



EVALUASI PENGGUNAAN *LARGE LANGUAGE MODEL (LLM)* DALAM PENGEMBANGAN *PERSONALIZED LEARNING* *SYSTEM (PLS)* DENGAN KERANGKA PEDAGOGIS

SHYFA KANAYA ZULKIFLI



**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Evaluasi Penggunaan *Large Language Model* (LLM) dalam Pengembangan *Personalized Learning System* (PLS) dengan Kerangka Pedagogis” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2026

Shyfa Kanaya Zulkifli
G6401221035

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

SHYFA KANAYA ZULKIFLI. Evaluasi Penggunaan *Large Language Model* (LLM) dalam Pengembangan *Personalized Learning System* (PLS) dengan Kerangka Pedagogis. Dibimbing oleh YENI HERDIYENI dan MEDRIA KUSUMA DEWI HARDHIENATA.

Perkembangan *Large Language Model* (LLM) membuka peluang baru dalam pengembangan *Personalized Learning Systems* (PLS), namun efektivitas pedagogisnya dalam pembelajaran *Computational Thinking* (CT) masih belum banyak dievaluasi secara sistematis. Penelitian ini bertujuan menganalisis dan mengevaluasi kerangka pedagogis pemanfaatan LLM untuk mendukung PLS pada pembelajaran CT. Penelitian ini menggunakan pendekatan evaluatif kuantitatif dengan *dataset* sintetis yang terdiri atas 144 respons, merupakan hasil kombinasi berbagai level pedagogis dan profil kognitif pengguna terhadap soal CT jenjang SD, SMP, dan SMA. Agen LLM menghasilkan 144 respons berbasis model Gemma 3 27B. Seluruh respons dievaluasi menggunakan metode *LLM-as-a-Judge* (GPT-5.2), lalu divalidasi oleh pakar manusia (*human expert judges*) terhadap 48 sampel terpilih menggunakan *purposive sampling*. Hasil penelitian menunjukkan bahwa seluruh respons yang dihasilkan sesuai dengan profil pengguna tanpa ditemukan respons kosong maupun tidak relevan. Evaluasi *LLM-as-a-Judge* menghasilkan 90,97% respons dengan skor pedagogis tertinggi pada keseluruhan *dataset*. Analisis kesesuaian antara *LLM-as-a-Judge* dan *human expert judges* pada 48 sampel menunjukkan tingkat kesepakatan yang tinggi, khususnya pada level pedagogis dasar hingga menengah. Hasil evaluasi mencatatkan tingkat akurasi berkisar antara 70,80%-95,80% dan koefisien Gwet's AC2 antara 0,65-0,95, dengan variasi yang dipengaruhi oleh acuan evaluator manusia. Temuan ini menunjukkan bahwa integrasi profil kognitif dan level pedagogis pada LLM mampu menghasilkan pembelajaran yang adaptif. Selain itu, penelitian ini juga menunjukkan bahwa pendekatan *LLM-as-a-Judge* berpotensi menjadi mekanisme evaluasi pedagogis yang reliabel dan memiliki skalabilitas tinggi, meskipun tetap memerlukan keterlibatan evaluator manusia untuk konteks penilaian dengan kompleksitas dan abstraksi tinggi.

Kata kunci: *Computational thinking*, evaluasi pedagogis, *large language model*, *personalized learning systems*, pembelajaran adaptif.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRACT

SHYFA KANAYA ZULKIFLI. Evaluation of Large Language Model (LLM) Utilization in the Development of a Personalized Learning System (PLS) Using a Pedagogical Framework. Supervised by YENI HERDIYENI and MEDRIA KUSUMA DEWI HARDHIENATA.

The advancement of Large Language Models (LLMs) has opened new opportunities for the development of Personalized Learning Systems (PLS). However, their pedagogical effectiveness in Computational Thinking (CT) education has not yet been systematically evaluated. This study aims to analyze and evaluate a pedagogical framework for leveraging LLMs to support PLS in CT learning. A quantitative evaluative approach was employed, utilizing a synthetic dataset of 144 responses generated from combinations of various pedagogical levels and cognitive profiles applied to CT problems spanning elementary, junior high, and senior high school levels. Responses were generated by an LLM agent using the Gemma 3 27B model, subsequently evaluated through an LLM-as-a-Judge method employing GPT-5.2 across the entire 144 responses, and further validated by human expert judges on a purposively selected subset of 48 samples. The results indicate that all generated responses were consistent with their corresponding user profiles, with no empty or irrelevant responses observed. The LLM-as-a-Judge evaluation yielded the highest pedagogical scores for 90.97% of the responses across the full dataset. The agreement analysis between LLM-as-a-Judge and human expert judges on the 48-sample subset revealed varying levels of agreement depending on which human evaluator served as the reference standard, with accuracy ranging from 70.8% to 95.8% and Gwet's AC2 coefficients ranging from 0.655 to 0.953, indicating an overall high level of alignment between LLM-as-a-Judge and human expert judgments, particularly at lower to intermediate pedagogical levels. These findings suggest that the integration of cognitive profiles and pedagogical levels within LLMs can produce adaptive learning experiences. Furthermore, this study demonstrates that the LLM-as-a-Judge approach holds strong potential as a reliable and scalable mechanism for pedagogical evaluation, although human evaluator involvement remains necessary for assessment contexts involving high complexity and abstraction.

Keywords: Adaptive learning, computational thinking, large language model, pedagogical evaluation, personalized learning systems.



@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

EVALUASI PENGGUNAAN *LARGE LANGUAGE MODEL (LLM)* DALAM PENGEMBANGAN *PERSONALIZED LEARNING* *SYSTEM (PLS)* DENGAN KERANGKA PEDAGOGIS

SHYFA KANAYA ZULKIFLI

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Evaluasi Penggunaan *Large Language Model* (LLM) dalam Pengembangan *Personalized Learning System* (PLS) dengan Kerangka Pedagogis

Nama : Shyfa Kanaya Zulkifli

NIM : G6401221035

Disetujui oleh

Pembimbing 1:

Prof. Dr. Yeni Herdiyeni, S.Si., M.Kom.

Pembimbing 2:

Dr. Medria Kusuma Dewi Hardhienata,
S. Kom., Ph. D.

Diketahui oleh

Ketua Program Sarjana Ilmu Komputer:

Dr. Sony Hartono Wijaya, S.Kom., M.Kom.
NIP. 19810809 200812 1 002

Tanggal Ujian:
13 Juli 2026

Tanggal Lulus:
4 Agustus 2026



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

PRAKATA

Puji dan syukur penulis panjatkan ke hadirat Allah Subhanahu wa Ta'ala atas segala rahmat, karunia, dan pertolongan-Nya sehingga karya ilmiah yang berjudul “Evaluasi Penggunaan *Large Language Model* (LLM) dalam Pengembangan *Personalized Learning System* (PLS) dengan Kerangka Pedagogis” ini dapat diselesaikan. Penyusunan karya ilmiah ini tidak terlepas dari doa, dukungan, bantuan, serta kehadiran banyak pihak yang telah menemani penulis sepanjang proses penelitian. Untuk itu, penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Yeni Herdiyeni, S.Si., M.Kom. dan Dr. Medria Kusuma Dewi Hardhienata, S.Kom., Ph.D., selaku dosen pembimbing atas segala arahan, bimbingan, masukan, kesabaran, serta dukungan yang diberikan selama proses penelitian dan penyusunan karya ilmiah ini.
2. Pak Ahmad Ridha, S.Kom., M.S. selaku dosen penguji yang telah memberikan waktu, ilmu, arahan, kritik, dan saran yang membangun selama proses penyusunan dan penyelesaian karya ini.
3. Bunda dan Papa, yang selalu menjadi rumah pertama dan tempat pulang terbaik bagi penulis. Terima kasih atas doa yang tidak pernah terputus, kasih sayang yang tidak pernah berkurang, pengorbanan yang tidak pernah terhitung, serta dukungan moral maupun material yang memungkinkan penulis sampai pada titik ini. Segala pencapaian yang diraih penulis tidak akan pernah lepas dari peran dan cinta Bunda serta Papa.
4. Adik Sheza dan Danish, yang dengan cara-cara sederhana selalu menghadirkan tawa, kehangatan, dan semangat di tengah berbagai kesulitan yang penulis hadapi. Terima kasih karena telah menjadi teman bertumbuh dan pelipur lara dalam hari-hari yang tidak selalu mudah.
5. Uti Hindun, Alm. Akung, Alm. Kakek Ayah, dan Alm. Nenek Mamah, yang telah menjadi bagian penting dalam perjalanan hidup penulis sejak masa kecil. Terima kasih atas kasih sayang, perhatian, doa, serta nilai-nilai kehidupan yang telah ditanamkan.
6. Tante Nur, yang selalu menjadi teman berdiskusi, pendengar yang baik, dan tempat penulis berbagi cerita.
7. Kak Aiz, yang senantiasa hadir menemani penulis dengan ketulusan, kesabaran, dan kasih sayang yang begitu besar.
8. Della dan Hanin, yang telah menjadi tempat pulang sejak masa SMA. Terima kasih karena selalu hadir dalam berbagai cerita, pertumbuhan, kegagalan, dan pencapaian penulis.
9. Alista, Yumna, Bulan, Intan, dan Haikal, yang telah menemani perjalanan penulis sejak tingkat pertama di IPB. Terima kasih atas kebersamaan, bantuan, dukungan, tawa, dan kenangan yang telah mewarnai masa perkuliahan.
10. Meena, teman seperjuangan skripsi penulis, yang telah menjadi rekan berdiskusi dan berbagi suka dan duka selama proses penyusunan skripsi ini.

Bogor, Juli 2026

Shyfa Kanaya Zulkifli



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR ISI

DAFTAR TABEL	xvi
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN	xvi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	3
1.5 Ruang Lingkup	3
II TINJAUAN PUSTAKA	5
2.1 <i>Generative AI (Gen AI) dan Large Language Model (LLM)</i>	5
2.2 <i>Personalized Learning System (PLS)</i>	5
2.3 Kerangka Pedagogis	6
2.4 <i>Computational Thinking</i>	7
III METODE	9
3.1 Data Penelitian	9
3.2 Tahapan Penelitian	11
3.3 Peralatan Penelitian	21
IV HASIL DAN PEMBAHASAN	23
4.1 Hasil Generasi <i>Dataset</i> Respons LLM LogiCT	23
4.2 Evaluasi Kualitas Pedagogis Menggunakan Pendekatan <i>LLM-as-a-Judge</i>	24
4.3 Evaluasi Kualitas Pedagogis Menggunakan Pendekatan <i>human expert judges</i>	26
4.4 Perbandingan Evaluasi <i>human expert judges</i> dan <i>LLM-as-a-Judge</i>	28
V KESIMPULAN DAN SARAN	39
5.1 Kesimpulan	41
5.2 Saran	42
DAFTAR PUSTAKA	43
LAMPIRAN	47
RIWAYAT HIDUP	72



DAFTAR TABEL

1	Aspek pedagogis dan kognitif	9
2	Karakteristik soal <i>computational thinking</i>	10
3	Struktur <i>dataset</i>	11
4	Skala penilaian <i>response score</i>	16
5	Skala penilaian <i>pedagogic score</i>	16
6	Contoh hasil generasi respons LogiCT	23
7	Analisis performa evaluasi <i>LLM-as-a-Judge</i> per level	25
8	Analisis performa evaluasi <i>human expert judges</i> per level	27
9	Statistik deskriptif evaluasi <i>human expert judges</i> dan <i>LLM-as-a-Judge</i>	29
10	Analisis <i>agreement</i> berdasarkan skor	39

DAFTAR GAMBAR

1	Prinsip <i>teaching for how students learn</i> (dimodifikasi dari Department for Education Children and Young People (2015) <i>Tasmanian Government</i> . p. 9)	6
2	Tahapan penelitian	12
3	Distribusi nilai pedagogis menggunakan <i>LLM-as-a-Judge</i>	25
4	Distribusi nilai pedagogis <i>human expert judges</i>	27
5	Perbandingan skor <i>human</i> dan LLM pada Level 1	30
6	Perbandingan skor <i>human</i> dan LLM pada Level 2	31
7	Perbandingan skor <i>human</i> dan LLM pada Level 3	33
8	Perbandingan skor <i>human</i> dan LLM pada Level 4	34
9	Perbandingan skor <i>human</i> dan LLM pada Level 5	35
10	Perbandingan skor <i>human</i> dan LLM pada Level 6	37

DAFTAR LAMPIRAN

1	<i>Chain-of-thought</i> generasi <i>dataset</i>	48
2	<i>System prompt</i> evaluasi <i>LLM-as-a-Judge</i>	50
3	Perhitungan statistika deskriptif	53
4	Perhitungan kuantitatif	55
5	Perhitungan <i>agreement</i>	59
6	Perhitungan <i>Inner-Annotator Agreement</i> (IAA)	61
7	Hasil penilaian <i>LLM-as-a-Judge</i> serta <i>human expert judges</i>	64