

Independence Program Recommendation For Autism Spectrum Disorder Using Hybrid Collaborative Filtering Model

Sitairesmi Wahyu Handani¹, Frengki Saputra^{*2}

¹Department of Computer Science and Information Engineering, College of Electrical Engineering and Computer Science, National Taiwan University, Taiwan

¹Department of Informatics, Universitas Amikom Purwokerto, Indonesia

²Department of Information System, Faculty of Computer Science, Universitas Amikom Purwokerto, Indonesia
e-mail: d09922018@csie.ntu.edu.tw¹, frengkis2809@gmail.com²

Article History	Abstract
Received : 29 June 2026 Revised : 27 August 2026 Accepted : 3 September 2026 Published : 16 September 2026 Keywords: <i>Autism Spectrum Disorder, Recommendation System, Alternating Least Squares, Patient Similarity Model, Self-Help Curriculum.</i>	Heterogeneity in the characteristics of children with Autism Spectrum Disorder (ASD) requires precise personalization of self-help skill programs. However, conventional Collaborative Filtering tends to suffer from popularity bias and struggles to capture clinical profile similarities between individuals, while patient-similarity approaches alone underuse collective success patterns, requiring a hybrid architecture. This study develops and evaluates a hybrid system integrating a Patient Similarity Model (PSM) based on cosine similarity and Collaborative Filtering based on Alternating Least Squares (ALS). The experiment used functional profile data from 50 students at a Special Education School (SLB), covering four core clinical aspects: motor, sensory, attention, and communication, along with the success history of 8 independence targets and 7 teaching strategies from teachers. An inverse frequency penalty mechanism was applied to address data popularity bias. Leave-One-Out Cross-Validation showed the hybrid model performed superiorly, achieving Hit Rate@3 of 0.8400, NDCG@3 of 0.8079, and MRR@3 of 0.7967, retaining PSM-Only's coverage advantage while empirically outperforming CF-Only (Hit Rate@3: 0.6200). The system also features a conditional rule engine based on pedagogical keyword mapping for accountable strategy recommendations. Key innovations include converting observational text into structured data, eliminating popularity bias, and producing explainable intervention outputs to support Individualized Learning Program (ILP) development.

^{*}Corresponding author : frengkis2809@gmail.com

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurobiological condition characterized by impairments in social interaction, atypical communication, and restricted, repetitive behaviors. The prevalence of ASD continues to increase globally, as confirmed by reports from the Centers for Disease Control and Prevention (CDC) in the United States as well as a global-scale three-level meta-analysis spanning the 1994–2019 period [1]. In Indonesia, one estimate presented at a national webinar puts the number of individuals with ASD at more than 2.4 million [2], which directly amplifies the burden on special education services in special education schools (SLB).

The core component of the ASD curriculum in SLB is the self-help program, which includes toilet training, independent eating, dressing, and functional communication, all of which have been proven to be strong predictors of the quality of life and social integration of individuals with ASD in adulthood [3]. However, a scoping review revealed low achievement in the acquisition of daily living skills (DLS) among adolescents and young adults with ASD, a condition further exacerbated by a lack of studies measuring long-term skill generalization and maintenance [4]. Due to the heterogeneity of

ASD clinical profiles wherein no two children are identical across motor, sensory, attention, and communication dimensions [5], one-size-fits-all protocols are ineffective, thereby necessitating granular mapping to recommend adaptive Individualized Education Programs (IEP) [6].

A preliminary study conducted through observations and interviews with SLB teachers in Central Java indicates that most educators encounter difficulties in prioritizing self-help programs without systematic, data-driven guidance. The high cognitive load resulting from managing multiple ASD students simultaneously causes the planning of Individualized Education Programs (IEP) to frequently rely on subjective intuition rather than empirical analysis [7]. A similar gap is identified within broader inclusive policy levels, where individualized transition planning for students with ASD still faces significant implementation disparities [8].

In the realm of general education, Educational Recommender Systems (ERS) have proven effective in recommending adaptive learning strategies [9], utilizing co-association patterns between independence targets to support intervention sequentiality [6]. However, the adaptation of such intelligent systems for the ASD population in SLB remains highly limited; a recent systematic review identified a lack of specific design guidelines for e-learning recommendation systems tailored to individuals with ASD [10]. Hybrid approaches that combine Collaborative Filtering (CF) with patient clinical information have been shown to yield higher recommendation performance compared to single methods, both in electronic health record-based systems [11] and ASD intervention recommendations [12]. Furthermore, patient similarity-based approaches using heterogeneous clinical data have proven to outperform conventional Euclidean distance methods in nearest neighbor identification [13].

Based on this evidence, this study designs a hybrid recommendation system that combines a cosine similarity-based Patient Similarity Model (PSM) with the Alternating Least Squares (ALS) algorithm, a combination previously shown to be effective in ASD intervention recommendation tasks [12]. The system is further equipped with a pedagogical rule engine powered by Explainable AI (XAI) to generate accountable clinical rationales, addressing the black-box limitations commonly associated with machine learning-based recommenders. The proposed approach was evaluated on real-world clinical data from students with ASD in special education schools (SLB) using a rigorous cross-validation scheme. Collectively, this design is expected to minimize the cognitive load of educators and optimize the efficiency of Individualized Education Program (IEP) development within special education environments.

2. METHOD

2.1. Research Design and Workflow

This study employs a design science research approach, which focuses on the design, implementation, and evaluation of a machine learning-based recommendation system artifact [16]. The developed artifact is a hybrid system that integrates a cosine similarity-based Patient Similarity Model (PSM) with Collaborative Filtering based on the Alternating Least Squares (ALS) algorithm to generate personalized independence targets and learning strategy recommendations for students with Autism Spectrum Disorder (ASD) in special education schools (SLB) [12]. This hybrid approach was selected to overcome the limitations of interaction data sparsity, which frequently poses a major obstacle in modeling intervention systems for specific clinical populations [17]. Data enrichment strategies through the integration of additional collaborative filtering techniques have also been empirically proven to reduce the sparsity level of the user-item interaction matrix while simultaneously improving the accuracy of matrix factorization-based collaborative filtering models [18].

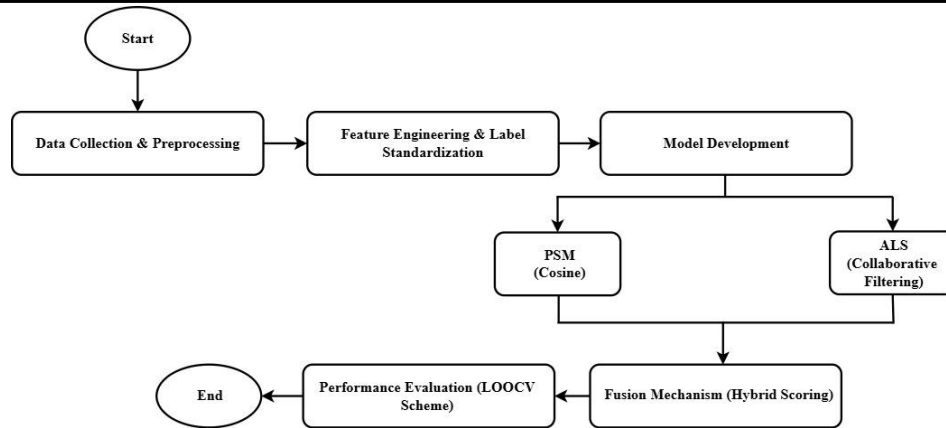


Figure 1. Research workflow of the hybrid system

Figure 1 illustrates the research workflow: starting from data collection and preprocessing, feature engineering and label standardization, parallel computation of PSM and ALS, hybrid score fusion (Hybrid Scoring), up to performance evaluation using the LOOCV scheme.

2.2. Data Sources and Participants

Primary data were gathered from the functional profiles of 54 students with ASD through in-depth observations and structured interviews with Special Education Teachers (GPK) across several special education schools (SLB) in the Central Java region. Four students were excluded because they generated zero-vectors after synchronization with the TARGET_WHITELIST, resulting in a final valid sample size of 50 students. The final dataset encompasses four main components of information parameters, as summarized in Table 1.

Table 1. Dataset Information Parameter Components

Component	Description
Demographic attributes	Quantitative data in the form of chronological age for each student as a baseline characteristic.
Clinical functional scores	Quantitative evaluation across four core developmental aspects motor, sensory, attention, and communication using an ordinal scale (1–4) based on direct assessments by Special Education Teachers (GPK).
Independence target history	Qualitative data records consisting of narratives regarding the types of self-help programs (toilet training, independent eating, dressing, etc.) actively integrated into interventions for the students.
Instructional strategy documentation	Teachers' narrative notes regarding variations in implemented learning strategies along with their corresponding achievement success rates.

2.3. Data Preprocessing

The data preprocessing stage is a crucial step to guarantee data quality prior to feeding it into the hybrid model architecture. This process is designed to handle the heterogeneity of raw data, which includes a combination of qualitative free text and clinical notes from the 54 initial observations. The preprocessing stage is divided into the following four linear phases:

2.3.1 Extraction of Ordinal Clinical Scores and Age Proximity Inspection

Clinical scores in the form of text value ranges are mapped onto a standard ordinal scale (Table 2) using a regular expression-based mapping function, applied simultaneously across four core features: motor, sensory, attention, and communication. One student aged ≤ 5 years was detected as a chronological outlier but was retained, given the already limited sample size ($n=50$), to preserve data availability; this may lower PSM recommendation quality for this

student due to fewer comparable neighbors, though it does not affect the aggregate LOOCV metrics.

Table 2. Standard Ordinal Scale

Score	Level	Description
10	4.0	Excellent
7-9	3.0	Fair/Good
4-6	2.0	Poor
1-3	1.0	Very Poor

2.3.2 Contextual Tokenization and Word-Boundary Protected Noise Audit

Free text in the *target_raw* and *strategi_raw* columns is tokenized using whitelist tokens protected by word boundaries (regex `\b`) to prevent partial match overlaps on short vocabulary words (≤ 5 characters). The system maps the text into 8 independent target labels and 7 learning strategy labels, while simultaneously eliminating noise outside the whitelist.

2.3.3 Zero-Vector Filtration

The total active targets per student are calculated using equation (1). Four students with *total_active_targets* = 0 were eliminated because they shared no matches with the standard self-help programs, leaving 50 valid operational students.

$$total_active_targets = \sum_{label \in Whitelist} T_{label} \quad (1)$$

2.3.4 Clinical Feature Normalization and CSR Sparse Matrix Construction

All clinical features are normalized to a range of [0,1] using the Min-Max Scaling formula in equation (2). They are then converted into the Compressed Sparse Row (CSR) format and multiplied by the confidence parameter (α), generating a Target Matrix (50×8) and a Strategy Matrix (50×7).

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

2.4. Patient Similarity Model (PSM)

The PSM computes student-to-student similarity based on cosine similarity applied to the normalized clinical feature vectors using equation (3). The system identifies the top $K = 3$ nearest neighbors that exceed the threshold $PSM_THRESHOLD = 0.5$, and subsequently calculates the PSM score for each item as a similarity-weighted average of the neighbors.

$$sim(u, v) = (u \cdot v) / (||u|| \times ||v||) \quad (3)$$

The single-model cosine similarity-based PSM approach was selected due to its computational efficiency and high interpretability for limited SLB dataset scales, despite its limitations regarding complex clinical data heterogeneity and dimensionality [19]. The effectiveness of patient similarity in objectively mapping ASD clinical profiles to generate personalized intervention targets has been confirmed [12], operating under the principle that students with similar clinical characteristics tend to share similar developmental trajectories [19]. The advantage of this approach in enhancing predictive accuracy compared to raw feature distance-based models has also been validated within heterogeneous electronic health record domains [13].

2.5. Collaborative Filtering based on Alternating Least Squares (ALS)

The CF module implements latent factor matrix decomposition using the ALS algorithm (via the implicit library). Two separate models were trained, one for independence targets and another for learning strategies utilizing CSR matrices with a binary confidence value multiplied by

ALS_ALPHA = 20, adapting the implicit feedback scheme [20]. The absence of explicit ratings from teachers was addressed by setting the confidence scaling parameter to $\alpha = 20$ [21]. The model hyperparameters are summarized in Table 3.

Table 3. ALS Model Hyperparameters

Parameter	Value	Description
ALS_FACTORS	4	Dimension of latent factors
ALS_REGULARIZATION	0.10	L2 regularization to prevent overfitting
ALS_ITERATIONS	30	Number of training iterations
ALS_ALPHA	20	Confidence factor for implicit interaction
TOP_K_SIMILAR	3	Number of nearest PSM neighbors
TOP_N_REKOMENDASI	3	Number of system output recommendations
PSM_THRESHOLD	0.5	Minimum similarity threshold for PSM

ALS_FACTORS = 4 was selected to be smaller than the total number of item labels to prevent overfitting on the limited dataset; low-dimensional linear models have proven to be more robust in terms of generalization compared to complex non-linear architectures [17]. L2 regularization (0,10) and 30 iterations complete the overfitting control scheme while allowing the model to capture co-association patterns between independent targets.

2.6. Inverse Frequency Penalty Mechanism

To mitigate popularity bias within the interaction data, this study implements an inverse frequency penalty mechanism inspired by the concept of Term Frequency-Inverse Document Frequency (TF-IDF). The recommendation score for each item is penalized proportionally to its frequency of occurrence in the dataset; items that are too frequently recommended to many students receive a penalty, thereby lowering their effective scores. This mechanism drives the system to prioritize recommendations that are highly specific and individually relevant, rather than merely replicating statistically dominant intervention patterns.

2.7. Hybrid Fusion Mechanism (PSM + ALS)

The final recommendation score of the hybrid system is calculated through an additive fusion mechanism that proportionally combines the PSM and ALS scores. For each item label i for a target student q , the fusion score is formulated as:

$$Score_{hybrid}(q, i) = \alpha \times Score_{PSM}(q, i) + (1 - \alpha) \times Score_{ALS}(q, i) \quad (4)$$

Prior to fusion, both scores are normalized to [0, 1] using per-student min-max normalization for comparable scaling. The fusion weight α was set to 0.5, equally weighting PSM (clinical similarity) and ALS (collaborative precision), and was applied consistently across all experiments, including LOOCV evaluation and the Student Index 0 case study. A systematic sensitivity analysis across other α values was not performed and is left for future work (see Limitations). Final recommendations are the Top-3 highest-scoring items not yet mastered by the target student (novelty filtering).

2.8. Pedagogical Rule Engine for Strategy Recommendations

The system incorporates a conditional rule-based pedagogical rule engine constructed from in-depth interviews with special education practitioners; for example, students with low communication scores are prioritized for PECS (Picture Exchange Communication System) strategies or speech therapy. This mechanism generates Explainable AI (XAI)-based clinical rationales that address the black-box nature of machine learning algorithms [22], similar to ontology-based approaches in pediatric clinical decision support systems, which have been proven to enhance practitioner trust [23].

2.9. Model Evaluation: Leave-One-Out Cross-Validation (LOOCV)

Model performance was evaluated using Leave-One-Out Cross-Validation (LOOCV), well-suited for small-scale datasets. This choice reflects two considerations: LOOCV maximizes data

use by training on $n-1$ samples per iteration, yielding a low-bias estimator important for small clinical datasets where conventional splits (e.g., 80:20) leave too few test samples for stable estimates [24]; and it aligns with common practice in machine learning research on small, rare clinical populations such as ASD, which typically involves limited sample sizes (tens of subjects) [25]. Accordingly, although $n=50$ is small by general machine learning standards, it falls within an acceptable range for exploratory studies on such specific clinical populations. In each iteration, one student's data was withheld as the test set while the remaining 49 were used for training. Evaluation was conducted independently across three model configurations: (1) PSM-Only, (2) CF-Only (ALS), and (3) Hybrid Fusion. Three top-K metrics were calculated at $K = 3$:

- 2.9.1 Hit Rate@3 (HR@3): The proportion of test students for whom at least one relevant item appears within the system's Top-3 recommendations.
- 2.9.2 Normalized Discounted Cumulative Gain@3 (NDCG@3): Measures the ranking quality of the recommendations by assigning higher weights to relevant items that appear at top positions; this metric serves as a formal standard for testing the ranking efficiency of autism intervention target recommendations [12].
- 2.9.3 Mean Reciprocal Rank@3 (MRR@3): Measures the average position of the first relevant item within the recommendation list.

All experiments were executed with a fixed *random seed* = 42 in Google Colaboratory (Python 3) using the *implicit*, *scikit-learn*, *scipy*, and *pandas/numpy* libraries.

3. RESULTS AND DISCUSSION

3.1. Dataset Characteristics, Preprocessing Findings, and Association Analysis

The data collection process in this study successfully gathered 54 initial observations from several special education schools (SLB) and special education institutions in the Central Java region, including PKBM Alam Baturaden, SLBN Purbalingga, SLBN Cilacap, SLBN Mandiraja, and PKBM Al-Izzah. The zero-vector filtration mechanism excluded 4 students whose teachers' intervention narratives shared no matches with the *TARGET_WHITELIST*. This exclusion indicates non-standard intervention practices that have not yet been mapped within the SLB self-help taxonomy, thereby yielding 50 valid students as the final operational dataset.

Label distribution analysis revealed that Independent Eating was the most widely mastered independence target among students (82%), followed by Independent Dressing (56%) and Toilet Training (52%). Conversely, Social Skills (4%) and Returning Items (12%) were the least achieved, indicating a higher level of difficulty within social-executive dimensions, as shown in Figure 2.

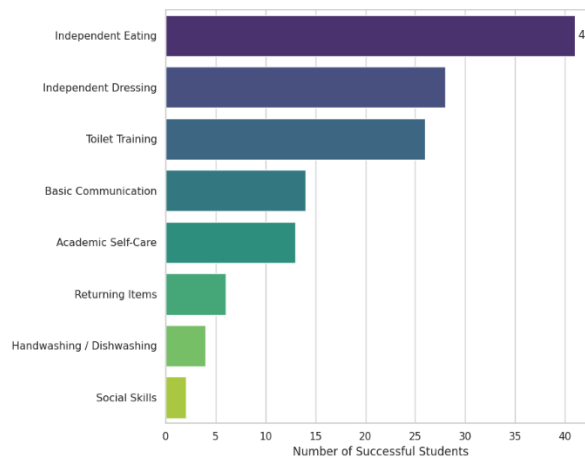


Figure 2. Achievement Frequency of ASD Student Independence Targets (n=50)

Regarding the analysis of learning strategies, the extraction results demonstrate an extreme dominance in the utilization of Verbal Instruction by teachers across almost the entire student sample. This method was followed by Direct Physical Prompting and Picture Card Schedules. This massive dominance of popularity bias in the Verbal Instruction strategy provides the primary empirical justification for implementing the inverse frequency penalty mechanism within the system. This phenomenon aligns with the findings of [26], which state that without specific intervention, recommendation algorithms inherently amplify popular items excessively (the rich-get-richer effect). Without such an inverse frequency-based penalty, the recommendation system would mathematically always suggest Verbal Instruction for every student profile without variation. This degradation of diversity ultimately reduces the practical utility for teachers in the field due to the loss of intervention personalization flexibility, as visualized in Figure 3.

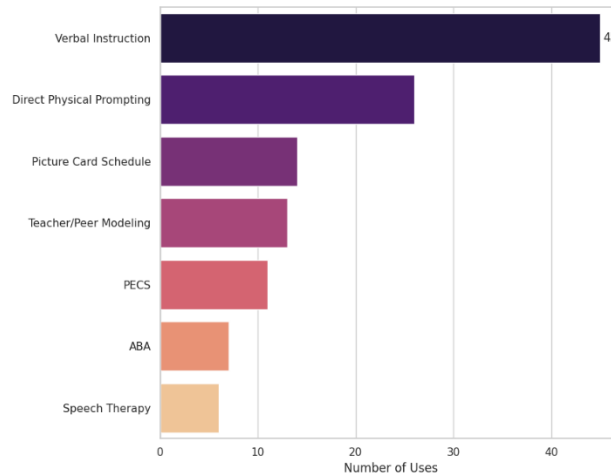


Figure 3. Frequency Distribution of Teacher Learning Strategy Implementation (n=50)

In addition to the single frequency distribution shown in Figure 2, the proximity of relationships between targets was explored through co-association matrix modeling, utilizing a dot product operation on the binary independence target matrix ($Y^T Y$). The resulting association heatmap visualization captures strong structural dependencies in the field, as presented in Figure 4.

A highly dense pairwise correlation was discovered among the targets Independent Eating, Independent Dressing, and Toilet Training. This high co-occurrence pattern empirically demonstrates that the cluster of students who have mastered basic self-care skills for eating exhibits a high probability of simultaneously or progressively mastering dressing and toileting independence as well. This implicit interconnectedness serves as foundational information for the Collaborative Filtering (ALS) component to extract latent factors for predicting the sequential order of subsequent learning targets.

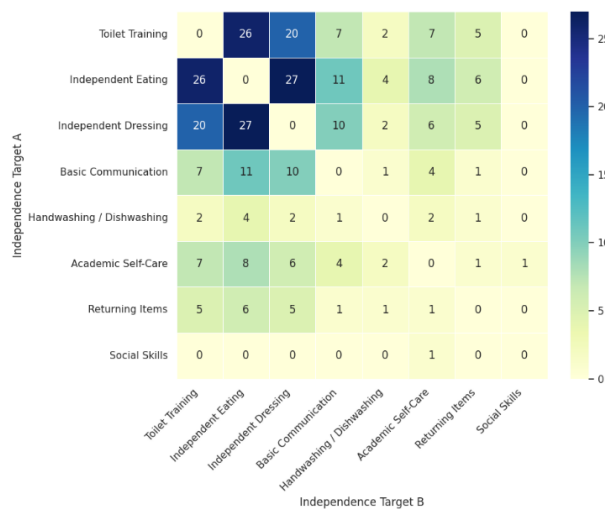


Figure 4. Heatmap of the Co-Association Matrix Between ASD Student Independence Targets

The age range of the students in this dataset spans from 5 to 18 years with highly heterogeneous clinical profiles, validating the use of the Patient Similarity Model (PSM), which is sensitive to individual deviations [1]. This approach aligns with a comprehensive review [14], which asserts that patient similarity yields more personalized predictions compared to conventional models based on population averages.

3.2. Model Performance: Comparison of Three Configurations

Model evaluation was conducted using a Leave-One-Out Cross-Validation (LOOCV) scheme to test the reliability of system predictions under clinical data constraints. The comparative performance results of the three architectural configurations were measured based on Hit Rate@3, Normalized Discounted Cumulative Gain@3 (NDCG@3), and Mean Reciprocal Rank@3 (MRR@3) metrics, as summarized in Table 4 and graphically visualized in Figure 5.

Table 4. LOOCV Evaluation Results of the Three Model Configurations (K=3)

Model Configuration	Hit Rate@3	NDCG@3	MRR@3
PSM-Only	0.8400	0.7883	0.7700
CF-Only (ALS)	0.6200	0.5279	0.4967
Hybrid (Fusion)	0.8400	0.8079	0.7967

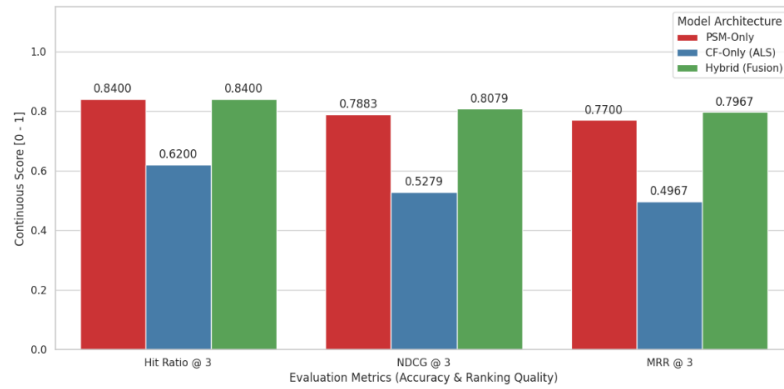


Figure 5. Comparative Performance of LOOCV Experiments Among PSM-Only, CF-Only (ALS), and Hybrid (Fusion) Models Across Various Accuracy Metrics (K=3)

Based on Table 4 and Figure 5, the Hybrid (Fusion) model achieves the highest performance with an HR@3 of 0.8400 (42 out of 50 students), an NDCG@3 of 0.8079, and an MRR@3 of 0.7967. This indicates that the relevant items are not merely included within the Top-3, but are effectively prioritized at the top positions. This attained NDCG@3 aligns with the efficacy range of 79%–81% developed by Kohli et al. [12].

Conversely, CF-Only (ALS) records the lowest performance (HR@3 = 0.6200). This is not due to mathematical deficiencies within the ALS algorithm itself, but rather a direct consequence of cold-start issues and data sparsity within the 50 users $\times$ 8 items matrix, which is too sparse to independently estimate latent factors. This phenomenon is consistent with the theoretical analysis by Rendle [17] and empirical findings within the autism clinical domain [12]. Consequently, this substantiates the architectural decision to integrate an explicit clinical component.

While both PSM-Only and the Hybrid model achieve an identical HR@3 of 0.8400, proving that the PSM serves as a robust recall anchor, the Hybrid model demonstrates superior ranking quality. The NDCG@3 increases from 0.7883 to 0.8079, and the MRR@3 improves from 0.7700 to 0.7967. This enhancement is contributed by the ALS component, which optimizes ranking precision by extracting collaborative latent factors. This fusion operates symbiotically: PSM

guarantees the accurate discovery of relevant items, while ALS positions them at the very top of the list.

Nevertheless, LOOCV log analysis detects prediction failures ($\text{Hit}@3 = 0$) for students with extreme clinical profiles, specifically those exhibiting either exceptionally high capabilities or severe deficits across all dimensions. These profiles fail to locate any neighbors that exceed the $\text{PSM_THRESHOLD} = 0,5$. This limitation could potentially be addressed through a multi-model-based Patient Similarity Network (PSN) that fuses static and dynamic data, which has been demonstrated by Navaz et al. [19] to significantly improve classification accuracy.

3.3. Inference Simulation and Individual-Level Case Study Analysis

An individual-level inference walkthrough was performed on Student Index 0 to validate the accountability of the system within the DSR paradigm. The normalized clinical feature profile of the subject is presented in Table 5.

Table 5. Normalized Clinical Feature Profile of Student Index 0

Feature Name	Normalized	Reality Interpretation / Description
usia_norm	0.6667	Representation of biological age ≈ 12 years old (data range: 5–18 years old)
motorik	0.6667	Moderate-to-high motor capacity (requires occasional supervision)
atensi	0.6667	Fairly good environmental responsiveness (rarely distracted by sensory stimuli)
komunikasi	0.3333	Independent focus sustained for 5–10 minutes

Through spatial distance matrix computation using Cosine Similarity on these clinical features, the system extracts $K = 3$ the students who share the most identical profile proximity within the database after excluding the target subject. The algorithm returns three nearest neighbor indices: Student 20, Student 40, and Student 26. The successful intervention tracks of these three peer students subsequently serve as the foundation for the prediction process. When the hybrid integrator function is executed with equal weighting between clinical and collaborative factors ($\alpha = 0,5$), the system detects the presence of active historical data and automatically routes the calculation into the normal Hybrid Fusion pathway. The calculation generates the Top-3 sequential recommendations for the next independence targets for Student Index 0, as detailed in Table 6.

Table 6. Top-3 Next Target Inference Output Results (Student Index 0)

Rank	Independence Target	Fused Prediction Score	Masking Mechanism Status
1	Toilet Training	0.6025	Passed (Never Achieved)
2	Independent Dressing	0.5004	Passed (Never Achieved)
3	Returning Items	0.2955	Passed (Never Achieved)

The recommendation output in Table 6 demonstrates high clinical validity and strict adherence to system logic. In terms of functional relevance, prioritizing Toilet Training and Independent Dressing is pedagogically highly appropriate, considering that the subject's deepest deficit lies within the communication dimension (0.3333). Both self-help programs demand intensive synchronization of adaptive instructions, consistent with the historical success records of the nearest neighbor cluster (Students 20, 40, and 26). In terms of logical compliance, the hard masking mechanism assigns a score of 0.0 to all targets already mastered by Student Index 0, ensuring that no previously achieved item reappears in the output list. This confirms that every recommendation generated by the system is a novel learning target not yet introduced in prior interventions, and is thus immediately ready for integration into the adaptive IEP (Individualized Education Program) curriculum by teachers in the field.

4. DISCUSSION

The experimental results fulfill the research objectives by demonstrating the superior performance of the Hybrid (Fusion) model that combines PSM and ALS, achieving an HR@3 of 0.8400, an NDCG@3 of 0.8079, and an MRR@3 of 0.7967. This performance outperforms the CF-Only baseline (HR@3: 0.6200) and enhances the ranking quality of the PSM-Only approach. This pattern aligns with medical recommender system literature, where hybrid architectures consistently record the highest precision compared to single algorithms [27]. It is also consistent with clinical physician recommender systems, where combining ALS with user-similarity-based collaborative filtering yields significantly higher accuracy and recall than conventional collaborative filtering methods [28].

Mechanistically, the PSM acts as a robust recall anchor that guarantees the inclusion of relevant items within the Top-3 based on functional-clinical profile similarities (motor, sensory, attention, and communication). Meanwhile, the ALS component optimizes the ranking order (reranking) by extracting latent co-association patterns between targets from the teachers' empirical history. The poor performance of the CF-Only configuration confirms the inherent limitations of pure collaborative filtering under cold-start conditions and small-scale clinical data sparsity. This is consistent with findings reported in medical service recommender systems, where sparse patient-doctor matrices lead to less reliable predictions, necessitating ALS matrix decomposition to overcome these constraints [28]. The validity of this mechanism is further substantiated by the case study of Student Index 0, where the hard masking mechanism successfully filters out already-mastered targets with high precision, thereby generating novel recommendations tailored to the student's specific communication deficit profile.

This study also addresses the challenge of system accountability through three integrated technical contributions. First, it converts qualitative teacher observation text into binary representations using a standardized whitelist tokens module. Second, it implements an inverse frequency penalty that eliminates popularity bias, ensuring that rare yet crucial independence targets can still be recommended. Third, it integrates an Explainable AI (XAI)-based pedagogical rule engine. This aligns with findings indicating that explainability techniques, such as LIME and SHAP, not only enhance transparency but can also drive improvements in recommendation precision under certain conditions, directly challenging the assumption of an absolute trade-off between accuracy and interpretability [27]. Clinical decision support system literature emphasizes that the feasibility of a system as a decision supporter rather than a decision dictator is heavily determined by its ability to explain the underlying rationale of its recommendations to end-users, in this case, SLB teachers and practitioners [29]. Collectively, these three mechanisms ensure that the system's outputs are not a black box, but instead possess a clinical and pedagogical reasoning foundation that supports more objective and personalized IEP formulation.

Despite the strong model evaluation performance, certain limitations remain that present strategic opportunities for future development. A fundamental constraint lies in the dataset scale, which currently encompasses only 50 students. This aligns with a common challenge in autism data science, where the majority of machine learning studies on ASD populations still rely on dozens of subjects per group, thereby limiting the generalizability of the findings [25]. This constraint causes prediction failures to concentrate on students with extreme clinical profiles who lack comparative neighbors. Another limitation lies in the fusion weight α , which was fixed ($\alpha=0.5$) without a systematic exploration of performance sensitivity across other α values; future work is encouraged to perform a grid search or cross-validation-based α optimization to identify a more optimal weighting combination. Additionally, the system does not yet feature a longitudinal progress tracking mechanism to dynamically update the recommendation matrix, a major opportunity that remains largely unexplored in data-driven ASD intervention literature [25]. Consequently, expanding the dataset through a national SLB consortium is a primary future agenda, alongside migrating the system to a user-friendly web/mobile platform and integrating continuous progress tracking features so that intervention recommendations can evolve adaptively alongside the child's real-world developmental progress over time.

5. CONCLUSION

This study successfully designed and validated a hybrid recommendation system that integrates a cosine similarity-based Patient Similarity Model (PSM) with ALS-based Collaborative Filtering to personalize self-help programs for students with Autism Spectrum Disorder (ASD). Using functional data from 50 SLB students in Central Java, LOOCV evaluation showed the Hybrid (Fusion) model outperformed single-method baselines, achieving a Hit Rate@3 of 0.8400, an NDCG@3 of 0.8079, and an MRR@3 of 0.7967, compared to an HR@3 of 0.6200 for the CF-Only baseline. The primary contributions of this research lie in its approach to converting qualitative teacher observations into structured input, mitigating popularity bias, and incorporating an XAI-based pedagogical rule engine to support clinically accountable recommendations. The study remains limited by its localized dataset scale and the absence of longitudinal tracking; future work should therefore prioritize expanding data collection through a nationwide SLB consortium, developing a web- or mobile-based platform, and integrating dynamic progress tracking to support adaptive, real-time intervention planning.

ACKNOWLEDGMENTS

The author sincerely thanks the advisors and faculty members for their guidance and feedback. Appreciation is also extended to PT Azki Inovasi Nusantara for providing a supportive Data Science internship environment. The author is grateful to the principals, teachers, and Special Education Teachers (GPK) at the participating institutions in Central Java for their cooperation and valuable data. Finally, sincere thanks are given to all students and parents for their trust and participation in the study.

BIBLIOGRAPHY

- [1] O. I. Talantseva, R. S. Romanova, E. M. Shurdova, T. A. Dolgorukova, P. S. Sologub, O. S. Titova, D. F. Kleeva, and E. L. Grigorenko, "The global prevalence of autism spectrum disorder: A three-level meta-analysis," *Front. Psychiatry*, vol. 14, 2023, Art. no. 1071181, doi: 10.3389/fpsyt.2023.1071181.
- [2] D. T. Kertapati, "Anak-anak Luar Biasa," *Mediakom*, no. 152, pp. 20–21, Mar. 2023. [Online]. Available: <https://mediakom.kemkes.go.id/>
- [3] J. A. Yon-hernández *et al.*, "Executive functions in daily living skills: A study in adults with autism spectrum disorder," no. April, pp. 1–10, 2023, doi: 10.3389/fpsyg.2023.1109561.
- [4] C. Auld and A. Cashin, "Daily living skills of autistic adolescents and young adults: A scoping review," no. April, pp. 456–474, 2022, doi: 10.1111/1440-1630.12806.
- [5] A. Masi, M. M. DeMayo, N. Glozier, and A. J. Guastella, "An overview of autism spectrum disorder, heterogeneity and treatment options," *Neurosci. Bull.*, vol. 33, no. 2, pp. 183–193, 2017, doi: 10.1007/s12264-017-0100-y.
- [6] T. Shu, K. Peng, Q. Liu, Y. Zhu, and J. Wang, "Personalized recommendation algorithm for rehabilitation intervention in children with autism spectrum disorder based on the cognitive diagnosis model," no. January, pp. 1–12, 2026, doi: 10.3389/fpsyg.2025.1696155.
- [7] N. C. Brunsting, M. M. Cumming, J. G. Garwood, and N. Urquiza, "Special Educators' Wellbeing, Burnout, and Engagement," in *Handbook of Research on Special Education Teacher Preparation*, 2nd ed., E. McCray, E. Bettini, M. Brownell, J. McLeskey, and P. Sindelar, Eds. New York: Routledge, 2024, pp. 296–318, doi: 10.4324/9781003297093-19.
- [8] M. M. Hughes *et al.*, "Individualized Education Programs and Transition Planning for Adolescents With Autism," vol. 152, no. 1, pp. 1–15, 2024, doi: 10.1542/peds.2022-060199.
- [9] M. Deschênes, "Recommender systems to support learners' Agency in a Learning Context: a systematic review," 2020.
- [10] V. N. Rathod, R. H. Goudar, A. Kulkarni, G. M. Dhananjaya, and G. S. Hukkeri, "A Survey on E-Learning Recommendation Systems for Autistic People," *IEEE Access*, vol. 12, no. November 2023, pp. 11723–11732, 2024, doi: 10.1109/ACCESS.2024.3355589.

-
- [11] Z. Ren, B. Peng, T. K. Schleyer, and X. Ning, "Hybrid collaborative filtering methods for recommending search terms to clinicians," *J. Biomed. Inform.*, vol. 113, p. 103635, 2021, doi: 10.1016/j.jbi.2020.103635.
 - [12] M. Kohli, A. K. Kar, A. Bangalore, and P. Ap, "Machine learning - based ABA treatment recommendation and personalization for autism spectrum disorder: an exploratory study," *Brain Informatics*, 2022, doi: 10.1186/s40708-022-00164-6.
 - [13] N. Wang *et al.*, "Study on the semi - supervised learning - based patient similarity from heterogeneous electronic medical records," *BMC Med. Inform. Decis. Mak.*, vol. 21, no. 2, pp. 1–14, 2021, doi: 10.1186/s12911-021-01432-x.
 - [14] A. Sharafoddini, J. A. Dubin, and J. Lee, "Patient similarity in prediction models based on health data: A scoping review," *JMIR Med. Inform.*, vol. 5, no. 1, e7, 2017, doi: 10.2196/medinform.6730.
 - [15] H. Sen *et al.*, "Patient similarity analytics for explainable clinical risk prediction," *BMC Med. Inform. Decis. Mak.*, pp. 1–12, 2021, doi: 10.1186/s12911-021-01566-y.
 - [16] J. vom Brocke, A. Hevner, and A. Maedche, "Introduction to Design Science Research," in *Design Science Research. Cases*, Progress in IS, Cham: Springer International Publishing, 2020, pp. 1–13, doi: 10.1007/978-3-030-46781-4_1.
 - [17] S. Rendle, W. Krichene, L. Zhang, and J. Anderson, "Neural Collaborative Filtering vs. Matrix Factorization Revisited," in *Proc. 14th ACM Conf. Recommender Systems (RecSys '20)*, 2020, pp. 240–248, doi: 10.1145/3383313.3412488. (Preprint: arXiv:2005.09683)
 - [18] L. J. Chew, S. C. Haw, S. Subramaniam, H. A. Santoso, D. S. Kusumo, and R. Kannan, "A hybrid recommender system based on data enrichment on the ontology modelling," *F1000Research*, vol. 10, p. 937, 2021, doi: 10.12688/f1000research.73060.1.
 - [19] A. N. Navaz, H. T. El-Kassabi, M. A. Serhani, A. Oulhaj, and K. Khalil, "A Novel Patient Similarity Network (PSN) Framework Based on Multi-Model Deep Learning for Precision Medicine," *J. Pers. Med.*, vol. 12, no. 5, p. 768, 2022, doi: 10.3390/jpm12050768.
 - [20] S. Rendle, "Item Recommendation from Implicit Feedback," arXiv:2101.08769, 2021.
 - [21] H. Liu, W. Wang, Y. Zhang, R. Gu, and Y. Hao, "Neural Matrix Factorization Recommendation for User Preference Prediction Based on Explicit and Implicit Feedback," vol. 2022, 2022.
 - [22] Y. Zhang and X. Chen, "Explainable recommendation: A survey and new perspectives," *Found. Trends Inf. Retr.*, vol. 14, no. 1, pp. 1–101, 2020, doi: 10.1561/15000000066.
 - [23] N. Ammar and A. Shaban-Nejad, "Explainable Artificial Intelligence Recommendation System by Leveraging the Semantics of Adverse Childhood Experiences: Proof-of-Concept Prototype Development," *JMIR Med. Inform.*, vol. 8, no. 11, p. e18752, 2020, doi: 10.2196/18752.
 - [24] R. H. Kirschen, E. A. O'Higgins, and R. T. Lee, "The Royal London space planning: An integration of space analysis and treatment planning," *Am. J. Orthod. Dentofac. Orthop.*, vol. 118, no. 4, pp. 456–461, 2000, doi: 10.1067/mod.2000.109032.
 - [25] P. Washington and D. P. Wall, "A Review of and Roadmap for Data Science and Machine Learning for the Neuropsychiatric Phenotype of Autism," *Annu. Rev. Biomed. Data Sci.*, vol. 6, pp. 211–228, 2023, doi: 10.1146/annurev-biodatasci-020722-125454.
 - [26] A. Klimashevskaja, D. Jannach, M. Elahi, and C. Trattner, "A survey on popularity bias in recommender systems," *User Model. User-adapt. Interact.*, vol. 34, no. 5, pp. 1777–1834, 2024, doi: 10.1007/s11257-024-09406-0.
 - [27] J. Govea, R. Gutierrez, and W. Villegas-ch, "Transparency and precision in the age of AI: evaluation of recommendation systems," no. September, 2024, doi: 10.3389/frai.2024.1410790.
 - [28] J. Wei, H. Yan, X. Shao, L. Zhao, L. Han, and S. Wang, "A machine learning-based hybrid recommender framework for smart medical systems," pp. 1–24, 2024, doi: 10.7717/peerj-cs.1880.
-

-
- [29] J. Amann *et al.*, "To explain or not to explain?—Artificial intelligence explainability in clinical decision support systems," *PLOS Digit. Health*, vol. 1, no. 2, p. e0000016, 2022, doi: 10.1371/journal.pdig.0000016.