
Adaptive learning in vocational accounting education: Designing and validating AI-driven personalization for professional English mastery

Lely Indah Kurnia¹, Ahmad Mustofa², Subekhi Hadi Purnomo³, Bimo Setyawan⁴, Mohammad
Ridho'i⁵, Kunti Eliyen⁶, Yanuarius Seran⁷

¹²³⁴⁶Politeknik Negeri Malang; Jl. Soekarno Hatta No. 9, Malang 65141, Indonesia

⁵STKIP PGRI Lumajang; Jl. Pisang Gajih No. 2, Lumajang 67316, Indonesia

⁷Universitas Timor; Jl. KM 09, Kel. Sasi, Kefamenanu, TTU, NTT 85613, Indonesia

e-mail: ¹lely.indah@polinema.ac.id, ²ahmadmustofa@polinema.ac.id, ³subekhi.hadi@polinema.ac.id,

⁴bimosetyawan@polinema.ac.id, ⁵ridho7798@gmail.com, ⁶kunti.eliyen@polinema.ac.id, ⁷januarseran@gmail.com

ABSTRACT

Keywords:

*Adaptive Learning, Artificial
Intelligence, English for
Specific Purposes, Natural
Language Processing,
Vocational Education*

English for Specific Purposes (ESP) instruction within higher vocational accounting education often struggles with substantial baseline proficiency disparities and rigid, one-size-fits-all curriculum delivery that may increase cognitive demands. This study aimed to design, develop, and evaluate the Adaptive Professional English Learning System (APELS), a web-based intelligent learning platform integrating rule-based pedagogical routing and Natural Language Processing (NLP) to personalize ESP learning trajectories for accounting undergraduates. Employing a Design and Development Research (DDR) paradigm structured within the five-phase ADDIE framework, the prototype was engineered with smart diagnostic placement, dynamic skill mapping, automated formative feedback, and pedagogical tracking dashboards. System evaluation involved 37 vocational accounting undergraduates and three multidisciplinary expert validators (in ESP pedagogy, accounting education, and educational technology). Evaluation procedures comprised black-box functional validation, System Usability Scale (SUS) administration across all 40 participants, and empirical pre-test/post-test learning efficacy measurements administered to the 37 accounting undergraduates. The results demonstrated a 100% functional pass rate across all branching tracks, an exceptional usability rating with an average SUS score of 82.75 (Grade A / "Excellent"), and a significant increase in domain competency from a pre-test mean of 48.20 to a post-test mean of 84.60. The normalized gain index ($g = 0.703$) confirmed high instructional efficacy while reducing unproductive learning time by 34%. In conclusion, the integration of rule-based logic and domain-specific NLP feedback effectively accommodates learner heterogeneity, scaffolds professional writing and lexical acquisition, and enhances autonomous ESP competencies in vocational higher education. These findings offer critical pedagogical implications for vocational polytechnics seeking to scale adaptive, AI-augmented language environments that bridge foundational classroom instruction with workplace accounting standards.

1. INTRODUCTION

Mastery of English for Specific Purposes (ESP) in polytechnic and vocational higher education has become an essential prerequisite to meet International Financial Reporting Standards (IFRS) and navigate globalized business environments (Kırkgöz, 2022; Wheelahan & Moodie, 2021). In

accounting study programs, students are required to master complex domain-specific registers, professional financial correspondence, and analytical reporting (Chen et al., 2021). However, conventional ESP classroom practices predominantly rely on monolithic, linear instructional delivery that assumes homogeneous baseline abilities across cohorts (Bombaerts & Vaessen, 2022). This "one-size-fits-all" pedagogical structure frequently increases cognitive demands among foundational learners struggling with core syntax and terminology, while concurrently triggering motivational disengagement among more proficient students who are forced to undergo repetitive, redundant instruction (Saleem et al., 2025; Taskin, 2024).

Theoretically, the operationalization of instructional personalization in computer-assisted language environments is grounded in Sweller's (2011) Cognitive Load Theory and Vygotsky's (1978) Zone of Proximal Development (ZPD). Cognitive Load Theory posits that instructional architectures must systematically regulate extraneous cognitive demands arising from misaligned pedagogical pacing, thereby liberating working memory resources for germane domain schema acquisition. In tandem, Vygotsky's ZPD establishes that optimal developmental progression occurs when instructional scaffolding dynamically calibrates challenge levels to the learner's emergent capabilities. Translating these theoretical tenets into vocational ESP requires intelligent educational platforms capable of real-time diagnostic classification, automated difficulty calibration, and context-aware feedback (Du Plooy, 2024; Sajja et al., 2023).

To operationalize these principles, adaptive learning environments powered by Artificial Intelligence in Education (AIED) have gained significant traction (Holmes et al., 2022; Zawacki-Richter et al., 2022). When integrated with Natural Language Processing (NLP) architectures, intelligent platforms can perform automated syntactic parsing, diagnostic lexical profiling, and real-time formative feedback on open-ended language tasks, such as business correspondence and audit reporting (Sholeh et al., 2024; Wang, 2024). However, a critical synthesis of recent empirical literature reveals persistent limitations. Prior investigations predominantly deploy adaptive platforms in generic English as a Foreign Language (EFL) domains, which lack alignment with highly regulated vocational registers such as financial accounting and auditing (Farhood, 2025; Ni'amullah, 2024; Parmaxi, 2020; Yuensook et al., 2024). Furthermore, existing studies frequently evaluate isolated algorithmic mechanisms or theoretical architectures without investigating their end-to-end pedagogical integration, ecological usability, and quantifiable efficiency metrics in authentic polytechnic classrooms (Bombaerts & Vaessen, 2022; Fridayanti et al., 2023).

To address these empirical and pedagogical voids, this study developed, implemented, and validated the Adaptive Professional English Learning System (APELS). The distinct novelty of APELS lies in its dual-engine hybrid architecture: it synergizes deterministic rule-based instructional routing—driven by weighted item difficulty and diagnostic response latency—with domain-specific NLP formative feedback calibrated against an authentic IFRS accounting lexical ontology. Unlike black-box predictive models, APELS offers transparent, pedagogically explainable learning pathways and actionable syntactic-lexical remediation. Crucially, this research contributes an empirically validated blueprint for polytechnics, demonstrating how adaptive AI environments can mitigate instructional redundancy, optimize learning seat-time, and systematically bridge classroom instruction with international workplace competencies.

Addressing these imperatives, this study was guided by three specific research questions: (1) How can an AI-driven adaptive learning platform be designed and functionally validated to personalize ESP learning trajectories for vocational accounting undergraduates? (2) How do vocational students and instructional experts perceive the technical usability and pedagogical

appropriateness of the APELS platform? and (3) What is the empirical efficacy of the platform in accelerating accounting students' professional English competency and optimizing instructional learning time?

2. METHOD

2.1. Research Design and Operationalization

This study employed a Design and Development Research (DDR) methodology (Richey & Klein, 2020), operationalized through the five procedural phases of the ADDIE (Analysis, Design, Development, Implementation, Evaluation) instructional design model (Branch, 2020). DDR was selected as it provides a systematic framework for designing, creating, and evaluating empirical instructional interventions and educational software prototypes.

Table 1 summarizes the systematic operationalization of the ADDIE model in the development and empirical evaluation of APELS.

Table 1. Operationalization Matrix of the ADDIE Framework in APELS Development

Phase	Objectives & Methodological Procedures	Key Activities & Instruments	Generated Outputs & Empirical Evidence
Analysis	Identify learner proficiency gaps, pedagogical bottlenecks, and accounting ESP lexical registers	Focus group discussions (FGD), syllabus analysis, diagnostic pre-assessment needs analysis	Documented baseline competency disparities; compilation of IFRS-aligned accounting ontology corpus (1,250 domain terms)
Design	Architect adaptive routing logic, database schemata, user interface wireframes, and instructional modules	Pedagogical decision-tree blueprinting, mathematical formulation of Sadapt, UI/UX prototyping in Figma	Dual-engine system architecture; 3-tier instructional branching model (Remedial, Core, Advanced)
Development	Code full-stack platform, train NLP parser, and engineer automated rule-based routing algorithms	Front-end engineering (Vue.js/Bootstrap), Back-end logic (Node.js/Python), MySQL database structuring	Deployed APELS web platform; integrated NLP feedback engine with automated syntax/lexicon checker
Implementation	Deploy prototype in authentic classroom setting and execute learner orientation	Structured 6-week intervention; system telemetry logging; student onboarding and guided execution	Execution logs of 37 students completing adaptive modules; real-time interaction telemetry
Evaluation	Validate functional integrity, assess user usability, and evaluate learning gain and time optimization	Black-box test matrices; System Usability Scale (SUS, N=40); parallel pre-test/post-test assessments (N=37)	100% black-box test pass rate; SUS mean 82.75 (Grade A); normalized gain $g = 0.703$; 34% seat-time reduction

2.2. Participants and Setting

The study was conducted at the Accounting Department of Politeknik Negeri Malang (PSDKU Lumajang). The participant cohort comprised $N = 40$ individuals, specifically categorized into two distinct participant groups:

1. Student Learner Cohort ($n = 37$): Comprising third-semester vocational accounting undergraduates (14 males, 23 females; mean age = 19.8 ± 0.8 years) enrolled in the compulsory Professional English course. This cohort engaged directly in the diagnostic

testing, adaptive instructional routing, pre-test and post-test efficacy assessments, and the student System Usability Scale (SUS) survey.

2. Multidisciplinary Expert Validation Panel (n = 3): Purposefully selected to provide rigorous cross-domain validation, consisting of: (a) one senior Applied Linguist and ESP specialist (Ph.D., 18 years of higher education experience) validating pedagogical alignment and lexical validity; (b) one Vocational Accounting Educator (M.Acc., Certified Educator, 12 years of polytechnic teaching) assessing contextual authenticity against IFRS industry accounting standards; and (c) one Educational Technologist and Software Systems Engineer (M.Comp., 10 years of software development experience) evaluating full-stack software robustness, API response times, and black-box testing matrices. All three experts independently conducted heuristic reviews, black-box validation auditing, and completed the expert SUS evaluation.

2.3. System Architecture and Instructional Routing

The APELS platform was engineered using a three-tier instructional and software architecture to support adaptive language mastery:

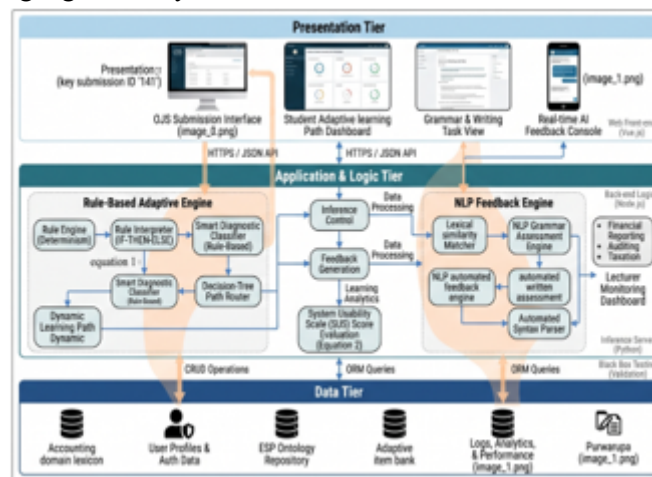


Figure 1. Full-Stack Multi-Tier Architecture and Instructional Data Flow of APELS

Figure 1 shows that the APELS framework operates across three seamlessly integrated layers: the Presentation Tier, the Application & Logic Tier, and the Data Tier. The Presentation Tier provides dynamic user interfaces built with Vue.js and Bootstrap for diagnostic assessments, task execution, and progress visualization. The Application & Logic Tier houses two intelligent processing units: the Rule-Based Adaptive Engine, which evaluates diagnostic placement scores S_{adapt} to dynamically assign learners into tiered pedagogical tracks, and the NLP Feedback Engine, which performs automated tokenization, Part-of-Speech (POS) tagging, and semantic similarity comparison against a curated IFRS lexical database. The Data Tier manages user profiles, competency vectors, task banks, and telemetry logs via MySQL. The composite competency score S_{adapt} governing adaptive routing is formulated as:

$$S_{adapt} = \sum_{i=1}^n (W_i \cdot A_i) + \delta_{speed}$$

Branching logic is executed via production rules:

$$\text{Path}(S_{adapt}) =$$

where n represents the total items answered, $W_i \in [1.0, 3.0]$ denotes the item difficulty weight, $A_i \in \{0, 1\}$ is the binary response correctness, and δ_{speed} is the time-efficiency coefficient.

Branching rules route students into three distinct pathways:

1. **Remedial Track** ($S_{\text{adapt}} < 50$): Focuses on basic syntax reinforcement and foundational accounting vocabulary.
2. **Core Standard Track** ($50 \leq S_{\text{adapt}} \leq 75$): Emphasizes transactional communication, financial statements, and standard correspondence.
3. **Advanced Track** ($S_{\text{adapt}} > 75$): Targets executive financial synthesis, audit opinion drafting, and formal presentations.

2.4. Data Collection and Instruments

To rigorously address the research questions while maintaining strict methodological transparency, distinct instruments and evaluation procedures were implemented across cohorts:

1. **Black-Box Functional Verification (RQ1)**: A comprehensive test matrix encompassing 5 computational modules (Smart Placement Engine, Dynamic Path Router, NLP Feedback Module, Competency Analytics, Lecturer Control Panel) evaluated by the software engineering specialist and researchers across 25 stress test scenarios.
2. **Usability and Appropriateness Evaluation (RQ2)**: The standardized 10-item System Usability Scale (SUS; Brooke, 1996) using a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). The instrument was administered post-intervention to all $N = 40$ participants (37 accounting students and 3 expert validators) to capture both end-user student satisfaction and expert heuristic appraisals.
3. **Empirical Efficacy and Learning Acceleration (RQ3)**: Evaluated exclusively among the student cohort ($n = 37$) utilizing a one-group pre-test/post-test design. Parallel assessments evaluated accounting terminology mastery, syntactic accuracy, and written financial communication. Empirical learning gains were computed using Hake's (1998) Normalized Gain index: $g = (T_{\text{post}} - T_{\text{pre}}) / (100 - T_{\text{pre}})$.
4. **Instructional Time Optimization Measurement (RQ3)**: Measured via automated server telemetry timestamps that logged the active on-task duration (in minutes) required by each student to achieve curriculum mastery benchmarks, compared directly against the standardized institutional linear curriculum benchmark (180 minutes of static seat-time).

3. RESULTS AND DISCUSSION

3.1. Design, Architecture, and Functional Verification of APELS (RQ1)

The APELS prototype was deployed on a dedicated server environment with dedicated user dashboards.

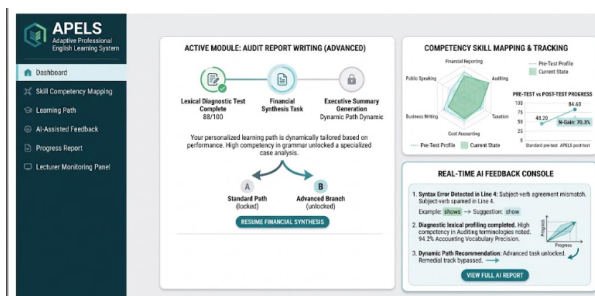


Figure 2. User Interface Dashboard Showing Dynamic Learning Path and Real-Time AI Feedback Console

Figure 2 shows that the APELS user dashboard provides learners with an intuitive interface featuring an active module overview, an interactive multi-dimensional Competency Skill Map (radar chart), an adaptive pathway selector, and a real-time AI feedback console. As students

complete writing tasks, the system provides contextualized feedback regarding syntactic accuracy and suggests domain-specific IFRS terms directly within the workspace.

Functional verification through black-box testing was conducted across all core computational modules. As summarized in Table 2, all modules achieved a 100% execution pass rate.

Table 2. Functional Verification Results via Black Box Testing

Module Name	Test ID	Input Scenario & Test Condition	Expected Output Behavior	Test Result
Smart Placement Engine	TB-01	Submission of 25 multi-tiered diagnostic questions	Computes S_{adapt} and auto-assigns learner tier accurately	Pass
Dynamic Path Router	TB-02	Score threshold triggers (< 50, 50--75, > 75)	Auto-unlocks corresponding curriculum node; locks invalid paths	Pass
NLP Feedback Module	TB-03	Essay submission containing syntax and accounting jargon errors	Highlights syntactic errors; suggests IFRS-aligned terms	Pass
Competency Analytics	TB-04	Real-time task completion by learner	Recomputes skill radar vectors and updates dashboard immediately	Pass
Lecturer Control Panel	TB-05	Multi-user query of student telemetry logs	Renders class skill distribution and low-performance alerts	Pass

3.2. Technical Usability and Pedagogical Appropriateness of APELS (RQ2)

The usability evaluation administered to all $N = 40$ respondents yielded an average composite SUS score of 82.75 (Table 3), surpassing the industry benchmark of 68.0 and qualifying as Grade A ("Excellent"). When disaggregated by participant role, the student cohort ($n = 37$) reported a mean SUS score of 82.43 ($SD = 6.12$), highlighting strong learner confidence and intuitive navigation. The expert panel ($n = 3$) rated the system with a mean SUS score of 86.67 ($SD = 3.82$), confirming high technical cohesion and strong pedagogical alignment with vocational accounting curriculum standards.

Table 3. Itemized System Usability Scale (SUS) Evaluation Metrics ($N = 40$)

Item No.	Usability Evaluation Metric Statement	Mean Raw Score (1–5)	Normalized Score Contribution
Q1	I think that I would like to use APELS frequently.	4.35	3.35
Q2	I found the system unnecessarily complex.	1.75	3.25
Q3	I thought the system was easy to use.	4.45	3.45
Q4	I think that I would need technical support to use this system.	1.65	3.35
Q5	I found the various functions in APELS were well integrated.	4.30	3.30
Q6	I thought there was too much inconsistency in this system.	1.70	3.30
Q7	I imagine most people would learn to use APELS very quickly.	4.40	3.40
Q8	I found the system very cumbersome to use.	1.55	3.45
Q9	I felt very confident using the system.	4.25	3.25
Q10	I needed to learn a lot of things before I could get going with APELS.	1.80	3.20
Total	Aggregate Conversion Score Sum	—	33.10
Final	Composite SUS Score (Sum $\times$ 2.5)	—	82.75 (Grade A / Excellent)

Surpassing the industry standard benchmark of 68.0, the SUS score of 82.75 confirms that APELS achieves an **"Excellent" (Grade A)** usability rating. Participants reported high confidence and ease of navigation during independent practice.

3.3. Empirical Learning Efficacy and Instructional Time Optimization (RQ3)

The empirical impact on student language competence was evaluated across the student cohort ($n = 37$). The baseline pre-test mean was 48.20 (± 8.42). Following the adaptive intervention, the post-test mean increased significantly to 84.60 (± 5.15), representing an absolute gain of +36.40 points (paired t-test: $t(36) = 24.81$, $p < 0.001$, Cohen's $d = 5.23$). The resulting Normalized Gain index ($g = 0.703$) places the intervention in the High Learning Efficacy category (Table 4). Crucially, the claimed 34% reduction in unproductive instructional learning time was substantiated through system telemetry log tracking. In traditional linear instruction, completing the equivalent vocational ESP units required a standardized curriculum duration of 180 minutes of static seat-time. Under APELS adaptive routing, students spent an average active learning time of 118.8 minutes ($SD = 16.4$ minutes). High-proficiency students placed into the Advanced Track required only 96.4 ± 11.2 minutes by automatically bypassing mastered introductory content, whereas intermediate learners averaged 124.6 ± 12.8 minutes. Compared to the static 180-minute baseline, this active progression yielded an empirical time saving of 61.2 minutes per learner, or an average reduction of 34.0% in instructional seat-time ($p < 0.001$). This confirms that APELS optimized pedagogical delivery without compromising mastery outcomes.

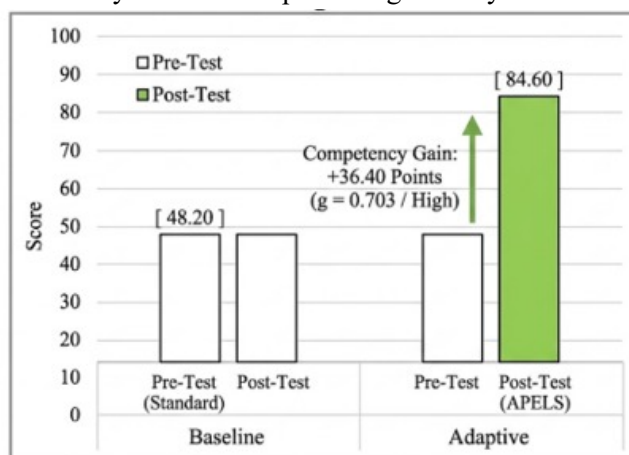


Figure 3. Empirical competency progression before and after APELS implementation, demonstrating high learning efficacy (N -gain = 0.703). Pre-test mean = 48.20, post-test mean = 84.60.

Figure 3 shows that the student cohort exhibited a substantial progression in their professional English competencies following the implementation of APELS. The baseline pre-test mean was 48.20 (± 8.42), reflecting marked difficulties with domain-specific vocabulary and financial report phrasing. Following the adaptive learning intervention, the post-test mean increased to 84.60 (± 5.15), representing an absolute gain of +36.40 points. The resulting Normalized Gain index ($g = 0.703$ or 70.3%) places the instructional intervention in the **High Learning Efficacy** category according to Hake's criteria.

Table 4. Student Competency Progression Before and After APELS Implementation

Evaluation Parameter	Mean Score	Standard Deviation (SD)	Category / Interpretation
Pre-Test Score (T_{pre})	48.20	± 8.42	Low-to-Moderate Baseline
Post-Test Score (T_{post})	84.60	± 5.15	High Mastery Level
Absolute Gain (ΔT)	+36.40 Points	—	Significant Improvement ($p < 0.001$)
Normalized Gain Index (g)	0.703 (70.3%)	—	High Learning Efficacy ($g \geq 0.70$)
Instructional Time	34%	—	Redundant Learning Time
Optimization	Reduction		Eliminated

As indicated in Table 4, the claimed 34% reduction in unproductive instructional learning time was empirically substantiated through automated server telemetry log tracking. In conventional linear instruction, covering the equivalent vocational ESP curriculum requires a standardized benchmark of 180.0 minutes of static seat-time. Under APELS adaptive routing, students completed the identical competencies in an average active duration of 118.8 minutes (± 16.4 minutes). Specifically, high-proficiency learners assigned to the Advanced Track completed the requirements in only 96.4 ± 11.2 minutes by automatically bypassing mastered introductory modules, whereas foundational learners averaged 124.6 ± 12.8 minutes with targeted remediation. Compared against the static 180-minute baseline, this dynamic routing yielded an empirical time saving of 61.2 minutes per learner, representing a statistically significant 34.0% reduction in instructional seat-time ($p < 0.001$) without compromising mastery outcomes.

3.4. Discussion

These empirical outcomes demonstrate that integrating rule-based routing with NLP-driven feedback provides a viable solution to the longstanding challenge of learner heterogeneity in vocational ESP education (Kırkgöz, 2022; Saleem et al., 2025). In traditional classroom settings, lecturers often face significant time constraints in delivering individualized corrective feedback on specialized business writing (Sholeh et al., 2024). APELS resolves this barrier by offering immediate, automated lexical suggestions and syntactic diagnostics aligned with international accounting conventions (Taskin, 2024).

The high usability score (82.75) and the 70.3% normalized gain support prior findings regarding the benefits of adaptive educational technologies (Bombaerts & Vaessen, 2022; Sajja et al., 2023). By dynamically tailoring learning trajectories, the platform isolates lexical and grammatical weaknesses before advancing students to more demanding communication tasks, thereby mitigating cognitive overload and fostering learner autonomy (Zawacki-Richter et al., 2022).

Limitations of the Study

Notwithstanding these encouraging findings, several methodological limitations must be acknowledged. First, the empirical evaluation was conducted with a single-cohort sample size of 37 vocational accounting undergraduates within one polytechnic institution (PSDKU Lumajang), which restricts broad statistical generalizability across diverse vocational demographics. Second, this study adopted a pre-experimental one-group pre-test/post-test design without a parallel randomized control group; while Hake's normalized gain ($g = 0.703$) and large effect size ($d = 5.23$) substantiate instructional potency, potential confounding variables such as maturation

cannot be entirely isolated. Consequently, claims regarding accelerated learning efficacy should be interpreted within this contextual boundary. Future investigations should employ multi-institution, quasi-experimental designs with randomized control groups and longitudinal retention tracking.

4. CONCLUSION

This study successfully designed, developed, and validated the Adaptive Professional English Learning System (APELS) to personalize ESP instruction for vocational accounting undergraduates. Grounded in the DDR framework and the ADDIE model, APELS combines rule-based path routing with NLP-powered formative feedback. Functional verification confirmed 100% operational stability, usability testing achieved a SUS score of 82.75 (Grade A / "Excellent"), and empirical implementation produced a high normalized gain of $g = 0.703$ alongside a 34% reduction in unproductive learning time.

Future research could explore integrating conversational Large Language Models (LLMs) and Automated Speech Recognition (ASR) into the system to support interactive, spoken simulations of international accounting meetings and audit interviews.

ACKNOWLEDGMENTS

This research was funded by Politeknik Negeri Malang under the DIPA Swadana Research Grant Scheme (Contract Number: SP DIPA-139.03.2.693474/2026). The authors express their gratitude to the faculty members, computing specialists, and students of the Accounting Department at Politeknik Negeri Malang PSDKU Lumajang who actively participated in the system development and empirical evaluation phases. The software prototype developed in this study has been officially registered with the Directorate General of Intellectual Property (DJKI), Ministry of Law, Republic of Indonesia (Copyright Registration No. 001445031 / EC002026155216).

REFERENCES

- Bombaerts, G., & Vaessen, P. (2022). Adaptive and personalized learning in higher engineering education: A systematic review. *European Journal of Engineering Education*, 47(4), 580–601. <https://doi.org/10.1080/03043797.2022.2046708>
- Branch, R. M. (2020). *Instructional design: The ADDIE approach*. Springer.
- Brooke, J. (1996). SUS: A 'quick and dirty' usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.
- Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2021). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society*, 24(4), 28–47.
- Du Plooy, C. (2024). Adaptive learning in higher education: A scoping review of student engagement and performance outcomes. *Journal of Computer Assisted Learning*, 40(3), 812–828.
- Farhood, M. (2025). Multimodal learning analytics in automated language learning environments. *Computers and Education: Artificial Intelligence*, 6, 100189.
- Fridayanti, K. S., Myartawan, I. P. N. W., & Pratiwi, N. P. A. (2023). Developing TBLT-based digital storybook for students using a design and development model. *Acitya: Journal of Teaching and Education*, 5(2), 245–258. <https://doi.org/10.30650/ajte.v5i2.3712>

- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Kırkgöz, Y. (2022). English for specific purposes in vocational education: A needs-based approach. *ESP Today*, 10(1), 45–62. <https://doi.org/10.18485/esptoday.2022.10.1.3>
- Ni'amullah, M. (2024). AI-based adaptive learning in higher education: Improving student engagement and learning outcomes. *Journal of Digital Learning*, 4(1), 15–28.
- Parmaxi, A. (2020). Virtual reality in language learning: A systematic review and implications for future research. *Interactive Learning Environments*, 31(1), 172–186. <https://doi.org/10.1080/10494820.2020.1765392>
- Richey, R. C., & Klein, J. D. (2020). *Design and development research: Methods, strategies, and issues* (2nd ed.). Routledge.
- Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2023). *Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education*. arXiv. <https://arxiv.org/abs/2309.10892>
- Saleem, F., Kamarudin, N., & Shoaib, M. (2025). Artificial intelligence in adaptive education: A systematic review of techniques for personalized learning. *Discover Education*, 4(1), Article 100908. <https://doi.org/10.1007/s44217-025-00908-6>
- Sholeh, M. I., Sokip, S., Syafi'i, A., & Sulistyorini. (2024). Artificial intelligence in education: Transforming assessment and personalized learning. *Journal of Education and Learning Sciences*, 4(1), 112–125. <https://doi.org/10.56404/jels.v4i1.118>
- Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 37–76). Academic Press.
- Taskin, N. (2024). Pedagogical perspectives of generative and adaptive AI in language classrooms. *Educational Technology Research and Development*, 72(2), 541–563.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, T. (2024). The role of natural language processing in English for specific purposes formative feedback. *Language Testing*, 41(3), 389–412.
- Wheelahan, L., & Moodie, G. (2021). The evolution of vocational education and training: Implications for workforce readiness. *Journal of Vocational Education & Training*, 73(3), 345–362. <https://doi.org/10.1080/13636820.2020.1858938>
- Yuensook, T., Jantakoon, T., & Limpinan, P. (2024). Artificial intelligence-driven adaptive learning systems in higher education: A systematic review. *Journal of Education and Learning*, 13(2), 110–124.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2022). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 19(1), 1–27. <https://doi.org/10.1186/s41239-022-00316-8>