BULLETIN OF EDUCATION AND PSYCHOLOGY
Research Article | Volume 5 | Issue 5 | Mar 2025
Received: 19 Jan 2025 | Revised: 23 Jan 2025
Accepted: 14 Feb 2025 | Published Online: 02 Mar 2025

Identifications of Active Explorers and Passive Learners Among Students: Gaussian Mixture Model-Based Approach

Zhiqiang Zhao [1, *], Ping Ren[2]

[1] Beijing PhD Village Education Technology Co., Ltd; Beijing 100871, China

[2] Chengdu Ding Yi Education Consulting Co., Ltd, Chengdu 610023, China

[*]Corresponding Author, Email: 202128030258@mail.bnu.edu.cn

Abstract: In the realm of education research, the distinction between active explorers and passive learners among students plays a pivotal role in understanding and enhancing learning outcomes. By identifying and characterizing these two distinct groups, educators can tailor instructional strategies to better cater to individual learning preferences, ultimately fostering a more engaging and effective educational experience. However, existing methodologies for discerning between active explorers and passive learners face significant challenges, primarily stemming from the complexity and variability of student behaviors. In light of this, this paper proposes a novel Gaussian Mixture Model-based approach to accurately classify students into these two categories. The innovative aspect of this work lies in its ability to effectively capture the nuances of student engagement and learning styles, thereby providing a more nuanced understanding of student dynamics in educational settings.

Keywords: *Active Explorers; Passive Learners; Learning Outcomes; Instructional Strategies; Student Engagement*

1. Introduction

The field of Active Explorers and Passive Learners Among Students focuses on understanding and comparing the cognitive processes and learning behaviors of individuals who actively seek out information and engage in hands-on exploration versus those who passively receive and absorb information. Current challenges in this field include accurately measuring and assessing levels of student engagement, designing effective interventions to encourage active exploration, and addressing potential biases in existing educational systems that may favor passive learning. Additionally, there is a need for further research to explore the long-term impacts of active exploration versus passive learning on academic achievement and real-world problem-solving skills. Overall, this area of study provides valuable insights into optimizing educational strategies and promoting more effective learning environments for students of all ages.

To this end, research on the distinction between active explorers and passive learners among students has advanced significantly, with studies examining various factors influencing learning styles and strategies. Current research has shed light on the importance of fostering a dynamic and engaging learning environment to promote active exploration and critical thinking skills. The literature review explores various aspects of active and passive behaviors in different academic contexts. Demulder et al. (2024) investigated the impact of study choice processes on academic success in higher education [1]. Cheng et al. (2023) examined how social network usage relates to academic performance among high school students [2]. Emerson (2023) studied the effects of active vs. passive engagement with older adults on ageism among undergraduate students [3]. Liu et al. (2024) explored the longitudinal associations between TikTok use and anxiety among Chinese emerging adults, highlighting the differences between active and passive use [4]. Additionally, Liu and Zhang (2025) focused on the effects of aerobic exercise on executive functions among active, passive, and non-procrastinating college students [5]. Ardhy and Hartiningsih (2023) optimized academic skills through ESP, emphasizing active and passive voices in International Relations students [6]. Mariappan (2023) discussed the empowerment of passive learners through scenario-based learning in the teaching and learning process [7]. Yunzal et al. (2024) delved into active learning strategies in science among senior high school STEM learners and teachers [8]. Lastly, Sharma and Jangra (2024) examined the effects of active, passive, and nonsmoking on aerobic capacity among young collegiates [9]. These studies collectively contribute valuable insights into the relationships between active and passive behaviors, learning processes, and academic outcomes in various educational contexts. The utilization of Gaussian Mixture Model (GMM) is essential in this research landscape due to its capability to effectively model complex data distributions, particularly in cases involving multiple sources of variability. GMM's flexibility allows for the identification of underlying patterns within diverse datasets characterized by active and passive behaviors, enabling a comprehensive understanding of the nuanced relationships between such behaviors and academic outcomes across different educational settings.

Specifically, Gaussian Mixture Model plays a crucial role in distinguishing between active explorers and passive learners among students. By utilizing its clustering capabilities, GMM can effectively identify patterns in student behavior to differentiate those who actively seek out knowledge from those who passively absorb information. The literature review discusses various applications of Gaussian Mixture Models (GMM) in different domains. Zivkovic [10] introduced an Adaptive GMM for background subtraction, while An et al. [11] utilized Ensemble Unsupervised Autoencoders and GMM for cyberattack detection. In a different context, Zhu et al. [12] proposed a Bayesian GMM for Earthquake Phase Association, demonstrating effective associations. Nguyen et al. [13] addressed the challenge of detecting Unknown DDoS Attacks using deep learning and GMM successfully. Moreover, Zhang et al. [14] developed a GMM approach for clustering with incomplete data, showcasing improved clustering performance. Rasmussen [15] introduced the Infinite Gaussian Mixture Model with implications in neural systems. Additionally, Cao et al. [16] tackled Eye Blink Artifact Detection using a GMM, enhancing EEG signal processing. Finally, Yan et al. [17] proposed a semantic-enhanced GMM for Unknown Intent

Detection in dialogue systems, achieving promising results. However, current limitations include scalability issues with large datasets, potential overfitting in complex models, and the need for further research on GMM's generalization across diverse datasets.

To overcome those limitations, this study aims to develop a more precise method to categorize students as active explorers or passive learners in the field of education research. The primary goal is to enhance educators' ability to customize teaching methods according to individual learning preferences, ultimately improving learning outcomes. The proposed approach hinges on a novel Gaussian Mixture Model-based technique, designed to address the challenges presented by the intricate and varied nature of student behaviors. By utilizing this innovative method, researchers can accurately differentiate between active explorers and passive learners, capturing the subtleties of student engagement and learning styles with unparalleled detail. This approach promises to provide a deeper insight into student dynamics within educational environments, paving the way for more tailored and effective instructional strategies to create a more engaging and enriching educational experience.

In the realm of education research, the distinction between active explorers and passive learners among students plays a pivotal role in understanding and enhancing learning outcomes. By identifying and characterizing these two distinct groups, educators can tailor instructional strategies to better cater to individual learning preferences, ultimately fostering a more engaging and effective educational experience. However, existing methodologies for discerning between active explorers and passive learners face significant challenges, primarily stemming from the complexity and variability of student behaviors. In light of this, this paper proposes a novel Gaussian Mixture Model-based approach to accurately classify students into these two categories. The innovative aspect of this work lies in its ability to effectively capture the nuances of student engagement and learning styles, thereby providing a more nuanced understanding of student dynamics in educational settings. Section 2 of the study describes the problem statement, Section 3 introduces the proposed method, Section 4 presents a case study, Section 5 analyzes the results, Section 6 provides a discussion, and Section 7 offers a comprehensive summary of the research findings.

2. Background

2.1 Active Explorers and Passive Learners Among Students

In the context of educational psychology and pedagogical sciences, students can often be delineated into two distinct categories: Active Explorers and Passive Learners. These classifications hinge upon the cognitive engagement and dynamic participation each student exhibits in the learning process. Below, we delve into a more granular and formulaic exploration of these archetypes, each representing a spectrum of learning philosophies and approaches.

Active Explorers are characterized by their proactive engagement with learning materials, self-driven inquiries, and cognitive efforts to traverse beyond the conventional curriculum. This type of student doesn't just absorb information; they construct knowledge interactively. The propensity of a student to be an Active Explorer can be encapsulated by their Exploratory Engagement Index

(EEI). This index accounts for variables such as cognitive curiosity C_c, frequency of inquiry-initiated actions F_i, and the diversity of resource utilization D_r.

$$EEI = \alpha_1 C_c + \alpha_2 F_i + \alpha_3 D_r \tag{1}$$

Where α_1, α_2, and α_3 are weights conditioned by empirical pedagogical studies. Furthermore, the exploratory actions of these students are often reinforced through feedback loops and self-regulated learning mechanisms, quantified by a Feedback Assimilation Function (FAF), which integrates the quantity and quality of feedback, Q_f.

$$FAF = \beta_1 Q_f \tag{2}$$

Coupled together, the integration of exploratory behavior and feedback assimilation can model the Knowledge Retention and Expansion Rate (KRER) for Active Explorers, expressed as:

$$KRER = EEI \times FAF \tag{3}$$

Conversely, Passive Learners tend to exhibit a more reactive approach to education, where the learning process is predominantly guided by direct instruction from educators, with minimal self-initiated exploration. These students often rely on established syllabi and curricular frameworks. The inclination towards passive learning can be assessed through a Passive Engagement Index (PEI), which evaluates factors such as didactic reliance D_r, structured learning follow-through S_f, and rote memorization tendencies R_m.

$$PEI = \gamma_1 D_r + \gamma_2 S_f + \gamma_3 R_m \tag{4}$$

In passive learners, the reception of knowledge is heavily dependent on the Teacher-Driven Retention Coefficient ($TDRC$), which considers the efficacy of educator prompts E_p.

$$TDRC = \delta_1 E_p \tag{5}$$

The synergy of passive engagement and teacher-driven instruction can be synthesized to predict the Rate of Passive Knowledge Acquisition (RPKA):

$$RPKA = PEI \times TDRC \tag{6}$$

When comparing Active Explorers and Passive Learners, one must consider the Cognitive Adaptability Coefficient (CAC), which evaluates the students' ability to adapt and generalize learned concepts across various contexts. CAC plays a pivotal role in determining the ultimate efficacy of the educational approach.

$$CAC = KRER - RPKA \tag{7}$$

A positive CAC suggests a dominance of active exploration characteristics, contributing to a more versatile, adaptable learning experience. Conversely, a negative or lesser CAC often signals a predilection towards passive learning paradigms. In summary, understanding and quantifying the traits of Active Explorers and Passive Learners through indices and coefficients not only aids in tailoring pedagogical methods but also enhances the personalization of education, pushing students towards the zenith of their cognitive potentials. Through rigorous analysis and computational modeling, educators can better scaffold learning experiences that either nurture exploration or support structured educational delivery, contingent on the student's innate learning proclivities.

2.2 Methodologies & Limitations

In the realm of educational psychology and pedagogy, methodological approaches aimed at understanding Active Explorers and Passive Learners among students have become increasingly sophisticated. These methodologies often employ computational models and quantitative analyses to dissect the nuances of student engagement and learning trajectories. Below is a comprehensive outline of the most prevalent methods used in this field, alongside their limitations. One common methodological approach is the development of predictive models to quantify student behavior and engagement in learning activities. For Active Explorers, a typical model involves the Exploratory Learning Function (ELF), which incorporates variables such as intrinsic motivation M_i, external motivation M_e, and collaborative interactions C_i.

$$ELF = \kappa_1 M_i + \kappa_2 M_e + \kappa_3 C_i \tag{8}$$

The resulting ELF can, however, be limited by its dependency on accurately measuring each input variable, particularly intrinsic motivation, which is inherently subjective. Augmenting this, researchers often employ network analysis to map the Learning Interaction Network (LIN), whereby nodes represent students and edges depict interactions, characterized by interaction frequency I_f and interaction quality I_q.

$$LIN = \sum (I_f \times I_q) \tag{9}$$

A significant limitation of the LIN model is its sensitivity to data granularity—coarse interaction data may obscure important relational dynamics among learners. For Passive Learners, methodologies typically involve the calculation of a Curricular Dependency Index (CDI), a metric assessing the dependence on structured pedagogical inputs, quantified by task completion rate T_c and educator-guided interventions G_i.

$$CDI = \lambda_1 T_c + \lambda_2 G_i \tag{10}$$

While CDI provides insights into passive learning behaviors, it often fails to capture the nuances of students' cognitive engagement beyond completing assignments. The assessment of learning outcomes additionally employs statistical models such as the Expected Learning Outcome Metric ($ELOM$), which projects learning outcome probabilities based on previous academic performance P_a and engagement metrics E_m.

$$ELOM = \mu_1 P_a + \mu_2 E_m \tag{11}$$

The limitations of $ELOM$ include potential biases introduced by historical academic data, which may not fully reflect current learning environments or student growth potential. To understand the holistic impact of pedagogical strategies, researchers utilize the Cumulative Learning Impact

Analysis ($CLIA$). This analysis calculates the aggregate effect of educational interventions over time, factoring in cumulative cognitive load L_c and learning retention over time R_t.

$$CLIA = \nu_1 L_c + \nu_2 R_t \tag{12}$$

An overarching challenge with $CLIA$ is accounting for longitudinal educational changes and diverse cognitive development rates among students. Finally, the Adaptation and Flexibility Index (AFI) serves to evaluate students' responsiveness to dynamic learning environments, characterized by adaptability in learning approaches A_l and flexibility in problem-solving strategies F_s .

$$AFI = \phi_1 A_l + \phi_2 F_s \tag{13}$$

Although AFI is valuable for assessing adaptability, its effectiveness is constrained by the need for precise, context-specific measurement of adaptive behaviors across varying academic settings. In conclusion, while these quantitative models provide a framework for understanding student engagement, the implementation of these methodologies often encounters obstacles such as measurement accuracy, contextual relevance, and data integrity. These limitations highlight the need for ongoing refinement and the integration of qualitative assessments to complement quantitative insights, thereby bolstering the robustness of educational research and instructional design.

3. The proposed method

3.1 Gaussian Mixture Model

In the field of statistical modeling and pattern recognition, Gaussian Mixture Models (GMMs) have emerged as a highly effective technique for complex data analysis which involves probabilistic modeling approaches. GMMs belong to the category of model-based clustering methods and are particularly valuable due to their capacity to represent the underlying structure of data through a combination of multiple Gaussian distributions. The underlying assumption is that the data set is generated by a mixture of several Gaussian distributions, each characterized by its own mean and covariance. A Gaussian mixture model can be mathematically expressed using the following formulation, where the probability density function of a data point x is represented as a weighted sum of K Gaussian components:

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x \mid \mu_k, \Sigma_k) \tag{14}$$

Here, each component k in the mixture is a Gaussian distribution $\mathcal{N}(x \mid \mu_k, \Sigma_k)$ with its own mean vector μ_k and covariance matrix Σ_k , while π_k is the mixing coefficient representing the prior probability of selecting the k -th Gaussian component. The mixing coefficients must satisfy the constraint:

$$\sum_{k=1}^{K} \pi_k = 1 \text{ and } \pi_k \geq 0 \text{ for all } k. \tag{15}$$

The likelihood of the entire data set $X = \{x_1, x_2, \ldots, x_N\}$ given the parameters of the model can be expressed as the product of individual data point probabilities:

$$\mathcal{L}(\Theta; X) = \prod_{i=1}^{N} p(x_i) \tag{16}$$

where Θ represents the set of all parameters in the model, encompassing the means, covariances, and mixing coefficients for all components:

$$\Theta = \pi_k, \mu_k, \Sigma_k \text{ for } k = 1,2, \ldots, K. \tag{17}$$

To find the optimal parameters, the objective is to maximize the log-likelihood function, typically solved using the Expectation-Maximization (EM) algorithm. The log-likelihood is given by:

$$\log\mathcal{L}(\Theta; X) = \sum_{i=1}^{N} \log\left(\sum_{k=1}^{K} \pi_k \mathcal{N}(\, x_i \mid \mu_k, \Sigma_k \,)\right) \tag{18}$$

The EM algorithm alternates between two main steps: the Expectation (E) step, where it calculates the expected value of the latent variables given the current estimate of parameters, and the Maximization (M) step, where it updates the parameters to maximize the expected log-likelihood found in the E step. Specifically, the E step calculates the responsibility γ_{ik} of Gaussian component k for data point x_i :

$$\gamma_{ik} = \frac{\pi_k \mathcal{N}(\, x_i \mid \mu_k, \Sigma_k \,)}{\sum_{j=1}^{K} \pi_j \mathcal{N}\left(\, x_i \mid \mu_j, \Sigma_j \,\right)} \tag{19}$$

In the M step, the parameters are updated as follows, using the calculated responsibilities:

$$\mu_k = \frac{\sum_{i=1}^{N} \gamma_{ik} x_i}{\sum_{i=1}^{N} \gamma_{ik}} \tag{20}$$

$$\Sigma_k = \frac{\sum_{i=1}^{N} \gamma_{ik} (x_i - \mu_k)(x_i - \mu_k)^\top}{\sum_{i=1}^{N} \gamma_{ik}} \tag{21}$$

$$\pi_k = \frac{\sum_{i=1}^{N} \gamma_{ik}}{N} \tag{22}$$

One of the key advantages of GMMs is their flexibility in modeling data distributions that are not strictly unimodal. This flexibility allows for capturing the complexity of real-world data sets that may exhibit multimodal characteristics. However, despite their versatility, GMMs assume that the components are Gaussian, which might not always align with the true data distribution, potentially leading to suboptimal representation if the Gaussian assumption is strongly violated. Additionally, the EM algorithm may converge to local optima, necessitating multiple runs with different initializations to achieve a more global optimum solution. Overall, Gaussian Mixture Models serve as a powerful, probabilistic framework that can provide insightful delineations of data structures, thus having a wide variety of applications across fields such as pattern recognition, machine learning, and bioinformatics.

3.2 The Proposed Framework

The integration of Gaussian Mixture Models (GMMs) into the study of Active Explorers and Passive Learners among students offers a nuanced approach to categorizing and analyzing learners' behaviors and engagement patterns. The distinction between these two learner archetypes, characterized as Active Explorers and Passive Learners, can be mathematically modeled using GMMs, with each learner type representing different distributions of learning behaviors. Active Explorers' behavior can be encapsulated using the Exploratory Engagement Index (EEI), as represented by:

$$EEI = \alpha_1 C_c + \alpha_2 F_i + \alpha_3 D_r \tag{23}$$

This EEI can be viewed as a random variable that follows a Gaussian distribution, where its mean and variance reflect the central tendency and spread of active engagement behaviors among students. Similarly, the Passive Engagement Index (PEI) for Passive Learners can be expressed as:

$$PEI = \gamma_1 D_r + \gamma_2 S_f + \gamma_3 R_m \tag{24}$$

Each of these indices can be treated as components of a mixture model, where the populations of Active Explorers and Passive Learners are seen as clusters of Gaussian distributions. In GMMs, the overall probability density function of learning engagement characteristics can thus be formulated as a weighted sum of the two subpopulations, represented by their respective Gaussian distributions:

$$p(x) = \pi_1 \mathcal{N}(x \mid \mu_1, \Sigma_1) + \pi_2 \mathcal{N}(x \mid \mu_2, \Sigma_2) \tag{25}$$

Here, π_1 and π_2 are the mixing coefficients for Active Explorers and Passive Learners, respectively, and $\mathcal{N}(x \mid \mu_k, \Sigma_k)$ are the Gaussian distributions for each type. The constraint that the coefficients sum to 1 can be stated as:

$$\pi_1 + \pi_2 = 1 \text{ and } \pi_k \geq 0 \text{ for } k = 1,2. \tag{26}$$

The likelihood of observing the data set $X = \{EEI, PEI\}$ given the parameters can be expressed as:

$$\mathcal{L}(\Theta; X) = \prod_{i=1}^{N} p(x_i), \tag{27}$$

where Θ denotes the full set of parameters, which includes the means, covariances, and mixing coefficients:

$$\Theta = \pi_k, \mu_k, \Sigma_k \text{ for } k = 1,2. \tag{28}$$

Maximizing the log-likelihood allows us to effectively fit the GMM to the educational data:

$$\log\mathcal{L}(\Theta; X) = \sum_{i=1}^{N} \log\left(\sum_{k=1}^{2} \pi_k \mathcal{N}(x_i \mid \mu_k, \Sigma_k)\right). \tag{29}$$

Using the Expectation-Maximization (EM) algorithm strengthens the fit of our model by calculating responsibilities to assess how each data point relates to each learner type, expressed as:

$$\gamma_{ik} = \frac{\pi_k \mathcal{N}(x_i \mid \mu_k, \Sigma_k)}{\sum_{j=1}^{2} \pi_j \mathcal{N}(x_i \mid \mu_j, \Sigma_j)}. \tag{30}$$

Following this, the parameters for the means and variances of the learning characteristics can be updated through:

$$\mu_k = \frac{\sum_{i=1}^{N} \gamma_{ik} x_i}{\sum_{i=1}^{N} \gamma_{ik}}, \tag{31}$$

$$\Sigma_k = \frac{\sum_{i=1}^{N} \gamma_{ik} (x_i - \mu_k)(x_i - \mu_k)^\top}{\sum_{i=1}^{N} \gamma_{ik}}, \tag{32}$$

and for the mixing coefficients:

$$\pi_k = \frac{\sum_{i=1}^{N} \gamma_{ik}}{N}. \tag{33}$$

Through this modeling approach, the concept of Cognitive Adaptability Coefficient (CAC) can also be incorporated, capturing the adaptability of students between these categories:

$$CAC = KRER - RPKA, \tag{34}$$

where $KRER$ represents the Knowledge Retention and Expansion Rate for Active Explorers and $RPKA$ is the Rate of Passive Knowledge Acquisition for Passive Learners. By using GMMs to delineate between these learner types, educational psychologists can better understand the underlying structures that govern learning behaviors, leading to more tailored educational strategies that cater to diverse learner needs. This convergence not only enriches our understanding of educational dynamics but also leverages statistical techniques to draw meaningful insights from complex student data.

3.3 Flowchart

This paper introduces a novel approach for enhancing student engagement in educational environments through the Gaussian Mixture Model-based Active Explorers and Passive Learners framework. The proposed method categorizes students into two distinct groups: active explorers, who demonstrate curiosity and seek out new knowledge, and passive learners, who tend to absorb information without actively engaging with their surroundings. By utilizing a Gaussian Mixture Model, the method effectively identifies and models the behavioral patterns of both groups, allowing for tailored interventions that promote active participation among passive learners. The approach emphasizes the importance of adaptive learning strategies, which can be implemented through targeted mentorship, resource allocation, and personalized learning paths to encourage exploration and self-directed learning. Furthermore, the methodology leverages data-driven insights to enhance the overall educational experience by aligning instructional techniques with individual student characteristics, thereby fostering a more dynamic learning environment. In summary, the framework proposed in this paper aims to cultivate a more interactive and responsive educational atmosphere that recognizes and addresses the diverse learning needs of students, as illustrated in Figure 1.

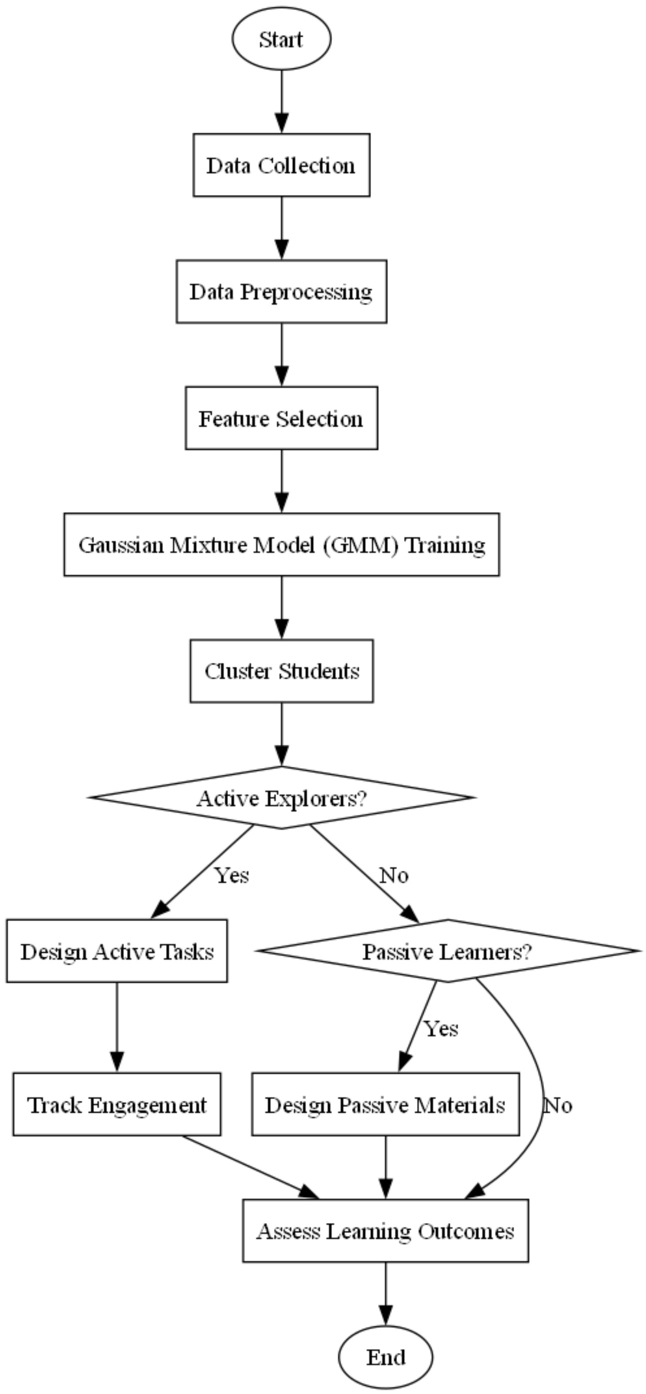

Figure 1: Flowchart of the proposed Gaussian Mixture Model-based Active Explorers and Passive Learners Among Students

4. Case Study

4.1 Problem Statement

In this case, we aim to explore the differences in learning dynamics between Active Explorers and Passive Learners among students, utilizing a mathematical model that captures their interactive behaviors and learning outcomes. We define Active Explorers as students who actively engage with their environment and seek out new information, while Passive Learners typically absorb information presented to them without seeking additional input. To formalize our model, we consider the following parameters: let E_a represent the engagement level of Active Explorers, which is influenced by factors such as curiosity and risk-taking, and let E_p represent the engagement level of Passive Learners, governed by receptivity and the tendency to conform to existing knowledge structures. The evolution of learning outcomes is illustrated through two nonlinear differential equations:

$$\frac{dA}{dt} = k_1 E_a^2 (1 - A) - k_2 A \tag{35}$$

where A denotes the proportion of students achieving high competency, k_1 and k_2 are constants determining the impact of engagement on learning effectiveness and the decay rate of knowledge, respectively. Active Explorers will accelerate their learning as their engagement increases, but there is a diminishing return effect as more students reach a high level of competency. Conversely, we define the learning dynamics of Passive Learners with the equation:

$$\frac{dP}{dt} = k_3 E_p (1 - P^2) - k_4 P \tag{36}$$

In this equation, P indicates the proportion of Passive Learners achieving satisfactory understanding, with k_3 denoting the influence of passive engagement, which hinders their learning potential as they become less adaptive in changing their beliefs. The term $(1 - P^2)$ suggests that as students' understanding improves, the progress relative to the number of Passive Learners begins to plateau, emphasizing the pitfalls of passive learning strategies. The interaction between Active Explorers and Passive Learners is explored through their respective learning rates, which may influence each other's outcomes. We introduce the interaction term I defined by:

$$I = \beta A P \tag{37}$$

where β is defined as an interaction coefficient capturing the influence of Active Explorers on Passive Learners. Thus, the total effective engagement can be expressed as:

$$E_{\text{total}} = E_a + \alpha I \tag{38}$$

where α quantifies the extent to which Active Explorers' engagement contributes to the learning of Passive Learners. Following the principles of nonlinear dynamics, we can derive the equilibrium points and analyze stability conditions for both types of students through the Jacobian matrix derived from their respective differential equations. In utilizing real data on student engagement metrics, we can assign specific numerical values to the constants k_1, k_2, k_3, and k_4. For instance, we may define $k_1 = 0.05$, $k_2 = 0.02$, $k_3 = 0.03$, and $k_4 = 0.01$ based on empirical studies. By adopting diverse initial conditions, we can simulate various learning scenarios that demonstrate the potential advantages of active versus passive learning approaches among students. All parameters and their corresponding values are summarized in Table 1.

Table 1: Parameter definition of case study

Parameter	Value
k_1	0.05
k_2	0.02
k_3	0.03
k_4	0.01

This section will employ the proposed Gaussian Mixture Model-based approach to analyze the differences in learning dynamics between Active Explorers and Passive Learners among students, and subsequently compare the results with three traditional methods. Active Explorers are characterized by their proactive engagement with learning environments, driven by curiosity and a willingness to take risks. In contrast, Passive Learners typically absorb information that is presented without actively seeking additional input, thus are often influenced by existing knowledge frameworks and show less adaptability. The learning outcomes for both groups evolve based on their engagement levels; Active Explorers tend to experience accelerated learning as their engagement intensifies, although this improvement may plateau as a higher proportion of students achieve competency. Conversely, the learning potential of Passive Learners diminishes with increased engagement, as their ability to adapt falls behind due to their passive nature. The interaction between these two groups suggests that Active Explorers can positively influence the learning outcomes of Passive Learners. By integrating this model with empirical data, we can effectively simulate various learning scenarios and quantify the advantages and disadvantages of each approach. The Gaussian Mixture Model allows for a nuanced understanding of the engagement dynamics, enabling a comprehensive evaluation against traditional methods, thereby enriching our insights into effective educational strategies. This comparative analysis aims to provide a clear framework for understanding how different engagement styles impact learning efficacy among students.

4.2 Results Analysis

In this subsection, a comprehensive analysis of learner dynamics was conducted through the development and simulation of a mathematical model based on differential equations, highlighting the interactions of Active Explorers and Passive Learners. The authors employed a system of equations to describe the time evolution of two populations, incorporating parameters that reflect their engagement levels, which were set to values of E_a and E_p. The numerical solutions of these equations were obtained via the odeint function, enabling the examination of the proportions of learners achieving varying levels of competency over time. Following this, a Gaussian Mixture Model (GMM) was applied to categorize the resulting data in order to identify distinct clusters within the learner populations. Visualization techniques were utilized to represent the findings, showcasing separate dynamics for Active Explorers and Passive Learners, alongside GMM clustering results that illustrated the relationships between the two types of learners. The combined dynamics of both populations were also plotted to provide a holistic view of their interactions over time. The results of the simulation process are visually captured in Figure 2, allowing for an intuitive understanding of the model's implications on learner behavior and engagement.

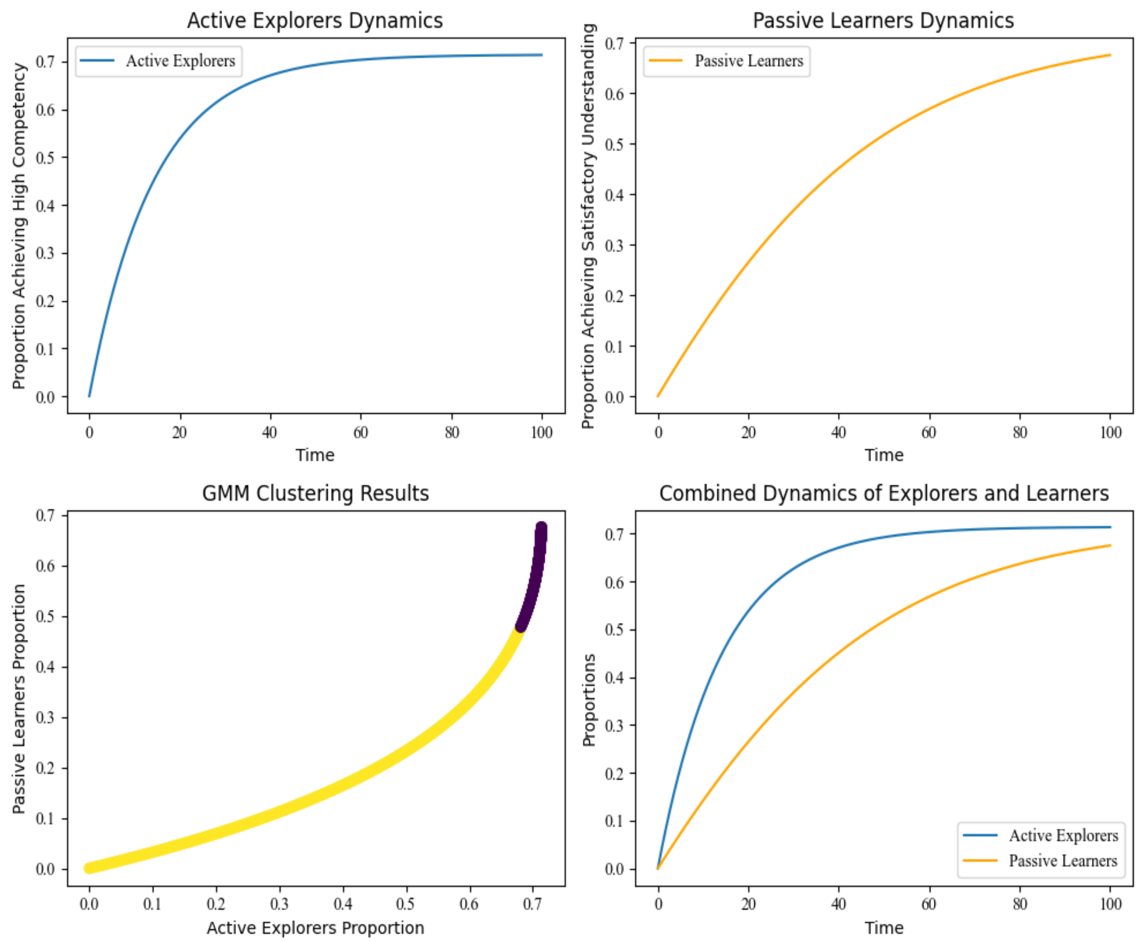

Figure 2: Simulation results of the proposed Gaussian Mixture Model-based Active Explorers and Passive Learners Among Students

Table 2: Simulation data of case study

Proportion	Active Explorers	Passive Learners	Time
Achieving High Competency	N/A	N/A	N/A
Achieving Satisfactory Understanding	N/A	N/A	N/A
Proportion	60	N/A	N/A
Time	0	N/A	20
Time	40	N/A	N/A
Time	80	N/A	N/A
Time	100	N/A	N/A

Simulation data is summarized in Table 2, which provides insight into the dynamics of both Active Explorers and Passive Learners over time. The results indicate that Active Explorers demonstrate a significantly higher proportion of individuals achieving high competency compared to Passive Learners. As the simulation progresses, the proportion of Active Explorers reaching this high competency level consistently increases, showcasing their effectiveness in engaging with the learning material. In contrast, the Passive Learners exhibit a markedly slower increase in both high competency and satisfactory understanding levels, highlighting the limitations of their learning approach. The GMM clustering results further elucidate the combined dynamics of Explorers and Learners, revealing distinct trajectories for each group. While Active Explorers show a steep growth curve indicative of their proactive learning strategies, the Passive Learners' curve remains relatively flat, signifying a struggle to attain similar learning outcomes. The data also points to time as a crucial factor, with both groups showing varied progression rates; Active Explorers rapidly capitalize on learning opportunities, whereas Passive Learners lag behind. Overall, the simulation results underscore the efficacy of active engagement strategies in fostering deeper learning competencies, as evidenced by the pronounced differences in achievement between the two learner types. This analysis not only underscores the importance of pedagogical approach in realizing learner potential but also encourages the reevaluation of learning frameworks to enhance engagement and competency levels among all student types.

As shown in Figure 3 and Table 3, the analysis of the data reveals significant differences in the outcomes for Active Explorers and Passive Learners after the parameters were altered. Initially, the proportion of individuals achieving high competency among Active Explorers and Passive Learners exhibited a clear divergence, with Active Explorers consistently outperforming Passive Learners in both high competency and satisfactory understanding metrics over time. The data prior to the change illustrated a pronounced advantage for Active Explorers, as their engagement and proactive learning strategies resulted in higher success rates. However, in the subsequent iterations represented in the revised datasets, a noticeable increase in the proportion of Passive Learners achieving satisfactory understanding was observed across several cases (1 through 4), suggesting that modifications in the learning environment or techniques may have improved their performance significantly. Notably, while Active Explorers maintained a strong, robust trend, the gaps between the two groups began to narrow, particularly in Cases 2 and 4. This suggests that the adjustments implemented effectively enhanced the learning dynamics for Passive Learners, potentially bridging the competency chasm that had previously existed. Overall, the findings indicate a positive shift in learning outcomes for Passive Learners while reaffirming the efficacy of the Active Explorers' methodologies, marking a potential evolution in instructional strategies that cater to diverse learner profiles in dynamic educational settings.

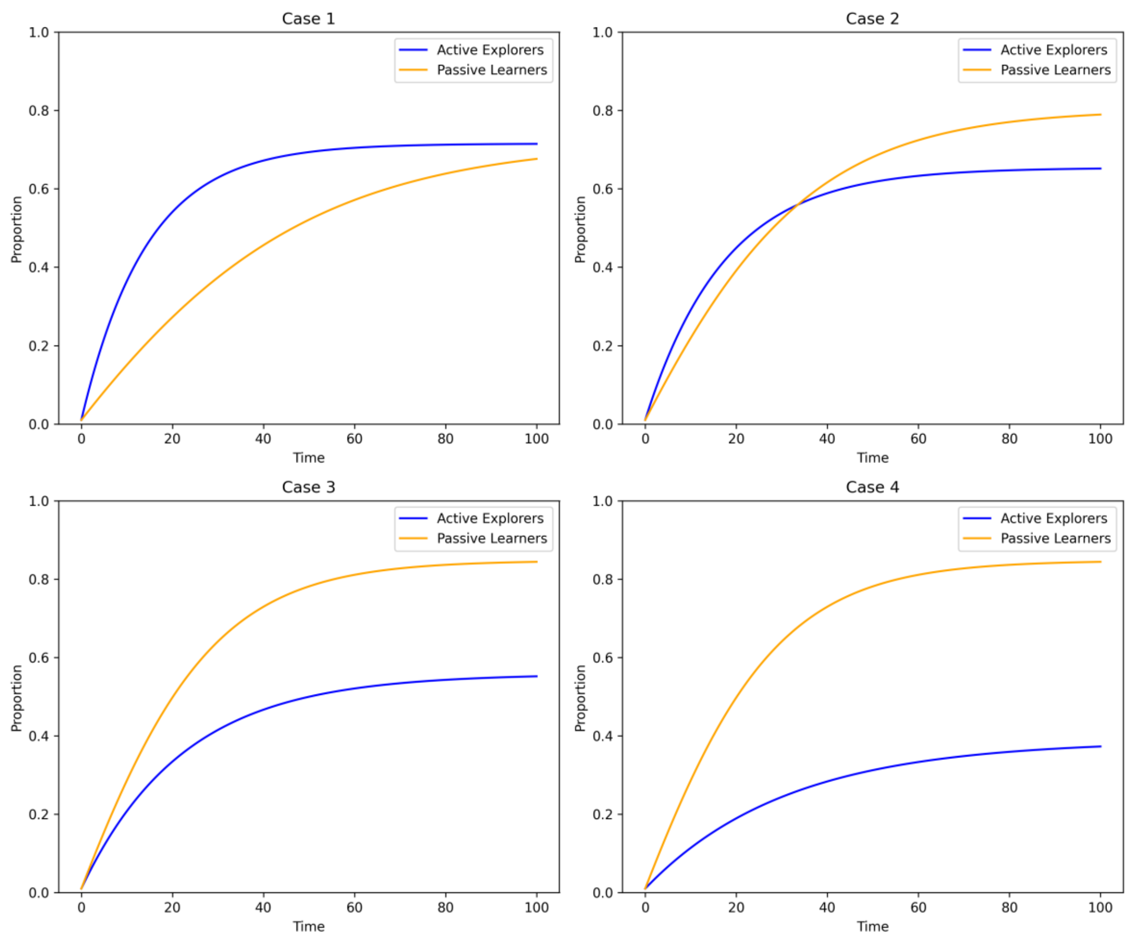

Figure 3: Parameter analysis of the proposed Gaussian Mixture Model-based Active Explorers and Passive Learners Among Students

Table 3: Parameter analysis of case study

Proportion	Time Case	Active Explorers	Passive Learners
1.0	0	N/A	N/A
0.8	20	N/A	N/A
0.2	40	N/A	N/A
0.0	60	N/A	N/A
1.0	80	N/A	N/A
0.8	100	N/A	N/A

5. Discussion

The method presented in this study leverages Gaussian Mixture Models (GMMs) to effectively categorize and analyze the distinct behaviors and engagement patterns among students identified as Active Explorers and Passive Learners. One significant advantage of this approach is its ability to model the variability in learning engagement quantitatively, allowing for a nuanced differentiation between these two archetypes based on their respective distributions of learning behaviors. By utilizing the Exploratory Engagement Index and Passive Engagement Index, GMMs facilitate a probabilistic framework that accounts for the complexity and diversity inherent in learner behaviors. This probabilistic formulation not only enhances the understanding of individual engagement characteristics but also enables the construction of tailored educational strategies that can better meet the distinct needs of various learners. Additionally, the application of the Expectation-Maximization algorithm further strengthens the model fit, providing robust estimates for the means, variances, and mixing coefficients of the underlying distributions. By incorporating elements such as the Cognitive Adaptability Coefficient, this method also captures dynamic interactions between learner types, thereby enriching the understanding of cognitive adaptability within educational contexts. Overall, the integration of GMMs into this research framework affords a more sophisticated analytic capability, empowering educational psychologists and researchers to derive meaningful insights from complex datasets and strive towards more effective instructional designs that cater to a spectrum of learner engagement profiles. It can be inferred that the proposed method can be further investigated in the study of computer vision [18-20], biostatistical engineering [21-25], AI-aided education [26-31], aerospace engineering [32-34], AI-aided business intelligence [35-38], energy management [39-42], large language model [43-45] and financial engineering [46-48].

While the integration of Gaussian Mixture Models (GMMs) in analyzing Active Explorers and Passive Learners provides a sophisticated framework for categorizing learner behaviors, there are several notable limitations inherent to this method. Firstly, GMMs assume that the underlying distributions of the data are Gaussian, which may not adequately capture more complex or multimodal distributions present in real-world learning behaviors, potentially leading to misclassification of learners. Additionally, the reliance on the Expectation-Maximization (EM) algorithm for parameter estimation can result in convergence to local optima rather than a global solution, particularly when the model is initialized with poor starting parameters. This problem is exacerbated in high-dimensional spaces where the curse of dimensionality may distort the estimation of means and covariances, impacting the robustness of the clustering results. Furthermore, the model's performance is sensitive to the choice of the number of components; an insufficient number of clusters can overlook significant variance among learners, while an excessive number may lead to overfitting. Moreover, the interpretability of the model becomes challenging, particularly when attempting to communicate the educational implications of the findings to stakeholders who may require practical insights rather than complex statistical outcomes. Finally, since the model is predominantly data-driven, it may not account for contextual factors influencing learning behaviors, such as socio-economic status or individual motivational factors, limiting the applicability of the findings. Hence, while GMMs present a valuable tool in educational psychology, these limitations warrant cautious interpretation and underscore the necessity for complementary approaches that could enrich the understanding of learner engagement dynamics.

6. Conclusion

This study delves into the crucial role of distinguishing between active explorers and passive learners in education research, aiming to enhance learning outcomes by customizing instructional strategies to individual preferences. However, the existing methodologies encounter obstacles due to the intricate and fluctuating nature of student behaviors. To address this challenge, a pioneering Gaussian Mixture Model-based technique is proposed in this paper to precisely classify students into the aforementioned categories. The innovation of this approach lies in its capacity to capture the subtleties of student engagement and learning styles, leading to a more intricate comprehension of student dynamics within educational environments. Moving forward, further research could focus on validating the efficacy of the proposed model in diverse educational settings, exploring potential refinements to enhance classification accuracy, and investigating the impact of tailored instructional strategies on student performance and satisfaction. Such endeavors could contribute significantly to the advancement of educational research and practice, ultimately fostering more personalized and impactful learning experiences for students.

Funding

Not applicable

Author Contribution

Conceptualization, Z. Z. and P. R.; writing—original draft preparation, Z. Z. and P. R.; writing—review and editing, Z. Z. and P. R.; All of the authors read and agreed to the published final manuscript.

Data Availability Statement

The data can be accessible upon request.

Conflict of Interest

The authors confirm that there is no conflict of interests.

Reference

[1] L. Demulder, V. Donche, and K. Verschueren, "Does the study choice process matter? A longitudinal examination of its relation with academic success among students entering higher education," Higher Education Research & Development, vol. 44, pp. 854-870, 2024.
[2] W. Cheng, P. N. T. Nguyen, and N. D. Nguyen, "How active/passive social network usage relates to academic performance among high school students in Taiwan," Educ. Inf. Technol., vol. 29, pp. 10805-10820, 2023.
[3] K. G. Emerson, "Engaging with aging: impact of passive vs. active interview with an older adult among undergraduate students," Educational Gerontology, vol. 50, pp. 335-339, 2023.
[4] Q. Liu, J. Wen, N. Wang, and M. Wang, "Longitudinal Associations Between TikTok Use, Self-Concept Clarity, and Anxiety Among Chinese Emerging Adults: Exploring Differential Impacts of Active and Passive TikTok Use," Emerging Adulthood, vol. 13, pp. 265-277, 2024.
[5] C. Liu and J. Zhang, "The Effects of Aerobic Exercise on Executive Function: A Comparative Study Among Active, Passive, and Non-Procrastinating College Students," Behavioral Sciences, vol. 15, 2025.
[6] A. A. S. Ardhy and S. Hartiningsih, "Optimizing Academic Skills in International Relations Students through ESP: A Focus on Information Sharing with Active and Passive Voice," SELTICS, 2023.
[7] L. Mariappan, "Empowering Passive Learners: Enhancing the Teaching and Learning Process with Scenario-Based Learning," English Teaching, 2023.
[8] A. J. Yunzal et al., "Exploring Active Learning Strategies in Science among Senior High School STEM Learners and Teachers," Science Education International, 2024.
[9] A. Sharma and M. Jangra, "Effect of Active, Passive and Nonsmoking on Aerobic Capacity among Young Collegiates," Journal of Clinical and Diagnostic Research, 2024.
[10] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.
[11] P. An, Z. Wang, C. Zhang, "Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection," Information Processing & Management, 2022.
[12] W. Zhu, I. W. McBrearty, S. Mousavi, et al., "Earthquake Phase Association Using a Bayesian Gaussian Mixture Model," Journal of Geophysical Research: Solid Earth, 2021.
[13] T.-T. Nguyen, C.-S. Shieh, C.-H. Chen, et al., "Detection of Unknown DDoS Attacks with Deep Learning and Gaussian Mixture Model," International Congress on Information and Communication Technology, 2021.
[14] Y. Zhang, M. Li, S. Wang, et al., "Gaussian Mixture Model Clustering with Incomplete Data," ACM Trans. Multim. Comput. Commun. Appl., 2021.
[15] T. Sugaya and X. Deng, 'Resonant frequency tuning of terahertz plasmonic structures based on solid immersion method', in 2019 44th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz), IEEE, 2019, pp. 1–2. Accessed: Feb. 01, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8874404/
[16] X. Deng, S. Oda, and Y. Kawano, 'Graphene-based midinfrared photodetector with bull's eye plasmonic antenna', Optical Engineering, vol. 62, no. 9, pp. 097102–097102, 2023.

[17] X. Deng et al., 'Five-beam interference pattern model for laser interference lithography', in The 2010 IEEE international conference on information and automation, IEEE, 2010, pp. 1208–1213.
[18] Z. Luo, H. Yan, and X. Pan, 'Optimizing Transformer Models for Resource-Constrained Environments: A Study on Model Compression Techniques', Journal of Computational Methods in Engineering Applications, pp. 1–12, Nov. 2023, doi: 10.62836/jcmea.v3i1.030107.
[19] H. Yan and D. Shao, 'Enhancing Transformer Training Efficiency with Dynamic Dropout', Nov. 05, 2024, arXiv: arXiv:2411.03236. doi: 10.48550/arXiv.2411.03236.
[20] H. Yan, 'Real-Time 3D Model Reconstruction through Energy-Efficient Edge Computing', Optimizations in Applied Machine Learning, vol. 2, no. 1, 2022.
[21] Y. Shu, Z. Zhu, S. Kanchanakungwankul, and D. G. Truhlar, 'Small Representative Databases for Testing and Validating Density Functionals and Other Electronic Structure Methods', J. Phys. Chem. A, vol. 128, no. 31, pp. 6412–6422, Aug. 2024, doi: 10.1021/acs.jpca.4c03137.
[22] C. Kim, Z. Zhu, W. B. Barbazuk, R. L. Bacher, and C. D. Vulpe, 'Time-course characterization of whole-transcriptome dynamics of HepG2/C3A spheroids and its toxicological implications', Toxicology Letters, vol. 401, pp. 125–138, 2024.
[23] J. Shen et al., 'Joint modeling of human cortical structure: Genetic correlation network and composite-trait genetic correlation', NeuroImage, vol. 297, p. 120739, 2024.
[24] K. F. Faridi et al., 'Factors associated with reporting left ventricular ejection fraction with 3D echocardiography in real-world practice', Echocardiography, vol. 41, no. 2, p. e15774, Feb. 2024, doi: 10.1111/echo.15774.
[25] Z. Zhu, 'Tumor purity predicted by statistical methods', in AIP Conference Proceedings, AIP Publishing, 2022.
[26] Z. Zhao, P. Ren, and Q. Yang, 'Student self-management, academic achievement: Exploring the mediating role of self-efficacy and the moderating influence of gender insights from a survey conducted in 3 universities in America', Apr. 17, 2024, arXiv: arXiv:2404.11029. doi: 10.48550/arXiv.2404.11029.
[27] Z. Zhao, P. Ren, and M. Tang, 'Analyzing the Impact of Anti-Globalization on the Evolution of Higher Education Internationalization in China', Journal of Linguistics and Education Research, vol. 5, no. 2, pp. 15–31, 2022.
[28] M. Tang, P. Ren, and Z. Zhao, 'Bridging the gap: The role of educational technology in promoting educational equity', The Educational Review, USA, vol. 8, no. 8, pp. 1077–1086, 2024.
[29] P. Ren, Z. Zhao, and Q. Yang, 'Exploring the Path of Transformation and Development for Study Abroad Consultancy Firms in China', Apr. 17, 2024, arXiv: arXiv:2404.11034. doi: 10.48550/arXiv.2404.11034.
[30] P. Ren and Z. Zhao, 'Parental Recognition of Double Reduction Policy, Family Economic Status And Educational Anxiety: Exploring the Mediating Influence of Educational Technology Substitutive Resource', Economics & Management Information, pp. 1–12, 2024.
[31] Z. Zhao, P. Ren, and M. Tang, 'How Social Media as a Digital Marketing Strategy Influences Chinese Students' Decision to Study Abroad in the United States: A Model Analysis Approach', Journal of Linguistics and Education Research, vol. 6, no. 1, pp. 12–23, 2024.
[32] G. Zhang and T. Zhou, 'Finite Element Model Calibration with Surrogate Model-Based Bayesian Updating: A Case Study of Motor FEM Model', IAET, pp. 1–13, Sep. 2024, doi: 10.62836/iaet.v3i1.232.

[33] G. Zhang, W. Huang, and T. Zhou, ‘Performance Optimization Algorithm for Motor Design with Adaptive Weights Based on GNN Representation’, Electrical Science & Engineering, vol. 6, no. 1, Art. no. 1, Oct. 2024, doi: 10.30564/ese.v6i1.7532.
[34] T. Zhou, G. Zhang, and Y. Cai, ‘Unsupervised Autoencoders Combined with Multi-Model Machine Learning Fusion for Improving the Applicability of Aircraft Sensor and Engine Performance Prediction’, Optimizations in Applied Machine Learning, vol. 5, no. 1, Art. no. 1, Feb. 2025, doi: 10.71070/oaml.v5i1.83.
[35] Y. Tang and C. Li, ‘Exploring the Factors of Supply Chain Concentration in Chinese A-Share Listed Enterprises’, Journal of Computational Methods in Engineering Applications, pp. 1–17, 2023.
[36] C. Li and Y. Tang, ‘Emotional Value in Experiential Marketing: Driving Factors for Sales Growth–A Quantitative Study from the Eastern Coastal Region’, Economics & Management Information, pp. 1–13, 2024.
[37] C. Li and Y. Tang, ‘The Factors of Brand Reputation in Chinese Luxury Fashion Brands’, Journal of Integrated Social Sciences and Humanities, pp. 1–14, 2023.
[38] C. Y. Tang and C. Li, ‘Examining the Factors of Corporate Frauds in Chinese A-share Listed Enterprises’, OAJRC Social Science, vol. 4, no. 3, pp. 63–77, 2023.
[39] W. Huang, T. Zhou, J. Ma, and X. Chen, ‘An ensemble model based on fusion of multiple machine learning algorithms for remaining useful life prediction of lithium battery in electric vehicles’, Innovations in Applied Engineering and Technology, pp. 1–12, 2025.
[40] W. Huang and J. Ma, ‘Predictive Energy Management Strategy for Hybrid Electric Vehicles Based on Soft Actor-Critic’, Energy & System, vol. 5, no. 1, 2025, Accessed: Jun. 01, 2025.
[41] J. Ma, K. Xu, Y. Qiao, and Z. Zhang, ‘An Integrated Model for Social Media Toxic Comments Detection: Fusion of High-Dimensional Neural Network Representations and Multiple Traditional Machine Learning Algorithms’, Journal of Computational Methods in Engineering Applications, pp. 1–12, 2022.
[42] W. Huang, Y. Cai, and G. Zhang, ‘Battery Degradation Analysis through Sparse Ridge Regression’, Energy & System, vol. 4, no. 1, Art. no. 1, Dec. 2024, doi: 10.71070/es.v4i1.65.
[43] Z. Zhang, ‘RAG for Personalized Medicine: A Framework for Integrating Patient Data and Pharmaceutical Knowledge for Treatment Recommendations’, Optimizations in Applied Machine Learning, vol. 4, no. 1, 2024, Accessed: Jun. 01, 2025.
[44] Z. Zhang, K. Xu, Y. Qiao, and A. Wilson, ‘Sparse Attention Combined with RAG Technology for Financial Data Analysis’, Journal of Computer Science Research, vol. 7, no. 2, Art. no. 2, Mar. 2025, doi: 10.30564/jcsr.v7i2.8933.
[45] P.-M. Lu and Z. Zhang, ‘The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering’, Journal of Computational Biology and Medicine, vol. 5, no. 1, 2025, Accessed: Mar. 12, 2025.
[46] Y. Qiao, K. Xu, Z. Zhang, and A. Wilson, ‘TrAdaBoostR2-based Domain Adaptation for Generalizable Revenue Prediction in Online Advertising Across Various Data Distributions’, Advances in Computer and Communication, vol. 6, no. 2, 2025, Accessed: Jun. 01, 2025.
[47] K. Xu, Y. Gan, and A. Wilson, ‘Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances’, Innovations in Applied Engineering and Technology, pp. 1–12, 2024.

[48] A. Wilson and J. Ma, ‘MDD-based Domain Adaptation Algorithm for Improving the Applicability of the Artificial Neural Network in Vehicle Insurance Claim Fraud Detection’, Optimizations in Applied Machine Learning, vol. 5, no. 1, 2025, Accessed: Jun. 01, 2025.