eL-Muhbib
Jurnal Pemikiran & Penelitian Pendidikan Dasar

P-ISSN: 2614-1051 & E-ISSN: 2716-5019

Volume 10, Nomor 1, Juni 2026

ARTIFICIAL INTELLIGENCE DALAM PENILAIAN MATEMATIKA: TINJAUAN PUSTAKA SISTEMATIS TENTANG TEKNOLOGI, FUNGSI, DAN DAMPAK PENDIDIKAN

Dewi Sartika[1*], Gede Ratnaya [2], I Dewa Ayu Made Budhyani [3]
Universitas Pendidikan Ganesha, Indoensia[1,2,3]

Corresponding Author: Dewi Sartika tikamamaurwa@gmail.com

ARTICLE INFO
Article history:
Received
19 Maret 2026
Revised
28 Mei 2026
Accepted
30 Juni 2026

ABSTRAK

Artificial Intelligence (AI) telah menjadi salah satu inovasi teknologi yang berpotensi mentransformasi asesmen matematika melalui otomatisasi penilaian, pemberian umpan balik, dan personalisasi pembelajaran. Namun, perkembangan pesat teknologi AI, khususnya Generative Artificial Intelligence (GenAI) dan Large Language Models (LLMs), menuntut adanya kajian komprehensif mengenai pemanfaatannya dalam asesmen matematika. Penelitian ini bertujuan untuk mengidentifikasi teknologi AI yang digunakan dalam asesmen matematika, menganalisis fungsi-fungsi AI yang diterapkan, mengevaluasi dampaknya terhadap proses dan hasil pembelajaran, serta mengkaji tantangan dan peluang implementasinya. Metode yang digunakan adalah Systematic Literature Review (SLR) dengan mengacu pada pedoman PRISMA 2020. Artikel diperoleh dari data Scopus dengan rentang publikasi tahun 2020–2026. Proses seleksi menghasilkan 20 artikel yang memenuhi kriteria inklusi untuk dianalisis. Hasil kajian menunjukkan bahwa teknologi AI yang dominan meliputi *Natural Language Processing*, *Intelligent Tutoring Systems*, *Learning Analytics*, *Automated Grading Systems, serta Generative AI*. Fungsi utama AI dalam asesmen matematika mencakup *automated grading, automated feedback, formative assessment, personalized assessment, dan predictive assessment*. Implementasi AI terbukti meningkatkan hasil belajar, kualitas umpan balik, efisiensi penilaian, dan personalisasi pembelajaran. Namun, tantangan terkait validitas asesmen, bias algoritma, integritas akademik, kesiapan guru, dan privasi data masih memerlukan perhatian. Temuan ini menunjukkan bahwa AI berpotensi menjadi mitra strategis dalam asesmen matematika yang lebih adaptif, efektif, dan berkelanjutan.

Kata Kunci: *Kecerdasan Buatan, Asesmen Matematika, Generative Artificial Intelligence, Tinjauan Literatur Sistematis*

How to Cit : Sartika, Dewi Sartika.,et.al. (2026). Kecerdasan Buatan Dalam Penilaian Matematika: Tinjauan Pustaka Sistematis Tentang Teknologi, Fungsi, Dan Dampak Pendidikan. *El. Muhbib: Jurnal Penelitian Dan Pemikiran Pendidikan Dasar*. Vol. 10, No. 1: 226-247
DOI : http:/doi.org/10.52266/el-muhbib.v10i1.6729
Journal Homepage: https://ejournal.iaimbima.ac.id/index.php/
This is an open access article under the CC BY SA license
: https://creativecommons.org/licenses/by-sa/4.0/

PENDAHULUAN

Perkembangan teknologi digital pada era Revolusi Industri 4.0 dan Society 5.0 telah mendorong transformasi signifikan dalam berbagai aspek pendidikan, termasuk proses asesmen pembelajaran. Salah satu inovasi yang berkembang

sangat pesat adalah *Artificial Intelligence* (AI), yaitu teknologi yang memungkinkan sistem komputer melakukan berbagai fungsi yang menyerupai kemampuan kognitif manusia, seperti analisis data, pengambilan keputusan, prediksi, dan pemberian rekomendasi secara otomatis (Chen et al., 2020). Dalam bidang pendidikan, AI tidak lagi dipandang sekadar sebagai alat bantu teknologi, tetapi sebagai sarana yang berpotensi mentransformasi proses pembelajaran, evaluasi, dan pengelolaan pendidikan secara lebih efektif dan efisien (Bond et al., 2024; Kamalov et al., 2023).

Asesmen merupakan komponen penting dalam proses pembelajaran karena berfungsi untuk mengukur pencapaian kompetensi, memantau perkembangan belajar, serta menyediakan informasi bagi pengambilan keputusan pembelajaran. Dalam konteks pembelajaran matematika, asesmen tidak hanya berfokus pada hasil akhir, tetapi juga menilai proses berpikir, kemampuan penalaran, strategi pemecahan masalah, dan pemahaman konseptual peserta didik (Hidayat et al., 2022; Hwang & Tu, 2021). Namun demikian, praktik asesmen matematika konvensional masih menghadapi berbagai kendala, seperti tingginya beban kerja guru dalam proses penilaian, keterlambatan pemberian umpan balik, serta keterbatasan dalam menyediakan asesmen yang sesuai dengan kebutuhan individual peserta didik (Hooda et al., 2022).

Perkembangan AI telah membuka peluang baru dalam pelaksanaan asesmen matematika. Berbagai teknologi seperti *Machine Learning (ML), Natural Language Processing (NLP), Intelligent Tutoring Systems (ITS), Learning Analytics,* dan *Generative Artificial Intelligence (GenAI)* mulai diintegrasikan ke dalam sistem asesmen pendidikan (Hwang & Tu, 2021; Awang et al., 2025). Teknologi tersebut memungkinkan proses penilaian dilakukan secara otomatis, penyediaan umpan balik secara real-time, analisis perkembangan belajar peserta didik secara berkelanjutan, hingga prediksi performa akademik berdasarkan data historis pembelajaran (Ahmed, 2024; Owan et al., 2023). Selain itu, kemunculan *Large Language Models* (LLMs) seperti ChatGPT semakin memperluas fungsi AI dalam asesmen melalui kemampuan menghasilkan soal, menyusun rubrik penilaian, dan memberikan umpan balik yang dipersonalisasi kepada peserta didik (Kooli & Yusuf, 2024; Xia et al., 2024).

Berbagai penelitian menunjukkan bahwa AI memiliki potensi besar dalam meningkatkan efektivitas asesmen pendidikan. González-Calatayud et al. (2021) mengemukakan bahwa AI mampu mendukung asesmen yang lebih adaptif dan efisien melalui otomatisasi proses evaluasi. Selanjutnya, Hooda et al. (2022) menegaskan bahwa penerapan AI dalam asesmen dan umpan balik dapat meningkatkan keberhasilan belajar peserta didik di pendidikan tinggi. Dalam pembelajaran matematika, penggunaan AI juga terbukti mampu meningkatkan kualitas umpan balik dan membantu guru dalam memonitor perkembangan belajar peserta didik secara lebih akurat (Botelho et al., 2023; Marwiang et al., 2025). Meskipun demikian, implementasi AI dalam asesmen matematika masih menghadapi berbagai tantangan. Xia et al. (2024) menemukan bahwa penggunaan Generative AI dalam asesmen menimbulkan persoalan terkait validitas penilaian dan integritas akademik. Selain itu, Yan et al. (2023) menyoroti berbagai tantangan etis dalam pemanfaatan model bahasa besar (Large Language Models), termasuk risiko bias algoritma, keamanan data, dan transparansi proses pengambilan keputusan. Temuan serupa juga dikemukakan oleh Clark et al. (2025),

yang menunjukkan bahwa sistem asesmen berbasis AI berpotensi menghasilkan bias terhadap kelompok peserta didik tertentu apabila data pelatihan yang digunakan tidak representatif.

Seiring perkembangan teknologi Generative AI yang sangat pesat sejak tahun 2023, lanskap penelitian mengenai asesmen pendidikan mengalami perubahan yang signifikan. Mao et al. (2023) menjelaskan bahwa Generative AI telah mengubah paradigma asesmen dari sekadar proses evaluasi hasil belajar menjadi proses yang mendukung pembelajaran secara berkelanjutan melalui pemberian umpan balik yang lebih personal dan interaktif. Namun demikian, sebagian besar penelitian sebelumnya masih berfokus pada implementasi AI dalam pendidikan secara umum (Chen et al., 2020; Bond et al., 2024) atau pendidikan matematika secara luas (Hidayat et al., 2022; Hwang & Tu, 2021), sehingga kajian yang secara khusus memetakan penggunaan AI dalam asesmen matematika masih relatif terbatas. Selain itu, penelitian yang ada cenderung mengkaji teknologi AI secara terpisah, sehingga belum memberikan gambaran komprehensif mengenai teknologi AI yang digunakan dalam asesmen matematika, fungsi-fungsi asesmen yang didukung, serta dampaknya terhadap proses dan hasil pembelajaran matematika. Padahal, pemahaman yang menyeluruh mengenai aspek-aspek tersebut sangat diperlukan untuk mendukung implementasi AI yang efektif, etis, dan berorientasi pada peningkatan kualitas pembelajaran matematika.

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk melakukan *Systematic Literature Review* (SLR) mengenai penggunaan Artificial Intelligence dalam asesmen matematika. Secara khusus, penelitian ini berupaya mengidentifikasi teknologi AI yang digunakan dalam asesmen matematika, menganalisis fungsi-fungsi utama AI dalam proses asesmen, mengevaluasi dampaknya terhadap pembelajaran matematika, serta mengidentifikasi tantangan dan peluang implementasinya pada masa depan. Hasil kajian ini diharapkan dapat memberikan kontribusi teoretis berupa pemetaan komprehensif perkembangan penelitian AI dalam asesmen matematika sekaligus menjadi dasar bagi pengembangan kebijakan dan praktik asesmen yang lebih adaptif, personal, dan berbasis data pada era kecerdasan buatan..

TINJAUAN TEORITIS

1. *Artificial Intelligence* dalam Pendidikan

Perkembangan *Artificial Intelligence* (AI) telah membawa perubahan signifikan dalam berbagai sektor, termasuk pendidikan. AI didefinisikan sebagai kemampuan sistem komputer untuk meniru proses kognitif manusia seperti belajar, bernalar, memecahkan masalah, dan mengambil keputusan. Dalam konteks pendidikan, AI digunakan untuk mendukung proses pembelajaran, personalisasi pengalaman belajar, analisis data pendidikan, serta asesmen hasil belajar siswa. Menurut Chen et al. (2020), AI telah berkembang menjadi salah satu teknologi paling berpengaruh dalam pendidikan karena mampu meningkatkan efektivitas pembelajaran melalui pemanfaatan data yang besar dan algoritma cerdas. AI memungkinkan sistem pendidikan memberikan layanan yang lebih adaptif, efisien, dan responsif terhadap kebutuhan peserta didik.

Berbagai penelitian menunjukkan bahwa penggunaan AI dalam pendidikan terus meningkat seiring berkembangnya teknologi machine learning, deep learning, *natural language processing (NLP), learning analytics, dan generative artificial intelligence*

(GenAI). Hwang dan Tu (2021) menjelaskan bahwa AI tidak hanya digunakan sebagai alat bantu pembelajaran, tetapi juga berfungsi sebagai sistem pendukung pengambilan keputusan dalam pendidikan. Perkembangan ini semakin dipercepat dengan hadirnya *Large Language Models* (LLM) seperti ChatGPT yang memungkinkan interaksi alami antara manusia dan mesin dalam lingkungan pembelajaran (Chiu, 2023).

Secara umum, implementasi AI dalam pendidikan mencakup berbagai fungsi, seperti sistem tutor cerdas *(Intelligent Tutoring Systems*), sistem rekomendasi pembelajaran, prediksi performa siswa, asesmen otomatis, serta pemberian umpan balik secara real-time. Oleh karena itu, AI dipandang sebagai teknologi yang mampu mentransformasi proses pembelajaran dari pendekatan yang bersifat seragam menjadi lebih personal dan adaptif terhadap karakteristik individu peserta didik (Paek & Kim, 2021).

2. Artificial Intelligence dalam Pendidikan Matematika

Pendidikan matematika merupakan salah satu bidang yang paling banyak mengadopsi teknologi AI karena karakteristik matematika yang membutuhkan latihan berulang, evaluasi berkelanjutan, serta analisis terhadap proses berpikir siswa. Hidayat et al. (2022) menyatakan bahwa AI dalam pendidikan matematika telah berkembang pesat melalui penerapan intelligent tutoring systems, adaptive learning systems, learning analytics, dan automated assessment systems. Hwang dan Tu (2021) mengidentifikasi bahwa penelitian AI dalam pendidikan matematika berfokus pada tiga area utama, yaitu peningkatan hasil belajar matematika, personalisasi pembelajaran, dan asesmen berbasis teknologi. AI memungkinkan sistem pembelajaran matematika menyesuaikan tingkat kesulitan soal berdasarkan kemampuan siswa sehingga proses belajar menjadi lebih efektif dan sesuai kebutuhan individu.

Perkembangan terbaru menunjukkan bahwa generative AI mulai digunakan dalam pembelajaran matematika untuk menghasilkan soal, menjelaskan konsep matematika, memberikan solusi langkah demi langkah, serta memberikan umpan balik otomatis kepada siswa. Menurut Walkington (2025), kehadiran generative AI membuka peluang baru dalam pendidikan matematika karena mampu membantu siswa memahami konsep abstrak melalui interaksi yang lebih fleksibel dan personal. Selain itu, berbagai penelitian menunjukkan bahwa AI mampu meningkatkan motivasi belajar matematika melalui penyediaan pengalaman belajar yang lebih interaktif. Egara dan Mosimege (2024) menemukan bahwa penggunaan ChatGPT dalam pembelajaran matematika dapat membantu siswa memperoleh penjelasan yang lebih cepat dan mudah dipahami, meskipun masih terdapat tantangan terkait akurasi jawaban dan ketergantungan siswa terhadap teknologi tersebut.

3. Konsep Asesmen dalam Pendidikan Matematika

Asesmen merupakan komponen penting dalam proses pembelajaran karena berfungsi untuk mengukur pencapaian kompetensi siswa, memberikan umpan balik terhadap proses belajar, serta mendukung pengambilan keputusan pendidikan. Dalam pendidikan matematika, asesmen tidak hanya digunakan untuk mengukur hasil akhir pembelajaran, tetapi juga untuk memahami proses berpikir matematis siswa. Cope et al. (2020) menjelaskan bahwa perkembangan teknologi digital telah mengubah paradigma asesmen dari pendekatan tradisional menuju asesmen yang lebih berpusat pada data. Dalam ekosistem pembelajaran berbasis AI, asesmen dapat dilakukan secara berkelanjutan melalui pengumpulan dan analisis data belajar siswa secara otomatis.

Asesmen matematika modern mencakup berbagai bentuk evaluasi seperti formative assessment, summative assessment, diagnostic assessment, dan adaptive assessment. Setiap bentuk asesmen memiliki tujuan yang berbeda dalam mendukung

proses pembelajaran. Oleh karena itu, integrasi AI dalam asesmen matematika dipandang sebagai solusi untuk meningkatkan kualitas evaluasi pembelajaran sekaligus mengurangi beban kerja guru dalam melakukan penilaian secara manual (González-Calatayud et al., 2021).

4. *Artificial Intelligence* untuk Asesmen Matematika

Penggunaan AI dalam asesmen pendidikan berkembang pesat dalam beberapa tahun terakhir. González-Calatayud et al. (2021) menjelaskan bahwa AI memungkinkan proses asesmen dilakukan secara lebih cepat, objektif, dan efisien dibandingkan metode konvensional. Sistem AI dapat mengotomatisasi proses penilaian, menganalisis pola kesalahan siswa, dan menghasilkan umpan balik yang bersifat personal. Dalam konteks matematika, AI digunakan untuk menilai jawaban siswa baik dalam bentuk pilihan ganda maupun jawaban terbuka. Botelho et al. (2023) menunjukkan bahwa teknologi NLP mampu menganalisis respons tertulis siswa pada soal matematika dan menghasilkan umpan balik otomatis yang relevan dengan kesalahan yang dilakukan siswa. Teknologi ini memberikan peluang untuk meningkatkan kualitas pembelajaran sekaligus mempercepat proses evaluasi.

Hooda et al. (2022) mengemukakan bahwa implementasi AI dalam asesmen pendidikan mencakup tiga fungsi utama, yaitu automated grading, automated feedback, dan student performance prediction. Ketiga fungsi tersebut berperan penting dalam mendukung proses pembelajaran yang lebih adaptif dan berbasis data. Perkembangan terbaru menunjukkan bahwa generative AI mulai digunakan dalam asesmen untuk menghasilkan soal, memeriksa jawaban siswa, serta memberikan rekomendasi pembelajaran yang sesuai dengan kebutuhan individu. Mao et al. (2023) menyatakan bahwa generative AI memiliki potensi besar dalam mentransformasi sistem asesmen pendidikan karena mampu memahami konteks jawaban siswa secara lebih mendalam dibandingkan sistem otomatis konvensional.

5. Teknologi *Artificial Intelligence* yang Digunakan dalam Asesmen Matematika

Berbagai teknologi AI telah diterapkan dalam asesmen matematika. Salah satu teknologi yang paling banyak digunakan *adalah Machine Learning* (ML). Ahmed (2024) menjelaskan bahwa algoritma ML mampu memprediksi performa akademik siswa berdasarkan data historis sehingga dapat digunakan sebagai alat asesmen prediktif. Selain ML, *Natural Language Processing* (NLP) menjadi teknologi penting dalam asesmen matematika modern. NLP memungkinkan sistem memahami jawaban tekstual siswa dan memberikan umpan balik secara otomatis. Penelitian Botelho et al. (2023) menunjukkan bahwa NLP efektif digunakan untuk mengevaluasi jawaban terbuka pada mata pelajaran matematika. Teknologi lain yang banyak digunakan adalah *Intelligent Tutoring Systems* (ITS). Menurut Marwiang et al. (2025), ITS mampu memberikan umpan balik real-time selama proses penyelesaian soal matematika sehingga siswa dapat memperbaiki kesalahan secara langsung.

Dalam beberapa tahun terakhir, Generative AI dan *Large Language Models* (LLM) menjadi fokus utama penelitian pendidikan. Xia et al. (2024) menemukan bahwa teknologi seperti ChatGPT mulai digunakan dalam proses asesmen untuk menghasilkan soal, mengevaluasi jawaban siswa, dan memberikan umpan balik yang lebih personal. Teknologi ini dianggap mampu meningkatkan efisiensi asesmen sekaligus memperluas akses terhadap layanan pendidikan berkualitas.

6. Dampak Pendidikan dari Penggunaan AI dalam Asesmen Matematika

Berbagai penelitian menunjukkan bahwa implementasi AI dalam asesmen matematika memberikan dampak positif terhadap proses dan hasil pembelajaran. Salah

satu dampak yang paling sering dilaporkan adalah peningkatan hasil belajar siswa. Meta-analisis yang dilakukan oleh N. dan Zhong (2025) menunjukkan bahwa penggunaan AI secara umum memberikan pengaruh positif terhadap capaian akademik peserta didik. Selain meningkatkan hasil belajar, AI juga mampu mempercepat pemberian umpan balik kepada siswa. Lee dan Moore (2024) menemukan bahwa sistem umpan balik otomatis berbasis generative AI dapat meningkatkan kualitas pembelajaran karena siswa memperoleh informasi mengenai kesalahan mereka secara lebih cepat dibandingkan metode tradisional.

Dampak positif lainnya adalah meningkatnya personalisasi pembelajaran. Farhood et al. (2025) menjelaskan bahwa AI memungkinkan penyusunan pengalaman belajar yang disesuaikan dengan kebutuhan, kemampuan, dan karakteristik individu siswa. Dalam pembelajaran matematika, personalisasi ini membantu siswa memperoleh materi dan latihan yang sesuai dengan tingkat pemahamannya. Namun demikian, penggunaan AI juga menimbulkan sejumlah tantangan. Yan et al. (2023) menyoroti berbagai isu etika yang muncul akibat penggunaan Large Language Models dalam pendidikan, seperti bias algoritma, privasi data, transparansi sistem, dan keandalan hasil asesmen. Selain itu, Zhai et al. (2024) menemukan bahwa ketergantungan yang berlebihan terhadap sistem AI berpotensi menurunkan kemampuan berpikir kritis dan pemecahan masalah siswa.

7. Tantangan dan Arah Pengembangan Penelitian

Berbagai penelitian menunjukkan manfaat AI dalam asesmen matematika, masih terdapat sejumlah tantangan yang perlu diperhatikan. Salah satu tantangan utama adalah validitas dan reliabilitas asesmen yang dihasilkan oleh sistem AI. Clark et al. (2025) menegaskan bahwa penggunaan AI dalam asesmen matematika harus tetap mempertimbangkan prinsip keadilan, aksesibilitas, dan akuntabilitas agar hasil penilaian dapat dipercaya.Tantangan lainnya berkaitan dengan kesiapan guru dalam mengintegrasikan teknologi AI ke dalam praktik asesmen. Wardat et al. (2024) menemukan bahwa banyak guru matematika masih menghadapi kendala dalam penggunaan teknologi AI karena keterbatasan kompetensi digital dan kurangnya pelatihan yang memadai. Selain itu, penelitian mengenai AI dalam asesmen matematika masih didominasi oleh studi yang berfokus pada efektivitas teknologi tertentu, sementara kajian yang membahas dampak jangka panjang terhadap kemampuan matematis siswa masih relatif terbatas. Oleh karena itu, diperlukan penelitian yang lebih komprehensif untuk memahami hubungan antara teknologi AI, fungsi asesmen, dan dampak pendidikan yang dihasilkan.

Berdasarkan berbagai temuan tersebut, dapat disimpulkan bahwa AI memiliki potensi besar untuk mentransformasi asesmen matematika melalui penerapan teknologi seperti machine learning, natural language processing, intelligent tutoring systems, dan generative AI. Namun, implementasi teknologi tersebut harus dilakukan secara hati-hati dengan mempertimbangkan aspek pedagogis, etika, dan kualitas asesmen. Tinjauan sistematis ini dilakukan untuk mensintesis bukti empiris mengenai teknologi AI yang digunakan dalam asesmen matematika, fungsi yang dijalankan, serta dampak pendidikan yang dihasilkan sehingga dapat memberikan arah pengembangan penelitian dan praktik pendidikan di masa mendatang.

METODE PENELITIAN

Penelitian ini menggunakan metode Systematic Literature Review (SLR) dengan mengacu pada pedoman Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020. Pendekatan SLR dipilih untuk mengidentifikasi, mengevaluasi, dan mensintesis secara sistematis berbagai penelitian terkait penerapan

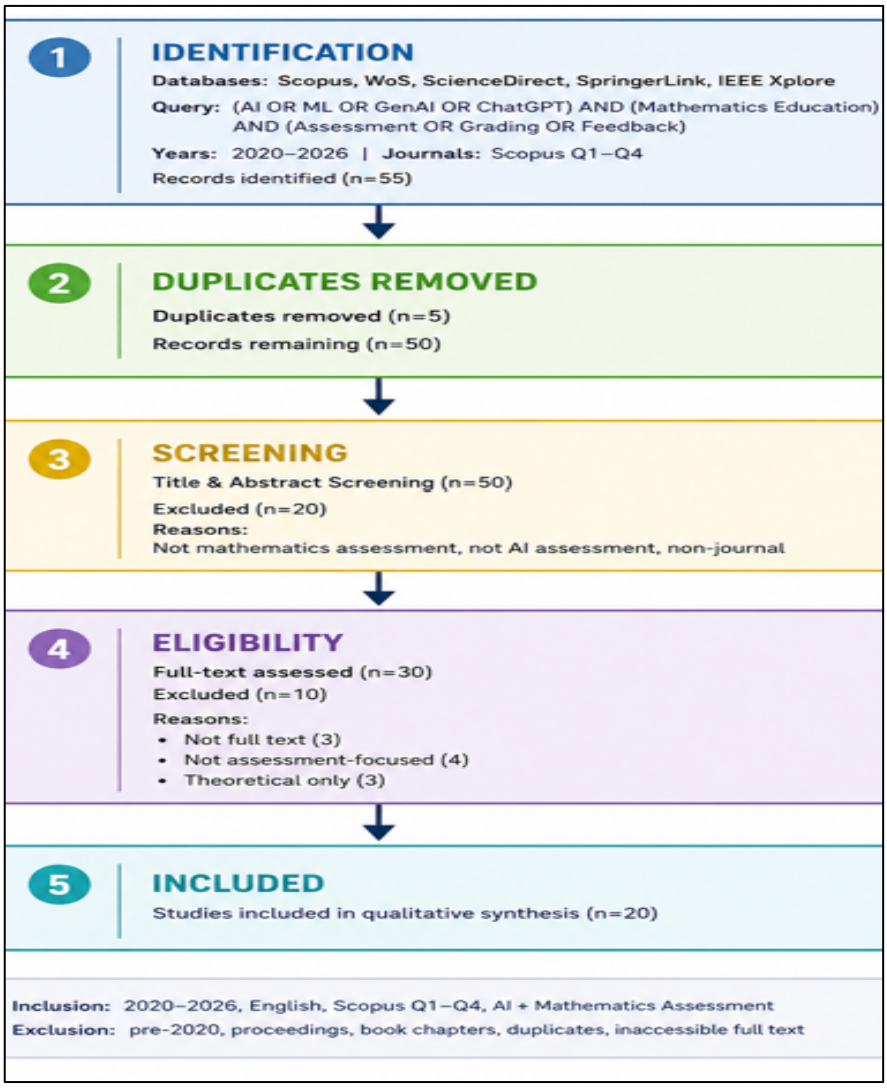
kecerdasan buatan (Artificial Intelligence/AI) dalam asesmen matematika. Metode ini memungkinkan peneliti memperoleh gambaran komprehensif mengenai teknologi AI yang digunakan, fungsi AI dalam proses asesmen, serta dampaknya terhadap pembelajaran matematika.

Strategi Penelusuran Literatur, Proses pencarian literatur dilakukan pada bulan Januari–Februari 2026 melalui beberapa basis data internasional bereputasi, yaitu Scopus. Strategi pencarian menggunakan kombinasi kata kunci yang disusun berdasarkan konsep utama penelitian, yaitu AI, matematika, dan asesmen pendidikan. String pencarian yang digunakan adalah ("Artificial Intelligence" OR AI OR "Machine Learning" OR ML OR "Generative AI" OR GenAI OR ChatGPT) AND ("Mathematics Education" OR "Mathematics Learning") AND (Assessment OR Grading OR Feedback) Pencarian dibatasi pada artikel yang diterbitkan antara 2020–2026, menggunakan bahasa Inggris, dan berasal dari jurnal yang terindeks Scopus Q1–Q4.

Kriteria Inklusi dan Eksklusi adalah untuk menjaga kualitas dan relevansi studi yang dianalisis, penelitian ini menerapkan kriteria inklusi dan eksklusi sebagai berikut. Kriteria Inklusi yaitu Artikel diterbitkan pada periode 2020–2026, artikel ditulis dalam bahasa Inggris, artikel merupakan publikasi jurnal ilmiah terindeks Scopus Q1–Q4, penelitian membahas penerapan AI dalam pendidikan matematika, penelitian secara khusus membahas fungsi AI dalam asesmen, penilaian, pemberian umpan balik (feedback), atau evaluasi hasil belajar matematika dan teks lengkap artikel dapat diakses. Sedangkan Kriteria Eksklusi adalah artikel yang diterbitkan sebelum tahun 2020, prosiding konferensi, buku, bab buku, editorial, atau laporan teknis dan artikel duplikat, artikel yang tidak memiliki akses teks lengkap, artikel yang membahas AI dalam pendidikan secara umum tanpa fokus pada asesmen matematika dan artikel yang bersifat konseptual tanpa pembahasan implementasi atau evaluasi asesmen.

Proses Seleksi Studi Berdasarkan PRISMA 2020

Proses seleksi literatur mengikuti empat tahapan utama dalam PRISMA 2020, yaitu identification, screening, eligibility, dan included. Pada tahap identification, diperoleh sebanyak 55 artikel dari lima basis data yang memenuhi kata kunci pencarian. Setelah dilakukan pemeriksaan duplikasi, ditemukan 5 artikel duplikat sehingga tersisa 50 artikel untuk tahap berikutnya. Pada tahap screening, peneliti melakukan penelaahan judul dan abstrak terhadap 50 artikel. Sebanyak 20 artikel dikeluarkan karena tidak relevan dengan asesmen matematika, tidak membahas penggunaan AI dalam penilaian, atau bukan artikel jurnal ilmiah. Dengan demikian, tersisa 30 artikel untuk tahap evaluasi kelayakan. Pada tahap eligibility, dilakukan penelaahan teks lengkap (full-text review) terhadap 30 artikel. Sebanyak 10 artikel dikeluarkan dengan rincian: 3 artikel tidak memiliki akses teks lengkap, 4 artikel tidak berfokus pada asesmen, dan 3 artikel hanya membahas aspek teoretis tanpa implementasi empiris. Tahap akhir (included) menghasilkan 20 artikel yang memenuhi seluruh kriteria inklusi dan digunakan dalam sintesis kualitatif.



Gambar: Diagram prisma 2020

Berdasarkan proses seleksi menggunakan PRISMA 2020, distribusi artikel pada setiap tahap dapat diringkas sebagai berikut:

Ekstraksi dan Analisis Data

Data dari artikel yang terpilih diekstraksi menggunakan lembar ekstraksi yang memuat informasi: penulis, tahun publikasi, negara penelitian, tujuan penelitian, jenis teknologi AI yang digunakan, fungsi AI dalam asesmen matematika, metode penelitian, serta dampak implementasi AI terhadap hasil belajar. Selanjutnya, data dianalisis menggunakan analisis tematik (thematic analysis) untuk mengidentifikasi pola dan tren penelitian. Hasil sintesis dikelompokkan ke dalam tiga tema utama: Technologies – jenis teknologi AI yang digunakan dalam asesmen matematika, seperti Machine Learning, Natural Language Processing (NLP), Intelligent Tutoring Systems (ITS), Learning Analytics, ChatGPT, dan Large Language Models (LLMs). Functions – fungsi AI dalam asesmen, meliputi automated grading, automated feedback, formative assessment, predictive assessment, dan adaptive assessment dan Educational Impacts – dampak penggunaan AI terhadap pembelajaran matematika, seperti peningkatan hasil belajar, efisiensi penilaian, personalisasi pembelajaran, motivasi belajar, dan tantangan etika penggunaan AI.

Validitas dan Reliabilitas

Untuk meningkatkan kredibilitas hasil kajian, proses seleksi artikel dilakukan secara sistematis berdasarkan pedoman PRISMA 2020 dan kriteria inklusi-eksklusi yang telah ditetapkan. Selain itu, proses ekstraksi data dilakukan secara konsisten menggunakan instrumen yang sama untuk seluruh artikel yang dianalisis sehingga dapat meminimalkan bias dalam interpretasi hasil penelitian.

HASIL DAN PEMBAHASAN

1.Teknologi AI Apa Saja yang Digunakan dalam Asesmen Matematika

Hasil sintesis terhadap 20 artikel yang memenuhi kriteria inklusi menunjukkan bahwa perkembangan teknologi Artificial Intelligence (AI) dalam asesmen matematika telah mengalami transformasi yang signifikan selama periode 2020–2025. Jika pada awal dekade AI lebih banyak digunakan untuk mendukung otomatisasi penilaian berbasis aturan *(rule-based assessment*), maka penelitian terbaru menunjukkan pergeseran menuju sistem yang lebih cerdas, adaptif, dan generatif. Secara umum, teknologi AI yang ditemukan dalam literatur dapat dikelompokkan menjadi lima kategori utama, yaitu *Natural Language Processing* (NLP), *Intelligent Tutoring Systems (ITS), Learning Analytics dan Predictive Analytics, Automated Grading Systems*, serta *Generative Artificial Intelligence* (GenAI) dan *Large Language Models* (LLMs).

a. Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan salah satu teknologi AI yang paling awal diterapkan dalam asesmen matematika. Teknologi ini memungkinkan komputer memahami, menganalisis, dan mengevaluasi jawaban siswa yang berbentuk teks atau uraian. Botelho et al. (2023) menunjukkan bahwa NLP dapat digunakan untuk menilai respons terbuka siswa dalam matematika secara otomatis tanpa harus bergantung pada jawaban pilihan ganda. Kemampuan ini menjadi terobosan penting karena selama bertahun-tahun sistem penilaian digital lebih efektif dalam menilai jawaban objektif dibandingkan jawaban yang membutuhkan penalaran matematis. Dalam konteks asesmen matematika, NLP tidak hanya berfungsi untuk memeriksa benar atau salahnya jawaban, tetapi juga mampu mengidentifikasi langkah-langkah penyelesaian yang dilakukan siswa. Dengan demikian, guru dapat memperoleh informasi yang lebih mendalam mengenai proses berpikir siswa dibandingkan hanya melihat hasil akhir. Namun demikian, hasil kajian menunjukkan bahwa NLP masih menghadapi keterbatasan dalam memahami variasi bahasa dan strategi penyelesaian masalah yang sangat beragam. Sistem sering kali mengalami kesulitan ketika siswa menggunakan pendekatan yang tidak umum atau menjelaskan solusi dengan bahasa yang ambigu. Oleh karena itu, meskipun NLP telah meningkatkan kemampuan asesmen otomatis, teknologi ini masih membutuhkan integrasi dengan model AI yang lebih canggih agar dapat menangkap kompleksitas penalaran matematis secara lebih akurat.

2. *Intelligent Tutoring Systems* (ITS)

Teknologi kedua yang banyak ditemukan adalah *Intelligent Tutoring Systems* (ITS). Hwang dan Tu (2021), serta Marwiang et al. (2025), menunjukkan bahwa ITS menjadi salah satu teknologi yang paling dominan dalam pendidikan matematika karena kemampuannya menggabungkan pembelajaran, asesmen, dan umpan balik dalam satu sistem terintegrasi. Berbeda dengan NLP yang berfokus pada analisis jawaban, ITS bekerja dengan memonitor seluruh proses belajar siswa secara real-time. Sistem ini mampu merekam pola kesalahan, tingkat kesulitan soal yang berhasil diselesaikan, serta waktu yang dibutuhkan siswa dalam menyelesaikan tugas. Berdasarkan data tersebut, ITS dapat menyesuaikan tingkat kesulitan soal berikutnya sehingga asesmen menjadi lebih adaptif dan personal.

Dari perspektif asesmen matematika, ITS menunjukkan pergeseran paradigma dari asesmen yang bersifat statis menuju asesmen yang bersifat dinamis. Dalam pendekatan tradisional, semua siswa mengerjakan soal yang sama dan dinilai menggunakan standar yang seragam. Sebaliknya, ITS memungkinkan setiap siswa memperoleh pengalaman asesmen yang berbeda sesuai dengan kebutuhan dan kemampuan masing-masing. Hal ini memperkuat validitas asesmen karena sistem dapat mengukur perkembangan individu secara lebih akurat. Meski demikian, implementasi ITS memerlukan infrastruktur teknologi yang memadai dan basis data pembelajaran yang besar. Selain itu, pengembangan ITS juga membutuhkan biaya yang relatif tinggi dibandingkan teknologi AI lainnya.

c. *Learning Analytics* dan *Predictive Analytics*

Teknologi berikutnya yang banyak digunakan adalah *Learning Analytics dan Predictive Analytics*. Penelitian oleh Liao et al. (2024) dan Owan et al. (2023) menunjukkan bahwa teknologi ini berperan penting dalam mengubah asesmen dari aktivitas yang hanya berorientasi pada hasil menjadi proses yang berorientasi pada data. Learning Analytics memanfaatkan data yang dihasilkan selama proses pembelajaran untuk mengidentifikasi pola perilaku belajar siswa. Data tersebut dapat berupa frekuensi akses materi, durasi belajar, tingkat keberhasilan menyelesaikan soal, hingga interaksi siswa dengan platform pembelajaran. Melalui analisis data tersebut, sistem dapat menghasilkan laporan visual yang membantu guru memahami perkembangan siswa secara lebih komprehensif. Sementara itu, Predictive Analytics menggunakan algoritma Machine Learning untuk memprediksi performa akademik siswa di masa depan. Dalam asesmen matematika, kemampuan prediktif ini sangat penting karena memungkinkan guru mengidentifikasi siswa yang berisiko mengalami kesulitan belajar sebelum masalah tersebut benar-benar terjadi. Dengan demikian, intervensi dapat dilakukan lebih awal. Namun demikian, penggunaan Learning Analytics juga menimbulkan sejumlah isu kritis. Akurasi prediksi sangat bergantung pada kualitas data yang tersedia. Jika data yang digunakan tidak lengkap atau mengandung bias, maka hasil prediksi dapat menyesatkan. Selain itu, penggunaan data siswa secara intensif menimbulkan pertanyaan mengenai privasi dan keamanan data pendidikan.

d. *Automated Grading Systems*

Automated Grading Systems merupakan teknologi AI yang secara khusus dikembangkan untuk melakukan penilaian otomatis terhadap hasil pekerjaan siswa.

Tan et al. (2025), Kooli dan Yusuf (2024), serta Deepshikha (2025) menemukan bahwa sistem ini menjadi salah satu teknologi yang paling banyak digunakan dalam asesmen matematika karena mampu mengurangi beban kerja guru secara signifikan. Pada tahap awal pengembangannya, Automated Grading Systems hanya mampu menilai jawaban yang memiliki satu solusi pasti. Namun, perkembangan AI memungkinkan sistem untuk menilai jawaban yang lebih kompleks, termasuk soal uraian dan tugas berbasis pemecahan masalah. Teknologi ini mampu menghasilkan skor secara cepat, konsisten, dan objektif. Keunggulan utama sistem ini adalah efisiensi. Dalam kelas dengan jumlah siswa yang besar, proses penilaian yang biasanya membutuhkan waktu berhari-hari dapat diselesaikan dalam hitungan menit. Selain itu, konsistensi penilaian juga meningkat karena sistem tidak dipengaruhi oleh faktor subjektivitas manusia. Meskipun demikian, sejumlah penelitian menunjukkan bahwa Automated Grading Systems masih menghadapi tantangan dalam menilai kreativitas, penalaran tingkat tinggi, dan strategi penyelesaian yang tidak sesuai dengan pola yang telah diprogram sebelumnya. Oleh karena itu, banyak peneliti merekomendasikan penggunaan sistem ini sebagai alat bantu guru, bukan sebagai pengganti penuh penilaian manusia.

e. *Generative Artificial Intelligence* (GenAI) *dan Large Language Models* (LLMs)

Temuan paling menarik dari literatur terbaru adalah meningkatnya penggunaan *Generative Artificial Intelligence (GenAI) dan Large Language Models (LLMs), seperti ChatGPT,* dalam asesmen matematika. Penelitian oleh Mao et al. (2023), Xia et al. (2024), Lee dan Moore (2024), Khlaif et al. (2024), serta Walkington (2025) menunjukkan bahwa teknologi ini telah mengubah cara asesmen dirancang dan dilaksanakan. Berbeda dengan teknologi AI generasi sebelumnya yang berfokus pada analisis data, GenAI mampu menghasilkan konten baru secara otomatis. Dalam asesmen matematika, GenAI dapat digunakan untuk menyusun soal, membuat rubrik penilaian, memberikan umpan balik yang dipersonalisasi, bahkan mengevaluasi jawaban siswa secara kontekstual. Kemampuan ini menjadikan GenAI sebagai teknologi yang paling revolusioner dalam asesmen matematika saat ini. AI tidak lagi sekadar berfungsi sebagai alat penilaian, tetapi mulai berperan sebagai mitra guru dalam seluruh siklus asesmen. Akan tetapi, perkembangan ini juga menimbulkan tantangan baru terkait validitas hasil penilaian, transparansi algoritma, serta integritas akademik siswa.

Berdasarkan keseluruhan temuan, dapat disimpulkan bahwa teknologi AI dalam asesmen matematika telah berkembang dari teknologi berbasis otomatisasi (NLP dan Automated Grading Systems) menuju teknologi berbasis adaptivitas (ITS dan Learning Analytics), dan akhirnya menuju teknologi berbasis generativitas (GenAI dan LLMs). Perubahan ini menunjukkan evolusi fungsi AI dari sekadar alat untuk mempercepat penilaian menjadi sistem cerdas yang mampu mendukung pengambilan keputusan pedagogis. Generative AI dan Large Language Models muncul sebagai teknologi yang paling menjanjikan sekaligus paling kontroversial. Teknologi ini menawarkan kemampuan yang jauh melampaui sistem AI sebelumnya, namun juga membawa risiko baru terkait reliabilitas, bias algoritma, dan etika penggunaan dalam asesmen. Oleh karena itu, masa depan asesmen matematika kemungkinan tidak akan bergantung pada satu teknologi tertentu, melainkan pada integrasi berbagai teknologi

AI yang saling melengkapi untuk menciptakan sistem asesmen yang lebih akurat, adaptif, personal, dan berpusat pada kebutuhan siswa.

2. Fungsi AI yang Paling Banyak Diterapkan dalam Asesmen Matematika

Hasil sintesis terhadap 20 artikel yang diinklusi menunjukkan bahwa implementasi Artificial Intelligence (AI) dalam asesmen matematika tidak hanya berfokus pada penggunaan teknologi tertentu, tetapi juga pada fungsi-fungsi pedagogis yang didukung oleh teknologi tersebut. Analisis literatur mengidentifikasi lima fungsi utama AI dalam asesmen matematika, yaitu automated grading, automated feedback, formative assessment, personalized assessment, dan predictive assessment. Dari kelima fungsi tersebut, automated grading dan automated feedback merupakan fungsi yang paling dominan, sedangkan fungsi lainnya berkembang sebagai konsekuensi dari meningkatnya kemampuan sistem AI dalam menganalisis data pembelajaran secara real-time.

a. *Automated Grading*: Fungsi yang Paling Dominan

Fungsi AI yang paling banyak ditemukan dalam literatur adalah automated grading atau penilaian otomatis. Penelitian oleh Tan et al. (2025), Kooli dan Yusuf (2024), Deepshikha (2025), serta Botelho et al. (2023) menunjukkan bahwa AI digunakan untuk mengevaluasi jawaban siswa secara otomatis, baik dalam bentuk pilihan ganda maupun jawaban uraian. Fungsi ini berkembang karena adanya kebutuhan untuk meningkatkan efisiensi proses penilaian, terutama pada kelas dengan jumlah siswa yang besar. Dalam asesmen matematika, sistem AI mampu mengidentifikasi jawaban benar dan salah, memberikan skor secara otomatis, serta mendeteksi pola kesalahan yang sering dilakukan siswa. Menurut Tan et al. (2025), sistem automated grading berbasis AI dapat mengurangi waktu penilaian secara signifikan tanpa menurunkan konsistensi hasil penilaian.

Temuan serupa dilaporkan oleh Kooli dan Yusuf (2024), yang menunjukkan bahwa Large Language Models (LLMs) seperti ChatGPT memiliki potensi besar dalam membantu proses grading pada tugas yang membutuhkan penalaran matematis. Namun demikian, dominasi automated grading juga memperlihatkan adanya kecenderungan bahwa sebagian besar penelitian masih berorientasi pada efisiensi administratif dibandingkan transformasi pedagogis. Banyak studi menempatkan AI sebagai alat untuk mempercepat pekerjaan guru, bukan sebagai sarana untuk memperdalam pemahaman siswa. Kondisi ini menunjukkan bahwa meskipun automated grading menjadi fungsi yang paling matang secara teknologi, kontribusinya terhadap peningkatan kualitas pembelajaran matematika masih memerlukan evaluasi lebih lanjut.

b. *Automated Feedback*: Pergeseran dari Penilaian Menuju Pembelajaran

Fungsi kedua yang paling banyak ditemukan adalah automated feedback. Lee dan Moore (2024), Steinert et al. (2024), Cosentino et al. (2025), serta Deepshikha (2025) menunjukkan bahwa AI tidak lagi hanya memberikan skor, tetapi juga mampu menghasilkan umpan balik secara otomatis dan personal. Dalam pembelajaran matematika, umpan balik merupakan komponen yang sangat penting karena membantu siswa memahami kesalahan konseptual maupun prosedural. Sistem AI

modern dapat memberikan penjelasan mengenai kesalahan yang dilakukan siswa serta menawarkan langkah-langkah perbaikan yang sesuai dengan kebutuhan individu. Cosentino et al. (2025) menemukan bahwa penggunaan GenAI dalam pembelajaran matematika memungkinkan pemberian umpan balik yang lebih kontekstual dibandingkan sistem berbasis aturan (rule-based systems).

Perkembangan automated feedback menunjukkan pergeseran paradigma dari assessment of learning menuju assessment for learning. Jika penilaian tradisional berfokus pada pengukuran hasil belajar, maka AI memungkinkan asesmen berfungsi sebagai bagian dari proses pembelajaran itu sendiri. Meskipun demikian, kualitas umpan balik yang dihasilkan AI masih bergantung pada kualitas model yang digunakan. Beberapa studi menunjukkan bahwa AI terkadang memberikan penjelasan yang kurang akurat atau terlalu umum sehingga tetap membutuhkan verifikasi dari guru.

c. *Formative Assessment*: Mendorong Asesmen Berkelanjutan

Fungsi berikutnya adalah formative assessment. Penelitian oleh Liao et al. (2024), Hooda et al. (2022), dan Marwiang et al. (2025) menunjukkan bahwa AI semakin banyak digunakan untuk memantau perkembangan belajar siswa secara berkelanjutan. Melalui Learning Analytics dan Intelligent Tutoring Systems (ITS), AI dapat mengumpulkan data mengenai aktivitas belajar siswa, menganalisis pola perkembangan kompetensi, dan memberikan rekomendasi pembelajaran secara real-time. Dalam konteks ini, asesmen tidak lagi dilakukan pada akhir pembelajaran, tetapi menjadi bagian yang terintegrasi dalam proses belajar sehari-hari. Temuan ini memperlihatkan perubahan mendasar dalam fungsi asesmen matematika. Asesmen tidak lagi hanya berfungsi sebagai alat evaluasi, melainkan sebagai mekanisme untuk mendukung pembelajaran adaptif. Namun demikian, implementasi formative assessment berbasis AI masih menghadapi tantangan terkait kesiapan guru dalam menginterpretasikan data yang dihasilkan sistem. Banyak guru belum memiliki kompetensi analitik yang memadai untuk memanfaatkan informasi tersebut secara optimal.

d. *Personalized Assessment*: Menuju Asesmen yang Berpusat pada Siswa

Fungsi lain yang semakin berkembang adalah personalized assessment. Hwang dan Tu (2021), Nguyen dan Pham (2025), serta Marwiang et al. (2025) menunjukkan bahwa AI memungkinkan asesmen disesuaikan dengan karakteristik dan kemampuan masing-masing siswa. Dalam sistem tradisional, semua siswa biasanya mengerjakan soal yang sama dengan tingkat kesulitan yang seragam. Sebaliknya, AI memungkinkan sistem memberikan soal yang berbeda berdasarkan performa sebelumnya. Siswa yang memiliki kemampuan tinggi dapat memperoleh soal yang lebih menantang, sedangkan siswa yang mengalami kesulitan dapat diberikan soal yang lebih sederhana disertai dukungan tambahan. Dari perspektif pedagogis, personalized assessment merupakan salah satu fungsi AI yang paling potensial karena sejalan dengan prinsip pembelajaran berdiferensiasi. Akan tetapi, implementasi asesmen personal juga menimbulkan pertanyaan mengenai keadilan dan kesetaraan. Jika setiap siswa menerima bentuk asesmen yang berbeda, maka diperlukan mekanisme yang jelas untuk memastikan bahwa hasil penilaian tetap dapat dibandingkan secara valid.

e. *Predictive Assessment*: Fungsi yang Mulai Berkembang

Fungsi terakhir yang ditemukan dalam literatur adalah predictive assessment. Owan et al. (2023), Clark et al. (2025), dan Liao et al. (2024) menunjukkan bahwa AI dapat digunakan untuk memprediksi performa akademik siswa berdasarkan data historis dan pola perilaku belajar. Fungsi ini memungkinkan guru mengidentifikasi siswa yang berisiko mengalami kesulitan belajar matematika sebelum masalah tersebut muncul secara nyata. Dengan demikian, intervensi dapat dilakukan lebih awal sehingga peluang keberhasilan belajar menjadi lebih besar. Meskipun menjanjikan, predictive assessment masih merupakan fungsi yang relatif baru dibandingkan automated grading dan feedback. Akurasi prediksi sangat dipengaruhi oleh kualitas data yang digunakan. Selain itu, penggunaan model prediktif juga berpotensi menimbulkan bias jika algoritma dilatih menggunakan data yang tidak representatif.

Berdasarkan keseluruhan temuan, dapat disimpulkan bahwa fungsi AI yang paling banyak diterapkan dalam asesmen matematika adalah automated grading dan automated feedback. Dominasi kedua fungsi ini menunjukkan bahwa fokus utama pengembangan AI masih berada pada peningkatan efisiensi dan efektivitas proses penilaian. Namun, tren penelitian terbaru memperlihatkan pergeseran menuju fungsi yang lebih pedagogis, seperti formative assessment, personalized assessment, dan predictive assessment.Perkembangan tersebut menunjukkan evolusi peran AI dari sekadar alat evaluasi menjadi mitra pembelajaran. Pada tahap awal, AI digunakan untuk menggantikan tugas administratif guru dalam memberikan nilai. Saat ini, AI mulai berperan dalam memberikan umpan balik, memonitor perkembangan belajar, menyesuaikan asesmen dengan kebutuhan individu, dan bahkan memprediksi performa siswa di masa depan. Dengan demikian, fungsi AI dalam asesmen matematika tidak lagi terbatas pada pengukuran hasil belajar, tetapi berkembang menjadi instrumen yang mendukung pembelajaran yang lebih adaptif, personal, dan berbasis data. Namun, keberhasilan implementasi fungsi-fungsi tersebut tetap bergantung pada kualitas algoritma, kesiapan guru, validitas asesmen, serta pengelolaan aspek etika dan privasi data pendidikan.

3. Dampak Penggunaan AI terhadap Proses dan Hasil Pembelajaran Matematika

Hasil sintesis dari artikel yang diinklusi menunjukkan bahwa implementasi Artificial Intelligence (AI) dalam asesmen matematika memberikan dampak yang signifikan terhadap proses maupun hasil pembelajaran. Dampak tersebut tidak hanya terlihat pada peningkatan efisiensi penilaian, tetapi juga pada perubahan cara siswa belajar, cara guru melakukan evaluasi, serta bagaimana asesmen digunakan untuk mendukung pembelajaran yang lebih personal dan adaptif. Secara umum, dampak penggunaan AI dapat dikelompokkan ke dalam lima aspek utama, yaitu peningkatan hasil belajar matematika, penguatan self-regulated learning, efisiensi dan produktivitas guru, personalisasi pembelajaran, serta munculnya tantangan terhadap kualitas dan integritas pembelajaran.

a. Peningkatan Hasil Belajar Matematika

Dampak yang paling konsisten ditemukan dalam literatur adalah peningkatan hasil belajar matematika. Penelitian oleh Marwiang et al. (2025), Cosentino et al. (2025),

Liao et al. (2024), serta Hidayat et al. (2022) menunjukkan bahwa siswa yang belajar melalui sistem berbasis AI cenderung memperoleh capaian akademik yang lebih baik dibandingkan pembelajaran konvensional. Peningkatan ini terjadi karena AI mampu menyediakan umpan balik secara cepat dan berkelanjutan. Dalam pembelajaran matematika, kesalahan yang tidak segera diperbaiki sering berkembang menjadi miskonsepsi yang lebih kompleks. Kehadiran AI memungkinkan siswa mengetahui kesalahan mereka segera setelah menyelesaikan tugas atau latihan. Marwiang et al. (2025) menemukan bahwa *Intelligent Tutoring System*s (ITS) berbasis umpan balik real-time mampu meningkatkan kemampuan pemecahan masalah matematika secara signifikan. Namun demikian, peningkatan hasil belajar tidak selalu berasal dari kecerdasan teknologi itu sendiri. Sebagian besar studi menunjukkan bahwa keberhasilan AI sangat dipengaruhi oleh desain pembelajaran yang mendukung pemanfaatannya. Dengan kata lain, AI bukan faktor tunggal yang menyebabkan peningkatan prestasi, melainkan berfungsi sebagai katalisator yang memperkuat kualitas proses pembelajaran. Temuan ini penting karena menunjukkan bahwa investasi pada teknologi tanpa diikuti strategi pedagogis yang tepat belum tentu menghasilkan peningkatan hasil belajar yang optimal.

b. Penguatan *Self-Regulated Learning*

Dampak kedua yang banyak ditemukan adalah meningkatnya kemampuan self-regulated learning (SRL) atau pembelajaran mandiri. Liao et al. (2024) menunjukkan bahwa penggunaan sistem asesmen berbasis AI yang dilengkapi visual report dan dashboard analitik membantu siswa memahami perkembangan belajar mereka secara lebih jelas. Melalui informasi yang disajikan secara real-time, siswa dapat mengetahui kompetensi yang telah dikuasai, area yang masih perlu diperbaiki, serta target belajar berikutnya. Dengan demikian, siswa tidak lagi bergantung sepenuhnya pada guru untuk mengevaluasi kemajuan belajar mereka. Dari perspektif teori konstruktivisme, dampak ini sangat penting karena pembelajaran matematika menuntut kemampuan refleksi dan pengaturan diri yang tinggi. AI membantu siswa menjadi lebih aktif dalam mengelola proses belajar mereka sendiri. Namun, terdapat perbedaan dampak berdasarkan karakteristik siswa. Beberapa penelitian menunjukkan bahwa siswa yang memiliki motivasi belajar rendah cenderung kurang memanfaatkan informasi yang diberikan sistem AI. Oleh karena itu, kemampuan self-regulated learning tidak muncul secara otomatis hanya karena adanya teknologi AI, tetapi juga dipengaruhi oleh kesiapan siswa dalam menggunakan informasi tersebut untuk memperbaiki strategi belajar mereka.

c. Meningkatkan Efisiensi dan Produktivitas Guru

Dampak berikutnya terlihat pada perubahan peran guru dalam proses asesmen. Kooli dan Yusuf (2024), Khlaif et al. (2024), Hooda et al. (2022), serta Deepshikha (2025) melaporkan bahwa AI mampu mengurangi beban administratif guru melalui otomatisasi proses grading dan pemberian umpan balik. Dalam praktik pembelajaran matematika, proses penilaian sering memerlukan waktu yang sangat besar, terutama ketika guru harus memeriksa soal uraian atau tugas pemecahan masalah. Kehadiran AI memungkinkan sebagian besar pekerjaan tersebut dilakukan secara otomatis sehingga

guru dapat lebih fokus pada aktivitas pedagogis seperti membimbing siswa, mengembangkan materi, dan merancang strategi pembelajaran. Meskipun demikian, efisiensi ini juga memunculkan dilema baru. Beberapa studi menunjukkan bahwa ketergantungan yang berlebihan terhadap sistem AI berpotensi mengurangi keterlibatan guru dalam memahami proses berpikir siswa secara mendalam. Jika guru hanya mengandalkan hasil analisis AI, maka ada risiko bahwa aspek-aspek penting yang tidak dapat dideteksi algoritma menjadi terabaikan. Oleh karena itu, literatur secara umum menegaskan bahwa AI seharusnya berfungsi sebagai alat pendukung (teacher augmentation), bukan sebagai pengganti peran profesional guru.

d. Mendorong Personalisasi Pembelajaran Matematika

Salah satu dampak paling transformasional dari penggunaan AI adalah kemampuannya mendukung personalisasi pembelajaran. Hwang dan Tu (2021), Nguyen dan Pham (2025), Marwiang et al. (2025), serta Clark et al. (2025) menunjukkan bahwa AI memungkinkan asesmen dan pembelajaran disesuaikan dengan kebutuhan individual siswa. Melalui analisis data performa belajar, sistem AI dapat menentukan tingkat kesulitan soal yang sesuai, memberikan rekomendasi materi tambahan, dan menyusun jalur pembelajaran yang berbeda untuk setiap siswa. Pendekatan ini sangat relevan dalam pembelajaran matematika karena kemampuan siswa sering kali sangat bervariasi meskipun berada dalam kelas yang sama. Dari perspektif pedagogi modern, personalisasi merupakan salah satu keunggulan utama AI dibandingkan asesmen tradisional. Sistem konvensional cenderung menerapkan pendekatan yang seragam untuk seluruh siswa, sedangkan AI memungkinkan asesmen yang lebih responsif terhadap kebutuhan individu. Namun demikian, beberapa penelitian mengingatkan bahwa personalisasi yang berlebihan dapat mengurangi kesempatan siswa untuk menghadapi tantangan yang lebih kompleks di luar zona nyaman mereka. Dengan demikian, keseimbangan antara adaptivitas dan tantangan belajar menjadi aspek penting dalam implementasi AI.

e. Dampak Negatif dan Tantangan terhadap Kualitas Pembelajaran

Meskipun sebagian besar penelitian melaporkan dampak positif, literatur juga menunjukkan adanya sejumlah konsekuensi negatif yang perlu diperhatikan. Walkington (2025), Xia et al. (2024), Khlaif et al. (2024), dan Clark et al. (2025) menyoroti beberapa tantangan utama yang muncul akibat penggunaan AI dalam asesmen matematika. Pertama, terdapat risiko ketergantungan berlebihan terhadap AI. Kemunculan Generative AI seperti ChatGPT memungkinkan siswa memperoleh jawaban dan solusi matematika secara instan. Jika tidak diimbangi dengan strategi pembelajaran yang tepat, kondisi ini dapat mengurangi kesempatan siswa untuk mengembangkan kemampuan berpikir kritis dan pemecahan masalah secara mandiri. Kedua, terdapat persoalan validitas asesmen. Meskipun AI mampu memberikan penilaian secara cepat, beberapa penelitian menunjukkan bahwa sistem masih mengalami kesulitan dalam mengevaluasi kreativitas matematis, penalaran tingkat tinggi, dan strategi penyelesaian yang tidak umum. Hal ini menunjukkan bahwa hasil asesmen AI belum sepenuhnya dapat menggantikan penilaian manusia. Ketiga, muncul isu etika dan integritas akademik. Xia et al. (2024) dan Walkington (2025) menemukan

bahwa penggunaan GenAI meningkatkan risiko plagiarisme, manipulasi tugas, dan ketidakjujuran akademik. Dalam konteks asesmen matematika, siswa dapat menggunakan AI untuk menghasilkan jawaban yang benar tanpa benar-benar memahami konsep yang dipelajari.

Berdasarkan keseluruhan temuan, dampak penggunaan AI terhadap proses dan hasil pembelajaran matematika bersifat transformasional sekaligus paradoksal. Di satu sisi, AI terbukti meningkatkan hasil belajar, mempercepat pemberian umpan balik, memperkuat self-regulated learning, meningkatkan efisiensi kerja guru, dan mendukung personalisasi pembelajaran. Temuan dari Marwiang et al. (2025), Liao et al. (2024), Cosentino et al. (2025), serta Hwang dan Tu (2021) menunjukkan bahwa AI berpotensi menciptakan pengalaman belajar matematika yang lebih adaptif dan berpusat pada siswa. Di sisi lain, perkembangan teknologi AI juga menghadirkan tantangan baru yang tidak dapat diabaikan. Risiko ketergantungan siswa terhadap AI, keterbatasan sistem dalam menilai penalaran matematis tingkat tinggi, serta meningkatnya isu integritas akademik menunjukkan bahwa dampak AI tidak selalu positif. Oleh karena itu, temuan utama dari literatur bukanlah bahwa AI secara otomatis meningkatkan kualitas pembelajaran matematika, melainkan bahwa AI dapat menjadi pengungkit (enabler) yang sangat kuat apabila diintegrasikan dengan desain pedagogis yang tepat dan didukung oleh kompetensi guru yang memadai. Secara keseluruhan, hasil sintesis menunjukkan bahwa masa depan asesmen matematika berbasis AI tidak terletak pada penggantian peran manusia, tetapi pada kolaborasi antara kecerdasan manusia dan kecerdasan buatan untuk menciptakan sistem pembelajaran yang lebih efektif, adaptif, dan berkelanjutan.

4. Tantangan dan Peluang Implementasi AI dalam Asesmen Matematika

Hasil sintesis menunjukkan bahwa implementasi Artificial Intelligence (AI) dalam asesmen matematika menghadirkan peluang yang sangat besar sekaligus tantangan yang kompleks. Seiring dengan berkembangnya teknologi seperti Natural Language Processing (NLP), Intelligent Tutoring Systems (ITS), Learning Analytics, Automated Grading Systems, dan Generative Artificial Intelligence (GenAI), peran AI tidak lagi terbatas pada otomatisasi penilaian, tetapi telah berkembang menjadi bagian integral dari ekosistem pembelajaran matematika. Namun, perkembangan tersebut juga menimbulkan persoalan terkait validitas asesmen, etika, kesiapan guru, hingga integritas akademik. Oleh karena itu, pembahasan mengenai implementasi AI perlu dilihat secara seimbang dengan mempertimbangkan peluang dan tantangan yang muncul secara bersamaan.

a. Tantangan Implementasi AI dalam Asesmen Matematika

1. Validitas dan Reliabilitas Penilaian

Tantangan yang paling banyak dibahas dalam literatur adalah validitas dan reliabilitas hasil asesmen yang dihasilkan AI. Xia et al. (2024), Kooli dan Yusuf (2024), serta Walkington (2025) menunjukkan bahwa meskipun AI mampu memberikan penilaian secara cepat dan konsisten, sistem masih menghadapi kesulitan dalam mengevaluasi aspek-aspek berpikir matematis yang kompleks. Dalam pembelajaran

matematika, jawaban yang benar tidak selalu diperoleh melalui prosedur yang sama. Siswa sering menggunakan strategi penyelesaian yang berbeda untuk mencapai hasil yang identik. AI cenderung lebih efektif dalam mengenali pola jawaban yang umum, tetapi kurang optimal dalam menilai kreativitas, argumentasi matematis, dan proses penalaran tingkat tinggi. Akibatnya, terdapat risiko bahwa sistem memberikan skor yang kurang mencerminkan kemampuan sebenarnya dari siswa. Tantangan ini menunjukkan bahwa keberhasilan asesmen matematika tidak hanya ditentukan oleh akurasi teknis algoritma, tetapi juga oleh kemampuannya memahami keragaman cara berpikir manusia. Hingga saat ini, kemampuan tersebut masih menjadi keterbatasan utama AI.

2. Bias Algoritma dan Keadilan Asesmen

Clark et al. (2025), Owan et al. (2023), serta Nguyen et al. (2022) menyoroti masalah bias algoritma sebagai salah satu tantangan serius dalam implementasi AI. Sistem AI belajar dari data yang digunakan selama proses pelatihan. Jika data tersebut tidak representatif atau mengandung bias tertentu, maka hasil asesmen yang dihasilkan juga berpotensi bias. Dalam konteks pendidikan matematika, bias dapat muncul terhadap kelompok siswa tertentu berdasarkan latar belakang sosial, bahasa, budaya, atau akses teknologi. Kondisi ini berpotensi menimbulkan ketidakadilan dalam proses asesmen. Analisis literatur menunjukkan bahwa isu bias tidak dapat diselesaikan hanya dengan meningkatkan kecanggihan teknologi. Diperlukan transparansi algoritma, audit sistem secara berkala, serta keterlibatan manusia dalam proses pengambilan keputusan untuk memastikan keadilan asesmen.

3. Integritas Akademik dan Ketergantungan pada AI

Kemunculan Generative AI seperti ChatGPT membawa tantangan baru yang belum banyak ditemukan pada generasi teknologi sebelumnya. Khlaif et al. (2024), Xia et al. (2024), Mao et al. (2023), dan Walkington (2025) menemukan bahwa siswa dapat menggunakan AI untuk menyelesaikan tugas matematika secara instan tanpa benar-benar memahami konsep yang dipelajari. Fenomena ini memunculkan kekhawatiran mengenai menurunnya integritas akademik. Jika siswa hanya berfokus pada memperoleh jawaban yang benar melalui AI, maka tujuan utama pembelajaran matematika, yaitu mengembangkan kemampuan berpikir logis, analitis, dan pemecahan masalah, dapat terabaikan. Secara kritis, masalah utama bukan terletak pada keberadaan AI, melainkan pada desain asesmen yang masih berorientasi pada produk akhir. Selama asesmen hanya mengukur jawaban, siswa akan selalu mencari cara tercepat untuk memperoleh jawaban tersebut. Oleh karena itu, muncul kebutuhan untuk merancang asesmen yang lebih menekankan proses berpikir daripada hasil akhir.

4. Kesiapan Guru dan Literasi AI

Tantangan berikutnya berkaitan dengan kesiapan sumber daya manusia. Khlaif et al. (2024), Hidayat et al. (2022), dan Hwang dan Tu (2021) menunjukkan bahwa banyak guru matematika belum memiliki kompetensi yang memadai dalam memanfaatkan AI untuk asesmen. Meskipun teknologi tersedia, pemanfaatannya sering kali terbatas pada fungsi-fungsi dasar seperti penilaian otomatis. Guru masih

mengalami kesulitan dalam menginterpretasikan data analitik, memverifikasi hasil AI, dan mengintegrasikan rekomendasi sistem ke dalam praktik pembelajaran. Temuan ini menunjukkan bahwa keberhasilan implementasi AI lebih banyak ditentukan oleh kesiapan pengguna dibandingkan kecanggihan teknologi itu sendiri. Tanpa peningkatan literasi AI di kalangan guru, potensi transformasi yang ditawarkan teknologi tidak akan tercapai secara optimal.

5. Privasi dan Keamanan Data

Learning Analytics dan Predictive Analytics memerlukan pengumpulan data siswa dalam jumlah besar. Liao et al. (2024), Owan et al. (2023), dan Xia et al. (2024) mengingatkan bahwa penggunaan data secara masif menimbulkan risiko terhadap privasi dan keamanan informasi pendidikan. Data perilaku belajar, performa akademik, dan interaksi siswa dengan sistem digital merupakan informasi yang sangat sensitif. Jika tidak dikelola dengan baik, data tersebut dapat disalahgunakan atau diakses oleh pihak yang tidak berwenang. Oleh karena itu, perlindungan data menjadi salah satu isu strategis dalam pengembangan asesmen berbasis AI.

b. Peluang Implementasi AI dalam Asesmen Matematika

1. Transformasi Asesmen Menuju Sistem yang Adaptif

Peluang terbesar yang ditemukan dalam literatur adalah kemampuan AI untuk mengubah asesmen dari sistem yang statis menjadi sistem yang adaptif. Hwang dan Tu (2021), Marwiang et al. (2025), serta Nguyen dan Pham (2025) menunjukkan bahwa AI memungkinkan setiap siswa memperoleh pengalaman asesmen yang sesuai dengan kemampuan dan kebutuhan masing-masing. Pendekatan ini mendukung pembelajaran berdiferensiasi yang selama ini sulit diwujudkan dalam kelas besar. AI memungkinkan asesmen yang lebih akurat karena mengukur perkembangan individu, bukan sekadar membandingkan siswa dengan standar yang sama.

2. Peningkatan Kualitas Umpan Balik

Lee dan Moore (2024), Cosentino et al. (2025), serta Steinert et al. (2024) menunjukkan bahwa AI mampu memberikan umpan balik secara cepat, berkelanjutan, dan personal. Dalam pembelajaran matematika, kualitas umpan balik merupakan faktor yang sangat menentukan keberhasilan belajar. Berbeda dengan sistem tradisional yang sering menunda pemberian umpan balik hingga beberapa hari atau minggu, AI memungkinkan siswa menerima informasi perbaikan secara langsung setelah menyelesaikan tugas. Hal ini mempercepat proses refleksi dan perbaikan pemahaman konsep.

3. Pengambilan Keputusan Berbasis Data

Learning Analytics dan Predictive Analytics membuka peluang baru dalam pengambilan keputusan pendidikan. Owan et al. (2023) dan Liao et al. (2024) menunjukkan bahwa AI mampu mengidentifikasi pola belajar siswa yang sulit diamati secara manual oleh guru. Informasi tersebut dapat digunakan untuk merancang intervensi yang lebih tepat sasaran, mengidentifikasi siswa berisiko, serta meningkatkan efektivitas strategi pembelajaran matematika.

4. Efisiensi dan Produktivitas Guru

Kooli dan Yusuf (2024), Tan et al. (2025), serta Deepshikha (2025) menunjukkan bahwa AI berpotensi mengurangi beban administratif guru secara signifikan. Waktu yang sebelumnya digunakan untuk memeriksa tugas dan memberikan nilai dapat dialihkan untuk aktivitas yang lebih bernilai pedagogis. Peluang ini sangat penting mengingat guru matematika sering menghadapi beban penilaian yang tinggi, terutama pada kelas dengan jumlah siswa yang besar.

5. Kolaborasi Manusia dan AI dalam Asesmen

Temuan paling menarik dalam literatur terbaru adalah munculnya paradigma baru yang memandang AI sebagai assessment partner, bukan sekadar alat bantu. Walkington (2025), Xia et al. (2024), dan Lee dan Moore (2024) menunjukkan bahwa AI mulai berfungsi sebagai mitra guru dalam merancang soal, memberikan umpan balik, menganalisis data pembelajaran, dan mendukung pengambilan keputusan. Paradigma ini membuka peluang untuk menciptakan model asesmen yang menggabungkan keunggulan manusia dalam memahami konteks dan nilai-nilai pendidikan dengan kemampuan AI dalam mengolah data secara cepat dan akurat.

Berdasarkan keseluruhan temuan, implementasi AI dalam asesmen matematika berada pada persimpangan antara inovasi dan risiko. Di satu sisi, AI menawarkan peluang besar berupa asesmen yang lebih adaptif, personal, efisien, dan berbasis data. Teknologi ini memungkinkan transformasi asesmen dari aktivitas administratif menjadi instrumen yang mendukung pembelajaran secara berkelanjutan. Di sisi lain, berbagai tantangan seperti validitas penilaian, bias algoritma, integritas akademik, kesiapan guru, dan perlindungan data menunjukkan bahwa implementasi AI tidak dapat dipandang sebagai solusi yang sepenuhnya bebas masalah.

Temuan literatur menunjukkan bahwa masa depan asesmen matematika bukanlah AI menggantikan guru, melainkan AI memperkuat kapasitas guru dalam melaksanakan asesmen yang lebih efektif dan bermakna. Dengan demikian, keberhasilan implementasi AI sangat bergantung pada keseimbangan antara inovasi teknologi dan prinsip-prinsip pedagogis. Tantangan yang ada bukan alasan untuk menolak AI, tetapi menjadi dasar untuk mengembangkan kebijakan, desain asesmen, dan program pengembangan profesional guru yang mampu memastikan bahwa teknologi digunakan secara etis, adil, dan berorientasi pada peningkatan kualitas pembelajaran matematika.

SIMPULAN

Systematic Literature Review (SLR) ini bertujuan untuk menganalisis perkembangan penelitian mengenai Artificial Intelligence (AI) dalam asesmen matematika melalui sintesis terhadap 20 artikel yang memenuhi kriteria inklusi berdasarkan pedoman PRISMA 2020. Hasil kajian menunjukkan bahwa AI telah berkembang menjadi salah satu inovasi paling signifikan dalam transformasi asesmen matematika, tidak hanya sebagai alat otomatisasi penilaian, tetapi juga sebagai teknologi yang mampu mendukung pembelajaran yang lebih adaptif, personal, dan berbasis data. Ditemukan bahwa teknologi AI yang paling banyak digunakan dalam asesmen matematika meliputi Natural Language Processing (NLP), Intelligent Tutoring Systems (ITS), Learning Analytics, Predictive Analytics, Automated Grading Systems, serta Generative Artificial Intelligence (GenAI) dan Large Language Models (LLMs).

Perkembangan teknologi tersebut menunjukkan adanya pergeseran dari sistem AI yang berorientasi pada otomatisasi penilaian menuju sistem yang mampu memahami, menganalisis, dan menghasilkan respons secara lebih cerdas dan kontekstual. Fungsi AI yang paling dominan dalam asesmen matematika adalah automated grading dan automated feedback. Selain itu, AI juga digunakan untuk mendukung formative assessment, personalized assessment, dan predictive assessment. Temuan ini menunjukkan bahwa peran AI telah berkembang dari sekadar alat evaluasi menjadi instrumen yang mendukung proses pembelajaran melalui pemberian umpan balik, pemantauan perkembangan belajar, dan prediksi performa siswa secara berkelanjutan.

Penggunaan AI memberikan dampak positif terhadap proses dan hasil pembelajaran matematika. AI terbukti mampu meningkatkan hasil belajar, memperkuat kemampuan self-regulated learning, meningkatkan efisiensi kerja guru, serta mendukung personalisasi pembelajaran sesuai kebutuhan individu siswa. Namun demikian, literatur juga menunjukkan bahwa manfaat tersebut sangat bergantung pada kualitas desain pembelajaran dan kompetensi guru dalam memanfaatkan teknologi AI secara efektif. Implementasi AI dalam asesmen matematika menghadirkan berbagai peluang sekaligus tantangan. Peluang utama meliputi terciptanya asesmen yang lebih adaptif, peningkatan kualitas umpan balik, pengambilan keputusan berbasis data, serta efisiensi proses evaluasi. Di sisi lain, tantangan yang masih perlu mendapat perhatian mencakup validitas dan reliabilitas penilaian, bias algoritma, integritas akademik, kesiapan guru, serta isu privasi dan keamanan data siswa.

Secara keseluruhan, hasil kajian ini menunjukkan bahwa perkembangan AI telah mendorong perubahan paradigma dari "AI for Assessment" menuju "AI as Assessment Partner". AI tidak lagi diposisikan hanya sebagai alat penilaian otomatis, tetapi sebagai mitra guru dalam merancang, melaksanakan, dan mengevaluasi asesmen matematika. Oleh karena itu, masa depan asesmen matematika tidak terletak pada penggantian peran guru oleh AI, melainkan pada kolaborasi yang sinergis antara kecerdasan manusia dan kecerdasan buatan untuk menciptakan sistem asesmen yang lebih akurat, adaptif, inklusif, dan berorientasi pada peningkatan kualitas pembelajaran.

DAFTAR PUSTAKA

Ahmed, E. (2024). Student performance prediction using machine learning algorithms. Applied Computational Intelligence and Soft Computing, 2024, 4067721. https://doi.org/10.1155/2024/4067721

Awang, L. A., Yusop, F., & Danaee, M. (2025). Current practices and future direction of artificial intelligence in mathematics education: A systematic review. International Electronic Journal of Mathematics Education. https://doi.org/10.29333/iejme/16006

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta-systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. International Journal of Educational Technology in Higher Education, 21, 1–41. https://doi.org/10.1186/s41239-023-00436-z

Botelho, A. F., Baral, S., Erickson, J. A., Benachamardi, P., & Heffernan, N. (2023). Leveraging natural language processing to support automated assessment and feedback for student open responses in mathematics. Journal of Computer Assisted Learning, 39, 823–840. https://doi.org/10.1111/jcal.12793

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. IEEE Access, 8, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Clark, A. K., Hirt, A., Whitcomb, D. H., Thompson, W., Wine, M., & Karvonen, M. (2025). Artificial intelligence in science and mathematics assessment for students with disabilities: Opportunities and challenges. Education Sciences. https://doi.org/10.3390/educsci15020233

González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. Applied Sciences. https://doi.org/10.3390/app11125467

Hidayat, R., Mohamed, M., Suhaizi, N. N. B., Sabri, N. B. M., Mahmud, M., & Baharuddin, S. K. B. (2022). Artificial intelligence in mathematics education: A systematic literature review. International Electronic Journal of Mathematics Education. https://doi.org/10.29333/iejme/12132

Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. Mathematical Problems in Engineering. https://doi.org/10.1155/2022/5215722

Hwang, G.-J., & Tu, Y. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. Mathematics. https://doi.org/10.3390/math9060584

Kamalov, F., Calonge, D. S., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. Sustainability. https://doi.org/10.3390/su151612451

Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M., Ayyoub, A., Hattab, M., & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. Education Sciences. https://doi.org/10.3390/educsci14101090

Kooli, C., & Yusuf, N. (2024). Transforming educational assessment: Insights into the use of ChatGPT and large language models in grading. International Journal of Human–Computer Interaction, 41, 3388–3399. https://doi.org/10.1080/10447318.2024.2338330

Lee, S. S., & Moore, R. L. (2024). Harnessing generative AI (GenAI) for automated feedback in higher education: A systematic review. Online Learning. https://doi.org/10.24059/olj.v28i3.4593

Marwiang, M., Prasertsang, M., & Junpeng, P. (2025). Enhancing students' learning outcomes in mathematics through intelligent tutoring systems based on real-time feedback. Journal of Education and Learning. https://doi.org/10.5539/jel.v14n6p186

Mao, J., Chen, B., & Liu, J. (2023). Generative artificial intelligence in education and its implications for assessment. TechTrends, 68, 58–66. https://doi.org/10.1007/s11528-023-00911-4

Owan, V. J., Abang, K. B., Idika, D., Etta, E. O., & Bassey, B. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. Eurasia Journal of Mathematics, Science and Technology Education. https://doi.org/10.29333/ejmste/13428

Xia, Q., Weng, X., Fan, O., Lin, T.-J., & Chiu, T. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. International Journal of Educational Technology in Higher Education, 21, 1–22. https://doi.org/10.1186/s41239-024-00468-z

Yan, L., Sha, L., Zhao, L., Li, Y., Maldonado, R. M., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. British Journal of Educational Technology, 55, 90–112. https://doi.org/10.1111/bjet.13370