RESEARCH PAPER

Personalized Learning Pathways in Engineering Education Driven by Large Language Models: Framework Development and Empirical Analysis

Nan Xie[1*], Gede Rasben Dantes[1], Dewa Gede Hendra Divayana[1], Made Hery Santosa[1]

[1]Universitas Pendidikan Ganesha (Undiksha), Bali 81116, Indonesia

Emails: djxnh421@163.com (Corresponding Author), rasben.dantes@undiksha.ac.id, hendra.divayana@undiksha.ac.id, mhsantosa@undiksha.ac.id

ABSTRACT

The rapid development of large language models (LLMs) can increase possibilities for personalized learning approaches in engineering education. However, there are insufficient empirical studies regarding theoretical approaches that would incorporate LLM features into a personalized learning approach based on the adaptive learning concept. The study develops a four-module framework for a learner-centric personalization learning path enabled by LLMs that consists of modules for learner profiling, knowledge graph creation, LLM-enabled learning paths creation, and feedback adaptation. To identify the basis for the development of the framework through establishing the necessary evidence, a large-scale secondary data analysis was carried out employing the Open University Learning Analytics Dataset. The dataset comprised 4,480 students undertaking science, technology, engineering, and mathematics (STEM) courses and classified into two groups, namely, High Adaptation group and Low Adaptation group in accordance with the alignment between the Virtual Learning Environment interaction pattern of students and the recommended course content pattern. Between-group differences were examined based on five dependent variables relating to evaluation results, pathway attainment, and learners' engagement through independent samples t-test and Cohen's d for measuring effect size. The results show that the High-Adaptation group significantly performed better than the Low-Adaptation group on all dependent variables (mean scores: 62.43 vs. 51.89; submission rates: 0.63 vs. 0.49; active learning days: 125.75 vs. 86.13), with the effect sizes varying from small ($d = 0.427$) to large ($d = 1.072$). These results serve to provide a baseline for the impact of structured pathway alignment that precedes the LLM framework design process; therefore, they contribute to the discussion, rather than being used to test its validity.

Keywords: Large Language Models; Personalized Learning Pathways; Engineering Education; Adaptive Learning; Learning Analytics; Secondary Data Analysis

How to cite this article: Xie N, Dantes GR, Divayana DGH, Santosa MH. Personalized Learning Pathways in Engineering Education Driven by Large Language Models: Framework Development and Empirical Analysis. Int J Drug Deliv Technol. 2026;16(51s): 58-70. DOI: 10.25258/ijddt.16.51s.7

Source of support: Nil.

Conflict of interest: None

1. Introduction

Rapid advances in digital technology have significantly changed the world of higher education, whereby engineering stands out as one of the major areas responsible for such changes. Engineering education faces many difficulties when it comes to dealing with different levels of existing knowledge among students enrolled in standard programs (Pham et al., 2023). Traditional modes of instruction that rely on standardization of curriculum and predetermined paths of learning often fail to address the varying needs of engineering students in an era when fast-paced technological development demands a different mode of instruction (Mukul & Büyüközkan, 2023).

With the emergence of the large language models (LLMs), there are new methods to approach these challenges. For example, ChatGPT and GPT-4 possess considerable capabilities in the areas of natural language processing, content creation, and conversation, which makes them available tools for educational use cases. Filippi and Motyl (2024) state that their review regarding the use of LLMs in engineering education shows significant possibilities in the use of LLMs in various fields including automated tutoring, assessment design, and personalized feedback. Kasneci et al. (2023) expand upon the discourse of possibilities and concerns regarding the use of LLMs in education, pointing out that the models enable personalized learning because the explanations can be made according to the requirements of the specific learner. In terms of institutions, the adoption of AI technology within higher education has achieved considerable advancement. According to the findings by Crompton and Burke (2023), intelligent tutoring systems, automated assessments, and adaptive learning have been the areas of development that have achieved significant development within the field of AI research in higher education.

In relation to engineering education specifically, the use of generative AI technologies has received both excitement and reservations. In this regard, Nikolic et al. (2023) undertook a multi-institution benchmarking experiment to measure how well ChatGPT could complete engineering assessment activities and determined that while it was capable in some domains, there was a need for careful critical analysis and pedagogic scaffolding. Mosly (2024) investigated the potential benefits and issues of incorporating AI technology within engineering curriculum development through a mixed approach involving a review and focus groups to show that the success of the application depends on matching AI functions to learning objectives.

The concept of learning path personalization has gained considerable attention in academia as an additional method for improving the educational process. Gligorea et al. (2023) studied AI-supported adaptation in e-learning and determined that machine learning techniques can accurately simulate behavior and tailor the content offered by the system (Okoye et al., 2023).

Although there is increasing literature on AI for education and adaptive learning, several gaps still remain. Bond et al. (2024) called for enhanced rigour and empiricism in the field by pointing out that most past studies have used only small scale implementation or reported results rather than using large behavioral data. Specifically, studies dealing with AI personalized engineering education curriculum are few (Zhai et al., 2021). According to Fariani et al. (2023), who conducted a systematic review on personalized learning within higher education, most of the current models operate on a more generic basis and do not account for the unique needs presented by field-specific learning environments like that of engineering.

The area of convergence between the LLM technology and the development of personal learning pathways in engineering education is one of the least explored ones. Despite the exceptional abilities of LLMs to process natural language and generate relevant information, very few scientific studies exist that present frameworks which combine the abilities of these models with traditional adaptive learning principles such as user profiling, knowledge graph-based content organization, and feedback loops. Furthermore, there is a lack of empirical research about the efficacy of systematic learning processes in the field of engineering, as most empirical literature focuses either on the technological structure of the learning process or on its pedagogic results in isolation.

To overcome the challenges, two related goals can be formulated. They include the development of a methodology of personalized learning path creation powered by LLMs involving the creation of the learner profile, constructing knowledge graphs, creating learning paths, and feedback adjustments dynamically. The study also aims at creating an empirically grounded basis for the model by investigating the effect of structured pathways alignment on learning outcomes using a large dataset from traditional learning analytics, thus setting up a benchmark of how the pathway structure affects learner behavior before LLM-enabled personalized education becomes a reality.

This study contributes to the state-of-the-art research in AI-driven engineering education through three unique contributions. This creates a theoretically sound approach towards incorporating the functionalities of LLMs into adaptive learning models. Moreover, this study introduces a modular approach that can be utilized by various branches of engineering. It also provides empirical evidence based on a data set of 4,480 STEM students to show that structuring learning experiences impacts academic success, course completion, and persistence. Therefore, the benchmark is set for future comparison if LLMs are used. The results imply that behavioral metrics of engagement are more sensitive to the structure of learning pathway than those obtained on the basis of assessment. It brings new insights into learning analytics tool design.

The rest of this study is structured as follows. Section 2 reviews the theoretical background, framework design, dataset, and methodology used in this study. Section 3 reports empirical findings from three perspectives on learning outcomes. Section 4 interprets findings relative to the related literature and their significance for practice in engineering education. Section 5 concludes this study and suggests future research directions.

2. Data and Methods

2.1 Theoretical Basis and Design Rationale

The recommended theoretical framework is based on the principles of constructivism, where the learner creates knowledge through interactions with the environment, experiences, and participation. Constructivism has been empirically proven to be highly effective in enhancing students' engagement and understanding concepts in the context of engineering studies (Ngo, 2024). This perspective provides a strong rationale for personalized education routes, particularly in subject areas that demand problem solving and reasoning processes and therefore require education that is tailored to individual characteristics.

Adaptive learning systems uses the constructivist theories in digital form. Essa et al. (2023) reported that accurate personalization requires the correct identification of the learners' cognitive style, knowledge state, and behavioral pattern to provide guidelines for designing the learner's profile and feedback components of the present system. The classification taxonomy suggested by Ismail et al. (2023) that classifies personalized learning system according to learning environment, content, and user models also played an important role in defining the modules adopted in this study.

2.2 Proposed Framework Architecture

Based on the aforementioned theoretical basis, the current research provides an LLM-based personalized pathway framework that is appropriate for educational engineering situations. The overall structure of the framework, as shown in Figure 1, comprises four interrelated modules that altogether enable the generation, delivery, and continuous optimization of individualized learning experiences.

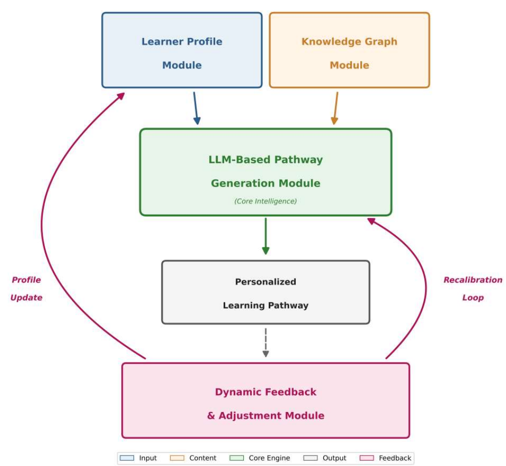

Figure 1. LLM-Driven Personalized Learning Pathway Framework Architecture

The Learner Profile Module acts as the first entry module, which consolidates multi-dimensional data about students such as previous performance, preferred styles of learning, mastery of knowledge level, and engagement behaviors. The profiling process uses the intelligent assistant model suggested by Sajja et al. (2024), showing that a combination of natural language processing and learner analytics provides dynamically changing models of students that can be used to generate paths.

The Knowledge Graph Module acts as the architecture behind the content domain. Since engineering curriculum requires several complex dependencies between different topics and prerequisite relationships, this module uses the graph structure for arranging the subject matter where nodes symbolize elements of knowledge and edges define prerequisite relationships. This approach further supports the research conducted by Liu et al. (2023), who found that using graph-based content representations improves the effectiveness of learning path recommendation.

The Pathway Generation Module, which relies on the use of LLMs, is the fundamental conceptual element. After receiving the information related to the learner profile and structured knowledge graph as an input, the module leverages the LLMs for generating unique learning pathways corresponding to the knowledge state of learners. This methodology builds on the model presented by Naseer et al. (2024), where the use of deep learning-based pathways created measurable improvements; large language models provide the system with the ability to comprehend natural language, understand the user's input, and generate learning activities based on their context.

The Dynamic Feedback and Adjustment module completes the cycle of adaptation. The feedback regarding assessment outcomes, level of engagement, and knowledge acquisition is provided back to the system to make adjustments to the pathway accordingly, following the principles of closed-loop learning analytics discussed in Ouyang et al. (2023), which were shown to greatly increase effectiveness of learning in online engineering classes.

As shown in Figure 1, all the four modules operate in a cyclic fashion as opposed to following a linear trend. The Learner Profile Module is constantly receiving updated information from the Dynamic Feedback Module while the Knowledge Graph Module provides a static but scalable content framework for the Pathway Generation Module. The interaction ensures that the recommended pathways are in accordance with both the developmental stage of the learner and the structure of the knowledge field, and thus allows for greater personalization with the accumulation of more learner data.

2.3 Data Source and Sample Selection

Considering that the large-scale utilization of LLMs-based learning platforms for engineering education is still in its early stages, experimentation to validate the suggested framework may not be possible at present. Alternatively, this study uses secondary data analysis method to provide basic evidence on whether structured learning pathways—a product of the framework that is under consideration—contribute towards improved learning outcomes in a pre-LLM setting. Such a baseline plays an important methodological role because, through identifying the extent of the impact of aligning structural pathways with respect to a traditional VLE, the study creates a benchmark for judging the relative advantage of future LLM applications.

In this regard, this study uses the Open University Learning Analytics Dataset (OULAD), which is a publically available benchmark dataset that was introduced by Kuzilek et al. (2017) in the journal Scientific Data. This dataset includes anonymous data from 32,593 students who were taking courses in 22 subjects at the Open University, one of the largest distance learning institutions worldwide. The database contains three types of information: student demographic information; assessment performances including scores varying from 0 to 100, and over 10 million records of daily aggregated click-streams logs reflecting students' activities in relation to the Virtual Learning Environment (VLE). The combination of all these factors makes the use of OULAD dataset relevant when analyzing the link between patterns of learning pathways and their impact on learning outcomes. Even though the data from OULAD was collected through course presentations that were held in 2013-2014 years, this time frame is still relevant to the research at hand. As the dataset reflects learner behaviors from within a traditional pre-LLM learning environment, it provides a pristine baseline for studying the impact of aligned structures in influencing learning outcomes without the help of generative AI. The dataset's relevance and learner behaviors in question have already been verified by its original authors and have since been extensively used in educational data mining studies.

The focus of this research on technology-related subjects made it necessary for the researcher to limit the sample to only those modules within the four STEM subject domains available in the data. These subject domains cover sciences, technology, engineering, and math. Social science modules were were excluded to ensure disciplinary relevance. After removing records with missing assessment scores or incomplete VLE interaction logs, the final analytical sample size was reduced to 4,480 students with data on all three dimensions to support meaningful comparison (see Table 1).

Table 1. Sample Characteristics of STEM-Domain Students

Characteristic	High-Adaptation (n = 2,240)	Low-Adaptation (n = 2,240)	Total (N = 4,480)
Course Module			
BBB	733 (32.7%)	737 (32.9%)	1,470 (32.8%)
CCC	345 (15.4%)	335 (15.0%)	680 (15.2%)
DDD	621 (27.7%)	629 (28.1%)	1,250 (27.9%)
FFF	541 (24.2%)	539 (24.0%)	1,080 (24.1%)
Gender			
Male	1,073 (47.9%)	1,082 (48.3%)	2,155 (48.1%)
Female	1,167 (52.1%)	1,158 (51.7%)	2,325 (51.9%)
Age Band			
0–35	1,223 (54.6%)	1,212 (54.1%)	2,435 (54.4%)
35–55	804 (35.9%)	791 (35.3%)	1,595 (35.6%)
55 and above	213 (9.5%)	237 (10.6%)	450 (10.0%)
Highest Education Level			
Lower than A Level	444 (19.8%)	452 (20.2%)	896 (20.0%)
A Level or Equivalent	901 (40.2%)	898 (40.1%)	1,799 (40.2%)
HE Qualification	676 (30.2%)	679 (30.3%)	1,355 (30.2%)
Post Graduate Qualification	219 (9.8%)	211 (9.4%)	430 (9.6%)
Pathway Alignment Score			
Mean (SD)	0.704 (0.097)	0.404 (0.107)	0.554 (0.181)
Range	0.555–0.980	0.050–0.554	0.050–0.980

Note. Sample consists of 4,480 students enrolled in the four STEM-domain modules of the Open University Learning Analytics Dataset (OULAD), which are anonymized in the dataset using the codes BBB, CCC, DDD, and FFF to protect institutional privacy. Students were classified into the High-Adaptation Group and Low-Adaptation Group based on a median split (median = 0.555) of the pathway alignment score, which measures the proportion of VLE materials accessed during their designated availability weeks. Percentages within each subgroup may not sum to 100 due to rounding.

To operationalize the term learning pathway adaptation, the learners in the identified STEM courses were divided into two categories depending on the degree of consistency between their virtual interaction patterns and the prescribed course content structure. The students that exhibited similar access pattern to the weekly designed progressive structure of the learning content were placed in the High Adaptation Group, meaning students whose behavior shows evidence of a learning sequence trajectory. Those who exhibited behavior deviant to the sequence prescribed in the study were placed in the Low Adaptation Group, indicating learners who engaged in non-linear routes. This method allows for a comparison approach that simulates the effects of individualized pathway guidance using a secondary data analysis perspective.

2.4 Experimental Design and Analytical Methods

Students were categorized into High-Adaptation Group and Low-Adaptation Group using an empirical procedure whereby the alignment of the students' interaction schedule on the VLE with recommended weekly schedule of contents delivery was measured using quantitative approach. A pathway alignment measure was calculated for each individual as the ratio between the number of VLE learning resources accessed during their availability period and the total number of resources accessed by that individual throughout the course. The students who scored above the median were considered to belong to the High Adaptation Group, while those scoring at the median or below the median were placed in the Low Adaptation Group. This splitting technique results in groups of almost equal size and provides a definite empirical base for making comparisons.

A comparative method is used in this research where the independent variable is described based on the level of adaptability found in the learning pathway process and is divided into two categories – high and low. The dependent variable is identified based on the three components of the learning outcome: assessment performance, pathway completion, and learning engagement. The assessment of baseline equivalence and possible effects of confounding factors was achieved through the analysis of demographic characteristics including age group, gender, and educational background using descriptive statistics. The results from Table 1 suggest that there is a close match regarding distribution in both sets of subjects for each demographic variable observed, with no differences being recorded beyond two percentage points, suggesting equality in baselines. However, it should be mentioned that more complex statistical methods such as propensity score matching, covariance analysis, and multiple regression have not been used in this research; therefore, the obtained effect sizes should be regarded as correlations rather than causations.

Statistics analyses were carried out to find any difference that exists between the two groups concerning the dependent variables. To establish whether the difference between means is statistically significant, an independent samples t-test was used at $p < 0.05$. The effect sizes for each comparison were calculated using Cohen's d to evaluate the practical significance of differences, considering the accepted values of small (d=0.2), medium (d=0.5), and large (d=0.8) effect sizes. This approach reflects the approach of closed-loop learning analytics suggested by Sailer et al. (2024). The above framework demonstrates the need for incorporation of performance and behavior measures in order to create an all-inclusive measure of learning outcomes. The following descriptive statistics were used in this research, namely, mean, standard deviation, and frequency distribution.

3. Results

3.1 Learning Performance Comparison

The main objective of this study was to examine whether learners whose development patterns resembled the recommended content structure of the course performed better academically compared to those following less structured approaches. According to the descriptive statistics, the mean test score of the High Adaptation group is considerably greater (M = 62.43, SD = 16.13) than that of the Low Adaptation group (M = 51.89, SD = 16.38). The difference between the two means is 10.54 points within 100 points. Such difference means that students who had access to information in accordance with the course schedule performed better in tests. The between-group comparison of mean assessment scores is visually depicted in Figure 2.

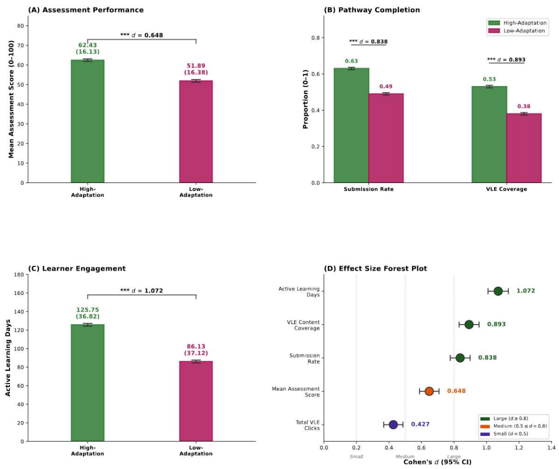

Figure 2. Between-group comparison across five learning outcome variables. (A) Mean assessment scores; (B) Pathway completion indicators (submission rate and VLE content coverage); (C) Active learning days; (D) Forest plot of Cohen's *d* effect sizes with 95% confidence intervals. Total VLE Clicks (not visualized in Panel C): High = 943.12 (SD = 1358.56); Low = 489.16 (SD = 643.59); t(4478) = 14.292, *** d = 0.427. Error bars in panels A–C represent 95% CIs. *** p < .001.

The results of an independent samples t-test showed that the difference was statistically significant, $t(4478) = 21.684$, $p < .001$. The corresponding effect size, calculated using Cohen's *d*, yielded a value of 0.648, indicating a medium effect. This result implies that pathway congruence is significantly related to academic achievement, but the medium effect size reveals considerable variance still exists for both groups. Figure 3 reveals that final course outcomes support this information. In relation to the High-Adaptation Group, 60.9% of the students received a Pass grade while 1.7% of the students got a Distinction grade and 8.6% did not complete the course. In contrast, the Low-Adaptation Group showed a pass rate of 42.2% where distinction was only 0.2% but the withdrawal rate was higher at 21.0%. Notably, the higher withdrawal rate in the Low-Adaptation Group is very significant since there is a 12.4%-point difference between the two groups, which denotes a large variation in course retention.

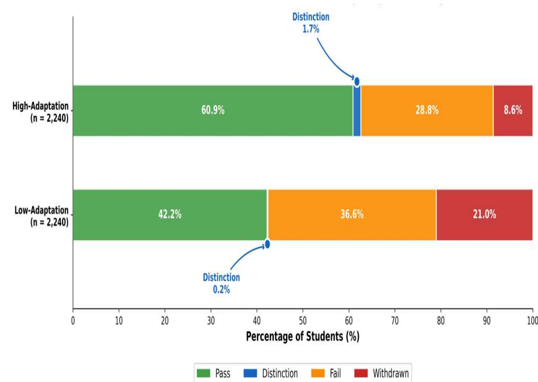

Figure 3. Distribution of Final Course Outcomes by Adaptation Group.The stacked horizontal bars display the percentage breakdown of four outcome categories (Pass, Distinction, Fail, and Withdrawn) for the High-Adaptation Group (n = 2,240) and the Low-Adaptation Group (n = 2,240). Distinction segments are highlighted with markers due to their small visual proportion. Outcome categories follow the OULAD standard final_result classification.

These differences between the groups should be understood against the background of the current study methodology, which is to compare behavioral pathways alignment within the framework of the traditional VLE. However, despite the obvious impossibility of comparing the current meta-analysis with those involving AI-integrated teaching such as Wang and Fan (2025), where the high impact of using the technology on student achievement (g = 0.867 for 51 studies) was found, the current effect size (d = 0.648) implies that following the structural pathway alone exerts considerable effects regardless of the AI-based personalization used. The full statistical findings on all dependent variables are shown in Table 2.

Table 2. Summary of Between-Group Comparisons across All Dependent Variables

Variable	High-Adaptation M (SD)	Low-Adaptation M (SD)	*t* (4478)	*p* value	Cohen's *d* (Effect Size)
Assessment Performance					
Mean Assessment Score	62.43 (16.13)	51.89 (16.38)	21.684	< .001	0.648 *(medium)*
Pathway Completion					

Variable	High-Adaptation M (SD)	Low-Adaptation M (SD)	t (4478)	p value	Cohen's d (Effect Size)
Assessment Submission Rate	0.63 (0.16)	0.49 (0.16)	28.050	< .001	0.838 *(large)*
VLE Content Coverage Rate	0.53 (0.17)	0.38 (0.17)	29.883	< .001	0.893 *(large)*
Learner Engagement					
Total VLE Clicks	943.12 (1358.56)	489.16 (643.59)	14.292	< .001	0.427 *(small)*
Active Learning Days	125.75 (36.82)	86.13 (37.12)	35.868	< .001	1.072 *(large)*

Note. Independent samples t-tests were conducted to compare the High-Adaptation Group (n = 2,240) and the Low-Adaptation Group (n = 2,240) across five dependent variables. Cohen's d effect sizes are interpreted using conventional thresholds: small (d = 0.2), medium (d = 0.5), and large (d = 0.8). All between-group differences are statistically significant at p < .001. Assessment scores range from 0 to 100; submission rate and coverage rate range from 0 to 1; active learning days range from 1 to 260.

3.2 Pathway Completion and Coverage Analysis

In addition to merely academic performance, the study sought to determine whether conformity with the path had an influence on different patterns in terms of course completion and course content exposure. These two variables are the assessment submission rate, which is the percentage of assessments completed; and the VLE content exposure rate, which is the percentage of available content exposed to.

The High-Adaptation Group demonstrated a mean assessment submission rate of 0.63 (SD = 0.16), compared to 0.49 (SD = 0.16) for the Low-Adaptation Group. An independent samples t-test indicated that this difference was statistically significant, $t(4478) = 28.050$, $p < .001$, with a large effect size (Cohen's $d = 0.838$). This shows that those learners who followed the instructional sequences were more likely to complete their formal tests, and the Low Adaptation group completed about 49% of their assessment tests.

A similar pattern emerged in VLE content coverage. The High-Adaptation Group accessed a significantly larger proportion of available learning materials (M = 0.53, SD = 0.17) compared to the Low-Adaptation Group (M = 0.38, SD = 0.17). This difference was also statistically significant, $t(4478) = 29.883$, $p < .001$, with a large effect size (Cohen's $d = 0.893$). The 15-percentage-point gap suggests that the learners who had their interaction with the materials matching the ideal weekly distribution used a broader range of course material than the students belonging to the Low-Adaptation Group, who limited their usage to fewer materials.

Collectively, these two measures reveal that there is an observable trend: pathway alignment is positively correlated to both assessments completed and content engagement. The large effect sizes observed for both predictors exceed those found for assessment results ($d = 0.648$). This would suggest that the behavioral aspect of learning might react more sensitively to the pathway's structure than the cognitive one as assessed using the score test alone (see Table 2).

3.3 Learner Engagement Analysis

The final aspect of between-groups analysis involved learner involvement measured by two behavioral factors namely: total VLE clicks and number of days when learners engaged in learning activity. The measures correspond to learner involvement intensity and frequency respectively.

The High-Adaptation Group accumulated a higher mean total of VLE clicks (M = 943.12, SD = 1358.56) compared to the Low-Adaptation Group (M = 489.16, SD = 643.59). An independent samples t-test confirmed that this difference was statistically significant, $t(4478) = 14.292$, $p < .001$. However, the effect size associated with this finding was relatively small (d=0.427), possibly due to a large variance within the groups regarding clicking behavior, particularly for the High Adaptation Group where the standard deviation exceeded the mean. This pattern indicates that even though high adaptation students accessed the VLE more often than low adaptation ones, they did so in highly varying ways depending on individual differences in the clicking activity.

A significant difference emerged during the active learning period. Members of the High-Adaptation group spent an average of 125.75 days active per week (SD = 36.82), whereas those in the Low-Adaptation group were active for 86.13 days (SD = 37.12) per week. Thus, the difference of 39.62 days between the groups was statistically significant, $t(4478) = 35.868$, $p < .001$, with a large effect size (Cohen's $d = 1.072$). The discrepancy between the impact of pathway fit on clicks and that on the number of active days suggests that the former has more of a connection with the time factor of engagement rather than with the volume of engagement activities.

The results of this engagement study conform to those observed by Rahiman and Kodikal (2024), whereby the use of AI-based learning platforms is known to enhance student engagement via adaptive learning techniques. In this study, consistent engagement along the learning path seems to have played a similar part in maintaining continuous engagement within the period.

4. Discussion

4.1 Interpretation and Comparison with Existing Studies

The results of this study contribute to a body of knowledge that aligns structured learning pathways with positive achievement in academic performance, course completions, content coverage, and engagement of STEM students. In all five measures of success used in the experiment, the High-Adaptation Group proved superior to the Low-Adaptation Group in achieving the results; the effect size varied between small (Cohen's d = 0.427 for total clicks) and large (Cohen's d = 1.072 for active learning days). The results obtained using the conventional pre-LLM approach match the basic hypothesis of the proposed model, which assumes that effective and properly constructed learning trajectories may help improve learning efficiency in the field of engineering. The results obtained using pre-LLM models are also valuable as they provide a benchmark to measure the contribution of the LLM-based solution.

The key insight that can be drawn from the findings above is that behavioral and cognitive measures have different sensitivities in response to the structure of a pathway. The effect sizes for assessment submission rate (d=0.838), VLE content coverage (d=0.893) and active learning days (d=1.072) were much higher than those found for assessment scores (d=0.648). This pattern suggests that the relationship between pathway alignment and learning behavior like commitment and exposure might be more direct than the relationship between pathway alignment and cognitive gains measured by formal assessments, which are likely mediated by other variables such as background knowledge and ability.

The results are in line with those from Mai et al. (2024), who found through a SWOT analysis of the ChatGPT integration into learning environments that personalized feedback and content sequencing were among the most important strengths of an AI-integrated environment. This study extends this point by illustrating that without an AI program, adhering to content sequencing leads to tangible benefits. Lo (2023) similarly highlighted ChatGPT's capabilities in promoting self-regulated learning, emphasizing the importance of establishing empirical evidence that goes beyond self-reported views, a limitation which this current evidence-based study aims to fill.

The difference between measures of user engagement based on clicks and measures based on time should be taken into consideration. The small effect size associated with total number of clicks compared to the large effect size of active days is consistent with the claim by Küchenmann et al. (2025) that successful user engagement requires multiple behavioral indicators. The findings provide an empirical support to this claim, suggesting that the time allocation for learning-related activities is a better predictor of meaningful participation compared to the quantity of engagement.

4.2 Implications for Engineering Education

Several applications can be derived from this research regarding how AI can be integrated into learning in engineering education. The strong connection observed between alignment of pathways and completion rate as well as engagement levels suggests that engineering courses conducted digitally can derive a lot of benefits by integrating structured pathway guidance. As opposed to using learning resources as fixed repositories of knowledge, education technology systems need to use knowledge graph architecture and learner profiling technologies (based on the framework proposed by this research) for sequencing learning resources based on hierarchical relationships and learner progress paths.

Implications for institutional policy makers include the realization that intelligent learning management system development is important because these need to move on from merely delivering content to actually orchestrating the process of learning. Behavioral engagement measures are more sensitive to the effect of pathways than test scores; this suggests that current performance metrics could be underestimating the influence of pathway-based learning support. Learning analytics dashboard integration which tracks adherence to learning paths, content exploration and time-based engagement can provide deeper insights into learning behaviors for the instructors.

The proposed LLM-driven framework offers an effective way to meet these demands due to the capabilities of LLMs to automate the process of assessing learners' progress, recommending content, and providing feedback that could have required significant input from instructors. According to Lyngdorf et al. (2024), who have conducted a systematic review about digital transformation framework for the purposes of engineering education, effective implementation of AI should consider matching technological potentials to the desired educational outcomes. The proposed framework is tailored in such a way as to consider the aforementioned correlation in relation to constructivist learning theories.

4.3 Limitations

Although several contributions of the study have been discussed above, there are also some limitations of this research which should be considered. There is the limitation posed by the fact that the OULAD dataset is obtained from a traditional VLE system rather than an LLM-native learning environment. Consequently, the approach takes behavioral pathway alignment as an approximation to personalize learning pathway, but does not measure the impact of content sequencing generated by LLMs directly. While this allows for empirical testing on a large scale, it is not possible to draw any causal conclusions about LLMs-based personalization.

The method of classifying pathways according to their alignment score through the use of the median-split classification method employed in the present research carries limitations inherent in dichotomizing continuous variables, such as information loss and reduced statistical power in the cut-off region. Furthermore, since the data is derived from observation, there exists the likelihood of selection bias because the students who might be motivated academically or possess better self-regulatory skills would have been more predisposed to following the prescribed order of content. Even though the baseline demographics were similar between the groups (see table 1), more stringent forms of statistical control measures such as propensity score matching, analysis of covariance or even multivariate regression using prior academic factors were not utilized in this study and thus, unmeasured individual differences such as motivational factors, prior knowledge and self efficacy could account for group differences observed. Future studies using causal inference techniques based on richer databases would help in untangling the effects of pathway alignment from those of the learner's inherent traits. Moreover, OULAD data were collected between 2013 and 2014, which is before the introduction of LLM systems. Even as this time-related placement helps serve the purpose of baseline setting through learning activity in the pre-LLM context, it also means that the current observations cannot directly answer to the question of the additional value of LLM-based personalization. The technological and pedagogical environment have gone through significant changes since the data collection period; hence there is a need for future research on platforms that are embedded with LLMs to determine how the generation of AI technology modifies or magnifies baseline pathway effects. Another limitation relates to the absence of subjective learner data. No data about learners' satisfaction level, perceived usefulness, and self-efficacy have been collected, factors that are considered critical elements of the process of learning. Future research needs to include validation of the suggested framework by means of prospective experiments that would be conducted using platforms with LLMs and would consider subjective and objective variables. Extending the study beyond a single organization to cover many organizations as well as extending the sub-disciplines of engineering covered would increase the validity of the results. This approach of solving the limitations through robust experiments and rich datasets is vital in future studies.

5. Conclusion

This study attempts to tackle the issue of addressing the widening discrepancy between the increasing ability of language models and the decreasing number of empirical tools available for designing individualized learning pathways for engineering students. To achieve this goal, a framework is proposed that includes four interdependent modules, namely, learner profiling, knowledge graph generation, pathway generation by LLM, and feedback adjustment. The research explored the impact of pathway alignment on learning outcomes in the light of the secondary data analysis based on 4,480 STEM learners from the Open University Learning Analytics Dataset. The results suggest that the High-Adaptation Group performed significantly better than the Low-Adaptation Group on all five dependent measures, with effect sizes ranging from small (d=0.427) to large (d=1.072). Notably, the measures of behavioral involvement were more sensitive to the structure of the pathways than the assessment-based measures traditionally used. This provides basic empirical evidence that the structure of the pathways affects learners' performance outcome before the LLM was used.

Several directions are worth exploring further. The application of the framework in an actual learning environment using LLMs to conduct causal inference through experiments should be conducted. Cross-institutional and cross-disciplinary validation that involve computer, mechanical, and chemical engineers would improve the generalizability of the results. The utilization of multimodal LLMs that can process text, images, and sound is an especially exciting prospect for engineering education due to the visual and math components involved. Further research should also incorporate subjective information related to learners' feelings of satisfaction, motivation, and emotionality, which would provide further insight on the impact of personalized learning pathways on overall learning experiences.

References

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*(1), 4.

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*(1), 1–22.

Essa, S. G., Celik, T., & Human-Hendricks, N. E. (2023). Personalized adaptive learning technologies based on machine learning techniques to identify learning styles: A systematic literature review. *IEEE Access, 11*, 48392–48409.

Fariani, R. I., Junus, K., & Santoso, H. B. (2023). A systematic literature review on personalised learning in the higher education context. *Technology, Knowledge and Learning, 28*(2), 449–476.

Filippi, S., & Motyl, B. (2024). Large language models (LLMs) in engineering education: A systematic review and suggestions for practical adoption. *Information, 15*(6), 345.

Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. *Education Sciences, 13*(12), 1216.

Ismail, H., Hussein, N., Harous, S., & Khalil, A. (2023). Survey of personalized learning software systems: A taxonomy of environments, learning content, and user models. *Education Sciences, 13*(7), 741.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274.

Küchemann, S., Avila, K. E., Dinc, Y., Hortmann, C., Revenga, N., Ruf, V., Stausberg, N., Steinert, S., Fischer, F., Fischer, M., Götz, U., Hartwig, F., Kasneci, E., Kasneci, G., Kuhr, T., Kutyniok, G., Lewalter, D., Lindl, A., Sailer, M., … Kuhn, J. (2025). On opportunities and challenges of large multimodal foundation models in education. *npj Science of Learning, 10*(1), 11.

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. *Scientific Data, 4*(1), 170171.

Liu, C., Zhang, H., Zhang, J., Zhang, Z., & Yuan, P. (2023). Design of a learning path recommendation system based on a knowledge graph. *International Journal of Information and Communication Technology Education, 19*(1), 1–18.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410.

Lyngdorf, N. E. R., Jiang, D., & Du, X. (2024). Frameworks and models for digital transformation in engineering education: A literature review using a systematic approach. *Education Sciences, 14*(5), 519.

Mai, D. T. T., Da, C. V., & Hanh, N. V. (2024). The use of ChatGPT in teaching and learning: A systematic review through SWOT analysis approach. *Frontiers in Education, 9*, 1328769.

Mosly, I. (2024). Artificial intelligence's opportunities and challenges in engineering curricular design: A combined review and focus group study. *Societies, 14*(6), 89.

Mukul, E., & Büyüközkan, G. (2023). Digital transformation in education: A systematic review of education 4.0. *Technological Forecasting and Social Change, 194*, 122664.

Naseer, F., Khan, M. N., Tahir, M., Addas, A., & Aejaz, S. H. (2024). Integrating deep learning techniques for personalized learning pathways in higher education. *Heliyon, 10*(11), e32628.

Ngo, T. T. A. (2024). Perception of engineering students on social constructivist learning approach in classroom. *International Journal of Engineering Pedagogy, 14*(1), 20–38.

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education, 48*(4), 559–614.

Okoye, K., Hussein, H., Arrona-Palacios, A., Quintero, H. N., Ortega, L. O. P., Sanchez, A. L., Ortiz, E. A., Escamilla, J., & Hosseini, S. (2023). Impact of digital technologies upon teaching and learning in higher education in Latin America: An outlook on the reach, barriers, and bottlenecks. *Education and Information Technologies, 28*(2), 2291–2360.

Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education, 20*(1), 4.

Pham, T., Nguyen, T. B., Ha, S., & Ngoc, N. T. N. (2023). Digital transformation in engineering education: Exploring the potential of AI-assisted learning. *Australasian Journal of Educational Technology, 39*(5), 1–19.

Rahiman, H. U., & Kodikal, R. (2024). Revolutionizing education: Artificial intelligence empowered learning in higher education. *Cogent Education, 11*(1), 2293431.

Sailer, M., Ninaus, M., Huber, S. E., Bauer, E., & Greiff, S. (2024). The end is the beginning is the end: The closed-loop learning analytics framework. *Computers in Human Behavior, 158*, 108305.

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education.

Information, 15(10), 596.

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications, 12*(1), 1–21.

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity, 2021*(1), 8812542.