The Journal of Education Insights Vol 3 Issue 4 2025

Original Research Article

Open Access

Omniscient Pte.Ltd.

Few-Shot Deep Learning Empowers Education Personalization: Technical Principles, Application Pathways, and Challenges

Zhen Jiang[1,*]

School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, 212013, Jiangsu Province, China

***Correspondence to:** Zhen Jiang, School of Computer Science and Communication Engineering, Jiangsu University, No.301, Xuefu Road, 212013, China; Email: jiangz@ujs.edu.cn

Abstract: The demand for personalized education is becoming increasingly urgent. However, artificial intelligence (AI) applications in education, which rely on big data paradigms, face fundamental limitations such as cold starts, and data scarcity. Few-shot learning (FSL), as a "bionic" learning paradigm, exhibits a high degree of theoretical alignment with the personalized needs of educational scenarios. It offers a promising path for building data-efficient personalized education systems. This paper systematically analyzes the core technical principles of FSL, explores its applicability to key educational scenarios, and constructs a technical implementation framework. It provides theoretical and architectural guidance for the application of FSL in the field of personalized education. Moreover, we prospectively identify potential risks and future directions.

Keywords: Few-shot learning; Personalized education; Technical architecture; Educational artificial intelligence

1. Introduction

"The teaching of students according to their aptitude," as the supreme ideal of human education, has run through the entire history of education, from Confucius and Socrates to the present. Its core lies in recognizing and respecting learners' differences in cognitive foundations, learning styles, interests, and motivations, and providing them with personalized teaching methods and resources[1]. AI is currently held in high hopes and regarded as the tool to achieve large-scale "individualized teaching". Deep learning models driven by big data, such as Deep Knowledge Tracing (DKT) and adaptive recommendation systems, aim to construct precise digital profiles for each student by analyzing massive amounts of learning behavior data, thereby realizing personalized education. However, this "big data + big model" paradigm faces two fundamental dilemmas in practical applications:

(1) Cold Start Problem: For a new student joining the system or newly launched course content, the system lacks historical behavior data, making it impossible to immediately conduct effective modeling and recommendations. This results in a prolonged blank period for personalized experiences.

(2) Data Scarcity and Long-Tail Problem: Personalized education involves "niche" demands with extremely

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

small sample sizes. Educational AI models rely heavily on large-scale, expertly annotated data, yet the cost of annotation is prohibitively high.

These dilemmas reveal a paradox: we pursue highly personalized education, but the tools to achieve it depend on large-scale and non-personalized data. Traditional educational AI models cannot, like human teachers, quickly understand the characteristics of a new student or the key teaching points of a new concept by observing just a few samples.

Few-shot learning (FSL)[2] aims to enable deep models to rapidly accomplish new tasks from very few samples. By mimicking the human cognitive ability of “drawing inferences from one instance,” FSL can transition educational AI from relying on “big data + general models” to requiring only “small data” for “personalized, agile models,” truly making the large-scale implementation of teaching according to aptitude possible. This paper deeply explores the theoretical feasibility, technical implementation pathways, and potential challenges of FSL technology empowering educational personalization.

2. Few-Shot Learning: Deconstructing Principles and Paradigms

To rapidly adapt to unseen classes with only a few samples, FSL methods typically learn prior knowledge from a large-scale auxiliary dataset, denoted as D^A. Based on how D^A is utilized, FSL methods can be broadly categorized into two groups:

2.1 Meta-Learning based methods

These methods train an FSC model in an episodic manner to acquire meta-knowledge from D^A. Subsequently, the meta-knowledge is utilized to adapt the model to target tasks. Within this category, metric-based FSC methods aim to learn a metric space with well-separated class boundaries. Prototype Networks[3] and Relation Networks[4] are typical examples. Another branch focuses on optimizing an embedding network to achieve effective initialization[5]. Optimization-based methods focus on optimizing a backbone network so that the model can quickly adapt to novel tasks using only a few labeled data. Others, however, emphasize refining the optimization process itself

2.2 Transfer learning-based methods

These methods pre-train a model on D^A and subsequently fine-tune it to adapt to novel tasks. Unlike meta-learning, this category of methods leverages the entire set of classes simultaneously during pre-training. The acquired prior knowledge is used to optimize the search in the hypothesis space, either by providing a favorable initialization or by guiding the search process[6]. Compared to meta-learning which employs episodic learning, this type of methods often shows a better performance, particularly with deeper neural networks.

2.3 Data augmentation methods

In addition to meta-Learning and transfer learning, another popular approach is to augment the support set, thereby transforming FSL into traditional learning problems. The first is to generate new samples by transforming original support samples. For image classification tasks, new images can be artificially created by flipping, cropping, and scaling. The second strategy focuses on synthesizing new samples using Generative Adversarial Networks (GAN) or autoencoders[7]. Moreover, pseudo-labeling unlabeled or support samples has proven to be an effective strategy[8]. Typically, this is achieved through self-training of FSC classifiers.

3. FSL and Personalized Education: Theoretical Mapping and Application Framework

By leveraging a small number of support samples, FSL enables rapid model training for new tasks, offering a novel paradigm for personalized education. This chapter proposes a systematic methodology to transform personalized education challenges into formal tasks solvable by FSL and designs concrete technical implementation pathways for them.

3.1 Transforming Educational Problems into FSL Tasks

The first and most critical step to successfully applying FSL is the formal definition of the educational problem.

• “Class”: In the “N-Way” setting of FSL, each “Way” corresponds to a class. In education, a “class” could represent a knowledge point (e.g., “the quadratic formula”), a learning style (e.g., “visual learner”), or a proficiency level (e.g., “essay score level B”).

• “Sample”: In the “K-Shot” setting of FSL, each “Shot” refers to a sample. In education, a sample is a data instance that characterizes the “class.” This could include: (1) A student’s response sequence to a set of questions related to the knowledge point (e.g., records of answering five relevant questions). (2) A behavioral segment reflecting a specific learning style (e.g., logs of clicking multiple video resources).

• "Task": An FSL task constitutes a complete "N-Way K-Shot" problem. In education, a task could represent a recommendation or an assessment. For example, "Based on five students' response records (Support Set), diagnose whether a new student has mastered three knowledge points (N=3)."

Through this mapping, complex educational personalization challenges are transformed into few-shot classification, regression, or matching problems—areas where FSL excels.

3.2 Core Application Scenarios

3.2.1 Agile Knowledge State Diagnosis

3.2.1.1 Problem Reframing:

Knowledge state diagnosis is the cornerstone of personalized education[1]. Traditional models require extensive training on response data from numerous students across a fixed set of knowledge points, making them ineffective for new or niche topics. FSL reframes this task: given a new knowledge point X, only a small number K of response samples from other students on X are needed to rapidly construct a diagnostic model. This model can then assess whether any new student has mastered X. For the target knowledge point X, collect response sequences from K students whose mastery status (mastered/unmastered) is known. Each sequence, along with its label, forms the support set.

3.2.1.2 Technical Approach: Implementation via Prototype Networks[3]

(1) An embedding function (e.g., Transformer) maps each student's response sequence into a low-dimensional vector.

(2) Prototype vectors for each class are computed: the mean of all embedding vectors for the "mastered" and "unmastered" classes yields two prototype vectors, c_m and c_un.

(3) For a new student S, their response sequence on knowledge point X is processed by the same embedding function to generate their representation vector V.

(4) The distances (e.g., Euclidean) between V and the two prototype vectors are calculated in the embedding space.

(5) The final prediction is based on a Softmax over distances: The student is assigned to the class whose prototype is closer.

3.2.2 Cold-Start Resource Recommendation

3.2.2.1 Problem Reframing:

The cold-start problem in resource recommendation includes two scenarios: new students and new resources. FSL excels at addressing new student cold-start: when a new student enters the system with only a few initial interactions (e.g., clicking two videos), how can the system immediately recommend suitable resources? This problem can be reframed as a matching task: in the embedding space, find the resources whose representations are most similar to that of this student.

3.2.2.2 Technical Approach: Implementation via Relation Networks[4]

During the meta-training phase, the model is exposed to a large number of "(student, resource)" paired examples.

(1) Two structurally identical, parameter-sharing embedding networks are constructed: one encodes student behavior sequences, and the other encodes resource content features.

(2) Learn a metric space where the distance between embeddings of positive pairs (a student likes a resource) is minimized, while the distance for negative pairs (a student dislikes a resource) is maximized.

(3) For a new student S, their limited initial behavior sequence is fed into the student embedding network to obtain their representation vector Vs.

(4) All candidate resources are processed by the resource embedding network to generate their representation vectors V_R.

(5) Compute the distance between Vs and each V_R, and recommend the Top-N resources with the smallest distances.

3.2.3 Agile Knowledge State Diagnosis

3.2.3.1 Problem Reframing:

Knowledge state diagnosis is the cornerstone of adaptive learning. Traditional models require extensive training on response data from numerous students across a fixed set of knowledge points, making them ineffective for new or niche topics. FSL reframes this task: given a new knowledge point X, only a small number K of response samples from other students on X are needed to rapidly construct a diagnostic model.

This model can then assess whether any new student has mastered *X*.

3.2.3.2 Technical Approach: Implementation via Prototypical Networks[3]

For the target knowledge point X, collect response sequences from K students whose mastery status (mastered/unmastered) is known. Each sequence, along with its label, forms the support set. Typically, K samples are prepared for each of the two classes ("mastered" and "unmastered"), resulting in a 2-Way K-Shot setup.

A shared embedding function (e.g., LSTM or Transformer) maps each student's response sequence into a low-dimensional vector (embedding).

Prototype vectors for each class are computed: the mean of all embedding vectors for the "mastered" and "unmastered" classes yields two prototype vectors, prototype vectors, c_m and c_un.

For a new student S, their response sequence on knowledge point X is processed by the same embedding function to generate their representation vector *Vs*.

The distances (e.g., Euclidean) between *Vs* and the two prototype vectors are calculated in the embedding space.

The final diagnosis is based on a Softmax over distances. The student is assigned to the class whose prototype is closer.

This approach eliminates the need to retrain the model for each new knowledge point. By simply updating the support set, it enables "plug-and-play" agile diagnosis.

3.3 A Conceptual FSL Personalized Education System Architecture

An intelligent education system supporting FSL should be a collaborative and integrated framework. Its conceptual architecture, as illustrated in the diagram below, can be divided into four core layers:

➢ **Data Layer:**

Function: Responsible for data collection, storage, cleaning, and management. This forms the foundation of the system.

Components: Includes a student behavior data warehouse (storing learning records based on xAPI/Caliper standards), an educational resource knowledge base (containing resources such as questions and videos, along with rich metadata and tags), and a crucial task support set management module. This module dynamically constructs the "N-Way K-Shot" tasks (episodes) required for FSL from raw data, which is a key differentiator between FSL systems and traditional systems.

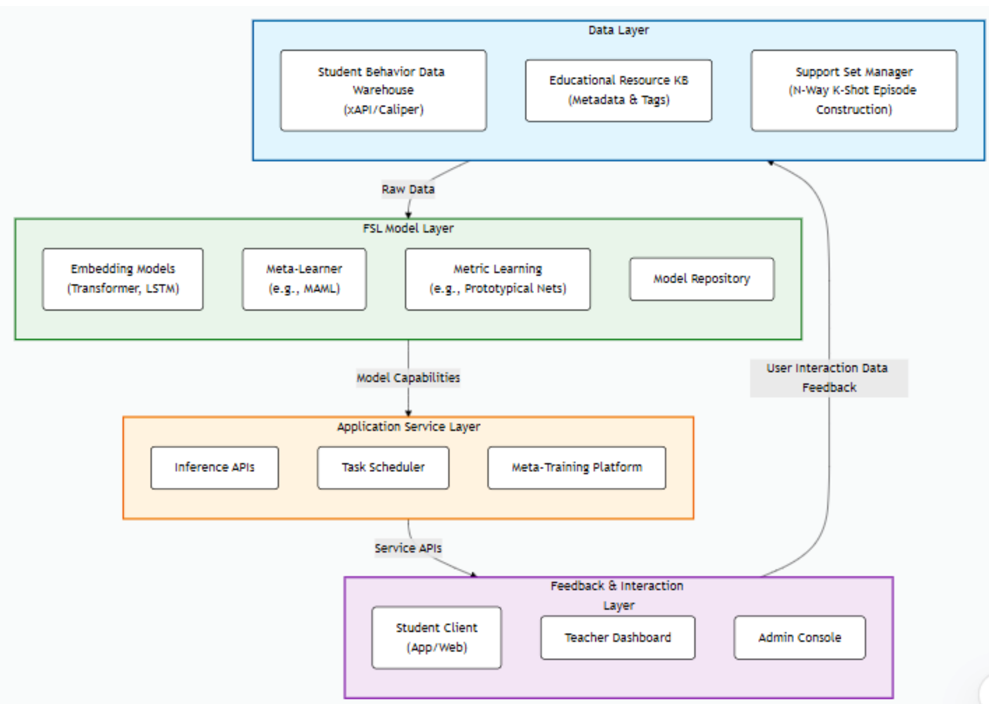

Figure 1: Personalized Education System Architecture Based on Few-Shot Learning

➢ **FSL Model Layer:**

Function: Serves as the core of the system, hosting various FSL algorithm models.

Components: Includes embedding models (e.g., Transformer, used for generating vector representations of students and resources), meta-learners (e.g., implementations of MAML), metric learning models (e.g., Prototypical Networks, Siamese Networks), and more. A model repository manages different versions of models tailored for various tasks, supporting A/B testing and rolling updates.

➢ **Application Service Layer:**

Function: Encapsulates model capabilities into callable services for upper-layer applications.

Components: Provides a range of inference APIs. A task scheduler organizes vast meta-tasks during the meta-training phase and specific support sets during the inference phase. The meta-training platform operates as an offline or near-online system, continuously leveraging new data for meta-training models and optimizing initial model parameters.

➢ **Feedback & Interaction Layer:**

Function: Directly interfaces with end-users (students, teachers, administrators) through interfaces and applications.

Components: Includes student-facing applications (e.g., apps, web platforms for receiving personalized recommendations and exercises), teacher dashboards (for viewing learning alerts and class knowledge mastery maps), and administrator consoles (for monitoring system operations and managing support set quality). New user interaction data is fed back into the data layer, forming a closed loop that enables the system to continuously evolve and improve.

3.4 Case Scenarios and Discussion

Few-Shot Learning (FSL) enhances personalized education through various classic applications. Typical scenarios include the application in a new subject or for a new student, providing personalized support for special learning needs. Additionally, FSL facilitates the launch of new courses by analyzing minimal data from early learners, allowing for instant personalization without extensive historical information. These scenarios demonstrate FSL's potential to revolutionize educational experiences by making them more responsive and individualized. FSL offers numerous application scenarios for personalizing education. Next, we will choose two classic applications for a more in-depth discussion.

3.4.1 Application in a New Subject or for a New Student

The cold-start problem is a major challenge in educational AI. When new subjects are added to a Learning Management System (LMS) or new students enroll, traditional data-driven models struggle due to a lack of historical interaction data. Few-Shot Learning (FSL) offers a transformative solution by enabling AI systems to function effectively with minimal data.

Consider an online platform introducing a module on "Python for Data Science." Without any student performance records, traditional adaptive systems would take weeks to gather sufficient data for personalization. In contrast, an FSL-powered system can utilize its meta-training experience, having learned from various programming courses (e.g., Java, C++). It generalizes how to map initial student attempts onto predictions of mastery in related concepts. For the Python course, the instructional designer provides a small "support set" for each new concept, such as "using Pandas for data filtering." When a new student, Alice, starts her exercises, the FSL model compares her problem-solving patterns against the support set. Based on this comparison, it can recommend personalized next steps—like hints or foundational resources—immediately bypassing the cold-start period.

The primary advantages of FSL are agility and scalability. As educational content evolves, FSL allows AI tutors to adapt quickly. The requirement for a small support set is manageable for educators compared to generating large datasets. However, challenges include ensuring the quality and representativeness of the support set; biased examples lead to flawed recommendations. Additionally, if the new subject is too different from those encountered during meta-training, the model's effectiveness diminishes—a phenomenon referred to as the "domain gap."

3.4.2 Providing Personalized Support for Special Learning Needs

FSL's application extends to addressing diverse student needs, particularly those with learning differences like dyslexia or ADHD, which don't conform to average profiles. Collecting large datasets for these specialized groups is challenging, making FSL especially valuable in data-scarce scenarios. Traditional models may struggle to accommodate the unique error patterns of dyslexic students. An FSL system can personalize support using minimal data from individual interactions. The meta-training utilizes a variety of reading comprehension tasks to build a comprehensive feature space.

For instance, after completing just a few passages, a student named Ben, who has dyslexia, provides a support set reflecting his reading speed and error patterns. The FSL model uses this information to fine-tune its understanding of Ben's profile and personalize further interactions. When Ben approaches a new passage, the system might:

➢ Dynamic Text Presentation: Adjust typography and spacing to reduce errors.

➢ Targeted Vocabulary Pre-teaching: Highlight similar words that he previously struggled with.

➢ Adaptive Questioning: Frame questions to focus on key ideas, lessening linguistic load.

This application emphasizes FSL's potential to create inclusive education by adapting to individual learners' needs. Ethical considerations include ensuring data privacy and avoiding algorithmic stereotyping. Collaboration among AI developers, special education experts, and learners is essential to ensure technology empowers rather than restricts.

In conclusion, these scenarios illustrate that FSL represents a significant shift in educational AI, enabling it to operate effectively under realistic conditions where data scarcity and diverse needs prevail, thus making truly personalized, scalable education achievable.

4. Potential Risks, Ethical Dilemmas, and Challenges

4.1 Technical Limitations

FSL models are built upon a series of idealized assumptions that are easily broken in real-world educational environments.

4.1.1 Sensitivity to Support Set Quality:

FSL models learn from a very small number of examples. This means a single mislabeled example, a biased sample, or a low-quality instance is enough to steer the model in the wrong direction. For example, in essay scoring, if a "model essay" in the support set is actually poorly structured, any subsequent essay with a similar style is likely to be incorrectly assigned a high score.

4.1.2 Inadequate Out-of-Distribution (OOD) Generalization:

FSL models excel at tasks similar to those seen during meta-training but suffer a significant performance drop when encountering novel, out-of-distribution tasks. The complex and ever-changing nature of personalized education amplifies this challenge.

4.1.3 Integration Costs with Existing Education Systems:

Seamlessly integrating the FSL system with existing workflows such as Learning Management Systems (LMS) and Student Information Systems (SIS) requires significant customization and API integration work. For many educational institutions, their IT infrastructure and budget may struggle to support such a complex deployment of cutting-edge technology.

4.2 Data and Ethical Challenges

4.2.1 Data Privacy and Security

The training of FSL models often requires aggregating data from multiple sources. Support sets may contain highly personalized student data (such as specific students' essays), which face the risk of leakage during invocation and transmission.

4.2.2 Algorithmic Bias and Perpetuation

Historical biases present in the training data (e.g., urban-rural education resource gaps, preferences for specific problem-solving styles) can be captured and amplified by the FSL model. This may lead the system to make poorer predictions and recommendations for students from under-resourced areas or those with atypical learning styles, thereby exacerbating educational inequality.

4.2.3 Reshaping the Teacher's Role and the Human-AI Collaboration Dilemma [9]:

FSL systems are positioned as "teacher's assistants," but how should the boundary between human and machine responsibilities be delineated? Do teachers possess sufficient digital literacy to comprehend, question, or even override the system's recommendations? This is not merely a technical issue, but a complex one involving pedagogy, psychology, and organizational management.

5. Conclusion

Few-shot learning (FSL) is inherently well-suited to address educational personalization needs, offering solutions to fundamental AI challenges like the cold-start problem and data scarcity. It represents a new paradigm for intelligent education based on human-AI collaboration. However, implementing FSL is a complex endeavor requiring a holistic architecture covering data, models, services, and feedback. This process faces technical limitations, ethical risks, and integration barriers. Moving beyond a "tech-first" approach is essential—educational equity, privacy protection, and algorithmic transparency must be central to FSL-driven personalization.

The ultimate goal of FSL should be to serve as a "super assistant" that rapidly interprets teaching intent and responds to individual needs. By alleviating teachers' repetitive workloads, it enables them to focus on higher-order educational responsibilities—fostering emotional well-being, inspiring creativity, and shaping values.

References

[1] Bernacki, Matthew L., Meghan J. Greene, and Nikki G. Lobczowski. "A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose (s)?" Educational Psychology Review 33.4 (2021): 1675-1715.

[2] Song Y, Wang T, Cai P, et al. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities[J]. ACM Computing Surveys, 2023, 55(13s): 1-40.

[3] Snell, J., Swersky, K., & Zemel, R. S. Prototypical networks for few-shot learning, in: Advances in Neural Information Processing Systems, MIT Press, 2017, pp. 4077–4087.

[4] Huang, Y., Hao, H., Ge, W., Cao, Y., Wu, M., Zhang, C., & Guo, J. Relation fusion propagation network for transductive few-shot learning, Pattern Recognition 151 (2024) 110367.

[5] Jiang, Z., Feng, Z., & Niu, B. Prototype-Neighbor Networks with Task-Specific Enhanced Meta-learning for Few-Shot Classification, Neural Networks (2025) 107761.

[6] Sun, Z., Wang, M., Ran, X., & Guo, P. Feature reconstruction-guided transductive few-shot learning with distribution statistics optimization, Expert Systems with Applications 270 (C) (2025) 1–14.

[7] Seo, J., Kang, J. S., & Park, G. M. (2023). Lfs-gan: Lifelong few-shot image generation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 11356-11366).

[8] Chen J, Deng S, Teng D, et al. APPN: An Attention-based Pseudo-label Propagation Network for few-shot learning with noisy labels[J]. Neurocomputing, 2024, 602: 128212.

[9] Chen, Feng. "Human-AI cooperation in education: human in loop and teaching as leadership." Journal of Educational Technology and Innovation 2.1 (2022).