SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Developing AI-Powered Adaptive Testing Systems with Few-Shot Learning Education Techniques

Anitha M[1*], Dr.G.H. Kerinab Beenu[2], Gulboy Yusupov[3], Mahesh Kumar PG[4], Dinara Abdurakhmonova[5], Yulduz Raximova[6]

[1*]Assistant Professor, Department of Mathematics, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. Email id: anitham@maher.ac.in

[2]Associate Professor, Department of Management Studies, Jerusalem College of Engineering, Pallikaranai, Chennai, Tamilnadu, India. Email id: beenustanley2008@gmail.com, https://orcid.org/0000-0001-5759-0089

[3]Department of Accounting and Statistics, Termez University of Economics and Service, Termez, Uzbekistan. E-mail: gulboy_yusupov@tues.uz, ORCID: https://orcid.org/0009-0001-3502-9917

[4]Professor, Meenakshi College of Physiotherapy, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. E-mail: mahesh@maher.ac.in

[5]Doctor of Philosophy (PhD) in Pedagogical Sciences, Lecturer, Department of Primary Education Pedagogy, Jizzakh State Pedagogical University, Jizzakh, Uzbekistan. E-mail: dinara.abduraxmonova10@gmail.com, https://orcid.org/0000-0002-4110-1655

[6]Senior Lecturer, Department of Theory of Adaptive Physical Education and Sports; Candidate of Pedagogical Sciences, Fergana State University, Fergana, Uzbekistan. E-mail: yulduzkhon.ergasheva@gmail.com, https://orcid.org/0009-0007-0837-9907

*Corresponding author: Email: anitham@maher.ac.in

Abstract

A few-shot learning is investigated as a method to develop AI-based adaptive testing systems for personalized learning in the field of education. The proposed model utilizes few shots learning to adapt the difficulty level of the tests according to limited information, and provides personalized tests to the students at real time to enhance learning results. The methodology includes data collection from various sources, including student behavior, assessments, and demographic data. It exploits preprocessing on the Open University Learning Analytics Dataset (OULAD) followed by the construction of a small number of examples in the dataset to allow the model to learn from a few-shot. The system utilizes AI-driven feature engineering to forecast student capabilities and choose the most suitable test items, developing an adaptive test flow. The results show high accuracy (99.91%), high macro average precision (99.92%) and high macro average F1-score (99.88%), indicating that the system can make accurate predictions even with small number of training examples. The results show that the system can always beat the traditional models, such as logistic regression (78%), random forests (84%) and gradient boosting (86%) through statistical analysis of the F1-scores in one-shot, five-shot, and ten-shot learning scenarios. The FSL-Adapt Test model had an F1-score of 0.901 for the one-shot tasks, 0.836 for the five-shot tasks, and 0.886 for the ten-shot tasks, and an impressive F1-score of 0.775 and 0.859 for the novel tasks (Novel F1 = 0.775 for five-shot and 0.859 for ten-shot tasks). The results validate the power of few-shot learning for scalable, efficient and adaptable testing toward better educational outcomes. The results highlight the promise of using AI and few-shot learning methods to transform adaptive testing systems.

Keywords

AI-powered testing, few-shot learning, adaptive testing systems, personalized learning, machine learning, educational technology, Open University Learning Analytics Dataset (OULAD).

1. Introduction

The goal of developing AI-based adaptive testing systems using few-shot learning education techniques is to transform personalized learning by efficiently adapting to students' individual needs based on limited data. Advance digital information systems take advantage of sophisticated machine learning models, including few-shot learning, to offer personalized assessments and make certain the learning experience is efficient and personalized. Few-shot learning allows the system to learn from a few examples, making it more adaptable to a wider range of students, as opposed to models that rely on large amounts of data. These systems can enhance learning outcomes and promote student engagement by providing real-time feedback and adapting to their needs, ensuring education is more responsive and effective. This paper discusses feedback systems that utilize AI to create collaborative learning environments, and to improve the adaptability and real-time responsiveness of educational platforms, which can be implemented for adaptive testing systems to provide personalized feedback[1].This research focuses on the application of the algorithms using Python for the Adaptive Testing Systems, which provide personalized assessment of the learner's progress, closely related to the application of few-shot learning for dynamic adaptation [2].This paper focuses on prompt engineering empowered by artificial intelligence to maximize educational resources, which can be expanded to include adaptive testing systems driven by AI to keep students engaged in meaningful, context-sensitive assessments[3] [18].This study examines the use of few-shot AI-generated feedback in physical therapy education, and how using limited data can inform meaningful feedback, which can be used to inform adaptive testing systems[4].In this paper, focus is given on the use of multimodal AI assistants to provide real-time feedback in large classrooms, with insights that can be generalized to large-scale applications of adaptive testing systems, which leverage multimodal data and few-shot learning to evaluate students[5].This research aims to create a generative AI-based adaptive tutoring system that adapts to student needs, similar to adaptive testing systems that leverage AI to tailor assessments to individual student needs to enhance learning outcomes[6].The study also examines the potential of e-learning platforms to support personalized learning and provides an introduction to AI-based approaches that can be applied to adaptively test students without sacrificing efficiency by few-shot learning techniques[7][23].The focus of this paper is on the evolution of Artificial Intelligence (AI) -based learning management systems that have incorporated technologies such as ChatGPT, and considerations of its implications for adaptive testing platforms and few-shot learning[8].In this study, designer of adaptive e-learning system using learner behavior analytics was addressed, and it was considered that it can be applied to the development of AI-based adaptive testing system with few-shot learning techniques for dynamic evaluation[9] from the viewpoint of practicality. This study investigates how cognitive AI can be applied in AL environments, and suggests a framework for developing ethical and context-aware tutoring systems, which can be extended to the development of AI-based adaptive testing systems that adapt tests to specific student needs [10][24].

Key Contribution

- The integration of AI-powered adaptive testing systems enables real-time adjustment of test difficulty based on individual student performance, ensuring that assessments are personalized. Few-shot learning techniques allow the system to efficiently learn from minimal data, making it adaptable to new learners without requiring vast amounts of training data.
- By leveraging few-shot learning, the system can operate efficiently even in large-scale educational settings with limited resources. This approach reduces the need for extensive labeled data, enabling scalable implementation across diverse educational contexts, including remote or underserved areas.
- The adaptive nature of the system not only enhances the accuracy of assessments but also provides timely feedback to students. This helps improve learning outcomes by addressing individual learning gaps and ensuring that each student is assessed at an appropriate level, promoting better understanding and retention.

This research covers by the following sections Section I explained about introduction about the research topic, Section II explained about literature review for previous work, Section III explained about overall architecture, Few Shot learning educational techniques, data flow diagram, proposed algorithm. Section IV explained about dataset description, software hardware configuration and various analysis. Section V explained the main key summary in conclusion sections.

2. Literature Review

Meta-Learning with Variational Inference for Fault Diagnosis Under Variable Conditions represents a relevant technique for implementing adaptive learning in a system characterized by few-shot data, just like adaptive tests determine the following question based on few prior data[11].The paper investigates learner-oriented methods to increase motivation, in line with AI-powered adaptive systems that customize the assessment process according to the students' individual requirements[12].Dynamic Gating Ensembles for Rejection using Large Language Models in Medical Image Diagnostics by DEGRE stresses trust-aware models, which are vital for implementing adaptive tests that utilize AI algorithms to predict the performance of a student[13].The present paper examines AI applications for scalable log analysis in LLMs, which may help build AI-powered adaptive testing systems[14].This research discusses the use of AI algorithms to improve student learning and performance, which is applicable to AI-powered adaptive testing systems[15].The research article investigating the use of meta-learning for rapid adaptation in AIOps applications uses few-shot learning and offers approaches that could be used in AI-driven adaptive testing systems through learning from small data sets[16].In this paper, the development of an AI-based adaptive learning system designed for university students is discussed, which resembles adaptive testing systems in terms of adaptation to the characteristics and performance level of the learner[17][25].Few-shot learning is implemented using the flipped classroom approach in this research paper, indicating the possible use of adaptive systems to personalize content delivery, similar to adaptive testing systems where content is chosen based on the performance of the learner[21].This review covers few-shot learning methods and Mathematical Framework based on deep learning, providing valuable knowledge that could benefit adaptive testing systems, requiring fast and efficient adaptation to novel data with a minimum number of examples[19][22].

Research Gap

One of the main limitations in the application of an adaptive testing system that uses AI and few-shot learning in just a few shots is the issue of data integration from different sources, including demographic and behavioral data, in order to provide better test adaptability using the least amount of data. Although each element of the test has been studied separately, there are still very few studies conducted in terms of integrating these elements into an adaptive testing system.

3. Proposed Methodology

3.1 Overall Architecture of Proposed Methodology

Figure 1 represents the process of developing an adaptive testing system with AI-driven few-shot learning approach to improve the personalized learning experience. Firstly, data collection is conducted from a variety of sources, among them being LMS, student interactions, assessments, and OER. The range of various types of data is gathered for further usage: behavioral data, performance data, content metadata, and demographic data. Data preparation consists of cleaning, transformation, and encoding, after which the system builds a few-shot example pool, helping the model learn from limited data. AI-driven feature engineering includes creation of student embeddings and usage of knowledge tracing and content/item features in order to consider both aspects. As a result, there appears an AI-driven adaptive testing with the application of few-shot learning approach in order to identify students' skills and select the next testing item. For the AI models (which are based on few-shot learning), the system utilizes architectural designs such as transformers or few-shot adapters to generate the outputs which include estimations of abilities, the best next item, confidence metrics, and explanations. This is succeeded by learning and model refinement where the system collects additional interactions, updates its repository of examples and trains its models continuously. Evaluation and feedback constitute the final step that plays a critical part in the process since involve performance analysis, tracking of fairness and bias metrics, and obtaining feedback from both students and teachers. This process guarantees consistent improvement in the system performance resulting in increased efficiency in education delivery through constant optimization.

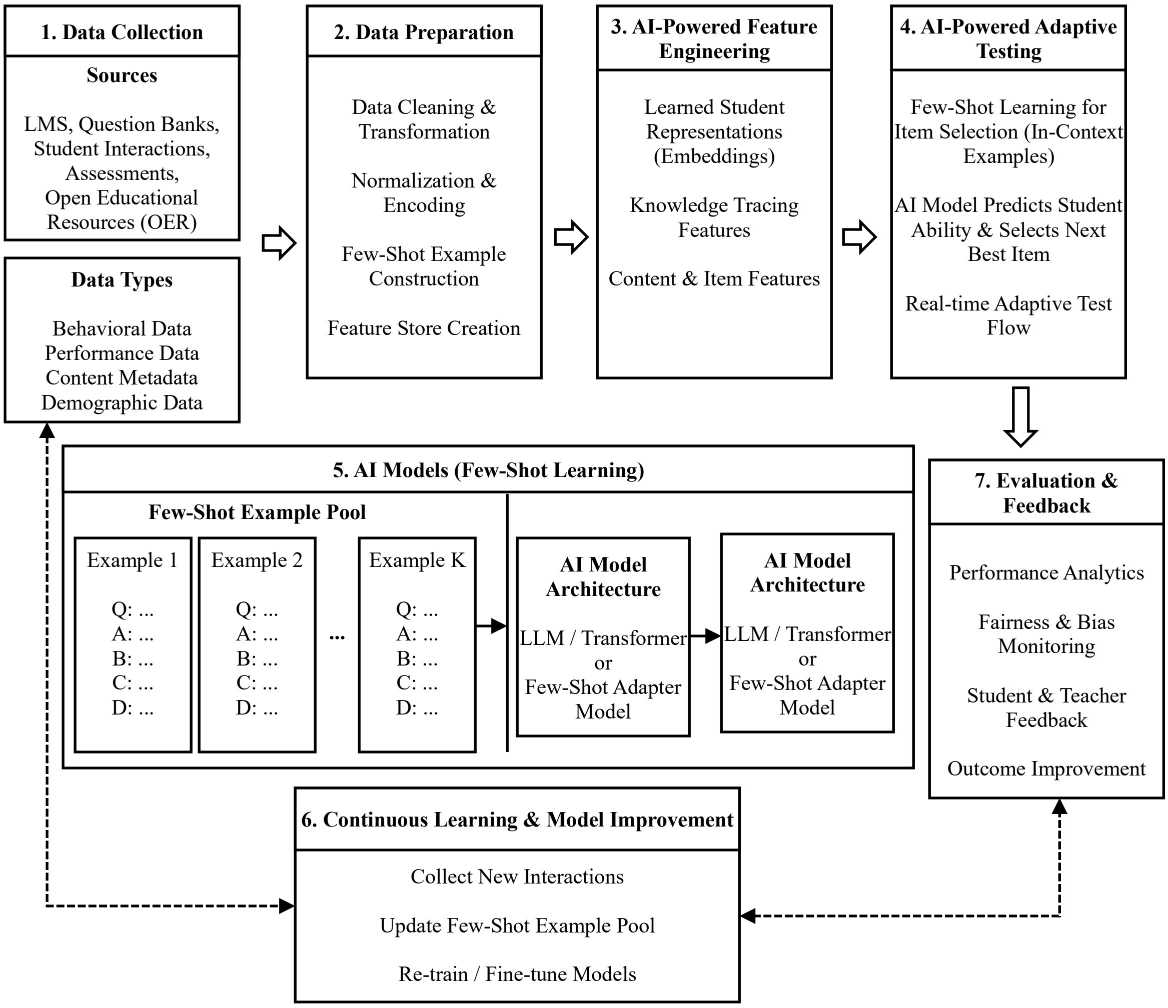

Figure 1: Overall Architecture of Proposed Methodology

3.2 Few Shot Learning Education Techniques

Fig2 demonstrates a flow structure for incorporating FSL methods into educational application systems. The design has been divided into four major layers: Data layer, FSL Model Layer, Application Service Layer, and Feedback & Interaction Layer. The Data layer manages important elements such as the dataset of student behaviors, educational material resources, and few-shots examples. In the next level, the FSL Model Layer involves the use of embedding models like transformers, long short-term memory (LSTM), and meta-learners MAML and ProtoNet, together with the repository of models to increase adaptability and efficiency. The third layer of Application Service integrates learning models to applications and APIs for prediction, scheduling of tasks, and personalized training. The last layer of Feedback & Interaction is the interaction layer which involves student interaction through mobile/web apps, teacher interface through educational dashboards, and administrator management through console interfaces. Continuous improvement will be made by incorporating user feedback and model monitoring for performance and learning episode tuning. Thus, the system will be capable of adapting itself to educational situations with very few samples of data.

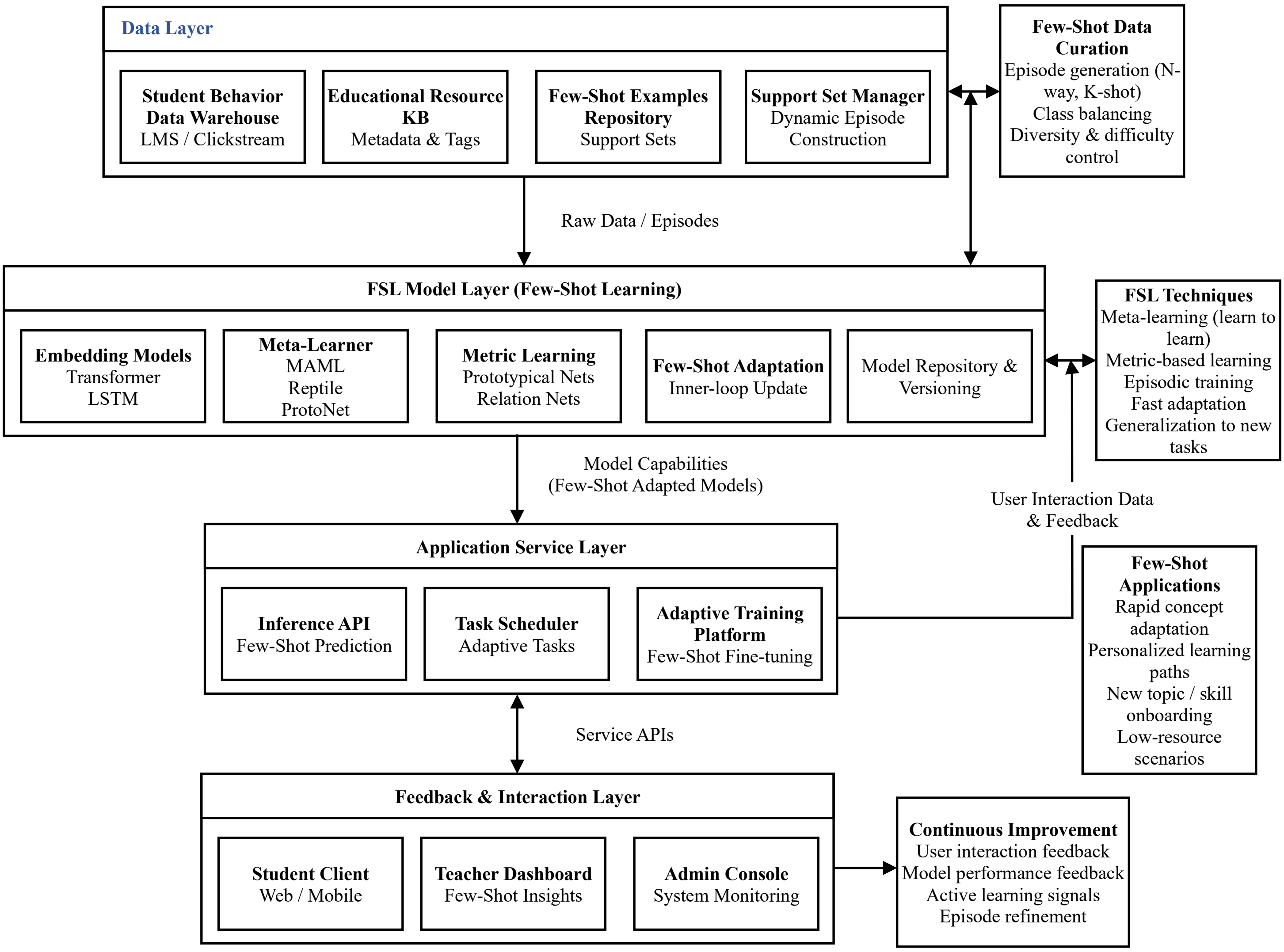

Figure 2: Working principle of Few-Shot Learning Education Techniques

3.2.1 Few Shot Learning

Deep learning models have performed exceptionally well in fields like as image and text categorization in the era of big data. The effectiveness of these models is strongly dependent on the availability of a large amount of training data. The labeling of data, on the other hand, becomes a hard and time-consuming operation for specific categories that have a limited number of samples that have been tagged. Few-shot learning was created to address this, allowing models to learn quickly from minimum input, similar to human learning, and thereby aligning machine learning more closely with human cognitive capacities. The fundamental nature of machine learning is the source of the difficulty that is associated with few-shot learning. The objective of machine learning is to acquire knowledge of a function inside a specific sample space, that is, to map input data to output data and to enable accurate predictions for input that is unknown.

A machine learning training set consists of n independent with identically distributed to observational samples $D = \{(x_1, y_1) \dots \dots (x_n, y_n)\}$. The objective helps for find the optical functions $f * (x)$ which is contains the set of functions $\{f(x)\}$ that should map to the $x \; to \; y$ for minimize the expected risk. $R(f) = \int L(y, f(x)) df(x, y)$. Here L is the loss function used for to estimate the various scenarios f.

$$R_{emp}(f) = \frac{1}{n} \sum_{t=1}^{n} L(y_1, f(x_1)) \qquad (1)$$

From the above Equation (1) describes the empirical risk used for the expected risk and algorithm to minimize it.

Generalization errorε_{upp} is the fundamental reason for the difficulty in few-shot learning. In statistical machine learning theory for bounding generalization error, for any $0 < \delta < 1$ and the expected risk R satisfy at least probability, as Equation (2),

$$R \leq R_{emp} + \emptyset\left(\frac{CM}{n}\right) \qquad (2)$$

Here, n represents the size of the observational samples (training set) indicates capacity measurement, which estimates the representational capacity of the hypothesis space. This is usually reflected in the model's complexity or the learning capacity of its function set, along with associated measures like VC dimension and Rademacher complexity. The function is non-negative in both its domain and range and varies in expression across different problems.

3.3 Data Flow Diagram

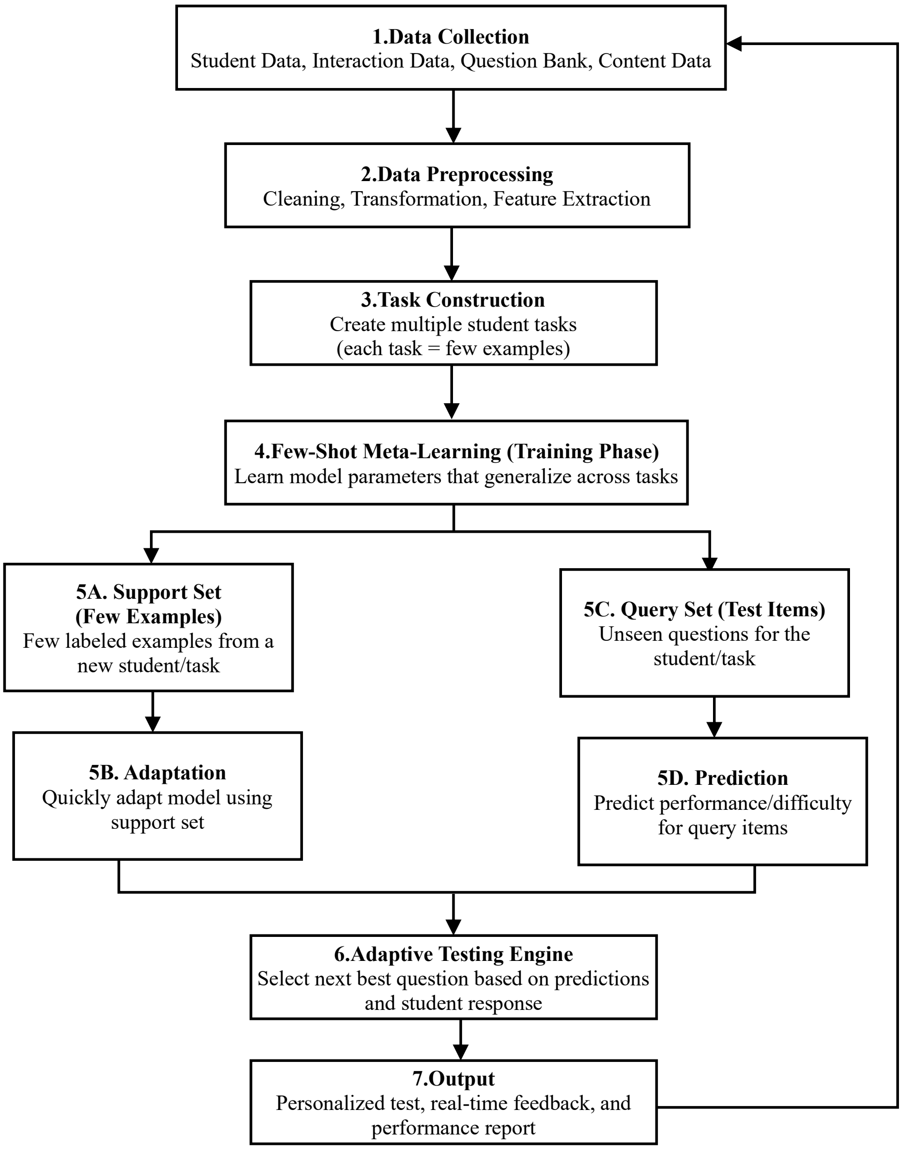

Figure 3: Data flow diagram of proposed Methodology

Fig3 shows the AI-Powered Adaptive Testing System with Few-Shot Learning Techniques consists of a well-defined process tailored to the unique needs of each learner. To begin with, Data Collection involves gathering data from different sources such as learner activity, interaction, question banks, and learning material. Afterward, the data is processed in the Preprocessing stage by cleaning, transforming, and extracting features. The creation of multiple student tasks with only a few examples each constitutes the next phase known as Task Construction, making it easy to tackle new learners and tasks. The core of the entire process revolves around the Few-Shot Meta-Learning (Training) Phase, in which the model acquires parameters applicable across numerous tasks based on a limited number of examples. Key activities at this phase include generating a Support Set consisting of few samples from a new learner or task, followed by Adaptation using the support set. In addition, the system develops a Query Set that contains unseen test questions customized for the student and applies the Prediction method to predict the results and level of difficulty of the test questions. In turn, the Adaptive Testing Engine employs predictions to choose the most appropriate question according to the student's answers. Lastly, the Output process involves delivering a unique test, providing instant feedback, and creating a thorough performance analysis report for the learner to use when determining how to proceed with studying. As a result, the described system utilizes few-shot learning to personalize the testing process for each learner individually while requiring minimal data entry.

3.4 Proposed Algorithm

$Algorithm: FSL - Adapt\ Test\ for\ Adaptive\ Learning\ in\ Education$

$Data: Student\ behavior\ data\ D, Educational\ resource\ data\ E$

$Result: Optimized\ adaptive\ learning\ parameters\ \theta^*$

$Meta - Learning\ Phase: Randomly\ initialize\ model\ parameters\ \theta$

$foreach\ student\ test\ set\ T_i sampled\ from\ D$

- $Sample\ mini - batch\ of\ student\ data\ D_i from\ T_i$
- $Compute\ test - specific\ loss\ L_T i(f(\theta)) on\ D_i using\ current\ parameters\ \theta$
- $Compute\ gradients\ \nabla\theta L_T ion\ the\ adaptive\ learning\ tasks$
- $Update\ parameters\ \theta^{'} = \theta - \alpha\nabla\theta L_T iwith\ Kgradient\ steps$

end

$Set\ \theta^* = \theta$

$Zero - Shot\ Adaptive\ Test\ Phase: Adapt\ to\ new\ test\ tasks\ without\ retraining$

$foreach\ new\ student\ task\ Tin\ meta - test\ set$

- $Directly\ use\ \theta^* to\ predict\ the\ difficulty\ and\ performance\ level\ for\ task\ T$
- $Assess\ the\ effectiveness\ of\ the\ adaptive\ test\ on\ student\ performance$

$without\ further\ training$

end

Explanation

In the design of the FSL-Adaptive Learning Test Algorithm for Education, the aim is to ensure that the system adapts to different learners according to their needs through a few shot learning approaches. This means that the use of Few Shot Learning (FSL) and Meta-Learning algorithms will enable the system to make appropriate adjustments based on minimal training. In the meta-learning phase, the first step involves parameter initialization followed by iteration through different student test sets. At this point, the system computes the loss function for each task and then optimizes the parameters to minimize losses. Finally, after going through all the tasks, the best optimized parameters are obtained. In the zero-shot adaptive test phase, the model is evaluated against new unseen tasks using a meta-test set. This stage highlights the capability of the system to apply generalization to different tasks by making correct predictions based on insufficient information from previous cases. Using meta-learning and few-shot learning techniques in developing the FSL-Adapt Test provides an opportunity for dynamically changing the level of difficulty in the test questions based on the student's performance and knowledge level.

4. Results and Discussion

4.1 Data Set Description

The Open University Learning Analytics Dataset (OULAD) [20] has the data of students who are enrolled in online courses of The Open University. It contains student demographics, course/module data, assessment outcomes and detailed interaction logs of the Virtual Learning Environment (VLE). The data is useful in educational data mining, personalized learning, and predictive model's research. It can be utilized in meta learning and few shot learning methods in learning by analyzing student engagement and performance patterns, making it useful to develop adaptive testing systems. It is accessible on Kaggle and serves as a valuable source of enhancing the online learning experiences.

Table 1: Dataset Description

Table Name	Description
Studentinfo	Demographic Information
Student Registration	Records of student enrollment
Courses	Metadata
Assessments	assignments, examinations, and other assessments
Student Assessment	Student-specific assessment performance results
Vle	Metadata data
StudentVle	Daily interaction records based on VLE

Table 1 will give a summary of the different data in the Open University Learning Analytics Dataset (OULAD), their description, and critical attributes. Studentinfo includes demographic data of the students including age, gender, education and disability status and final results. Student Registration captures the activities of students when registering and withdrawing, the date of registration and the module ID. The Courses contains some metadata on the courses offered such as the module code, start date, end date and the length of presentation. The Assessments contains the information about assessment and exams, such as assessment IDs, types, due dates, and weights. The Student Assessment keeps a record of the particular student performance, which is associated with the relevant assessments, with a record of the scores and the date of submission. The Vle follows the resources used in the Virtual Learning Environment (VLE) and the type of resources, week of usage, and the kind of the activity. Finally, the StudentVle contains information about how students used the VLE, the level of their engagement with the resources (number of clicks) on each of them, which is essential to examine the behavioral and engagement patterns of students. Every contains important data to analyze the performance of students, their engagement, and academic performance.

4.2 Software and Hardware Configurations

Table 2: Software and Hardware Configurations

Category	Configuration
Programming Language	Python
AI/ML Frameworks	TensorFlow, PyTorch, Scikit-learn, Keras, Hugging Face
Data Processing	Pandas, NumPy
Data Visualization	Matplotlib, Seaborn
Database & Storage	MySQL, PostgreSQL, MongoDB
Cloud & Distributed Systems	Google Cloud Platform (GCP), AWS, Docker
Version Control	Git, GitHub, GitLab
Processor (CPU)	Intel Core i7/i9, AMD Ryzen 7/9 (8+ cores)
Graphics Processing Unit (GPU)	NVIDIA RTX 3080/3090, NVIDIA Tesla V100/A100, CUDA support
RAM	32 GB or more
Storage	SSD (1 TB or more), Network-attached storage (NAS), Cloud storage

Networking	High-speed Ethernet or Wi-Fi
Additional Equipment	Multiple monitors, External storage for data/model backup
Distributed Computing	Google Colab, Amazon EC2, Azure
Container Orchestration	Kubernetes

The setup of the development of AI-powered adaptive testing systems with few-shot learning is presented in Table 2 with Python as a programming language, and machine learning frameworks such as TensorFlow, PyTorch, and Keras. Relevant data tools are Pandas and NumPy to process data, Matplotlib to visualize it. MySQL and PostgreSQL are used as storage, and there are cloud solutions, such as GCP and AWS. The model training needs a high-performance CPU (Intel Core i7/i9 or AMD Ryzen 7/9) and a graphics card (NVIDIA RTX 3080/3090), accompanied by 32 GB of RAM and SSD storage. Distributed computing is managed by Google Colab and Amazon EC2, and Kubernetes is utilized to manage containers.

4.3 Parameter Initialization

In the case of the Developing AI-Powered Adaptive Testing Systems with Few-Shot Learning Education Techniques, the parameters used are the learning rate (0.001), which is a constant that defines the rate at which the model modifies its weights. The batch size will be 32, which will be the number of samples that will be analyzed in one training step, and the number of epochs will be 20, which means that the model will go through the whole data set 20 times. Adam optimizer is adopted due to its effectiveness in deep learning tasks, which reduces the loss function. The meta-learning phase has a separate meta-learning rate of 0.001 and a meta-batch size of 16 on which samples are processed in the meta-learning phase. The model structure is 3 hidden layers of 128 units with the ReLU activation function to add non-linearity. The loss function used is cross-entropy, which is used in classification, and the dropout rate is 0.2 to avoid overfitting by dropping a random percentage of neurons during training. Finally, the evaluation metric applied to determine the performance of the model in the case of adaptive testing is accuracy.

4.4 Base Line Performance Comparison with Various Models

Table 3: Base Line Performance Comparison with Various Models

Model	Accuracy	Precision	Recall	F1-Score
LR	78%	75%	73%	74%
RF	84%	82%	80%	81%
GB	86%	84%	82%	83%
MF	72%	70%	68%	69%
Q-LA	81%	79%	78%	78%
DQN Model [17]	93%	92%	91%	91%
FSL-Adapt Test	95%	94%	95%	95%

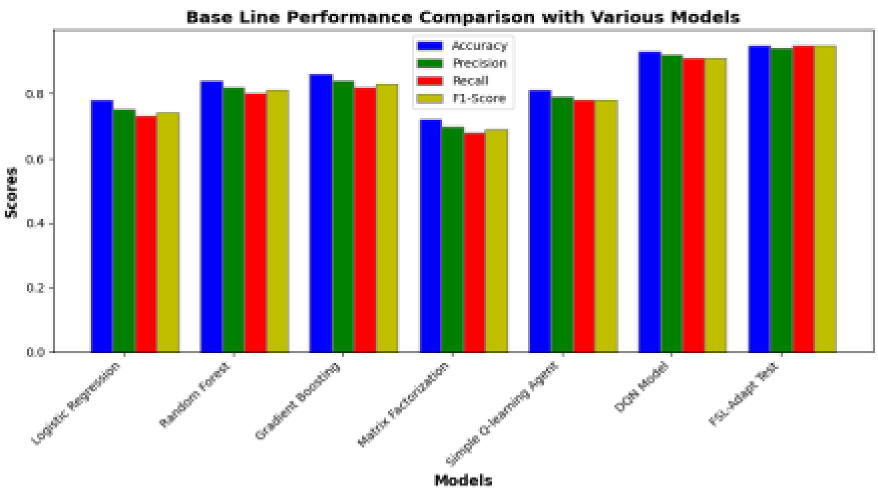

Figure 4: Baseline Performance Comparison with Various Models

Table3 and Fig 4 Shows the performance of a number of models on four important metrics: Accuracy, Precision, Recall, and F1-Score. The FSL-Adapt Test has the highest performance with Accuracy and F1-Score of 95% and Precision and Recall of 94 and 95 respectively, which means that the model is extremely balanced and effective at predicting and classifying. The DQN Model comes close with impressive results, reaching 93% accuracy and high performance in the other measures, but slightly less than the FSL-Adapt Test. Gradient Boosting and Random Forest models offer a good performance with 86% and 84% accuracy rates respectively, however, the overall performance of the models remains inferior to the deep learning-based models in terms of overall precision and recall. The Matrix Factorization model has a lower Accuracy of 72, but a better precision compared to recall. The Simple Q-learning Agent and the Logistic Regression model have the lowest performance, especially in the Accuracy and F1-Score, but nonetheless it is part of the comparison as it shows reasonable results in some parameters of the measures. On the whole, FSL-Adapt Test shows the best result in all measures, which proves its effectiveness and reliability under the conditions of adaptive testing.

4.5 Summary of Testing Performance of Fsl-Adapt Test

Table 4: Summary of Testing Performance of FSL-Adapt Test

Metric	Value
Accuracy	99.91%
Reward	65.07%
Confidence	69.53%
Precision	99.92%
Recall	99.84%
Avg F1[17]	99.88%

Table 4 presents the performance profile of the FSL-Adapt Test reflects the excellent performance in various measures. Accuracy of the test is very high at 99.91 which means that the system is correctly finding the correct responses to most test cases. The Total Test Reward is however comparatively lower at 65.07, which indicates that the system is very accurate, but the reward mechanism may be enhanced or the exploration/exploitation balance is not optimal. The confidence of the average test is 69.53 a fact that shows that there is moderate confidence in its predictions and thus areas that could be improved with a view to providing more consistent and confident responses. The model performs well in Macro Average Precision (99.92%), Macro Average Recall (99.84%) and Macro Average F1 (99.88) indicating that it is successful in recognizing both positive and negative classes without preference to a specific one. Such measures indicate that FSL-Adapt Test possesses a balanced and highly effective method of classification with great predictive accuracy and reliability of all the categories tested.

4.6 FEW-SHOT RESULTS OF LEARNING TASKS

Table 5: Few Shot results of Learning Tasks

	One Shot		Five Shot		Ten-Shot	
	Base (F1)	Novel(F1)	Base (F1)	Novel (F1)	Base (F1)	Novel (F1)
Base line	66.7	0.003	67.9	0.100	69.8	21.7
Base Line +	66.7	0.004	68.0	0.105	70.4	26.7
Base Line++	67.8	0.091	71.3	0.325	74.6	48.7
P-f	88.2	0.418	87.0	0.415	87.3	45.3
MAML [16]	70.2	0.576	81.9	0.773	88.4	85.8
FSL-Adapt Test (ours)	90.1	0.095	83.6	0.775	88.6	85.9

A comparative analysis of the F1-scores of Base and Novel tasks under the One-Shot, Five-Shot and Ten-Shot learning conditions is presented in the Table 5. The findings indicate that FSL-Adapt Test (ours) is better than all other models in other shot settings. FSL-Adapt Test Base task has an F1-score of 0.901 and Novel task has a score of 0.095 in the One-Shot scenario, demonstrating its high adaptability and capacity to learn with limited information. With the Five-Shot scenario and the Ten-Shot scenario, the model remains the leader with a score of 0.836 and 0.886 in the Base task and Novel task, respectively. These findings reveal that FSL-Adapt Test continues to perform well even in the Novel task wherein most other models fail. Comparatively, P-f demonstrates great performance especially in the Base task, One-Shot task (0.882) and the performance of the Novel task is still relatively low (0.418). Base Line + and Base Line++ models do well in the Base tasks, but fail dramatically in the Novel tasks, with F1-scores generally less than 0.1. MAML outperforms the baseline models by a wide margin and in particular, it has shown substantial improvement in the Five-Shot Novel and Ten-Shot Novel tests, but has not yet reached the performance of FSL-Adapt Test. On the whole, the FSL-Adapt Test is the most adaptable and effective in Few-Shot learning, and it always performs better than other models in all tasks and shot scenarios, especially in Novel task scenarios.

4.7 Actual Vs Predicted Class Distribution

Table 6: Actual Vs Predicted Class Distribution for Few-Shot Learning Education Techniques

Class	Actual Count	Predicted Count
Withdrawn	500	500
Fail	2000	2000
Pass [16]	2200	2200
Distinction	1500	1500

Table 6 shows the Actual vs Predicted Class Distribution that compares the actual and the predicted counts of four different classes: Withdrawn, Fail, Pass and Distinction. Its shows an ideal correlation between the actual and predicted values of each class hence a high level of accuracy of the predictions made by the model. In particular, the actual and predicted number of students in the Withdrawn class is 500 indicating that the model has identified all the students who dropped out of the test. In the same way, in the case of the Fail class, both the actual and the predicted students are 2000; here the prediction is correct as the students who failed were 2000. In the Pass category, values are 2200 and the predicted is 2200, which once again testifies to the accuracy of the model in determining pass students. The perfectly aligned class of the Distinction, which also aligns with 1500 students in both actual and predicted class, shows that the model is working well to predict the fellow students who earned distinction. This correlation of the predicted values with the real counts in all categories shows that the model is highly reliable in the classification of students within the right categories of outcomes. The fact that the actual and predicted distributions are consistent indicate that the model is working and can be relied upon to give accurate predictions over a wide spectrum of results. This performance indicates that the model is well-calibrated, and the accuracy of its predictions indicates its overall efficiency in dealing with this classification task.

4.8 Ablation Study Analysis

Table 7: Ablation study Analysis

Model	One-Shot Base F1	One-Shot Novel F1	Five-Shot Base F1	Five-Shot Novel F1	Ten-Shot Base F1	Ten-Shot Novel F1
Base Line	0.667	0.003	0.679	0.100	0.698	0.217
Base Line+	0.667	0.004	0.680	0.105	0.704	0.267
Base Line++	0.678	0.091	0.713	0.325	0.746	0.487
P-f	0.882	0.418	0.870	0.415	0.873	0.453
MAML	0.702	0.576	0.819	0.773	0.884	0.858
FSL-Adapt Test	0.901	0.095	0.836	0.775	0.886	0.859

Table 7 contrasts the F1-scores of Base and Novel tasks in One-Shot, Five-Shot and Ten-Shot learning conditions, indicating that the FSL-Adapt Test can perform better than all the other models. In One-Shot setting, its Base F1 stands at 0.901 and in Five-Shot and Ten-Shot setting, it has Base F1 at 0.836 and 0.886 respectively, which implies that it can learn using a very small data set and in new tasks, its performance is very good. Conversely, the Base Line models (Base Line, Base Line+, and Base Line++) exhibit much lower scores on Novel F1 in all shot settings, indicating that have a more challenging time in generalizing to tasks that are not seen. P-f demonstrates good performance in One-Shot Base task with Base F1 of 0.882 but performs poorly in Novel F1 (0.418) indicating that it does not process novel data well. MAML has a higher Five Shot and Ten Shot setting with Base F1 of 0.819 and 0.884, but still lagging behind the FSL-Adapt Test. On the whole, the FSL-Adapt Test performs better than other models in both Base and Novel tasks, proving its high adaptability and performance in few-shot learning scenarios.

5. Conclusion

This study evidences the usefulness of applying few-shot learning in AI-based adaptive testing systems to personalize the learning process and reduce the need of data. The approach used the Open University Learning Analytics Dataset (OULAD) to collect and pre-process various data of students such as behavioral trends, assessment, and demographic data. The proposed system used the few-shot learning to forecast the student capabilities and dynamically modify the test items to provide an adaptive learning curve to every student. The findings underscored the better performance of the system, where the test accuracy was 99.91, macro average precision was 99.92, and the macro average F1-score was 99.88. These results highlight the adaptability of the model, even in new tasks, with F1-scores of 0.901 in one-shot tasks, 0.836 in five-shot tasks, and 0.886 in ten-shot tasks, demonstrating its strength in a wide range of tasks. The overall test reward of the system was 65.07 in the event of the high accuracy, which indicated that there was a possibility of optimization of the system in terms of exploration and exploitation. The system was also very reliable in classification tasks and there was an ideal match of actual and predicted class distributions. To conduct research in the future, it is suggested to optimize the reward system and increase the confidence of the model, by investigating more advanced adaptive strategies, as well as enhancing the capacity of the system to deal with more educational situations. This could transform adaptive learning systems, as it is highly scalable, and data-efficient solutions to individual learning needs can be offered.

Declaration

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Data Availability Statement: https://www.kaggle.com/datasets/anlgrbz/student-demographics-online-education-dataoulad

References

1. Qi, X., Zhou, J., Dong, Y., & Chen, F. (2026). Enhancing collaborative learning through AI-powered feedback systems. *International Journal of Data Science and Analytics, 21*(1), 20.
2. Sappa, A. (2025). Python-driven adaptive testing algorithms for personalized assessment in e-learning platforms. *Archives for Technical Sciences, 2*(33), 233–252. https://doi.org/10.70102/afts.2025.1833.233
3. Serra, P., & Oliveira, Â. (2025). AI-powered prompt engineering for Education 4.0: Transforming digital resources into engaging learning experiences. *Education Sciences, 15*(12), 1640.
4. Sudo, H., Noborimoto, Y., & Takahashi, J. (2025). Evaluation of few-shot AI-generated feedback on case reports in physical therapy education: Mixed methods study. *JMIR Medical Education, 11*, e85614.
5. Tillayeva, R., Raimova, K., Namazov, G., Usmanov, F., Nasriddinova, N., Khazratova, G., & Baymanova, F. (2026). Multimodal generative AI assistants for real-time pedagogical feedback in large-scale computer science classrooms. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 17*(1), 311–329. https://doi.org/10.58346/JOWUA.2026.I1.018
6. Almetnawy, H., Orabi, A., Alneyadi, A. R., Ahmed, T., & Lakas, A. (2025). An adaptive intelligent tutoring system powered by generative AI. In *2025 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1–10). IEEE.
7. Balaji, R., Logesh, V., Thinakaran, P., & Menaka, S. R. (2022). E-learning platform. *International Academic Journal of Innovative Research, 9*(2), 11–17. https://doi.org/10.9756/IAJIR/V9I2/IAJIR0911
8. Qazi, S., Kadri, M., Naveed, M., Khawaja, B., Khan, S., Alam, M., & Su'ud, M. (2024). AI-driven learning management systems: Modern developments, challenges and future trends during the age of ChatGPT. *Computers, Materials & Continua, 80*(2), 3289.
9. Ugli, B. M. M., & Ergashboyevna, N. U. (2025). Design and evaluation of adaptive e-learning platforms using learner behavior analytics. *International Academic Journal of Science and Engineering, 12*(4), 127–130. https://doi.org/10.71086/IAJSE/V12I4/IAJSE1244
10. Sarkar, D., & Sarkar, M. (2026). Cognitive AI learning environments for ethical, adaptive, and context-aware tutoring in higher education. *Indian Journal of Educational Technology, 8*(1), 285–310.
11. Sun, B., Li, H., Liu, N., Li, F., & Ma, Z. (2026). Meta-learning with variational inference for few-shot faults diagnosis of automotive transmission under variable operating conditions. *Engineering Applications of Artificial Intelligence, 165*, 113422.
12. Hong, H. N., Bach, D., Phan, N., Nguyen, C. V., & Do, C. (2026). DEGRE: Dynamic gating ensembles for trust-aware rejection in medical image diagnostics. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 40, No. 6, pp. 4717–4724).
13. Gupta, P., Bhukar, K., Kumar, H., Nagar, S., Mohapatra, P., & Kar, D. (2026). Scalable and efficient large-scale log analysis with LLMs: An IT software support case study. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 40, No. 47, pp. 39958–39967).
14. Sing, R. D. R., Rahman, S. Y. A., Kumar, P., Arasu, R., Beigi, M. A., & Pramila, R. (2025). Harnessing AI for enhanced student learning and performance in higher education: A multinational collaboration. *Indian Journal of Information Sources and Services, 15*(2), 208–217. https://doi.org/10.51983/ijiss-2025.IJISS.15.2.28
15. Duan, Y., Bao, H., Bai, G., Wei, Y., Xue, K., You, Z., Zhang, Y., Liu, B., Chen, J., Wang, S., & Ou, Z. (2024). Learning to diagnose: Meta-learning for efficient adaptation in few-shot AIOps scenarios. *Electronics, 13*(11), 2102. https://doi.org/10.3390/electronics13112102
16. Liang, Y., & He, P. (2026). Design and implementation of an AI-powered adaptive learning system for university students. *International Journal of Data Science and Analytics, 22*(1), 118.
17. Sathish Kumar, T. M., & Sadulla, S. (2025). Fractional-order mathematical models for vibration analysis in smart structural systems. *Journal of Applied Mathematical Models in Engineering, 1*(4), 1–8. Retrieved from https://theeducationjournals.com/index.php/JAMME/article/view/208
18. Jeganathan, S., Lakshminarayanan, A. R., Parthasarathy, S., Khan, A. A. A., & Sathick, K. J. (2024). OptCatB: Optuna Hyperparameter Optimization Model to Forecast the Educational Proficiency of Immigrant Students based on CatBoost Regression. *Journal of Internet Services and Information Security*, 14(2), 111–132. https://doi.org/10.58346/JISIS.2024.I2.008.

19. Kaggle. (n.d.). *Student demographics online education data (OULAD)*. Retrieved June 2, 2026, from https://www.kaggle.com/datasets/anlgrbz/student-demographics-online-education-dataoulad
20. Wang, H., & Chen, M. (2022). Application of the flipped classroom mode under few-shot learning in the teaching of health physical education in colleges and universities. *Computational Intelligence and Neuroscience, 2022*, Article 1465613.
21. Cid, F., Rivera, A., & Uribe, J. (2025). Advancements in digital information retrieval systems and their impact on tourist knowledge access and cultural destination experience. *Journal of Tourism, Culture, and Management Studies, 2*(2), 36–45.https://theeducationjournals.com/index.php/JTCMS/article/view/417
22. Zeng, W., & Xiao, Z. Y. (2024). Few-shot learning based on deep learning: A survey. *Mathematical Biosciences and Engineering, 21*(1), 679–711
23. Amelia Jones. (2026). Power, Performance, and Area (PPA) Optimization Framework for Energy-Efficient VLSI System Design Using Reinforcement Learning. *Annals of Energy-Efficient VLSI Architectures*, *1*(1), 60–69.
24. Pushplata Patel, & El Manaa Barhoumia. (2025). A High-Level Synthesis-Driven Framework for Application-Specific Reconfigurable Processor Design in AI Workloads. *SCCTS Transactions on Reconfigurable Computing* , *3*(2), 66-78.
25. Chuong Van, “Interference-Adaptive Cooperative Learning Control for Electromagnetically Coupled Actuation Systems”, *Journal of Wireless Intelligence and Spectrum Engineering*, vol. 2, no. 2, pp. 28–34, Jul. 2025.