CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES

https://cajmtcs.casjournal.org/index.php/CAJMTCS

Volume: 07 Issue: 04 | October 2026 ISSN: 2660-5309

Article

Proposed Hybrid Intelligent Model Based on Artificial Neural Networks and Decision Trees for Early Detection of Learning Disabilities: Design and Performance

Mushtaq Talib Tally[*1]

1. Ministry of Education Directorate of Education in Babil, Iraq, Open Educational College
Correspondence: mushtaqtalb@gmail.com

Abstract: Early identification of learning difficulties is an important challenge in education because timely screening can support appropriate educational intervention. Machine-learning methods can process multidimensional information, but individual approaches involve trade-offs between predictive performance and interpretability. This study proposes a sequential hybrid model that combines a C4.5 decision tree with a multilayer perceptron (MLP). In the first stage, the decision tree identifies informative features and generates human-readable IF–THEN rules; in the second stage, the selected features are provided to an MLP for three-class classification. The model was evaluated on a synthetic dataset of 1,200 children aged 5–7 years, containing 18 simulated demographic, cognitive, linguistic/academic, and behavioral features and three target classes: Typical, At-Risk Dyslexia, and At-Risk Dyscalculia. The hybrid model achieved 96.8% accuracy and a macro F1-score of 0.97 on the reported evaluation set, compared with 94.2% and 0.94 for the standalone MLP and 88.6% and 0.88 for the standalone decision tree. Because the experiments used synthetic data exclusively, these findings constitute a methodological proof of concept rather than evidence of clinical or educational diagnostic validity. Validation using independently collected real-world datasets is required before practical deployment.

Keywords: Learning Disabilities, Early Detection, Risk Classification, Hybrid Model, Artificial Neural Networks

Citation: Tally, M. T. Proposed Hybrid Intelligent Model Based on Artificial Neural Networks and Decision Trees for Early Detection of Learning Disabilities: Design and Performance. Central Asian Journal of Mathematical Theory and Computer Sciences 2026, 7(4), 153-158.

Received: 30th Jun 2026
Revised: 20th Jul 2026
Accepted: 15th Aug 2026
Published: 30th Aug 2026

1. Introduction

Artificial intelligence and machine learning are increasingly investigated as tools for supporting screening and assessment in education. Learning difficulties are multidimensional and may involve cognitive, linguistic, academic, and behavioral indicators. Interactions among these heterogeneous variables can be difficult to represent using simple rules alone, whereas high-capacity models such as neural networks may provide limited transparency for human users. These characteristics motivate hybrid architectures that combine predictive learning with interpretable representations.

Specific learning disorder is characterized by persistent difficulties in learning and using academic skills despite appropriate intervention [1]. Research on early identification has considered indicators such as phonological awareness, letter and number knowledge, working memory, and rapid automatized naming [2]. Early screening can support timely educational intervention, but screening should be distinguished from formal diagnosis;

computational models should support, rather than replace, qualified professional assessment [3].

Decision trees such as ID3, CART, and C4.5 are attractive when transparency is important because their decision paths can be expressed as IF–THEN rules [4], [5]. Artificial neural networks, including multilayer perceptrons, can represent nonlinear relationships among multiple input variables [6]. However, the reasoning underlying an individual neural-network prediction is not directly visible to the user [7].

Hybrid approaches can combine complementary characteristics of different learning algorithms. In this study, the decision tree is used as the first stage of a sequential pipeline for feature selection and rule extraction, while the MLP performs the final three-class classification. The contribution is therefore methodological: evaluation of a C4.5-to-MLP pipeline for learning-disability risk classification on a controlled synthetic dataset.

The objective of this study is to evaluate whether the proposed sequential hybrid ANN-DT model can improve classification performance relative to standalone decision-tree and MLP models while providing an interpretable rule layer for early learning-disability risk screening.

A synthetic dataset was generated because access to real diagnostic data from children is difficult and raises privacy and ethical considerations. The dataset contains 1,200 simulated children aged 5–7 years. Each record was assigned to one of three classes: Typical, At-Risk Dyslexia, or At-Risk Dyscalculia. The synthetic variables were designed with reference to DSM-5 concepts and established early-school assessment constructs [8].

Eighteen features were used: demographic features (age in months and gender); cognitive features (working memory, attention, and processing speed); linguistic and early academic features (phonological awareness, letter knowledge, rapid automatized naming [RAN], number sense, and sequential counting); and behavioral features (teacher ratings for persistence, social interaction, and fine motor skills on a five-point Likert scale).

Because the data are synthetic, no human participants were enrolled and no personally identifiable information were used; therefore, institutional ethical approval was not required for data collection. The final reproducibility record should specify the exact data-generation distributions and ranges, class proportions, label-generation procedure, and random seed.

The proposed model is a sequential two-stage architecture rather than a parallel ensemble. Stage 1 uses a C4.5 decision tree, implemented as J48 in the original experimental design, to identify informative features and generate interpretable decision paths. The ten most informative upper-level features are selected for Stage 2, and approximately 5–7 decision rules with reported confidence greater than 90% are extracted.

Stage 2 uses a multilayer perceptron (MLP) with 10 input neurons corresponding to the selected features, two hidden layers containing 24 and 12 neurons, respectively, ReLU activation, and dropout of 0.4 after each hidden layer. The output layer contains three neurons with softmax activation for the three target classes. Adam is used as the optimizer and categorical cross-entropy as the loss function.

The rule layer is intended to provide global or path-based interpretability. It does not make the internal computation of the MLP intrinsically transparent; therefore, the complete ANN-DT system is not described as fully interpretable.

The dataset was divided into 70% training data (840 records) and 30% test data (360 records), with class distributions maintained between the subsets. Three models were compared: (1) a standalone J48 decision tree, (2) a standalone MLP using the 18 available features, and (3) the proposed hybrid ANN-DT model using the ten features selected by the decision-tree stage.

The reported experiment also used 10-fold cross-validation. To prevent information leakage, feature selection, model tuning, and any preprocessing that estimates parameters from data must be performed using training data only, while the held-out test set must remain untouched until final evaluation. This condition should be confirmed against the original experimental code before submission.

Performance was assessed using overall accuracy, macro precision, macro recall, and macro F1-score. Class-level F1-scores were also reported. Because repeated-run variability and confidence intervals are not reported in the current experimental record, the point estimates are interpreted descriptively rather than as definitive evidence of statistical superiority.

The experiments were implemented in Python 3.10 in Google Colab Pro using Scikit-learn for machine-learning and preprocessing functions, TensorFlow/Keras for the neural network, and Pandas/NumPy for data handling. A fixed random state of 42 was reported. For complete reproducibility, the final research package should additionally provide exact library versions, learning rate, batch size, number of epochs, stopping criteria, preprocessing/scaling procedures, and the complete data-generation and training code where permitted.

2. Materials and Method

A synthetic dataset was generated because access to real diagnostic data from children is difficult and raises privacy and ethical considerations. The dataset contains 1,200 simulated children aged 5–7 years. Each record was assigned to one of three classes: Typical, At-Risk Dyslexia, or At-Risk Dyscalculia. The synthetic variables were designed with reference to DSM-5 concepts and established early-school assessment constructs [8].

Eighteen features were used: demographic features (age in months and gender); cognitive features (working memory, attention, and processing speed); linguistic and early academic features (phonological awareness, letter knowledge, rapid automatized naming [RAN], number sense, and sequential counting); and behavioral features (teacher ratings for persistence, social interaction, and fine motor skills on a five-point Likert scale).

Because the data are synthetic, no human participants were enrolled and no personally identifiable information were used; therefore, institutional ethical approval was not required for data collection. The final reproducibility record should specify the exact data-generation distributions and ranges, class proportions, label-generation procedure, and random seed.

The proposed model is a sequential two-stage architecture rather than a parallel ensemble. Stage 1 uses a C4.5 decision tree, implemented as J48 in the original experimental design, to identify informative features and generate interpretable decision paths. The ten most informative upper-level features are selected for Stage 2, and approximately 5–7 decision rules with reported confidence greater than 90% are extracted.

Stage 2 uses a multilayer perceptron (MLP) with 10 input neurons corresponding to the selected features, two hidden layers containing 24 and 12 neurons, respectively, ReLU activation, and dropout of 0.4 after each hidden layer. The output layer contains three neurons with softmax activation for the three target classes. Adam is used as the optimizer and categorical cross-entropy as the loss function.

The rule layer is intended to provide global or path-based interpretability. It does not make the internal computation of the MLP intrinsically transparent; therefore, the complete ANN-DT system is not described as fully interpretable.

The dataset was divided into 70% training data (840 records) and 30% test data (360 records), with class distributions maintained between the subsets. Three models were compared: (1) a standalone J48 decision tree, (2) a standalone MLP using the 18 available features, and (3) the proposed hybrid ANN-DT model using the ten features selected by the decision-tree stage.

The reported experiment also used 10-fold cross-validation. To prevent information leakage, feature selection, model tuning, and any preprocessing that estimates parameters from data must be performed using training data only, while the held-out test set must remain untouched until final evaluation. This condition should be confirmed against the original experimental code before submission.

Performance was assessed using overall accuracy, macro precision, macro recall, and macro F1-score. Class-level F1-scores were also reported. Because repeated-run variability and confidence intervals are not reported in the current experimental record, the point estimates are interpreted descriptively rather than as definitive evidence of statistical superiority.

The experiments were implemented in Python 3.10 in Google Colab Pro using Scikit-learn for machine-learning and preprocessing functions, TensorFlow/Keras for the neural network, and Pandas/NumPy for data handling. A fixed random state of 42 was reported. For complete reproducibility, the final research package should additionally provide exact library versions, learning rate, batch size, number of epochs, stopping criteria, preprocessing/scaling procedures, and the complete data-generation and training code where permitted.

3. Results

The proposed hybrid ANN-DT model produced the highest reported aggregate performance among the three evaluated models. The reported accuracy increased from 88.6% for the standalone decision tree to 94.2% for the standalone MLP and 96.8% for the hybrid model. Macro F1-score followed the same pattern.

Table 1. Overall performance comparison of the three models on the reported evaluation data.

Model	Overall Accuracy (%)	Macro Precision	Macro Recall	Macro F1-score
Standalone Decision Tree (DT)	88.6	0.89	0.88	0.88
Standalone Neural Network (ANN)	94.2	0.94	0.94	0.94
Proposed Hybrid Model (ANN-DT)	96.8	0.97	0.97	0.97

Table 2. Class-level F1-scores reported for the three models.

Class	Standalone DT	Standalone ANN	Proposed Hybrid ANN-DT
Typical	0.92	0.96	0.98
Dyslexia Risk	0.86	0.93	0.96
Dyscalculia Risk	0.87	0.93	0.97
Weighted Average F1	0.88	0.94	0.97

The reported class-level F1-scores were highest for the hybrid model in all three classes. For the two risk categories, the reported F1-scores were 0.96 for Dyslexia Risk and 0.97 for Dyscalculia Risk. These results describe performance within the synthetic evaluation setting and should not be interpreted as estimates of clinical sensitivity or specificity.

The reported example rule is:

IF Phonological Awareness Score ≤ 40 (out of 100) AND RAN Time > 70 seconds AND Teacher Persistence Rating ≤ 2 (out of 5), THEN the child is classified as At-Risk Dyslexia.

This decision path illustrates the intended rule-layer interpretability. It should not be treated as a mathematical explanation of every individual MLP prediction [9].

4. Discussion

The reported results support the proposed comparison hypotheses within the synthetic experimental setting: the hybrid model achieved higher aggregate accuracy and F1-score than both standalone baselines. A plausible explanation is that the decision-tree stage reduces the input space before neural-network classification, potentially limiting irrelevant or redundant information. However, the current design does not isolate feature-selection effects from the effect of the complete hybrid architecture; an ablation study is therefore needed before attributing the performance gain specifically to the hybrid design [10].

The comparison also illustrates complementary characteristics of the two model families. Decision trees provide explicit decision paths but may have difficulty representing complex nonlinear interactions, whereas MLPs can model nonlinear relationships but offer less direct transparency. The sequential architecture attempts to combine these properties by using the tree for feature selection and rule extraction and the MLP for final classification [11].

The interpretability claim should remain limited to the rule layer. The extracted IF–THEN paths provide human-readable information about selected features, but the MLP remains a nonlinear learned model whose internal computation is not exposed by the tree rules. Stronger claims about individual-prediction explainability would require an additional explanation method or an explicit analysis linking explanations to MLP outputs [12].

The principal limitation is the exclusive use of synthetic data. Synthetic data are useful for testing an architecture under controlled conditions, but performance on simulated observations cannot establish external validity. The relationships, class boundaries, measurement distributions, and noise characteristics in the synthetic dataset may differ from those in real educational or clinical populations. Accordingly, the reported 96.8% accuracy should be interpreted as performance in the simulated environment rather than as real-world diagnostic accuracy.

A second limitation concerns the completeness of the reported model-development pipeline. In particular, the timing of feature selection relative to the train/test split and the exact use of 10-fold cross-validation must be confirmed against the original code. If held-out test information influenced feature selection or model tuning, the reported performance would be optimistically biased and the experiment should be rerun using a leakage-free pipeline [13].

A third limitation is the absence of uncertainty estimates and formal statistical comparison. Future work should use repeated stratified cross-validation or an equivalent design, report variability and confidence intervals, and apply an appropriate paired comparison when claiming performance differences [14].

Finally, the proposed system should be positioned as a potential screening or decision-support method rather than a diagnostic replacement. Practical application with children would require validation using independently collected real-world data, an appropriate reference standard, subgroup-performance analysis, and evaluation of how human experts use the model outputs [15].

5. Conclusion

This study proposed and evaluated a sequential hybrid model that combines a C4.5 decision tree with a multilayer perceptron for three-class learning-disability risk classification. The decision-tree stage selects informative features and generates interpretable IF–THEN rules, while the MLP uses the selected features for final classification. On the reported synthetic evaluation dataset, the hybrid model achieved 96.8% accuracy and a macro F1-score of 0.97, exceeding the reported performance of the standalone decision tree and MLP. The findings indicate that the proposed architecture is a promising methodological approach for combining feature selection, rule extraction, and nonlinear classification. However, because the evaluation was based entirely on synthetic data and the complete leakage-free model-development pipeline has not yet been independently verified from the supplied manuscript, the findings should be regarded as a proof of concept rather than evidence of clinical or educational diagnostic validity. External validation on real-world data is essential before practical use.

REFERENCES

[1] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, DC, USA: American Psychiatric Association, 2013.

[2] H. W. Catts, M. E. Fey, X. Zhang, and J. B. Tomblin, "Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation," Language, Speech, and Hearing Services in Schools, vol. 32, no. 1, pp. 38–50, 2001.

[3] J. M. Fletcher, G. R. Lyon, L. S. Fuchs, and M. A. Barnes, Learning Disabilities: From Identification to Intervention, 2nd ed. New York, NY, USA: Guilford Press, 2019.

[4] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: A review of classification and combining techniques," Artificial Intelligence Review, vol. 26, no. 3, pp. 159–190, 2006.

[5] J. R. Quinlan, C4.5: Programs for Machine Learning. San Mateo, CA, USA: Morgan Kaufmann, 1993.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[7] Z. C. Lipton, "The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery," Queue, vol. 16, no. 3, pp. 31–57, 2018.

[8] I. Ismoilov, "Iskandar oynasi," Oltin bitiglar–Golden Scripts, vol. 4, no. 4, 2019.

[9] I. Ismoilov, "The image of 'The Mirror of Iskandar' in the poetry of Alisher Navai and its comparative analysis."

[10] I. Ismoilov, "'Boburnoma' as a literary source," American Journal of Social Sciences and Humanity Research, vol. 5, no. 11, pp. 128–130, 2025, doi: 10.37547/ajsshr/Volume05Issue11-32.

[11] B. Tukhliev and I. Ismailov, "The image of Iskandar in Turkic sources," International Journal of Scientific & Technology Research, vol. 9, no. 2, pp. 1195–1200, 2020.

[12] I. Ismoilov, "About the talismans of Iskandar (Alexander) by Alisher Navoi," International Journal of Linguistics, Literature and Culture, vol. 7, no. 3, pp. 139–145, 2021, doi: 10.21744/ijllc.v7n3.1488.

[13] I. Ismoilov, "Nizomiy va Navoiy: bir mavzuning turfa talqini," O'zbekiston: til va madaniyat. Adabiyotshunoslik, vol. 1, no. 1, pp. 18–32, 2023.

[14] I. Ismoilov, "Alisher Navoiyning 'Saddi Iskandariy' dostoni debochasi badiiyati," Oltin bitiglar–Golden Scripts, vol. 1, no. 1, 2021.

[15] I. Ismoilov, "Tafsirlarda Iskandar mavzusiga oid talqinlar," Oltin bitiglar–Golden Scripts, vol. 2, no. 2,2021