PENGEMBANGAN SISTEM BANTU PENYELESAIAN MASALAH MATEMATIS SEKOLAH BERBASIS KECERDASAN BUATAN DENGAN PENDEKATAN *SCAFFOLDING*

SKRIPSI

Untuk memenuhi sebagian persyaratan

mencapai derajat Sarjana S-1

Program Studi Pendidikan Matematika



Diajukan Oleh:

ALIFIA KANIA RAMADHANI

NIM. 22104040040

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PROGRAM STUDI PENDIDIKAN MATEMATIKA

FAKULTAS ILMU TARBIYAH DAN KEGURUAN

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2026

HALAMAN PENGESAHAN

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS ILMU TARBIYAH DAN KEGURUAN
Jl. Marsda Adisucipto Telp. (0274) 513056 Fax. (0274) 586117 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-2131/Un.02/DT/PP.00.9/07/2026

Tugas Akhir dengan judul : PENGEMBANGAN SISTEM BANTU PENYELESAIAN MASALAH MATEMATIS SEKOLAH BERBASIS KECERDASAN BUATAN DENGAN PENDEKATAN *SCAFFOLDING*

yang dipersiapkan dan disusun oleh:

Nama : ALIFIA KANIA RAMADHANI
Nomor Induk Mahasiswa : 22104040040
Telah diujikan pada : Rabu, 08 Juli 2026
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR

Ketua Sidang

Iqbal Ramadani, M.Pd.
SIGNED

Valid ID: 6a572bc4a379e

Penguji I

Prof. Dr. Ibrahim, S.Pd., M.Pd.
SIGNED

Valid ID: 6a578c180f539

Penguji II

Burhanuddin Latif, M.Si.
SIGNED

Valid ID: 6a572a2cdd0ae

Yogyakarta, 08 Juli 2026
UIN Sunan Kalijaga
Dekan Fakultas Ilmu Tarbiyah dan Keguruan

Prof. Dr. Sigit Purnama, S.Pd.I., M.Pd.
SIGNED

Valid ID: 6a586b0d7a65c

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSETUJUAN SKRIPSI

Universitas Islam Negeri Sunan Kalijaga FM-STUINSK-BM-05-01/R0

SURAT PERSETUJUAN SKRIPSI

Hal : Persetujuan Skripsi
Lamp : 3 Eksemplar Skripsi

Kepada Yth.
Dekan Fakultas Ilmu Tarbiyah dan Keguruan
UIN Sunan Kalijaga Yogyakarta
Di Yogyakarta

Assalamu'alaikum Warahmatullahi Wabarakatuh

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi, serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Alifia Kania Ramadhani
NIM : 22104040040
Judul Skripsi : Pengembangan Sistem Bantu Penyelesaian Masalah Matematis Sekolah berbasis Kecerdasan Buatan dengan Pendekatan Scaffolding

sudah dapat diajukan kembali kepada Program Studi Pendidikan Matematika Fakultas Tarbiyah dan Keguruan UIN Sunan Kalijaga Yogyakarta untuk memperoleh gelar Sarjana Strata Satu dalam Pendidikan Matematika.

Dengan ini, kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqosyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Yogyakarta, 24 Juni 2026
Pembimbing

Iqbal Ramadani, M.Pd.
NIP. 19940125 202012 1 004

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

SURAT PERNYATAAN KEASLIAN SKRIPSI

SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini:

Nama : Alifia Kania Ramadhani

NIM : 22104040040

Prodi : Pendidikan Matematika

Fakultas : Ilmu Tarbiyah dan Keguruan

Dengan ini menyatakan bahwa skripsi saya yang berjudul "PENGEMBANGAN SISTEM BANTU PENYELESAIAN MASALAH MATEMATIS SEKOLAH BERBASIS KECERDASAN BUATAN DENGAN PENDEKATAN SCAFFOLDING" merupakan hasil penelitian saya sendiri, tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 23 Juni 2026

Yang menyatakan,



Alifia Kania Ramadhani

NIM 22104040040

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

"Dan bahwasanya seorang manusia tidak memperoleh selain apa yang telah diusahakannya."

(QS. An-Najm [53]: 39)

"*It always seems impossible until it's done*."

(Nelson Mandela)

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Alhamdulillahi Rabbil'alamin

Puji syukur senantiasa tercurahkan kepada Allah SWT, atas segala limpahan nikmat dan karunia-Nya sehingga skripsi ini dapat terselesaikan dengan baik. Shalawat dan salam senantiasa tercurahkan kepada Nabi Muhammad SAW.

Karya Tugas Akhir atau Skripsi ini penulis persembahkan kepada:

Almarhum Bapak Ahmadi Dwi Susanto dan Almarhumah Ibu Sri Tekani

yang telah memberikan doa, kasih sayang, dukungan, dan pengorbanan semasa hidup.

Adikku Jasmin Kania Nafizhah

dan Keluarga besar

yang selalu mendukung dan menyemangati.

Bapak ibu dosen

yang telah mendidik, membimbing, dan memberikan doa.

Serta almamaterku,

Program Studi Pendidikan Matematika

Fakultas Ilmu Tarbiyah dan Kependidikan

Universitas Islam Negeri Sunan Kalijaga Yogyakarta

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Alhamdulillahi Rabbil'alamin, puji syukur kepada Allah SWT yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penelitian dan penyusunan skripsi ini dengan baik. Shalawat serta salam tak lupa selalu teriring kepada Nabi Muhammad SAW yang senantiasa menjadi teladan manusia dan kita nantikan syafaatnya di hari kiamat nanti. Skripsi ini disusun untuk memenuhi persyaratan memperoleh gelar Sarjana Program Studi Pendidikan Matematika. Penyusunan skripsi ini tentu tidak terlepas dari dukungan, dorongan, bimbingan dan doa dari berbagai pihak. Oleh karena itu, dengan segala syukur dan kerendahan hati, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Noorhaidi, S.Ag., M.A., M.Phil., Ph.D., selaku Rektor Universitas Islam Negeri Sunan Kalijaga.
2. Ibu Prof. Dr. Sigit Purnama, S.Pd., M.Pd., selaku Dekan Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga.
3. Bapak Buhanuddin Latif, M.Si., selaku Ketua Program Studi Pendidikan Matematika Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga.
4. Ibu Nidya Ferry Wulandari, M.Pd., selaku Dosen Pembimbing Akademik.
5. Bapak Iqbal Ramadani, M.Pd., selaku Dosen Pembimbing Skripsi yang selalu memberikan dukungan, bimbingan, semangat, dan motivasi sehingga skripsi ini dapat terselesaikan.
6. Bapak Ibu Dosen Pendidikan Matematika Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga yang telah memberikan ilmu dan doa kepada penulis.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

7. Seluruh dosen dan karyawan Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga yang telah memberikan ilmu dan membantu kelancaran studi selama perkuliahan.
8. Pemilik *dataset* di platform Kaggle dan Hugging Face yang telah menyediakan *dataset* secara terbuka sehingga dapat dimanfaatkan sebagai sumber data dalam penelitian ini.
9. Orang tua tercinta yang semasa hidup telah memberikan doa, dukungan, dan motivasi kepada penulis dalam menjalankan amanah menuntut ilmu di UIN Sunan Kalijaga, dan adik penulis yang telah memberikan doa, dorongan, dan semangat agar penulis menyelesaikan skripsi, serta tak lupa kucing penulis Cancan yang selalu menemani dalam mengerjakan skripsi.
10. Sahabat penulis Nabila Jasmin N.H. yang menemani pada masa akhir perkuliahan, memberi semangat, dan membersamai penulis dalam menyelesaikan skripsi.
11. Pakde, bude, om, tante keluarga besar, nenek, dan kakek yang selalu memberikan doa dan semangat selama penulis menyelesaikan skripsi.
12. Arifa, Qissya, Laili, Maulyna, dan teman-teman pendidikan matematika Angkatan 2022 yang telah menyertai penulis sepanjang masa perkuliahan.
13. Segenap pihak yang tidak dapat disebutkan satu persatu.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

HALAMAN PENGESAHAN ii
HALAMAN PERSETUJUAN SKRIPSI iii
SURAT PERNYATAAN KEASLIAN SKRIPSI iv
MOTTO v
HALAMAN PERSEMBAHAN vi
KATA PENGANTAR vii
DAFTAR ISI ix
DAFTAR GAMBAR xi
DAFTAR TABEL xii
DAFTAR LAMPIRAN xiii
ABSTRAK xiv
BAB I PENDAHULUAN 15
A. Latar Belakang 15
B. Batasan Masalah 21
C. Rumusan Masalah 21
D. Tujuan Pengembangan 21
E. Manfaat Pengembangan 22
F. Spesifikasi Produk 22
G. Batasan Penelitian 23
H. Definisi Operasional 23
BAB II KAJIAN PUSTAKA 25
A. Landasan Teori 25
1. Masalah Matematis Semi-Rutin 25
2. *Technological Pedagogical Content Knowledge* (TPACK) 27
3. *Natural Language Processing* (NLP) 29
4. *Supervised Fine Tuning* (SFT) 31
5. *Direct Preference Optimization* (DPO) 33
6. *Scaffolding* dalam Pendidikan Matematika 34
B. Penelitian yang Relevan 35
C. Kerangka Konseptual 37
BAB III METODOLOGI PENELITIAN 39
A. Desain Penelitian 39
B. Prosedur Pengembangan 41

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

C. Instrumen Penelitian 46
D. Teknik Pengumpulan Data 46
E. Teknik Analisis Data 46
BAB IV HASIL DAN PEMBAHASAN 50
A. Hasil 50
1. *Problem Scoping* 50
2. *Data Acquisition* 51
3. *Data Exploration* 54
4. *Data Preparation* 55
5. *Modelling* 59
6. *Evaluation* 69
B. Pembahasan 80
BAB V PENUTUP 85
A. Kesimpulan 85
B. Saran 86
DAFTAR PUSTAKA 87
LAMPIRAN 94

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR GAMBAR

Gambar 2. 1. Framework TPACK 28
Gambar 2. 2. Diagram Kerangka Konseptual 38
Gambar 3. 1. CDAC AI Life Cycle 39
Gambar 3. 2. Tahapan adaptasi AI Life Cycle 40
Gambar 3. 3. Diagram Alur Pengembangan 45
Gambar 3. 4. Ilustrasi Diagram Evaluasi Deep Learning 48
Gambar 4. 1 Tampilan *dataset* SFT yang telah dikumpulkan 53
Gambar 4. 2 Tampilan *dataset* DPO yang telah dikumpulkan 53
Gambar 4. 3 Diagram Distribusi Persentase Materi *Dataset* SFT 54
Gambar 4. 4 Diagram Distribusi Persentase Materi *Dataset* DPO 56
Gambar 4. 5 Tampilan *dataset* SFT yang sudah diterjemahkan 55
Gambar 4. 6 Tampilan *dataset* DPO yang sudah diterjemahkan 56
Gambar 4. 7 Tampilan *Dataset* SFT setelah prapemrosesan 58
Gambar 4. 8 Tampilan *Dataset* DPO setelah prapemrosesan 58
Gambar 4. 9 *Library* yang sudah di-*install* atau tersedia 59
Gambar 4. 10 Tampilan kode SFT untuk import *library* 60
Gambar 4. 11 Kode *load base* model Llama 8B 60
Gambar 4. 12 Tampilan kode SFT untuk *load* LoRA dan penyesuaian GPU 61
Gambar 4. 13 Tampilan kode SFT untuk *load dataset* dan penyesuaian format *input* 61
Gambar 4. 14 Tampilan kode SFT untuk memulai *training* 62
Gambar 4. 15 Tampilan kode SFT untuk menyimpan model 62
Gambar 4. 16 Alur tahap *modelling* SFT 63
Gambar 4. 17 Kode untuk mengkonversi menjadi model SFT penuh 63
Gambar 4. 18 *Library* yang sudah di-*install dan terse*dia 64
Gambar 4. 19 Tampilan kode DPO untuk import *library* 64
Gambar 4. 20 Tampilan kode untuk *load base* model SFT 65
Gambar 4. 21 Tampilan kode untuk *load* LoRA 65
Gambar 4. 22 *System Prompt* untuk *output* model 66
Gambar 4. 23 *Load dataset* dan penyesuaian format *input* 66
Gambar 4. 24 Tampilan kode DPO untuk memulai *training* 67
Gambar 4. 25 Tampilan *save checkpoint* dan *save output* model 68
Gambar 4. 26 Alur tahap *modelling* DPO 68
Gambar 4. 27 Grafik Evaluasi Model SFT 69
Gambar 4. 28 Grafik Evaluasi Model DPO 71

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR TABEL

Tabel 3. 1. AI Project Scoping 41
Tabel 3. 2. Komponen *Library* 44
Tabel 3. 3. Interpretasi Performa Model 48
Tabel 3. 4. Interpretasi *Output* Metrik Evaluasi NLP 49
Tabel 4. 1 Jumlah perolehan *dataset* 53
Tabel 4. 2 Pembagian *dataset* SFT dan DPO 59
Tabel 4. 3 *Output* model SFT dalam menjawab soal 70
Tabel 4. 4 *Output* model DPO dalam menjawab soal 73

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR LAMPIRAN

Lampiran 1 Contoh *Dataset* SFT Setiap Materi 98
Lampiran 2 Contoh *Dataset* DPO Setiap Materi 102
Lampiran 3 Surat Keterangan Tema Skripsi 109
Lampiran 4 Surat Penunjukkan Pembimbing 110
Lampiran 5 Bukti Seminar Proposal 111
Lampiran 6 Curriculum Vitae 112
Lampiran 7 QR Code dan Tautan konfigurasi model pada Huggingface 113

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PENGEMBANGAN SISTEM BANTU PENYELESAIAN MASALAH MATEMATIS SEKOLAH BERBASIS KECERDASAN BUATAN DENGAN PENDEKATAN *SCAFFOLDING*

Oleh
Alifia Kania Ramadhani
22104040040

ABSTRAK

Pembelajaran matematika di Indonesia berfokus pada kemampuan siswa menyelesaikan masalah secara sistematis melalui pemahaman konsep dan penalaran. Namun, siswa masih mengalami kesulitan dalam menyelesaikan masalah matematis, terutama pada soal semi-rutin yang menuntut pemahaman konsep dan penyusunan strategi penyelesaian. Perkembangan kecerdasan buatan membuka peluang untuk mengembangkan sistem bantu pembelajaran yang mampu memberikan *scaffolding* atau petunjuk penyelesaian secara bertahap. Oleh karena itu, penelitian ini bertujuan mengembangkan prototipe sistem bantu penyelesaian masalah matematis semi-rutin tingkat menengah pertama berbasis kecerdasan buatan yang mampu memberikan petunjuk bertahap dalam bentuk teks.

Penelitian ini merupakan penelitian pengembangan dengan prosedur *AI Life Cycle*, meliputi *problem scoping, data acquisition, data exploration, data preparation, modelling,* dan *evaluation*. Sistem dikembangkan menggunakan model bahasa besar LLaMA 3.1 8B yang dilatih melalui dua tahap, yaitu *Supervised Fine-Tuning* (SFT) dan *Direct Preference Optimization* (DPO). *Dataset* SFT yang digunakan sebanyak 6723 dan *dataset* DPO sebanyak 3184. Evaluasi dilakukan terhadap keluaran model berdasarkan ketepatan langkah penyelesaian dan kualitas *scaffolding* yang dihasilkan.

Hasil penelitian menunjukkan bahwa evaluasi model mempertimbangkan beberapa metrik, yaitu BLEU sebesar 0,072, ROUGE sebesar 0,3126, dan Perplexity sebesar 3,4329, berdasarkan karakteristik *input* dan *output* yang dihasilkan. Analisis keluaran model menunjukkan bahwa sistem mampu memberikan petunjuk yang mengarah pada *conceptual scaffolding*, yaitu membantu pengguna memahami informasi dan konsep yang diperlukan dalam penyelesaian masalah, serta *procedural scaffolding*, yaitu membimbing urutan langkah penyelesaian secara bertahap. Namun, model masih mengalami kesulitan dalam menghasilkan langkah penyelesaian yang konsisten hingga memperoleh jawaban akhir yang benar. Hasil evaluasi menunjukkan akurasi sebesar 36,89% pada tahap SFT dan 27,13% pada tahap DPO. Dengan demikian, alat bantu pemecahan masalah matematika berbasis kecerdasan buatan yang dikembangkan belum mampu memberikan *scaffolding* secara optimal tetapi menunjukkan potensi sebagai pendukung pembelajaran matematika melalui pemberian petunjuk penyelesaian secara bertahap.

Kata Kunci: Kecerdasan Buatan (AI), masalah matematis semi-rutin, *scaffolding*

BAB I

PENDAHULUAN

A. Latar Belakang

Penyelesaian masalah matematis adalah komponen penting dalam pembelajaran matematika karena memuat penerapan konsep, prosedur, dan penalaran untuk memeroleh solusi dari suatu permasalahan matematis. Penyelesaian masalah matematis melibatkan pemilihan strategi yang tepat, seperti membuat model matematika, menggambar ilustrasi, melakukan pendekatan *trial and error*, berpikir analitis, atau menggunakan rumus yang relevan (Schoenfeld, 1985). Selaras dengan pernyataan NCTM (2020), penyelesaian masalah merupakan bagian yang tidak terpisahkan dari pembelajaran matematika sehingga siswa perlu memperoleh kesempatan untuk mengembangkan dan menerapkan berbagai strategi dalam menyelesaikan masalah matematika. Oleh karena itu, proses penyelesaian masalah tidak hanya berorientasi pada jawaban akhir, tetapi juga pada bagaimana siswa membangun pemahaman dan menempuh langkah-langkah penyelesaian secara bertahap.

Namun, dalam praktiknya, pembelajaran matematika di sekolah masih sering berpusat pada penyelesaian masalah matematis dalam soal dengan prosedur rutin tanpa mengintegrasikan strategi dari pemikiran siswa. Siswa belum terbiasa menghadapi soal yang menuntut perencanaan strategis, karena proses belajar yang terlalu terfokus pada penggunaan rumus sederhana atau jawaban akhir, bukan proses berpikirnya. Schoenfeld (2022) menegaskan bahwa pembelajaran matematika tidak semestinya berorientasi semata pada pencapaian jawaban akhir,

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

melainkan perlu memberikan ruang bagi peserta didik untuk mengembangkan kemampuan metakognitif, yaitu kesadaran dan pengendalian terhadap proses berpikir yang mereka lakukan. Tidak hanya menekankan pencapaian hasil akhir, tetapi juga menuntut siswa untuk mengatur diri, berdiskusi, dan merefleksikan proses penyelesaian masalah yang mereka lakukan. Namun, tantangan tetap ada, baik bagi guru yang harus berperan sebagai fasilitator maupun bagi siswa yang perlu beradaptasi dengan pola belajar yang lebih mandiri (Ssali et al., 2025).

Hasil penilaian Program for International Student Assessment (PISA) menunjukkan bahwa Indonesia menduduki peringkat ke-68 dari 81 negara dalam literasi matematika. Dengan skor 379, yang menunjukkan kesenjangan signifikan dibandingkan dengan rata-rata skor global yang sebesar 472 (OECD, 2023). Skor rata-rata Indonesia pada bidang matematika adalah 366, sedangkan untuk membaca sebesar 359 dan sains 383. Angka ini masih tertinggal dibandingkan rata-rata OECD yang masing-masing mencapai 472 untuk matematika, 476 untuk membaca, dan 485 untuk sains. Kesenjangan tersebut tampak jelas, misalnya dalam bidang matematika, hanya sekitar 18% siswa Indonesia yang mampu mencapai Level 2 atau lebih tinggi, jauh lebih rendah dibandingkan rata-rata OECD sebesar 69%. Lebih jauh lagi, hampir tidak ada siswa Indonesia yang mencapai Level 5 atau 6, yakni level yang menunjukkan kemampuan berpikir kritis dalam menerapkan matematika pada situasi yang kompleks.

Kondisi ini mengindikasikan bahwa sebagian besar siswa masih berada pada level dasar, di mana mereka hanya mampu menyelesaikan prosedur rutin dan soal sederhana. Rendahnya hasil ini tidak hanya menunjukkan keterbatasan kognitif

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

siswa, tetapi juga mencerminkan perlunya transformasi pendekatan pembelajaran matematika yang lebih mendalam dan strategis. Kesulitan ini tidak hanya tercermin dari hasil PISA, tetapi juga diperkuat oleh berbagai penelitian yang menemukan bahwa siswa mengalami hambatan dalam menyelesaikan soal-soal matematika non-rutin karena tidak terbiasa menghadapi tipe soal pemecahan masalah dalam proses belajar (Enlisia et al., 2020; Faizzah & Sutarni, 2023; Fitri & Abadi, 2021). Dalam konteks ini, soal semi-rutin menjadi penting sebagai jembatan antara soal rutin dan non-rutin.

Soal semi-rutin digunakan sebagai tahap transisi yang memungkinkan siswa mengembangkan kemampuan berpikir secara bertahap, khususnya dalam memahami konteks, menyesuaikan strategi, serta mengintegrasikan beberapa konsep yang telah dipelajari (Romberg, 1992; Silver, 1994). Namun, siswa sering mengalami kesulitan karena belum terbiasa menghadapi variasi soal yang membutuhkan penyesuaian strategi tersebut. Selain itu, siswa dengan kemampuan kognitif tinggi cenderung mampu mengelola beban kognitif dengan mengintegrasikan informasi secara efisien, sedangkan siswa dengan kemampuan kognitif sedang memerlukan proses bertahap untuk mengembangkan kemampuan tersebut (O. Chen et al., 2023). Oleh karena itu, diperlukan solusi inovatif yang mampu memberikan dukungan dalam proses penyelesaian masalah siswa secara bertahap, agar mereka dapat memahami dan menyelesaikan soal semi-rutin secara lebih sistematis dan mandiri.

Seiring berkembangnya teknologi, kecerdasan buatan (AI) menawarkan solusi yang menjanjikan untuk membantu siswa mengembangkan kemampuan berpikir

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

logis dan sistematis dalam memahami dan menyelesaikan masalah semi-rutin secara mandiri. Siswa menganggap bahwa AI dapat menjadi fasilitas pembelajaran yang efektif jika mampu menyajikan visualisasi yang mudah dipahami dan memberikan panduan dalam menyelesaikan masalah matematika (Soesanto et al., 2022). Penelitian lain menunjukkan bahwa AI dapat menjadi alat pembelajaran yang efektif dalam memperkuat kesadaran suatu topik melalui interaksi berbasis *chatbot* yang bersifat responsif dan dialogis (Harahap et al., 2025). Hal ini mengindikasikan bahwa penerapan AI dalam pendidikan tidak hanya berfungsi sebagai penyedia informasi, tetapi juga sebagai fasilitator yang mampu menumbuhkan pemahaman mendalam dan kemandirian belajar siswa.

Dalam konteks integrasi teknologi pendidikan, mengacu pada kerangka *Technological Pedagogical Content Knowledge* (TPACK) Mishra & Koehler (2006) yang merefleksikan pentingnya kolaborasi antara pengetahuan konten, pendekatan pedagogis, dan teknologi. Dalam konteks ini, kecerdasan buatan dikembangkan dan diintegrasikan sebagai inovasi teknologi yang mendukung peningkatan kualitas pembelajaran. Untuk memastikan integrasi kecerdasan buatan dalam pendidikan sesuai dengan elemen TPACK, guru dan calon pendidik harus dilatih dengan aktivitas pembelajaran yang secara eksplisit menghubungkan AI dengan pengetahuan pedagogis, konten, dan teknologi (Goldman et al., 2024). Hal itu sejalan juga dengan penelitian oleh Tillman et al. (2024) mengenai AI sebagai alat bantu dalam pendidikan harus dilandasi pemahaman yang kuat terhadap kerangka TPACK, agar teknologi tersebut dapat diterapkan secara efektif dan

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

bertanggung jawab. Penting untuk melihat lebih dekat penerapan kecerdasan buatan (AI) dalam pembelajaran matematika.

D. Zhang et al. (2025), menginvestigasi bagaimana tingkat literasi dan kepercayaan terhadap AI memengaruhi ketergantungan calon guru matematika terhadap *Generative AI* (Gen-AI), serta dampaknya terhadap keterampilan abad ke-21 seperti berpikir kritis, kemampuan pemecahan masalah, kreativitas, kolaborasi, komunikasi, dan kepercayaan diri. Temuan lain mengenai kepercayaan tinggi tetapi literasi rendah menunjukkan kecenderungan lebih besar untuk terlalu bergantung pada kecerdasan buatan, yang berimplikasi negatif pada keterampilan seperti berpikir kritis, pemecahan masalah, kreativitas, serta kolaborasi (Wijaya et al., 2024). Oleh karena itu, diperlukan strategi pembelajaran yang tidak hanya mengenalkan AI sebagai alat bantu, tetapi juga membekali calon guru dengan literasi AI yang memadai agar mampu menggunakan teknologi secara kritis. Dengan begitu, AI dapat difungsikan untuk memperkuat kemampuan berpikir, kreativitas, dan pemecahan masalah tanpa menimbulkan ketergantungan berlebihan.

Penggunaan pendekatan dalam pembelajaran matematika, seperti pembelajaran adaptif, umpan balik instan, visualisasi konsep abstrak, sistem rekomendasi soal yang disesuaikan dengan kemampuan siswa, dan sebagainya, AI memiliki potensi besar untuk membantu perkembangan kemampuan berpikir siswa (Lee et al., 2024; Sanusi et al., 2022). Melihat potensinya dalam mendukung efektivitas pembelajaran, pengembangan AI dalam konteks pendidikan matematika menjadi langkah penting untuk menghadirkan proses belajar yang lebih relevan dan

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

bermakna bagi siswa (Muawanah et al., 2024; Ramadhani & Ramadani, 2025). Seiring dengan pesatnya perkembangan teknologi, studi lebih lanjut mengenai pengembangan, implementasi, dan penerapan AI dalam pendidikan, khususnya pada pembelajaran matematika, menjadi semakin penting. Pemanfaatan sistem berbasis AI ini perlu diposisikan sebagai sistem bantu yang mampu memberikan dukungan bertahap sesuai kebutuhan siswa, guna membantu mereka mencapai pemahaman yang lebih mendalam dan mandiri dalam belajar.

Berbagai pengembangan sistem kecerdasan buatan sebelumnya umumnya berfokus pada aspek teknis, seperti ketepatan jawaban dan efisiensi algoritma, namun masih terbatas dalam memperhatikan dimensi pedagogis dalam interaksi belajar. Welleck et al. (2022) telah melakukan pengembangan menggunakan teknik *constrained decoding* untuk menghasilkan langkah-langkah penyelesaian yang relevan dan valid. Teknik ini mendukung proses penalaran dan pembuktian otomatis dengan mengandalkan referensi pengetahuan, seperti teori dan definisi matematika, sebagai dasar berpikir sistem. Beragam strategi seperti pembelajaran agen, evaluasi kinerja berbasis *benchmark*, dan teknik augmentasi data telah digunakan dalam pengembangan LLM untuk pemecahan masalah matematika (Fang et al., 2024; Tabib & Deedar, 2024; Xie et al., 2024). Namun, sebagian besar pengembangan tersebut masih terbatas pada penggunaan bahasa asing dan belum mengakomodasi konteks bahasa Indonesia serta kebutuhan pedagogis lokal. Berdasarkan kondisi tersebut, diperlukan pengembangan sistem bantu pembelajaran berbasis teks dengan kecerdasan buatan dalam bahasa Indonesia yang tidak hanya menyelesaikan soal, tetapi juga memberikan petunjuk secara bertahap.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Penelitian ini merupakan langkah awal dalam pengembangan prototipe sistem bantu berbasis kecerdasan buatan pada penyelesaian masalah matematis semi-rutin. Pengembangan sistem ke depan akan mencakup pemberian petunjuk penyelesaian serta memunculkan pertanyaan yang mendorong siswa untuk mengembangkan pengetahuan dan strategi penyelesaian mereka sendiri. Oleh karena itu, diharapkan sistem AI tidak hanya berfungsi sebagai alat untuk menemukan jawaban, tetapi juga membantu siswa belajar berpikir secara mandiri dengan membantu mereka berinteraksi dengan lebih adaptif dan berfokus pada proses belajar.

B. Batasan Masalah

1. Rendahnya kemampuan siswa dalam menyelesaikan soal-soal matematika semi-rutin yang menuntut pemahaman konsep, analisis informasi, dan penalaran strategis.
2. Pembelajaran yang berfokus pada soal rutin serta kurangnya dukungan belajar yang membantu proses berpikir siswa secara sistematis dan mandiri.

C. Rumusan Masalah

Bagaimana mengembangkan prototipe sistem bantu penyelesaian masalah matematis semi-rutin tingkat menengah pertama berbasis kecerdasan buatan yang mampu memberikan petunjuk bertahap dalam bentuk teks?

D. Tujuan Pengembangan

Mengembangkan prototipe sistem bantu penyelesaian masalah matematis semi-rutin tingkat menengah pertama berbasis kecerdasan buatan yang mampu memberikan petunjuk bertahap dalam bentuk teks.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

E. Manfaat Pengembangan

1. Kegunaan Teoritis

Memberikan kontribusi pada pengembangan AI edukatif dalam bidang pendidikan matematika, khususnya sebagai landasan awal bagi pengembangan sistem bantu berbasis AI yang memberikan bimbingan bertahap untuk mendukung proses berpikir siswa.

2. Kegunaan Praktis

a. Menyediakan prototipe sistem bantu pembelajaran berbasis kecerdasan buatan yang mampu memberikan petunjuk penyelesaian secara bertahap dalam menyelesaikan masalah matematis.

b. Memberikan referensi bagi guru dan peneliti dalam pemanfaatan kecerdasan buatan sebagai sistem bantu pembelajaran tanpa menggantikan peran guru.

F. Spesifikasi Produk

1. Produk berupa prototipe AI atau kecerdasan buatan berbasis teks.
2. *Input*: *Dataset* soal masalah matematis dalam format teks yang berfokus pada materi matematika tingkat menengah pertama atau SMP bersifat semi-rutin yang disertai penyelesaian per langkah.
3. *Output*: Petunjuk yang disusun berdasarkan urutan logis penyelesaian masalah.
4. Basis teknologi yang digunakan adalah *Natural Language Processing* (NLP), *Supervised Fine Tuning* (SFT), dan *Direct Preference Optimization* (DPO).
5. Sistem evaluasi berbasis eksperimen komputasi (*Computer Experiment*).

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

G. Batasan Penelitian

1. Fokus pada soal materi matematika sekolah tingkat menengah pertama dengan penyelesaian yang bersifat semi-rutin.
2. Tidak mencakup pengolahan *input* visual berupa grafik, gambar, atau tulisan tangan.
3. Tidak mengukur dampak penggunaan pada siswa secara langsung.

H. Definisi Operasional

1. Soal Semi-Rutin

Dalam penelitian ini, soal semi-rutin diartikan sebagai soal matematika yang tidak dapat diselesaikan melalui penerapan langsung rumus atau prosedur yang telah diajarkan, melainkan menuntut siswa untuk memahami situasi masalah, merumuskan strategi penyelesaian secara mandiri, dan menggunakan penalaran matematis untuk menemukan solusi. Soal jenis ini biasanya tidak menyebutkan secara eksplisit apa yang harus dicari, dapat diselesaikan dengan berbagai pendekatan, serta memuat tantangan konseptual dan strategis yang mendorong keterampilan berpikir tingkat tinggi seperti analisis, evaluasi, dan refleksi.

2. Kecerdasan Buatan atau Artificial Intelligence (AI)

Dalam penelitian ini, kecerdasan buatan merujuk pada sistem berbasis komputer yang dirancang untuk meniru proses berpikir manusia, khususnya dalam memahami, menganalisis, dan memberikan respons terhadap *input* berbasis teks. Berbasis teknologi pembelajaran mesin dan NLP, AI ini dirancang untuk membimbing siswa menyelesaikan masalah matematika secara bertahap dan kontekstual.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

3. *Scafolded Hints*

Scaffolded hints adalah bentuk bantuan bertahap yang diberikan kepada siswa dalam proses penyelesaian masalah matematika. Dalam konteks penelitian ini, *scaffolded hints* diartikan sebagai petunjuk berbasis teks yang dihasilkan oleh sistem AI, yang disusun secara berjenjang untuk membantu siswa memahami masalah, mengidentifikasi langkah-langkah penyelesaian, dan membangun solusi secara mandiri, sesuai dengan prinsip *scaffolding* dalam pembelajaran.

4. *Dataset*

Dataset dalam penelitian ini merupakan kumpulan soal matematika sekolah tingkat menengah pertama yang dikembangkan atau dikurasi untuk merepresentasikan *problem solving* yang sesuai dengan kurikulum tingkat menengah. *Dataset* digunakan sebagai *input* utama untuk pelatihan dan pengujian model kecerdasan buatan dalam mengidentifikasi dan memberi petunjuk penyelesaian soal.

5. *Computer Experiments*

Eksperimen komputasi (*computer experiments*) dalam penelitian ini merujuk pada metode pengujian performa model AI melalui simulasi berbasis komputer. Eksperimen dilakukan dengan memberikan berbagai variasi *input* teks kepada sistem, untuk mengevaluasi keandalan, efisiensi, dan kemampuan adaptif AI dalam memberikan petunjuk penyelesaian masalah. Pendekatan ini mengacu pada definisi yang dikemukakan oleh Sacks et al. (1989) mengenai penggunaan simulasi komputer sebagai pengganti eksperimen fisik dalam meneliti sistem kompleks.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

BAB V
PENUTUP

A. Kesimpulan

Berdasarkan hasil penelitian, sistem bantu pemecahan masalah matematika berbasis kecerdasan buatan yang dikembangkan belum mampu menghasilkan petunjuk pemecahan masalah secara optimal. Akurasi model pada tahap akhir sebesar 27.13% menunjukkan bahwa model masih mengalami kesulitan dalam menghasilkan langkah penyelesaian yang benar, bahkan mengalami penurunan dibandingkan tahap sebelumnya yang mencapai 36.89%. Hasil analisis keluaran model juga menunjukkan bahwa *scaffolding* yang dihasilkan masih belum baik dalam memberikan petunjuk pemecahan masalah matematika yang tepat dan konsisten. Sementara itu, hasil evaluasi menggunakan metrik BLEU, ROUGE-L, dan *Perplexity*. Nilai BLEU sebesar 0.072, ROUGE sebesar 0.3126, menunjukkan adanya kesesuaian struktur dan informasi dengan jawaban acuan, sedangkan nilai *Perplexity* sebesar 3.4329 mengindikasikan bahwa model mampu menghasilkan teks yang sesuai dengan pola data yang dipelajari atau petunjuk bertahap yang diinginkan. Dengan demikian, model yang telah dikembangkan telah menunjukkan kemampuan model dalam memberikan petunjuk bertahap, meskipun belum sepenuhnya diikuti oleh kemampuan *reasoning* yang benar secara logis. Penelitian ini menghasilkan prototipe AI *scaffolding* matematika berbahasa Indonesia yang memfasilitasi proses berpikir mandiri siswa melalui pertanyaan penuntun, sekaligus menjadi landasan bagi pengembangan sistem bantu edukasi selanjutnya.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

B. Saran

Beberapa hal yang dapat dijadikan sebagai saran dari penelitian yang telah dilakukan adalah sebagai berikut.

1. Pengembangan lebih lanjut diharapkan dapat menggunakan *dataset* yang lebih besar.
2. Pengembangan lebih lanjut diharapkan dapat menggunakan metode yang lebih baik dan sesuai dengan pemecahan masalah matematika.
3. Pengembangan lebih lanjut diharapkan dapat mengembangkan sistem AI yang tidak hanya memberikan petunjuk penyelesaian, tetapi juga mampu menghasilkan pertanyaan-pertanyaan penuntun (*guiding questions*) yang mendorong siswa berpikir secara mandiri.
4. Pengembangan lebih lanjut diharapkan dapat menghasilkan akurasi yang lebih tinggi dan dapat disebarluaskan.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR PUSTAKA

Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2019). Software Engineering for Machine Learning: A Case Study. *Proceedings - 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice, ICSE-SEIP 2019*, 291–300. https://doi.org/10.1109/ICSE-SEIP.2019.00042

Amini, A., Vieira, T., & Cotterell, R. (2024). Direct Preference Optimization with an Offset. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 9954–9972. https://doi.org/10.18653/V1/2024.FINDINGS-ACL.592

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching and Assessing A Revision of Bloom's Taxonomy of Educational Objectives* (Complete Edition). Longman. https://www.scirp.org/reference/referencespapers?referenceid=1223916

Boros, G. (2023). *Do you know overfitting and underfitting?* Medium | Data Science Europe. https://medium.com/@datascienceeurope/do-you-know-overfitting-and-underfitting-f27f87ac2f37

Brookhart, S. M. (2010). How to Assess Higher-Order Thinking Skills in Your Classroom. In *Ascd*. Association for Supervision and Curriculum Development (ASCD). http://www.ascd.org/Publications/Books/Overview/How-to-Assess-Higher-Order-Thinking-Skills-in-YourClassroom.aspx

Cappuccio, E., Kathirgamanathan, B., Rinzivillo, S., Andrienko, G., & Andrienko, N. (2025). Integrating human knowledge for explainable AI. *Machine Learning* , *114*(11), 250-. https://doi.org/10.1007/S10994-025-06879-X

Chen, G., Liao, M., Li, C., & Fan, K. (2024). Step-level Value Preference Optimization for Mathematical Reasoning. *EMNLP 2024 - 2024 Conference on Empirical Methods in Natural Language Processing, Findings of EMNLP 2024*, 7889–7903. https://doi.org/10.18653/v1/2024.findings-emnlp.463

Chen, O., Paas, F., & Sweller, J. (2023). A Cognitive Load Theory Approach to Defining and Measuring Task Complexity Through Element Interactivity. *Educational Psychology Review* , *35*(2), 63-. https://doi.org/10.1007/S10648-023-09782-W

Chollet, F. (2023). *Training & evaluation with the built-in methods*. TensorFlow Core. https://www.tensorflow.org/guide/keras/training_with_built_in_methods

Chowdhary, K. R. (2020). Natural Language Processing. In *Fundamentals of Artificial Intelligence* (pp. 603–649). Springer India. https://doi.org/10.1007/978-81-322-3972-7_19

Church, K. W., Chen, Z., & Ma, Y. (2021). Emerging trends: A gentle introduction to fine-tuning. *Natural Language Engineering*, *27*(6), 763–778. https://doi.org/10.1017/S1351324921000322

Davies, P. M., Passonneau, R. J., Muresan, S., & Gao, Y. (2022). Analytical Techniques for Developing Argumentative Writing in STEM: A Pilot Study.

IEEE Transactions on Education, *65*(3), 373–383. https://doi.org/10.1109/TE.2021.3116202

De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, *3*(6), 100489. https://doi.org/10.1016/J.PATTER.2022.100489

Enlisia, A. P., Rahardjo, S., & Sisworo, S. (2020). Kesulitan Siswa Dalam Menyelesaikan Masalah Matematika Berdasarkan Langkah Polya. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, *5*(12), 1820. https://doi.org/10.17977/jptpp.v5i12.14347

Faizzah, S. N., & Sutarni, S. (2023). Investigasi Kesulitan Siswa dalam Menyelesaikan Masalah HOTS Matematika. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *7*(2), 1963–1975. https://doi.org/10.31004/cendekia.v7i2.2438

Fang, M., Wan, X., Lu, F., Xing, F., & Zou, K. (2024). MathOdyssey: benchmarking mathematical problem-solving skills in large language models using odyssey math data. *ArXiv: Computation and Language*. https://arxiv.org/pdf/2406.18321

Fitri, A., & Abadi, A. M. (2021). Kesulitan siswa SMA dalam menyelesaikan soal matematika pada materi peluang. *Jurnal Riset Pendidikan Matematika*, *8*(1), 96–105. https://doi.org/10.21831/jrpm.v8i1.17004

Forootani, A. (2025). A Survey on Mathematical Reasoning and Optimization with Large Language Models. *IEEE Transactions on Artificial Intelligence*, *0*(0). https://arxiv.org/pdf/2503.17726

Gardner, J. (2012). *Assessment and Learning*. SAGE Publications Ltd. https://doi.org/10.4135/9781446250808

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media, Inc.

Gliksman, Y., Berebbi, S., & Henik, A. (2022). Math Fluency during Primary School. *Brain Sciences*, *12*(3), 371. https://doi.org/10.3390/brainsci12030371

Goldman, S., Carreon, A., & Smith, S. (2024). Exploring the integration of artificial intelligence into special education teacher preparation through the TPACK framework. *Journal of Special Education Preparation*, *4*(2), 52–64. https://doi.org/10.33043/6zx26bb2

Goodfellow, Ian., Bengio, Yoshua., & Courville, Aaron. (2016). *Deep learning*. The MIT Press.

Gugger, Sylvain., & Howard, Jeremy. (2020). *Deep Learning for Coders with fastai and PyTorch*. O'Reilly Media, Inc.

Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques. In *Data Mining: Concepts and Techniques* (3rd ed., 3). Elsevier. https://doi.org/10.1016/B978-0-12-381479-1.00003-4

Hao, Y., Pan, X., Zhang, H., Ye, C., Pan, R., & Zhang, T. (2025). Understanding Overadaptation in Supervised Fine-Tuning: The Role of Ensemble Methods. *Artificial Intelligence*. https://arxiv.org/pdf/2506.01901

Harahap, S. S., Salim, R. R. M., & Harjatanaya, T. Y. (2025). Pemanfaatan AI (Artificial Intelligence) dalam Pengenalan Kebhinekaan Indonesia di Sekolah.

Jurnal PkM (Pengabdian Kepada Masyarakat), *7*(6), 825. https://doi.org/10.30998/jurnalpkm.v7i6.26411

Hogan, K., & Pressley, M. (1997). *Scaffolding student learning: Instructional approaches and issues*. Brookline Books.

Huan, M., Li, Y., Zheng, T., Xu, X., Kim, S., Du, M., Poovendran, R., Neubig, G., & Yue, X. (2025). Does Math Reasoning Improve General LLM Capabilities? Understanding Transferability of LLM Reasoning. *Artificial Intellegence*. https://doi.org/10.9

Im, S., & Li, S. (2024). Can DPO Learn Diverse Human Values? A Theoretical Scaling Law. *Machine Learning*. https://arxiv.org/pdf/2408.03459

Jiang, Y., & Wang, Z. (2025). Scaffolding Students' Problem-Solving Process in Case-Based Teaching: A Multimodal Analysis. *Sage Open*, *15*(4). https://doi.org/10.1177/21582440251410652

Kachroo, D., Caraeni, A., Anbazhagan, A. P., Lagasse, B., & Zhu, K. (2026). HiPO: Hierarchical Preference Optimization for Adaptive Reasoning in LLMs. *Artificial Intelegence*. https://arxiv.org/pdf/2604.20140

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. https://doi.org/10.1016/J.LINDIF.2023.102274

Kim, M. C., & Hannafin, M. J. (2011). Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. *Computers & Education*, *56*(2), 403–417. https://doi.org/10.1016/J.COMPEDU.2010.08.024

Krulik, Stephen., & Rudnick, J. A. (1996). *The new sourcebook for teaching reasoning and problem solving in junior and senior high school*. Allyn and Bacon. https://books.google.com/books/about/The_New_Sourcebook_for_Teaching_Reasonin.html?hl=id&id=DqcLAQAAMAAJ

Lee, K.-H., Shin, D., Park, J., Kim, Y., & Lee, S. H. (2024). Exploring the use of ChatGPT in mathematical problem solving competency education: An analysis of error diagnosis and feedback performance in solving word problems on division of fractions. *Korean School Mathematics Society*, *27*(4), 639–666. https://doi.org/10.30807/ksms.2024.27.4.008

Li, F.-F., Johnson, J., & Yeung, S. (2019). *Course CS231n: Convolutional Neural Networks for Visual Recognition*. Course CS231n Spring. https://cs231n.stanford.edu/2019/

Li, L., Liu, H., Zhou, Y., Gui, Z., Weng, X., Yuan, Y., Wei, Z., & Li, Z. (2025). Uncertainty-Aware Iterative Preference Optimization for Enhanced LLM Reasoning. *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 23996–24012. https://doi.org/10.18653/v1/2025.acl-long.1169

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Lin, C. Y. (2004). ROUGE: A Package for Automatic Evaluation of Summaries. *Text Summarization Branches Out*, 74–81. https://aclanthology.org/W04-1013/

Lu, C., & Chen, W. (2025a). Alignment of Teachers' TPACK and Practices. In *Unpacking Technological Pedagogical Content Knowledge for Classroom Practice* (pp. 163–197). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-8193-8_8

Lu, C., & Chen, W. (2025b). Teachers' TPACK Knowledge. In *Unpacking Technological Pedagogical Content Knowledge for Classroom Practice* (pp. 117–132). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-8193-8_5

Lu, S., Cheng, J., Xu, Y., Yu, Y., Sheng, L., Wang, P., Jiang, S., Hu, Y., Ling, R., Shao, Y., Ma, A., Feng, W., He, L., Wang, M., Xie, Q., Wang, X., Sebe, N., Liang, J., He, L.-I., & He, R. (2026). Do MLLMs Really Understand Space? A Mathematical Reasoning Evaluation. *Proceedings of the 34th ACM International Conference on Multimedia (ACM MM '26)*, *1*. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Lu, Z., Zhou, A., Wang, K., Ren, H., Shi, W., Pan, J., Zhan, M., & Li, H. (2024). Step-Controlled DPO: Leveraging Stepwise Error for Enhanced Mathematical Reasoning. *Computation and Language*, null. https://arxiv.org/pdf/2407.00782

Marchis, L. (2012). Non-routine problems in primary mathematics workbooks from Romania. *Acta Didactica Napocensia: Cluj-Napoca*, *5*(3), 49–56.

Mavrepis, P., Makridis, G., Fatouros, G., Koukos, V., Separdani, M. M., & Kyriazis, D. (2024). *XAI for All: Can Large Language Models Simplify Explainable AI?* https://arxiv.org/pdf/2401.13110

Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record: The Voice of Scholarship in Education*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Muawanah, U., Marini, A., & Sarifah, I. (2024). The interconnection between digital literacy, artificial intelligence, and the use of E-learning applications in enhancing the sustainability of Regional Languages: Evidence from Indonesia. *Social Sciences & Humanities Open*, *10*, 101169. https://doi.org/10.1016/j.ssaho.2024.101169

NCTM. (2020). Catalyzing Change in Early Childhood and Elementary Mathematics: Initiating Critical Conversations. In *National Council of Teachers of Mathematics*. National Council of Teachers of Mathematics.

OECD. (2022). *PISA 2022: Mathematics Framework*. OECD Publishing. https://pisa2022-maths.oecd.org/ca/index.html?

OECD. (2023). *PISA 2022 Results (Volume I)*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Oh, J., Yoo, H., Haznitrama, F. G., & Oh, A. (2026). OLA: Output Language Alignment in Code-Switched LLM Interactions. *Computation and Language*. https://arxiv.org/pdf/2601.03589

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, *35*. https://arxiv.org/pdf/2203.02155

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2001). BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*, 311. https://doi.org/10.3115/1073083.1073135

Provost, Foster., & Fawcett, Tom. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.

Ramadhani, A. K., & Ramadani, I. (2025). AI in mathematics education: Potential ranging from automation to personalized learning. *LINEAR: Journal of Mathematics Education*, *6*(1), 85–99. https://doi.org/10.32332/46xc9p64

Romberg, T. A. (1992). Perspectives on scholarship and research methods. In *Handbook of research on mathematics teaching and learning* (3rd). Macmillan Publishing Co, Inc.

Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical Science*, *4*(4). https://doi.org/10.1214/ss/1177012413

Santner, T. J., Williams, B. J., & Notz, W. I. (2018). *The Design and Analysis of Computer Experiments*. Springer New York. https://doi.org/10.1007/978-1-4939-8847-1

Sanusi, I. T., Olaleye, S. A., Agbo, F. J., & Chiu, T. K. F. (2022). The role of learners' competencies in artificial intelligence education. *Computers and Education: Artificial Intelligence*, *3*, 100098. https://doi.org/10.1016/j.caeai.2022.100098

Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Elsevier. https://doi.org/10.1016/C2013-0-05012-8

Schoenfeld, A. H. (2022). Why Are Learning and Teaching Mathematics So Difficult? In *Handbook of Cognitive Mathematics* (pp. 1–35). Springer International Publishing. https://doi.org/10.1007/978-3-030-44982-7_10-1

Scott-Emuakpor, C. (2025). *Maximizing Scalable AI: Efficient Language Model Adaptation Using Fine-Tuning, Direct Preference Optimization, and Online Reinforcement*. https://doi.org/10.36227/TECHRXIV.175427207.75884899/V1

Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of Data Warehousing*, *5*(4), 13–22.

Silver, E. A. (1994). On mathematical problem posing. *For the Learning of Mathematics*, *14*(1), 19–28.

Soesanto, R. H., Dirgantoro, K. P. S., & Priyanti, N. (2022). Indonesian students' perceptions towards AI-based learning in mathematics. *Journal on Mathematics Education*, *13*(3), 531–548. https://doi.org/10.22342/jme.v13i3.pp531-548

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Ssali, C., Kabuye Batiibwe, M. S., Dahl, B., Ampeire Kariisa, H., Magero, J. M., & Mayende, G. (2025). Problem-based learning in secondary school mathematics: a review. *Educational Research*, *67*(2), 212–230. https://doi.org/10.1080/00131881.2025.2493255

Stein, M. Kay., Smith, M. S., Henningsen, M., & Silver, E. A. (2009). *Implementing standards-based mathematics instruction : a casebook for professional development* (2nd ed.). Teachers College Press.

Sundaram, S. S., Gurajada, S., Fisichella, M., P, D., & Abraham, S. S. (2022). Why are NLP Models Fumbling at Elementary Math? A Survey of Deep Learning based Word Problem Solvers. *Computation and Language*. https://arxiv.org/pdf/2205.15683

Tabib, S., & Deedar, J. A. (2024). End-to-End bangla aI for solving math olympiad problem benchmark: leveraging large language model using integrated approach. *International Journal on Natural Language Computing (IJNLC)*, *13*(5), 49–59. https://doi.org/10.5121/ijnlc.2024.13604

Tian, C., Lu, Z., Qian, S., Liu, N., Li, P., Jin, L., Hu, L., Zeng, Z., Wang, S., Zeng, K., & Cas, G. (2026). Rectify Evaluation Preference: Improving LLMs' Critique on Math Reasoning via Perplexity-aware Reinforcement Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *40*(39), 33241–33249. https://doi.org/10.1609/AAAI.V40I39.40609

Tillman, D. A., Cochran, C., Soto, T. J., Robertson, W. H., An, S. A., & Tinajero, J. V. (2024). Use of artificial intelligence to improve teacher preparation. *2024 Artificial Intelligence x Humanities, Education, and Art (AIxHEART)*, 41–45. https://doi.org/10.1109/AIxHeart62327.2024.00014

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models*. https://arxiv.org/pdf/2302.13971

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *LIX*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Urrutia, F., & Araya, R. (2023). Automatically Detecting Incoherent Written Math Answers of Fourth-Graders. *Systems 2023, Vol. 11, Page 353*, *11*(7), 353. https://doi.org/10.3390/SYSTEMS11070353

Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E. H., Le, Q. V., & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Advances in Neural Information Processing Systems*, *35*. https://arxiv.org/pdf/2201.11903

Welleck, S., Liu, J., Lu, X., Hajishirzi, H., & Choi, Y. (2022). NaturalProver: Grounded Mathematical Proof Generation with Language Models. *Advances in Neural Information Processing Systems*, *35*. https://arxiv.org/pdf/2205.12910

Wenbo, W., Chong, N., Le, J., Heng, L. C., & Ling, F. (2025). *Multiple Descents in Deep Learning as a Sequence of Order-Chaos Transitions*. https://arxiv.org/pdf/2505.20030

Wijaya, T. T., Yu, Q., Cao, Y., He, Y., & Leung, F. K. S. (2024). Latent Profile Analysis of AI Literacy and Trust in Mathematics Teachers and Their Relations with AI Dependency and 21st-Century Skills. *Behavioral Sciences*, *14*(11), 1008. https://doi.org/10.3390/bs14111008

Witten, I. H., Frank, E., & Hall, M. A. (2011). Data Mining: Practical Machine Learning Tools and Techniques, Third Edition. In *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition* (3rd ed.). Elsevier. https://doi.org/10.1016/C2009-0-19715-5

Wood, D. (1998). *How children learn*. Blackwell Publishers.

Wulff, P., Kubsch, M., & Krist, C. (2025). *Natural Language Processing and Large Language Models* (pp. 117–142). https://doi.org/10.1007/978-3-031-74227-9_7

Xie, W., Liu, D., Yan, H., Wu, W., & Liu, Z. (2024). MathLearner: a large language model agent framework for learning to solve mathematical problems. *ArXiv: Computation and Language*. https://arxiv.org/pdf/2408.01779

Yeliz, Y. (2015). Sixth graders and non-routine problems: Which strategies are decisive for success? *Educational Research and Reviews*, *10*(13), 1807–1816. https://doi.org/10.5897/ERR2015.2230

Yuan, Z., Yuan, H., Li, C., Dong, G., Lu, K., Tan, C., Zhou, C., & Zhou, J. (2023). Scaling Relationship on Learning Mathematical Reasoning with Large Language Models. *Computation and Language*. https://arxiv.org/pdf/2308.01825

Zhang, D., Wijaya, T. T., Wang, Y., Su, M., Li, X., & Damayanti, N. W. (2025). Exploring the relationship between AI literacy, AI trust, AI dependency, and 21st century skills in preservice mathematics teachers. *Scientific Reports*, *15*(1), 14281. https://doi.org/10.1038/s41598-025-99127-0

Zhang, Y., & Teng, Z. (2022). Natural Language Processing: A Machine Learning Perspective. *Computational Linguistics*, *48*(1), 233–235. https://doi.org/10.1162/coli_r_00423

Zhou, Y., & Zhao, T. (2024). Dual Instruction Tuning with Large Language Models for Mathematical Reasoning. *Computation and Language*. https://arxiv.org/pdf/2403.18295

Zhou, Z., Yuhao, T., Li, Z., Yao, Y., Guo, L.-Z., Ma, X., & Li, Y.-F. (2025). Bridging Internal Probability and Self-Consistency for Effective and Efficient LLM Reasoning. *Machine Learning*. https://arxiv.org/pdf/2502.00511

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA