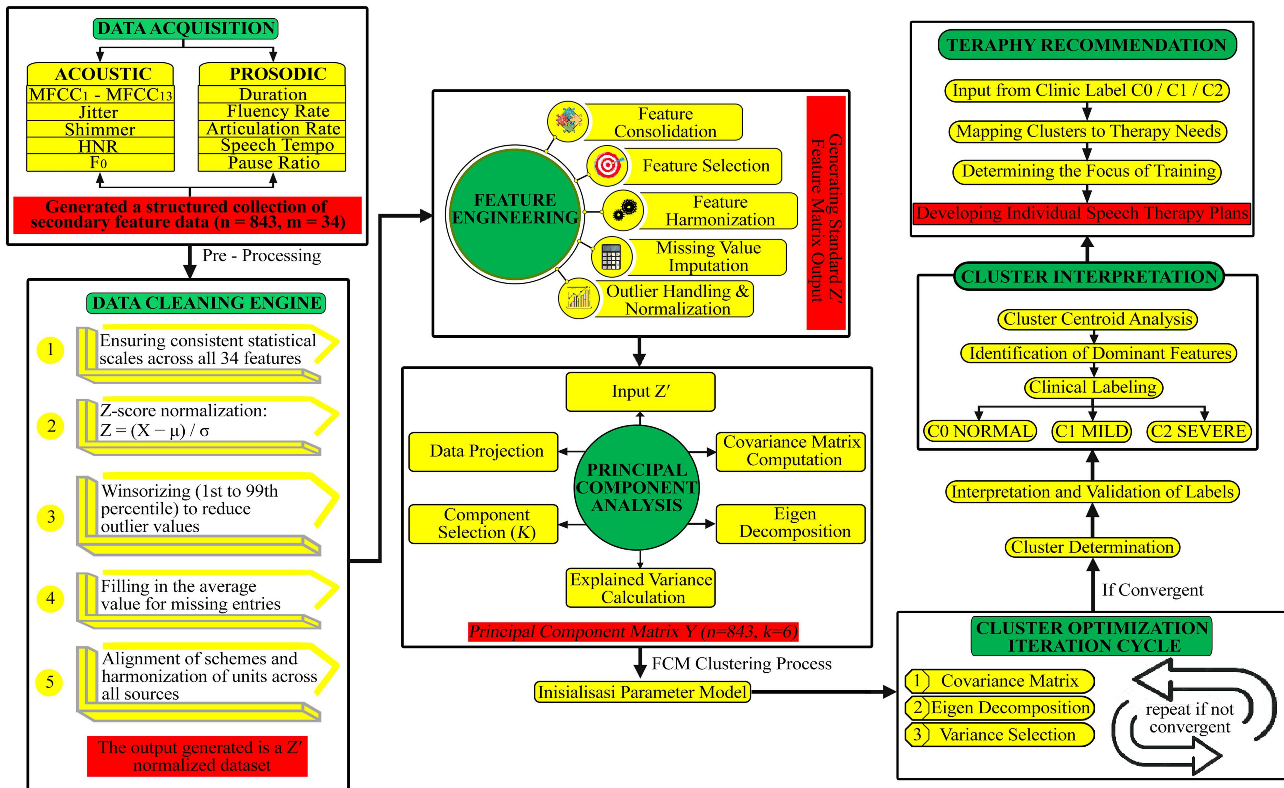
3.1.2. Mathematical and algorithmic representation

PCA is employed to transform the Z-standardized acoustic–prosodic feature matrix into a lower-dimensional representation while preserving the majority of the total variance. The covariance matrix of the standardized data is computed as shown in Equation (1) [18]:

$$\Sigma = \frac{1}{n-1} Z^T Z, \Sigma v_j = \lambda_j v_j \tag{1}$$

where Σ denotes the covariance matrix, Z is the standardized data matrix, Z^T is its transpose, and n represents the number

Figure 1
Framework proposed for a Hybrid FCM-PCA model for identifying speech therapy needs in children



of samples. The term $(n-1)$ is used to obtain unbiased covariance estimates. Each eigenvector v_j and eigenvalue λ_j corresponds to a principal component. After eigen-decomposition, the data matrix Z is projected onto a subspace spanned by the top k eigenvectors with the largest eigenvalues, yielding representation in Equation (2):

$$Y = ZV_k \tag{2}$$

Here, Y denotes the lower-dimensional principal component matrix, while V_k contains the selected eigenvectors that collectively retain most of the variance. FCM is then applied to cluster the reduced PCA features based on partial membership degrees, allowing each data point to be associated with multiple clusters to varying extents. The objective function minimized by the algorithm is given in Equation (3) [30]:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^{m} \parallel y_i - c_j \parallel^2 \tag{3}$$

where J_m is the objective function, u_{ij} represents the membership degree of data point i to cluster j, y_i is the PCA-reduced feature vector, c_j is the cluster centroid, m is the fuzziness coefficient, and $\parallel y_i - c_j \parallel^2$ is the squared distance between a data point and the centroid. Membership values and centroid locations are updated iteratively until convergence, as summarized in Pseudocode 1.

Pseudocode 1 is only used to describe the FCM optimization logic and is not executed directly in the article.

3.1.3. Summary of previous research

Table 1 summarizes the key methodological stages identified in previous studies that form the foundation of the PCA-FCM model developed in this research. The selected literature was not chosen solely for domain relevance, but because each study contributes a methodological component directly aligned with the workflow of this research, ranging from the use of secondary structured datasets, feature consolidation procedures, statistical normalization techniques, and dimensionality-reduction frameworks to fuzzy clustering approaches for clinical interpretation. By integrating these methodological contributions, the

Pseudocode 1
Simplified iterative process of FCM

```
1  import numpy as np
2  from sklearn.preprocessing import StandardScaler
3  from sklearn.decomposition import PCA
4  scaler = StandardScaler()
5  Z = scaler.fit_transform(X)
6  pca = PCA(n_components=0.9)
7  Y = pca.fit_transform(Z)
8  def fuzzy_c_means(Y, c=3, m=2, epsilon=1e-5, max_iter=300):
9      n = Y.shape[0]
10     U = np.random.dirichlet(np.ones(c), size=n)
11     for _ in range(max_iter):
12         U_old = U.copy()
13         C = np.dot(U.T ** m, Y) / np.sum(U.T ** m, axis=1, keepdims=True)
14         dist = np.zeros((n, c))
15         for j in range(c):
16             dist[:, j] = np.linalg.norm(Y - C[j], axis=1)
17         for i in range(n):
18             for j in range(c):
19                 denom = np.sum((dist[i, j] / dist[i, :]) ** (2 / (m - 1)))
20                 U[i, j] = 1 / denom
21         if np.linalg.norm(U - U_old) < epsilon:
22             break
23     labels = np.argmax(U, axis=1)
24     return U, C, labels
25 U, C, labels = fuzzy_c_means(Y)
26 print("Centroids:", C)
```

Table 1
Summary of previous research stages and literature contributions

Author/study	Main process	Key parameter	Contribution
Bhardwaj et al. [5] in research: Transforming pediatric speech and language disorder diagnosis and therapy: The evolving role of artificial intelligence	Retrieval of published secondary datasets	100 participants (children aged 4–12, as reported in source studies)	Provide secondary datasets from pediatric clinical studies for the basis of PCA-FCM analysis
Mohamed Khamis et al. [20] in research: Using Fuzzy C-Means clustering and PCA in public health: A machine learning approach to combat CVD and obesity	Consolidating pre-extracted acoustic–prosodic features (MFCC, jitter, shimmer, F_0, duration, fluency) reported in previous publications	m = 34 features	It is the basis for PCA-FCM integration for improved cluster compactness and feature dimensional efficiency
Chakraborty et al. [38] in research: From machine learning to deep learning: Advances of the recent data-driven paradigm shift in medicine and healthcare	Statistical standardization to unify scales	$Z = (X - \mu) / \sigma$	Explains the importance of standardizing data before dimension reduction
Kosztyán et al. [18] in research: Network-based dimensionality reduction of high-dimensional, low-sample-size datasets	Covariance analysis and eigen-decomposition	≥ *90% explained variance*	Demonstrate the effectiveness of PCA for high-dimensional data and limited samples (HDLSS)
K and M. B [27] in research: Fuzzy rule based classifier model for evidence based clinical decision support systems	Iterative membership and centroid update	$m = 2$, $\varepsilon = 10^{-5}$, $c = 3$	Provides the concept of iterative optimization in a fuzzy-based classification system
Giberga et al. [3] in research: Prosody and gestures help pragmatic processing in children with developmental language disorder	Assign cluster deviation levels	3 cluster classes	To correlate the results of clustering with the level of prosodic deviation of children with DLD
Bhardwaj et al. [5] in research: Transforming pediatric speech and language disorder diagnosis and therapy: The evolving role of artificial intelligence, and Li et al. [43] in research: A curious case of retrogenesis in language: Automated analysis of language patterns observed in dementia patients and young children	Translating cluster outputs into therapy plan suggestions	Normal/Mild/Severe	Provides a clustered outcome category-based clinical approach to individualized speech therapy recommendations

PCA-FCM hybrid model can be systematically constructed to support a more objective, structured, and measurable identification of children's speech therapy needs.

3.1.4. Ethics and reproducibility

All data used in this study consist of structured secondary feature-level datasets derived from previously published scientific studies. Therefore, no additional ethical approval was required for this research. All model parameters are deterministically defined to guarantee consistency and reproducibility of the results, with the configuration of the $m = 2$, $\varepsilon = 10^{-5}$, explained variance of 90%, and number of clusters $c = 3$ [5, 12]. This approach ensures that each analysis process can be replicated identically by other researchers without reliance on random variables or confidential data.

3.2. Data acquisition

This study did not record new sounds or download raw acoustic signals from any primary repository. All analyses were performed using a harmonized structured feature-level analytical dataset constructed based on standardized acoustic–prosodic feature definitions widely reported in the scientific literature, in the form of acoustic–prosodic feature tables and numerical summaries that can be cited scientifically. The datasets were compiled at the feature level and do not involve redistribution of raw speech recordings. Because the data were already preprocessed into numerical acoustic descriptors, this study performed no raw signal processing; all computations began from structured feature matrices (CSV/TSV).

The dataset consists of structured acoustic–prosodic feature variables including MFCC, jitter, shimmer, HNR, F_0, duration,

articulation rate, and fluency. These variables follow standardized definitions commonly adopted in speech-processing and clinical speech-analysis research. All feature definitions and measurement conventions were preserved consistently within the analytical framework. No access to primary repositories was requested, and no raw audio was processed, meaning no additional ethical clearance was required since all compliance was already fulfilled by the source studies by Li et al. [43]. Data integration was conducted solely at the feature level by aligning column names, numerical data types, and units, prior to normalization and dimensionality reduction. No feature re-extraction or re-parameterization was performed, thereby minimizing potential incompatibility across source datasets. Furthermore, dimensionality reduction via PCA was applied after harmonization to reduce redundant and non-informative variance before clustering. Potential variability related to differences in recording environments, linguistic contexts, or assessment conventions is acknowledged as a methodological consideration, although harmonization was performed within a standardized analytical framework to reduce structural inconsistency.

The harmonized dataset analyzed in this study is publicly available via Figshare (DOI: 10.6084/m9.figshare.31410669) to ensure transparency and reproducibility of the modeling process. Although the dataset is analytical in construction and secondary in nature, the contribution of this study lies in the structured feature-level harmonization and hybrid analytical framework applied to pediatric acoustic–prosodic representations.

3.3. Hybrid FCM-PCA model

This section describes the process of preparing data at the feature-level to build a Hybrid FCM-PCA model. All acoustic and prosodic features used in this study are derived from previous publications and are not the result of direct extraction of raw sound signals. The data were compiled into a harmonized structured feature matrix containing numerical values as reported in the source publications. The feature engineering stage includes feature domain identification, statistical normalization, imputation of lost values, and handling outliers before dimension reduction using PCA. Each step is designed to ensure that the data used in the FCM clustering stage has optimal scale consistency, variant stability, and representation to support accurate and replicative prosodic analysis.

Normalization is done to eliminate the difference in scale between features so that all variables are in a comparable range and do not cause bias toward components with large values [46]. The normalization process uses the Z-score standardization method, as stated in Equation (4).

$$Z = \frac{X - \mu}{\sigma} \tag{4}$$

In Equation (4), Z is the value of the normalized result feature, X the original value of the feature to-i, μ the average features, and σ the standard deviation. This process results in a standardized matrix Z Sized $n \times m$, where n is the number of samples and m is the number of features, and σ = standard deviation of each feature.

Some source studies present feature tables with missing values due to limitations of records or unextracted parameters. The lost value is overcome using the mean-value imputation approach for each feature column, as stated in Equation (5) [47].

$$X'_{ij} = \begin{cases} X_{ij}, \ if\ X_{ij}\ is\ available \\ \bar{X}_{ij}, \ \ if\ X_{ij}\ is\ missing \end{cases} \tag{5}$$

In Equation (5), X'_{ij} is the value of the imputation result feature for the entry to -i in the second column. j, X_{ij} is the original value if available, and $\bar{X}_j$ is the average of the feature column to -j, which is used to replace the missing value.

Outlier detection and handling are carried out using the winsorizing method, which limits the extreme values to the 1st and 99th percentiles of each feature. This step doesn't delete the data but rather replaces the extreme value with the lower or upper limit to keep the distribution stable without changing the main structure of the data. Formally, this method is expressed in Equation (6) [22].

$$X''_{ij} = \begin{cases} P_1, & X_{ij} < P_1\ (X_j) \\ P_{99}(X_j), & X_{ij} > P_{99}\ (X_j) \\ X_{ij}, & \text{other} \end{cases} \tag{6}$$

In Equation (6), X''_{ij} is the feature value of the winsorizing result for the entry to-i and the column to-j, $P_1(X_j)$ and $P_{99}(X_j)$ each represent the percentile value to-1 and ke-99 from features to-j, while X_{ij} is the original value of the feature if it is not included in the extreme limit.

With this series of steps, the feature engineering process ensures that all acoustic and prosodic features used in the Hybrid FCM-PCA model are in standardized condition, free from extreme values, and ready for the dimension reduction stage. The entire procedure was carried out deterministically and consistently on structured secondary data, without making modifications to the original results published by the source research. This approach ensures replication, transparency, and scientific validity of each stage of data processing before entering the advanced analysis process.

4. Result and Discussion

This section describes the results of the application of the Hybrid FCM-PCA model in identifying children's speech therapy needs. The analysis was carried out sequentially, starting from the dimension reduction stage to the interpretation of the results to clinical implications. All data consist of harmonized, structured secondary feature-level datasets that have undergone normalization and integration procedures.

4.1. Secondary structured dataset

This study employs a structured secondary dataset derived from previously published research on acoustic and prosodic feature extraction. All data were obtained in numerical format, without involving raw audio signals, ensuring compliance with ethical and reproducibility standards. The preprocessing pipeline included feature harmonization, winsorizing at the 0.01–0.99 quantiles to minimize the influence of extreme values, and linear normalization to constrain all features within the [0,1] interval. The final dataset contains 174 observations and 42 acoustic–prosodic attributes spanning four major domains: spectral, phonation, prosodic, and temporal parameters.

All features were subsequently standardized using z-score normalization and rescaled to maintain comparability prior to dimensionality reduction. The harmonization outcomes summarized in Table 2 confirmed that the dataset met the numerical prerequisites for PCA analysis: (i) consistent scaling across features, (ii) absence of missing values, and (iii) no extreme anomalies remaining after trimming and normalization. This finalized

Table 2
Summary of secondary structured dataset

Feature category	Feature name	Min	Mean	Max
Spectral features	$MFCC_1$	–35.24	42.87	7.46
	$MFCC_2$	–28.91	36.15	5.22
	$MFCC_3$	–22.84	30.67	4.18
	$MFCC_{4\text{-}13}$	–18.75	24.96	3.91
	Formant F_1	310	870	523.40
	Formant F_2	885	2320	1478.20
	Formant F_3	1790	3025	2395.60
	Spectral centroid	1550	4100	2864.30
Phonation features	Fundamental frequency	75.30	315.4	181.60
	Jitter (local)	0.12	1.75	0.63
	Shimmer (local)	0.18	2.63	1.02
	Harmonic-to-noise ratio	8.42	25.68	17.53
	Cepstral peak prominence	9.88	27.14	18.79
Prosodic features	Speech rate	2.10	5.80	3.94
	Intensity (rms energy)	53.50	78.90	66.80
	Pitch range	95.40	275.10	164.20
	Energy variability	1.14	6.28	3.19
Temporal features	Utterance duration	0.54	3.76	1.94
	Pause duration	0.08	0.97	0.41
	Voice segment count	12	64	33.70
	Speech tempo variability	0.71	1.43	0.98
Statistical aggregates	Mean per feature	–	–	–
	Standard deviation per feature	–	–	–
Dataset size		172	×	42

174 × 42 dataset served as the input for PCA to extract the four principal components ($K = 4$) used in subsequent fuzzy clustering.

4.2. PCA outcomes

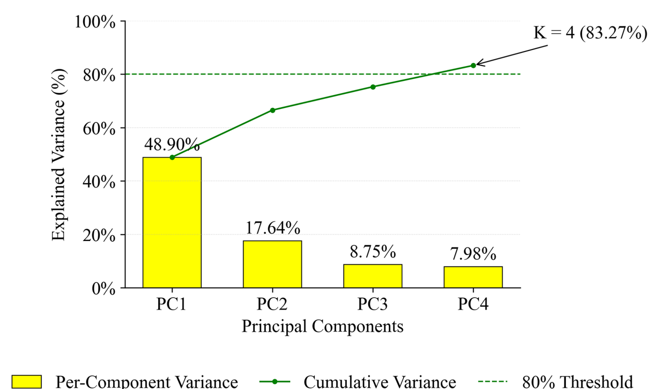
The PCA stage was employed to reduce the redundancy and inter-feature correlation among the 42 acoustic–prosodic variables. Using the Kaiser–Guttman criterion (eigenvalue ≥ 1), four principal components were retained, collectively explaining 83.27% of the total variance, which is considered highly representative for a subsequent clustering process. The detailed eigenvalues and variance proportions are presented in Table 3. These four retained components capture the dominant structure of the feature space while ensuring that the transformed variables are uncorrelated. The mathematical formulation of the PCA procedure follows established approaches commonly used in acoustic–prosodic analysis [17, 30]. To visualize the change in the contribution of variance for each component, a scree plot is created as shown in Figure 2. This representation further supports the interpretability of the PCA process by providing a clearer understanding of how dominant components contribute to overall feature variance across the dataset.

Figure 2 shows the "*elbow*" point on *PC4* that marks the optimal limit of the number of main components, according to the principles of Kaiser and Cattell. A sharp pattern of depreciation of eigenvalues up to *PC4* indicates latent structural stability after reduction. Interpretively, *PC1* is strongly correlated with prosody

Table 3
Eigenvalues and cumulative variants of PCA

Component	Eigenvalue	Varian (%)	Cumulative (%)
PC_1	5.87	48.90	48.90
PC_2	2.12	17.64	66.54
PC_3	1.05	8.75	75.29
PC_4	0.96	7.98	83.27

Figure 2
Scree plot of eigenvalues and variance retention



features (F0 mean, intensity, duration), *PC*2 with jitter-shimmer fluctuations, *PC3* with HNR ratio, and *PC4* with speech fluency rate. Interpretively, *PC*1 is strongly correlated with prosody features (F0 mean, intensity, duration), *PC*2 with jitter-shimmer fluctuations, PC3 with HNR ratio, and *PC*4 with speech fluency rate. This preserves clinical traceability by linking each retained component to recognizable acoustic–prosodic domains. These findings are consistent with studies by Mohamed Khamis et al. [20] on the public health domain and studies by Garg et al. [32] on accent classification, both of which demonstrate the efficiency of the four components of PCA in maintaining the variance > 80%.

The results of the two-dimensional projection (*PC1–PC2*) show clear differences between groups: the severe cluster is in the *range of PC1* > +0.8, while the normal cluster is < −0.7. This confirms that PCA reduction not only lowers the complexity from 24 features to 4 components but also increases data separation before the fuzzy clustering process.

4.3. Clustering result

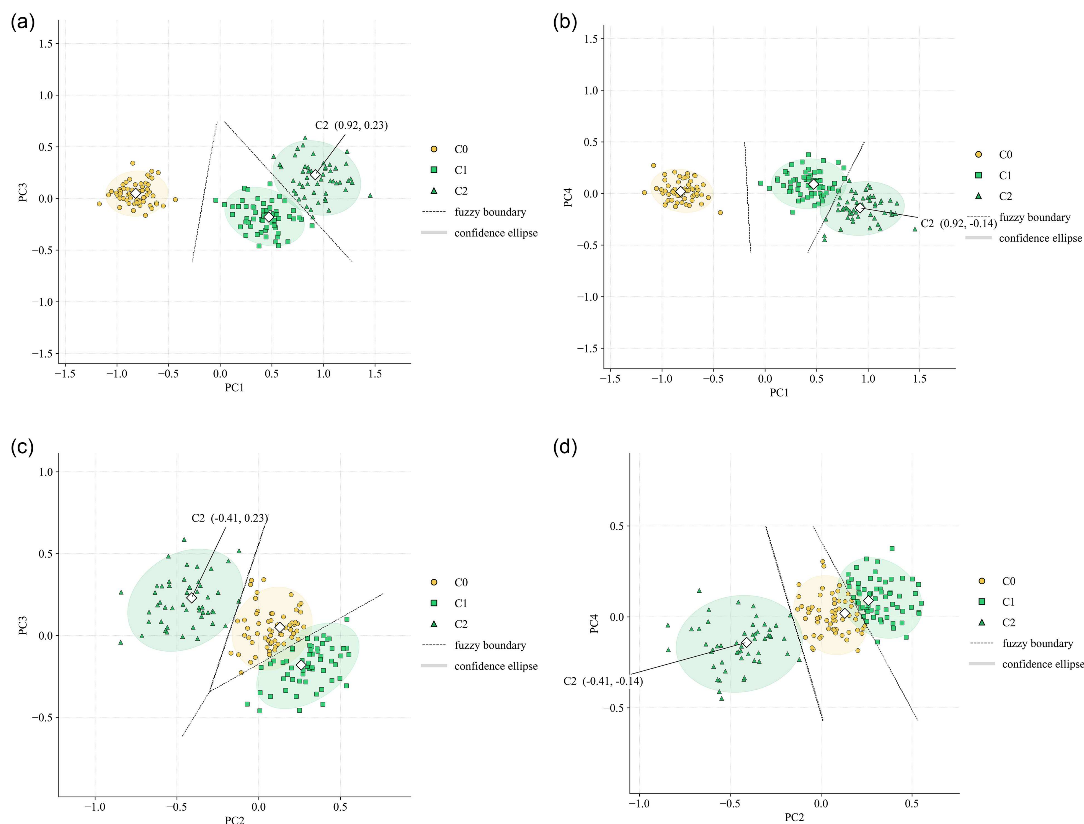
The FCM algorithm was applied to the four principal components extracted from the PCA stage, which retained 82.27% of the total cumulative variance. The model was configured to produce three clusters—Normal, Mild, and Severe—with a fuzzifier value of *m = 2.0*, cluster count *c = 3*, and a maximum of 300 iterations. Table 4 presents the centroid coordinates of each cluster in the PCA-transformed feature space, along with the dominant membership proportions (hard assignment based on the argmax of fuzzy membership values). The results indicate stable and well-separated centroid structures, particularly in the *PC*1–*PC*2 plane. The severe cluster (*C*2) is positioned at a highly positive *PC*1 (0.92) and negative *PC*2 (−0.41), while the normal cluster (*C*0) lies at a strongly negative *PC*1 (−0.82). The cluster membership distribution is relatively balanced across groups (*C*0 = 33.33%, *C*1 = 36.21%, *C*2 = 30.46%). A multidimensional visualization of the FCM clustering results is provided in Figure 3. Panel (a) displays the *PC*1–*PC*3 projection with iso-membership contours (threshold: 0.5), whereas panels (b–d) illustrate additional projection pairs (*PC*1–*PC*4, *PC*2–*PC*3, and *PC*2–*PC*4), including fuzzy boundary regions and trust ellipses centered around the respective centroids.

Figure 3 shows the multidimensional projection of the results of the FCM-PCA clustering on various combinations of the main components (*PC*1–PC4). The distribution pattern shows that C_0

Table 4
Centroid PCA and member distribution (FCM-PCA)

Cluster	*PC*1	*PC*2	*PC*3	*PC*4	Member (*n*)	Propose (%)
*C*0	–0.82	0.13	0.05	0.02	58	33.33
*C*1	0.47	0.26	–0.18	0.09	63	36.21
*C*2	0.92	–0.41	0.23	–0.14	53	30.46

Figure 3
Multidimensional PCA projections and fuzzy cluster distributions: (a) *PC*1–*PC*3 projection, (b) *PC*1–*PC*4 projection, (c) *PC*2–*PC*3 projection, and (d) *PC*2–*PC*4 projection



(Normal) and C_2 (Severe) have a clear spatial separation, while $C1$ (Mild) forms a transition area with a *fuzzy overlap* between them. The contours of *the iso*-membership boundary *(μ = 0.5)* show a zone of uncertainty at the relatively narrow inter-cluster edge, indicating stable membership. The 95% confidence ellipse surrounding each centroid reinforces the interpretation that the clustering results have a controlled internal distribution. These results are in line with the findings in Yan and Xie [30], where PCA was used to improve cluster separability prior to the application of FCM. The cluster color represents *the hard* assignment resulting from the fuzzy membership maximum value, while the dotted line depicts the membership limit $\mu = 0.5$. The position of the centroid is taken from Table 4 with the coordinates: $C0$ (−0.82, 0.13, 0.05, 0.02), $C1$ (0.47, 0.26, −0.18, 0.09), and $C2$ (0.92, −0.41, 0.23, −0.14).

4.4. Evaluation model

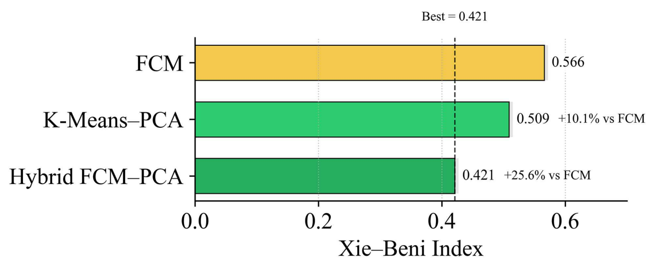
The evaluation of the clustering quality in this study was carried out using the Xie-Beni (XB) index, which is specifically designed to assess the performance of *the FCM* [48]. The XB index measures the ratio between the total intra-cluster dispersion and the minimum inter-centroid distance; therefore, a lower XB value indicates more compact clusters and greater separation across centroids.

To examine the effect of dimensionality reduction on clustering stability, three models were compared: (i) FCM-only (baseline model without PCA), (ii) K-Means-PCA (hard clustering after PCA), and (iii) the proposed Hybrid FCM-PCA model. As shown in Table 5, the Hybrid FCM-PCA model achieved the lowest XB value (0.421), indicating the most compact clusters and the largest centroid separation among the three approaches. This improvement reflects the contribution of PCA in reducing feature redundancy and decreasing fuzzy membership dispersion by decorrelating the acoustic–prosodic features. These findings are consistent with previous studies by Mohamed Khamis et al. [20] as well as Yan and Xie [30], which also report enhanced cluster separability following PCA integration. Overall, the proposed model demonstrates a performance improvement of 25.6% relative to the FCM-only baseline. A visual comparison of XB values across the models is provided in Figure 4.

Figure 4 shows a 25.6% decrease in *the* XB index value from the FCM-only model to the Hybrid FCM-PCA, indicating a significant increase in cluster compaction and separation. These results indicate that dimension reduction using PCA is able to improve the fuzzy space structure, resulting in a denser membership distribution and more assertive inter-cluster boundaries.

The Hybrid FCM-PCA model shows an *XB value = 0.421*, representing the optimal balance between *compactness* and *separability* of the cluster. In contrast, the comparator models

Figure 4
Comparative vierbeins index among clustering models



(FCM-only and K-Means-PCA) produce looser *fuzzy* structures. Thus, PCA integration has been shown to improve the efficiency of acoustic–prosodic feature representation without sacrificing fuzzy division stability.

4.5. Implication for speech therapy

The results of the Hybrid FCM-PCA model reveal a strong correspondence between children's acoustic–prosodic patterns and their levels of speech therapy need. The fuzzy clustering approach used in this study reflects clinical reality, wherein speech deviations occur along a continuous spectrum rather than as strictly discrete categories [49]. The three clusters identified—C0 (Normal Speech Pattern), $C1$ (Mild Deviation), and $C2$ (Severe Deviation)—represent distinct levels of prosodic disruption commonly observed in pediatric speech therapy assessments. The XB validity index (XB = 0.421) confirms that the model produces compact clusters with clear centroid separation, indicating its suitability for early-stage clinical decision support.

Table 6 provides a detailed interpretation of these clusters, including typical acoustic–prosodic characteristics and their corresponding therapy implications. The fuzzy membership values allow transitional conditions to be detected, such as a child displaying partial membership between mild and severe categories (e.g., $\mu(C1) = 0.63$ and $\mu(C2) = 0.37$), suggesting an increasing likelihood of requiring more intensive intervention [35]. This supports the view that speech disorder severity progresses continuously, not categorically.

Building on these interpretations, Figure 5 presents a conceptual mapping between the fuzzy cluster outcomes and recommended therapeutic intervention levels. The horizontal axis represents the degree of acoustic–prosodic deviation from normal to severe, while the vertical axis illustrates the corresponding therapy intensity, ranging from routine monitoring to structured therapy and intensive rehabilitation. This mapping demonstrates how fuzzy cluster membership can be translated into practical clinical recommendations.

Figure 5 illustrates the relationship between fuzzy model output and speech therapy results. The higher the prosodic

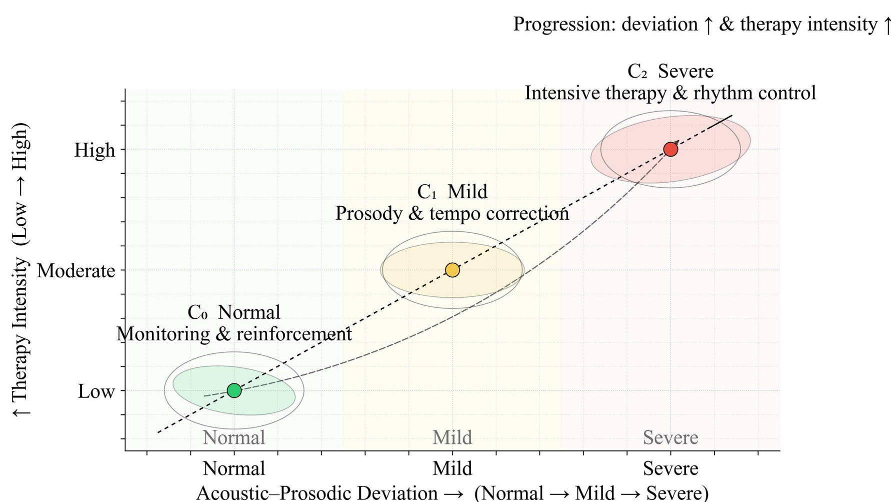
Table 5
Comparison of Xie-Beni between clustering models

Model	Xie-Beni index	Downside to FCM	Interpretation
FCM	0.566	–	Loose clusters, fuzzy borders intersect
K-Means-PCA	0.509	10.07%	Clusters are more compact after dimension reduction
Hybrid FCM-PCA (proposed)	0.421	25.6%	The most stable and distinct clusters are clear

Table 6
Interpretation of fuzzy clusters and recommended therapy actions

Cluster	Acoustic–prosodic characteristics	Indicative speech pattern	Recommendation
Normal ($C0$)	High pitch and tempo stability, low jitter and shimmer, uniform phoneme duration	Normal speech pattern without phonation or prosody deviation	No intensive therapy is required, just regular monitoring
Mild ($C1$)	Tempo variation and intonation are slightly limited; the shimmer increases slightly ($\approx +0.02$ from normal)	Mild disturbances in fluency and intonation	Mild phonetic interventions, focusing on tempo control and prosodic expression
Severe ($C2$)	Pitch and formant unstable, high jitter (> 0.09), inconsistent tempo	Severe disturbances in phonation and articulation	Intensive therapy: breathing phonation exercises, speech rhythm control, and structured prosodic stimulation

Figure 5
Conceptual mapping between fuzzy clusters and speech therapy levels



deviation (right direction), the greater the intensity of the intervention required (upward). This model can be used as the initial framework for clinical decision support systems to determine therapy priorities based on automatically measured acoustic parameters. The Hybrid FCM-PCA approach not only successfully classifies acoustic data statistically but also has practical implications in children's speech therapy. By providing clear inter-cluster fuzzy boundaries, the model allows for early identification of children at risk of mild prosodic disorders so that interventions can be made before symptoms worsen. In addition, the results of the clustering can be used as a reference map for therapists to adjust the intensity of the exercise according to the level of disturbance detected [5, 44]. Overall, the integration of artificial intelligence at this stage of diagnosis has the potential to speed up the assessment process and improve the efficiency of data-driven speech therapy.

5. Conclusion

This study demonstrates that the Hybrid PCA-FCM model is effective in identifying children's speech therapy needs based on acoustic–prosodic patterns within the analyzed structured dataset. By reducing dimensional redundancy through PCA and modeling continuous speech deviation through fuzzy membership, the model successfully classifies children into three clinically interpretable groups: Normal, Mild Deviation, and Severe Deviation. The lowest XB index value (XB = 0.421) confirms strong compactness and separation of clusters, supporting the model's suitability for exploratory early-stage clinical pattern analysis. These findings highlight the potential of hybrid unsupervised learning as a foundation for automated early-screening tools in pediatric speech assessments. The reported results reflect structural clustering validity within the analyzed dataset rather than inferential statistical comparison between predefined clinical groups. Formal validation involving certified speech-language professionals is required before clinical deployment of the proposed framework. Further validation using independent and prospectively collected datasets is recommended to strengthen the generalizability and robustness of the proposed framework.

Recommendations

Based on the findings of this study, several recommendations can be proposed. First, the Hybrid PCA-FCM framework may serve as a supportive analytical framework for preliminary screening in pediatric speech-language assessments to help clinicians prioritize children who are likely to require therapeutic intervention. Second, clinicians are encouraged to complement perceptual evaluations with acoustic–prosodic analysis, as fuzzy membership values can provide additional insight into subtle or

transitional deviations. Third, the system should be expanded with automated feature extraction and real-time processing modules to improve its practicality in clinical environments. Finally, future research is recommended to incorporate larger multilingual datasets, raw signal analysis, and expert-labeled clinical data and to explore XAI techniques to enhance model transparency and support wider clinical adoption.

Acknowledgment

The authors would like to express their deepest gratitude to the Department of Informatics of Universitas Muhammadiyah Makassar for its invaluable guidance, constructive insights, and continuous supervision throughout the development of this research. Its expertise in artificial intelligence, data science, and scientific writing has significantly contributed to the refinement and completion of this study.

Ethical Statement

This study did not involve the collection of new human subject data, clinical interventions, or direct interaction with children. All analyses were conducted using anonymized analytical feature-level data within a structured dataset; therefore, ethical approval and informed consent were not required.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The analytical dataset supporting this study is publicly available in Figshare at https://doi.org/10.6084/M9.FIGSHARE.31410669.

Author Contribution Statement

Muhammad Rizal Haris: Conceptualization, Methodology, Validation, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Muhammad Faisal:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Supervision. **Rizki Yusliana Bakti:** Methodology, Software, Validation, Data curation, Project administration. **Muhyiddin AM Hayat:** Validation, Formal analysis, Investigation, Resources, Project administration. **Titik Khawa Abd Rahman:** Conceptualization, Methodology, Software, Validation, Writing – review & editing, Project administration. **Muhammad Syafaat S. Kuba:** Validation, Writing – review & editing, Supervision. **Titin Wahyuni:** Data curation, Software, Project administration.

References

[1] Liu, K.-C., Huang, J.-J., Pei, Y.-C., Huang, S.-C., & Chung, C.-Y. (2026). The improvement following speech and language therapy in children with mixed receptive-expressive language disorder across different age groups: A retrospective study. *Pediatrics and Neonatology*, *67*(1), 38–44. https://doi.org/10.1016/j.pedneo.2025.03.007

[2] Blake, A., Dąbrowska, E., & Riches, N. (2025). Exploring cognitive predictors of language in children with developmental language disorder: The role of nonverbal IQ, working memory, implicit statistical learning, and speed of automatization. *Journal of Communication Disorders*, *116*, 106541. https://doi.org/10.1016/j.jcomdis.2025.106541

[3] Giberga, A., Guerra, E., Ahufinger, N., Igualada, A., Aguilera, M., & Esteve-Gibert, N. (2025). Prosody and gestures help pragmatic processing in children with Developmental Language Disorder. *Journal of Communication Disorders*, *115*, 106525. https://doi.org/10.1016/j.jcomdis.2025.106525

[4] Kisjes, J., van der Schaaf, A. L., Noordstar, J. J., Mombarg, R., Gerrits, E., Wijnen, F., & Luinge, M. R. (2025). A systematic review of language and motor skills in children with developmental coordination disorder (DCD) and developmental language disorder (DLD). *Research in Developmental Disabilities*, *161*, 104994. https://doi.org/10.1016/j.ridd.2025.104994

[5] Bhardwaj, A., Sharma, M., Kumar, S., Sharma, S., & Sharma, P. C. (2024). Transforming pediatric speech and language disorder diagnosis and therapy: The evolving role of artificial intelligence. *Health Sciences Review*, *12*, 100188. https://doi.org/10.1016/j.hsr.2024.100188

[6] van Witteloostuijn, M., & Blom, E. (2025). Parental language mixing and its association with language outcomes of children with (a suspicion of) Developmental Language Disorder. *Journal of Communication Disorders*, *117*, 106557. https://doi.org/10.1016/j.jcomdis.2025.106557

[7] Kempe, V., Ota, M., & Schaeffler, S. (2024). Does child-directed speech facilitate language development in all domains? A study space analysis of the existing evidence. *Developmental Review*, *72*, 101121. https://doi.org/10.1016/j.dr.2024.101121

[8] Kobayashi, A. (Ed.). (2019). *International encyclopedia of human geography* (2nd ed.). UK: Elsevier.

[9] Tritschler, T., Patel, A., Kraaijpoel, N., di Santo, P., Hibbert, B., Le Gal, G., & Castellucci, L. (2021). Abstracts. *Research and Practice in Thrombosis and Haemostasis*, *5*, e12591. https://doi.org/10.1002/rth2.12591

[10] Setiawan, K. E., Kurniawan, A., Chowanda, A., & Suhartono, D. (2023). Clustering models for hospitals in Jakarta using fuzzy c-means and k-means. *Procedia Computer Science*, *216*, 356–363. https://doi.org/10.1016/j.procs.2022.12.146

[11] Georgiou, G. P. (2025). Transforming speech-language pathology with AI: Opportunities, challenges, and ethical guidelines. *Healthcare*, *13*(19), 2460. https://doi.org/10.3390/healthcare13192460

[12] Deepa, P., & Khilar, R. (2022). Speech technology in healthcare. *Measurement: Sensors*, *24*, 100565. https://doi.org/10.1016/j.measen.2022.100565

[13] Shi, M., Cheung, G., & Shahamiri, S. R. (2023). Speech and language processing with deep learning for dementia diagnosis: A systematic review. *Psychiatry Research*, *329*, 115538. https://doi.org/10.1016/j.psychres.2023.115538

[14] Zhang, Y.-D., Dong, Z., Wang, S.-H., Yu, X., Yao, X., Zhou, Q., . . . , & Gorriz, J. M. (2020). Advances in multimodal data fusion in neuroimaging: Overview, challenges, and novel orientation. *Information Fusion*, *64*, 149–187. https://doi.org/10.1016/j.inffus.2020.07.006

[15] Nudelman, C. J., Tardini, V., & Bottalico, P. (2025). Artificial intelligence to detect voice disorders: An AI-supported systematic review of accuracy outcomes. *Journal of Voice*. Advance online publication. https://doi.org/10.1016/j.jvoice.2025.09.021

[16] Yang, G., Ye, Q., & Xia, J. (2022). Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. *Information Fusion*, *77*, 29–52. https://doi.org/10.1016/j.inffus.2021.07.016

[17] Polver, S., Cantiani, C., Bulf, H., Piazza, C., Turati, C., Molteni, M., & Riva, V. (2025). Gamma oscillations and auditory perception: A cluster-based statistic investigation in infants at higher likelihood of autism and developmental language disorder. *Journal of Experimental Child Psychology*, *252*, 106132. https://doi.org/10.1016/j.jecp.2024.106132

[18] Kosztyán, Z. T., Kurbucz, M. T., & Katona, A. I. (2022). Network-based dimensionality reduction of high-dimensional, low-sample-size datasets. *Knowledge-Based Systems*, *251*, 109180. https://doi.org/10.1016/j.knosys.2022.109180

[19] de Silva, U., Madanian, S., Olsen, S., Templeton, J. M., Poellabauer, C., Schneider, S. L., . . . , & Rubaiat, R. (2025). Clinical decision support using speech signal analysis: Systematic scoping review of neurological disorders. *Journal of Medical Internet Research*, *27*, e63004. https://doi.org/10.2196/63004

[20] Mohamed Khamis, G. S., Alqahtani, N. S., Munadi Alanazi, S., Alruwaili, M. M., Alenazi, M. S., & Alrawaili, M. A. (2025). Using Fuzzy C-Means clustering and PCA in public health: A machine learning approach to combat CVD and obesity. *Informatics in Medicine Unlocked*, *57*, 101666. https://doi.org/10.1016/j.imu.2025.101666

[21] Kumar, L. A. (2024). *Predictive analytics using MATLAB® for biomedical applications*. USA: Academic Press.

[22] Wang, S., Bai, L., Chen, X., Wang, Z., & Shao, Y.-H. (2022). Divergent projection analysis for unsupervised dimensionality reduction. *Procedia Computer Science*, *199*, 384–391. https://doi.org/10.1016/j.procs.2022.01.047

[23] Zhang, L., & Gu, X. (2025). A novel career prediction method based on fuzzy model—Fuzzy clustering approach. *Acta Psychologica*, *259*, 105475. https://doi.org/10.1016/j.actpsy.2025.105475

[24] Murugavel, K. D., R, P., Mathivanan, S. K., Srinivasan, S., Shivahare, B. D., & Shah, M. A. (2025). A multimodal machine learning model for bipolar disorder mania classification: Insights from acoustic, linguistic, and visual cues. *Intelligence-Based Medicine*, *11*, 100223. https://doi.org/10.1016/j.ibmed.2025.100223

[25] Górriz, J. M., Álvarez-Illán, I., Álvarez-Marquina, A., Arco, J. E., Atzmueller, M., Ballarini, F., ..., & Ferrández-Vicente, J. M. (2023). Computational approaches to Explainable Artificial Intelligence: Advances in theory, applications and trends. *Information Fusion*, *100*, 101945. https://doi.org/10.1016/j.inffus.2023.101945

[26] Deng, C., Ji, X., Rainey, C., Zhang, J., & Lu, W. (2020). Integrating machine learning with human knowledge. *iScience*, *23*(11), 101656. https://doi.org/10.1016/j.isci.2020.101656

[27] K, N., & M. B, M. K. (2024). Fuzzy rule based classifier model for evidence based clinical decision support systems. *Intelligent Systems with Applications*, *22*, 200393. https://doi.org/10.1016/j.iswa.2024.200393

[28] Goyal, J., Khandnor, P., & Aseri, T. C. (2020). Classification, prediction, and monitoring of Parkinson's disease using computer assisted technologies: A comparative analysis. *Engineering Applications of Artificial Intelligence*, *96*, 103955. https://doi.org/10.1016/j.engappai.2020.103955

[29] Kanagaraj, R., Krishnaraj, N., Selvakumar, J., & Ramprasath, J. (2025). Implementation of multi-label fuzzy classification system using topic detection data set. *Procedia Computer Science*, *252*, 15–24. https://doi.org/10.1016/j.procs.2024.12.003

[30] Yan, B., & Xie, Z. (2026). A robust fuzzy cluster validity index based on local distances. *Expert Systems with Applications*, *295*, 128883. https://doi.org/10.1016/j.eswa.2025.128883

[31] Joudar, S. S., Albahri, A. S., & Hamid, R. A. (2022). Triage and priority-based healthcare diagnosis using artificial intelligence for autism spectrum disorder and gene contribution: A systematic review. *Computers in Biology and Medicine*, *146*, 105553. https://doi.org/10.1016/j.compbiomed.2022.105553

[32] Garg, A., Aribi, Y., Althobaiti, T., & Bhowmik, T. (2025). An analysis of acoustic features for accented speech classification. *Egyptian Informatics Journal*, *31*, 100743. https://doi.org/10.1016/j.eij.2025.100743

[33] Anwar, A., Khalifa, Y., Coyle, J. L., & Sejdic, E. (2025). Transformers in biosignal analysis: A review. *Information Fusion*, *114*, 102697. https://doi.org/10.1016/j.inffus.2024.102697

[34] Saeedi, S., Hetjens, S., Grimm, M. O. W., & Latoszek, B. B. V. (2024). Acoustic speech analysis in Alzheimer's disease: A systematic review and meta-analysis. *The Journal of Prevention of Alzheimer's Disease*, *11*(6), 1789–1797. https://doi.org/10.14283/jpad.2024.132

[35] Islam, M. A., Majumder, M. Z. H., Hussein, M. A., Hossain, K. M., & Miah, M. S. (2024). A review of machine learning and deep learning algorithms for Parkinson's disease detection using handwriting and voice datasets. *Heliyon*, *10*(3), e25469. https://doi.org/10.1016/j.heliyon.2024.e25469

[36] Gilholm, P., Mengersen, K., & Thompson, H. (2020). Identifying latent subgroups of children with developmental delay using Bayesian sequential updating and Dirichlet process mixture modelling. *PLoS ONE*, *15*(6), e0233542. https://doi.org/10.1371/journal.pone.0233542

[37] Gohider, N., & Basir, O. A. (2024). Recent advancements in automatic disordered speech recognition: A survey paper. *Natural Language Processing Journal*, *9*, 100110. https://doi.org/10.1016/j.nlp.2024.100110

[38] Chakraborty, C., Bhattacharya, M., Pal, S., & Lee, S.-S. (2024). From machine learning to deep learning: Advances of the recent data-driven paradigm shift in medicine and healthcare. *Current Research in Biotechnology*, *7*, 100164. https://doi.org/10.1016/j.crbiot.2023.100164

[39] Au, Y. C., & Ng, M. L. (2025). Developing a smart system for binary classification of disordered voices using machine learning. *American Journal of Otolaryngology*, *46*(4), 104672. https://doi.org/10.1016/j.amjoto.2025.104672

[40] Walha, A., Ksibi, A., Zakariah, M., Ayadi, M., Alshalali, T., Saidani, O., ..., & Almujally, N. A. (2025). Deep learning and machine learning architectures for dementia detection from speech in women. *Computer Modeling in Engineering & Sciences*, *142*(3), 2959–3001. https://doi.org/10.32604/cmes.2025.060545

[41] Mulyadi, I., Faisal, M., Aryasa, K., Thamrin, S. M., Darniati, & Adnan, M. (2023). Intelligent security system based on biometric face recognition. *Journal of Theoretical and Applied Information Technology*, *101*(17), 6953–6964.

[42] Faisal, M., Irmawati, Abd Rahman, T. K, Jufri, Sahabuddin, Herlinah, & Mulyadi, I. (2025). A hybrid MOO, MCGDM, and sentiment analysis methodologies for enhancing regional expansion planning: A case study Luwu—Indonesia. *International Journal of Mathematical, Engineering and Management*

Sciences, *10*(1), 163–188. https://doi.org/10.33889/IJMEMS.2025.10.1.010

[43] Li, C., Solinsky, J., Cohen, T., & Pakhomov, S. (2024). A curious case of retrogenesis in language: Automated analysis of language patterns observed in dementia patients and young children. *Neuroscience Informatics*, *4*(1), 100155. https://doi.org/10.1016/j.neuri.2023.100155

[44] Faisal, M., Abd Rahman, T. K., Zainal, D., Mubarak, H., Shabir, F., Anwar, N., & Asrowardi, I. (2025). Utilizing machine learning-based decisionmaking to align higher education curriculum with industry requirements. *International Journal of Modern Education and Computer Science*, *17*(4), 1–25. https://doi.org/10.5815/ijmecs.2025.04.01

[45] M., S., Murugappan, A., & T., M. (2023). Cognitive computing technological trends and future research directions in healthcare—A systematic literature review. *Artificial Intelligence in Medicine*, *138*, 102513. https://doi.org/10.1016/j.artmed.2023.102513

[46] van Lingen, H. J., Suarez-Diez, M., & Saccenti, E. (2024). Normalization of gene counts affects principal components-based exploratory analysis of RNA-sequencing data. *Biochimica et Biophysica Acta (BBA)—Gene Regulatory Mechanisms*, *1867*(4), 195058. https://doi.org/10.1016/j.bbagrm.2024.195058

[47] Sebastian, A. M., Peter, D., & Sebastian, R. A. (2025). An optimal imputation algorithm for reducing bias and errors in missing data handling for AI models. *Decision Analytics Journal*, *16*, 100627. https://doi.org/10.1016/j.dajour.2025.100627

[48] Faisal, M., & Abdul Rahman, T. K. (2023). Determining rural development priorities using a hybrid clustering approach: A case study of South Sulawesi, Indonesia. *International Journal of Advanced Technology and Engineering Exploration*, *10*(103), 696–719. https://doi.org/10.19101/IJATEE.2023.10101215

[49] Patel, S. P., Landau, E., Martin, G. E., Rayburn, C., Elahi, S., Fragnito, G., & Losh, M. (2023). A profile of prosodic speech differences in individuals with autism spectrum disorder and first-degree relatives. *Journal of Communication Disorders*, *102*, 106313. https://doi.org/10.1016/j.jcomdis.2023.106313

How to Cite: Haris, M. R., Faisal, M., Bakti, R. Y., Hayat, M. A. M., Rahman, T. K. A., Syafaat S. Kuba, M., & Wahyuni, T. (2026). Leveraging a Hybrid Fuzzy C-Means-PCA Model for Identifying Speech Therapy Needs in Children. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS62028311