

Tingkat Akurasi Model AI ChatGPT dalam menyelesaikan Evaluasi Pembelajaran pada *Massive Open Online Course*

Fitrah Izul Falaq

Bikinkarya Creative Labs, Kota Probolinggo, Jawa Timur, Indonesia

Abstract

Pembelajaran berbasis web melalui *Massive Open Online Course* (MOOC) telah terbukti menjadi solusi efektif untuk menyediakan akses pembelajaran yang lebih fleksibel tanpa batasan waktu, tempat dan jumlah peserta. Namun, tantangan terbesar dalam penerapan MOOC terletak pada evaluasi pembelajaran yang digunakan. Evaluasi sumatif berupa soal pilihan ganda menjadi instrumen yang paling banyak digunakan untuk mengukur hasil belajar pada platform pembelajaran MOOC. Pebelajar akan mendapatkan sertifikat kelulusan setelah mengerjakan soal diakhir kursus. Padahal, kehadiran kecerdasan buatan seperti ChatGPT membawa ancaman dan membuka potensi kecurangan. Saat ini, pebelajar dapat dengan mudah mengerjakan evaluasi akhir tanpa harus mempelajari materi terlebih dahulu. Penelitian ini bertujuan untuk mengukur tingkat akurasi ChatGPT-4o dalam menyelesaikan evaluasi pembelajaran pada MOOC dengan area subjek materi manajerial *Core Value* ASN. Topik tersebut menarik karena menjadi salah satu syarat wajib yang harus ditempuh CASN untuk menjadi ASN. Metode penelitian menggunakan metode eksperimental dimana peneliti melakukan pengujian secara langsung. Peneliti menggunakan 3 akun Chat GPT-4o versi gratis untuk mengerjakan 8 kursus hanya dengan 1x percobaan. Tidak ada perintah khusus, soal dan jawaban disalin kedalam aplikasi kemudian menginputkan jawabannya pada platform MOOC. Hasil penelitian menunjukkan bahwa tingkat akurasi ChatGPT-4o adalah sebesar 87,5%. Prosentase tersebut tergolong tinggi, artinya setiap pebelajar dapat menyelesaikan pembelajaran cukup dengan 1x percobaan saja. Kesimpulannya, ChatGPT-4o terbukti memiliki potensi cukup signifikan untuk menjawab soal. Rekomendasi temuan ini adalah dengan melakukan tindakan diantaranya 1) menggunakan soal berupa gambar, 2) membatasi navigasi serta mencegah perpindahan tab browser laman MOOC untuk menghindari kecurangan, dan 3) gunakan soal studi kasus yang dilengkapi dengan materi konseptual. Selain itu, peneliti menyarankan bahwa perlu dilakukan penelitian lanjutan terkait desain evaluasi pembelajaran yang mampu meminimalisir kecurangan pemanfaatan model AI ChatGPT-4o dalam pembelajaran MOOC.

Keywords: ChatGPT, MOOC, Evaluasi Pembelajaran, AI dalam Pendidikan, Kecerdasan Buatan

Corresponding author:

Fitrah Izul Falaq
Bikinkarya Creative Labs
Kota Probolinggo, Jawa Timur, Indonesia
Email: fitrahizulfalaq@gmail.com

Article history

Received 22 April, 2025
Revised 1 Mei, 2025
Accepted 9 Mei, 2025
Published Online 15 Mei, 2025

Cite this article:

Falaq, Fitrah Izul. (2025). Tingkat Akurasi Model AI ChatGPT dalam menyelesaikan Evaluasi Pembelajaran pada *Massive Open Online Course*. *Jurnal Pendidikan Prapanca* 1(1), 30-39

PENDAHULUAN

Pembelajaran berbasis web melalui *Massive Open Online Courses* (MOOC) menjadi solusi efektif untuk menyediakan akses pembelajaran tanpa batasan jarak, waktu, tempat dan jumlah peserta. Secara historis, sejarah perkembangan konsep MOOC pertama kali diperkenalkan oleh David Cormier dan Bryan Alexander pada tahun 2008 untuk menggambarkan kursus "*Connectivism and Connective Knowledge*" oleh Stephen Downes dan George Siemens yang diikuti oleh lebih dari 2.200 peserta dari berbagai belahan dunia (Despujol et al., 2022). Hingga saat ini, penerapan pembelajaran MOOC telah banyak diadopsi dan diimplementasikan oleh lembaga pendidikan, pemerintah bahkan perusahaan. Bukan tanpa alasan, MOOC berhasil menjadi alternatif inovatif dalam dunia pendidikan dan pelatihan, khususnya meningkatkan aksesibilitas dan fleksibilitas. Secara eksplisit, keberhasilan MOOC terletak

pada aspek biaya, kemudahan akses, keragaman materi, jangkauan peserta dan kecepatan perkembangan pengetahuan (Milligan et al., 2017; N. et al., 2025; Adedoyin & Famolu, 2020)

Di Indonesia, implementasi MOOC telah diterapkan di berbagai institusi maupun perusahaan. Pertama, pada institusi pendidikan, Universitas Indonesia (UI) melalui Program Pendidikan Vokasi menyelenggarakan kursus berbasis MOOC yang dirancang untuk memenuhi kebutuhan industri. Selain itu, Universitas Terbuka (UT) juga telah lama menerapkan sistem pembelajaran jarak jauh dan terus berinovasi dalam menyediakan akses pendidikan tinggi yang terbuka bagi masyarakat Indonesia. Ketiga, Universitas Ciputra melalui platform Universitas Ciputra Entrepreneurship Online (UCEO) juga telah mengembangkan MOOC untuk mendukung pembelajaran kewirausahaan secara daring (Rokhim & Rusydiyah, 2021). Tidak hanya institusi pendidikan, platform SPADA Indonesia, IndonesiaX dan Sekolah Pintar menjadi bukti nyata penerapan MOOC pada perusahaan (Santoso, 2018). Penerapan MOOC oleh institusi-institusi tersebut menunjukkan komitmen dalam memperluas akses pendidikan dan meningkatkan kualitas pembelajaran di era digital.

Trend implementasi MOOC saat ini juga merambah ke instansi pemerintah yang juga telah menerapkan MOOC untuk menyediakan program pelatihan bagi Aparatur Sipil Negara (ASN). Pada masa pandemi COVID-19, Badan Pengawasan Keuangan dan Pembangunan (BPKP) menyelenggarakan pendidikan dan pelatihan melalui MOOC untuk meningkatkan efektivitas program pelatihan bagi ASN, terutama kondisi keterbatasan anggaran (Irianto et al., 2023). Selanjutnya, diikuti oleh Kementerian Kesehatan (Kemenkes) yang turut mengembangkan pelatihan jabatan fungsional bidan berbasis cMOOC untuk meningkatkan kompetensi tenaga kesehatan di Indonesia (Dewi Wijayanti et al., 2023). Terbaru, terdapat program pelatihan ASN Berpijar oleh Pijar *Foundation* berkerjasama dengan Lembaga Administrasi Negara (LAN RI) yang menyediakan akses MOOC bagi seluruh ASN maupun NON ASN secara gratis tanpa batasan jarak, tempat dan waktu. Platform MOC tersebut dirancang untuk meningkatkan kompetensi ASN dalam menghadapi tantangan era digital dengan menyediakan berbagai topik pelatihan yang relevan dan praktis.

Sayangnya, pembelajaran menggunakan MOOC tidak selalu berjalan efektif dan berkualitas. Sebagian besar MOOC menggunakan soal evaluasi pilihan ganda sebagai metode evaluasi utama karena kemudahan dalam otomatisasi penilaian. Namun, pendekatan evaluasi pembelajaran ini memiliki keterbatasan mengukur pemahaman mendalam dan kemampuan berpikir kritis. Hal tersebut diperkuat oleh penelitian Costello (2018) yang menunjukkan bahwa banyak soal pilihan ganda dalam MOOC memiliki kelemahan penulisan item yang dapat memengaruhi validitas dan reliabilitas evaluasi. Selain itu, evaluasi MOOC asesmen non-otomatis juga perlu dipertimbangkan, karena cenderung memiliki tingkat penyelesaian yang lebih rendah karena beban kerja yang lebih besar (Jordan, 2014).

Evaluasi sumatif berupa pilihan ganda pada MOOC juga rentan terhadap praktik kecurangan. Salah satu teknik kecurangan yang populer yaitu "*Copying Answers using Multiple Existences Online*", (CAMEO). Dalam metode ini, seorang peserta membuat beberapa akun; akun "*harvester*" digunakan untuk menemukan jawaban yang benar, kemudian jawaban tersebut digunakan oleh akun utama untuk menyelesaikan evaluasi dengan skor tinggi. Penelitian oleh Northcutt, Ho, dan Chuang (2015) menemukan bahwa sekitar 1,3% sertifikat dalam 69 MOOC diperoleh melalui metode CAMEO, dengan prevalensi lebih tinggi di kalangan peserta internasional dan pria muda. Temuan ini menunjukkan bahwa meskipun MOOC menawarkan akses luas, mereka juga membuka peluang baru untuk kecurangan akademik.

Tak hanya itu, kehadiran kecerdasan buatan (AI) yang semakin terlatih berpotensi mengurangi kualitas pembelajaran pada platform MOOC yang hanya mengandalkan evaluasi soal pilihan ganda. Pemanfaatan AI memungkinkan peserta menyelesaikan ujian tanpa harus memahami seluruh materi. Lebih lanjut, AI dapat melakukan personalisasi materi pembelajaran, adaptasi terhadap kebutuhan individu, dan peningkatan interaktivitas dalam proses belajar mengajar (Utami et al., 2024). Berdasarkan pengujian, AI dapat digunakan untuk analisis teks jawaban esai atau simulasi interaktif. Salah satu model AI yang populer dan terbukti mampu menjawab soal pilihan ganda dari berbagai macam subjek materi dengan benar adalah ChatGPT (Alseddiqi et al., 2023). Bahkan, ChatGPT bisa mengerjakan soal evaluasi dengan menjawab soal dari berbagai mata pelajaran dan tingkat kesulitan

di platform EdX dan Coursera dengan rata-rata akurasi sebesar 83,5% (Gilson et al., 2023). Dimana pada platform tersebut, lebih banyak kursus yang bersifat teknis daripada manajerial.

Kemampuan ChatGPT dalam menjawab soal pilihan ganda memang telah mengalami peningkatan yang signifikan seiring dengan perkembangan versinya. Dalam sebuah tinjauan oleh (Newton & Xiromeriti, 2024) ChatGPT menunjukkan tingkat kelulusan sebesar 93% pada ujian berbasis pilihan ganda dengan subjek materi terkait topik kesehatan. Hasil penilaian tersebut melampaui performa mahasiswa yang hanya mampu mendapatkan tingkat kelulusan sebesar 35% dari ujian yang dianalisis. Studi lain oleh Kochanek et al., (2024) menunjukkan bahwa ChatGPT memiliki kemampuan dengan tingkat akurasi sebesar 65-69% dalam menyelesaikan soal materi diagnosis pendengaran. Penelitian tersebut menunjukkan konsistensi jawaban yang tinggi dari waktu ke waktu. Keberhasilan ChatGPT dalam menjawab soal pilihan ganda memiliki implikasi penting dalam pembelajaran MOOC.

Berdasarkan uraian diatas, dapat disimpulkan bahwa ChatGPT memiliki kemampuan menyelesaikan soal pilihan ganda dimana hal tersebut juga berpeluang menjadi potensi kecurangan. Namun, diantara banyak uji tingkat akurasi ChatGPT yang sudah dipublikasi, masih belum banyak penelitian yang dilakukan pada MOOC berbahasa Indonesia, khususnya untuk subjek materi terkait manajerial dalam rumpun ilmu sosial. Seperti yang disampaikan (Grassini, 2023) yang mencatat bahwa sebagian besar evaluasi ChatGPT masih berfokus pada konteks berbahasa Inggris dan belum mempertimbangkan tantangan kultural serta linguistik di Indonesia. Ditinjau dari aspek etika, Penelitian (Oravec, 2023) juga menekankan bahwa AI seperti ChatGPT dapat secara signifikan merusak integritas akademik jika tidak digunakan secara etis.

Oleh karena itu, penelitian ini akan berfokus untuk melakukan eksperimen pengujian tingkat akurasi penggunaan ChatGPT dalam menyelesaikan soal evaluasi pembelajaran MOOC pada subjek materi manajerial *Core Value ASN*. Topik tersebut menarik karena menjadi salah satu syarat wajib yang harus ditempuh CASN untuk menjadi ASN. Artinya, topik ini tidak hanya sekedar memiliki peran penting dalam pembelajaran, melainkan juga memiliki pertimbangan administratif bagi seorang CANS. Harapannya, penelitian ini dapat membantu dalam merumuskan sistem evaluasi pada pembelajaran MOOC yang mampu mengidentifikasi dan mengatasi potensi kecurangan akademik melibatkan AI. Dengan pendekatan holistik demikian, MOOC dapat terus berkembang sebagai platform pembelajaran yang efektif dan terpercaya kualitasnya.

METODOLOGI

Penelitian ini menggunakan metode eksperimen dengan pendekatan kuasi-eksperimental, dimana peneliti melakukan pengujian langsung terhadap kinerja model AI ChatGPT-4o versi gratis dalam menyelesaikan evaluasi pembelajaran pada platform MOOC ASN Berpijar yang dikelola oleh LAN RI. Penelitian ini dilakukan dengan menggunakan tiga akun ChatGPT-4o yang berbeda untuk mengerjakan delapan kursus online dengan materi *Core Values ASN*. Setiap kursus hanya dikerjakan sebanyak satu kali, tanpa memberikan perintah tambahan atau modifikasi *prompt*. Secara teknis, peneliti hanya menyalin soal dan jawaban, kemudian memasukkan jawaban yang dihasilkan ke dalam platform MOOC.

Metode eksperimen ini dipilih karena memungkinkan peneliti untuk mengamati langsung pengaruh penggunaan AI dalam konteks pembelajaran, sebagaimana dijelaskan oleh (Creswell, 2014), bahwa pendekatan ini digunakan untuk mengukur efek perlakuan tertentu terhadap variabel yang diamati. Dalam konteks ini, ChatGPT diposisikan sebagai perlakuan yang diuji efektivitasnya dalam menyelesaikan soal-soal yang tersedia pada platform MOOC.

Fraenkel, Wallen, dan Hyun (2012) menekankan bahwa eksperimen sangat relevan dalam penelitian pendidikan, terutama saat menilai efektivitas alat bantu pembelajaran seperti teknologi kecerdasan buatan. Sejalan dengan itu, Luckin et al. (2016) mengungkapkan bahwa AI memiliki potensi besar dalam mendukung proses belajar, dan eksperimen langsung merupakan metode yang valid untuk mengevaluasi peranannya dalam pendidikan. Tak hanya itu, studi sistematis oleh Zawacki-Richter et al. (2019) juga menunjukkan bahwa pendekatan eksperimental banyak digunakan dalam

penelitian terkait aplikasi AI dalam pendidikan tinggi, yang memperkuat validitas metode yang digunakan dalam penelitian ini.

HASIL DAN PEMBAHASAN

Peneliti mendeskripsikan hasil dan pembahasan kedalam 4 pokok bahasan. Dalam pengujian tingkat akurasi sebuah model AI, pembagian ini penting dilakukan untuk menjelaskan secara detail terkait prosedur eksperimen serta hasil yang telah diperoleh. Pertama, pembagian meliputi desain evaluasi pembelajaran MOOC yang diuji, dalam hal ini adalah ASN Berpijar yang dikelola oleh LAN RI. Kedua, prosedur *prompting* yang dilakukan untuk mendapatkan jawaban dari setiap soal yang diujikan pada MOOC. Ketiga, hasil analisis pada materi *Core Value ASN*. Terakhir, mendeskripsikan potensi, ancaman dan rekomendasi etis.

Desain Evaluasi Pembelajaran pada MOOC



Gambar 1 Tampilan MOOC ASN Berpijar

Pengujian dilakukan pada menggunakan soal evaluasi pembelajaran yang terdapat pada platform MOOC ASN Berpijar yang dikelola oleh LAN RI. Terdapat 2 kategori peserta dalam MOOC ini yaitu ASN dan NON ASN. Dalam pengujian ini, pelatihan yang diambil adalah *Core Value ASN Berakhlak* yang terdiri dari delapan kursus dengan syarat pelatihan bebas dan terbuka bagi NON ASN.

Secara umum, alur pembelajaran belajar platform MOOC adalah 1) melakukan pembuatan akun, 2) mendaftar kursus yang telah disediakan, 3) melakukan pembelajaran secara berurutan dengan mempelajari materi setiap kursus berupa dokumen dan/atau video, 4) melakukan kuis diakhir sesi pembelajaran. Pada penelitian ini, materi tidak diujikan. Penelitian langsung berfokus pada pengerjaan kuis masing-masing kursus. Setiap pelatihan memiliki sertifikat kelulusan sebagai bukti telah berhasil menyelesaikan MOOC. Dari aspek evaluasi pembelajaran, setiap kuis memiliki batasan waktu pengerjaan, namun tidak dibatasi untuk jumlah melakukan tes, sehingga peserta dapat melakukan tes secara berulang.

Alur pembelajaran dan evaluasi ASN Berpijar ini sangatlah sesuai dengan karakteristik dari MOOC. Metode kuantitatif meliputi pre-test dan post-test, kuis pilihan ganda, serta penilaian otomatis yang mudah diskalakan untuk ribuan pengguna (Wang et al., 2021). Dalam hal ini, pengujian yang dilakukan adalah evaluasi pembelajaran atau post-test.

Prompting untuk Analisis Tingkat Akurasi ChatGPT

Poin paling penting dalam pengujian tingkat akurasi ChatGPT adalah *prompting*. Prompting pada ChatGPT merujuk pada teks masukan (input) atau instruksi yang diberikan kepada model bahasa besar (*Large Language Model*) untuk memicu dan mengarahkan respons yang dihasilkan. Kualitas dan kejelasan prompt sangat memengaruhi relevansi, akurasi, dan koherensi jawaban yang dihasilkan oleh

ChatGPT (OpenAI Help Center, 2025). Dalam penelitian ini, peneliti tidak melakukan *prompting* tambahan. Hal ini bertujuan untuk menghasilkan jawaban dengan tingkat paling minimum yang dihasilkan dengan asumsi semakin baik hasil jawabannya, semakin baik pula tingkat akurasi.

Ujicoba Pengerjaan Evaluasi Pembelajaran Materi CoreValue ASN

Terdapat 4 jenis pengetahuan dalam konteks pembelajarannya yaitu pengetahuan yang bersifat faktual, konseptual, prosedural dan meta-kognitif. Kategorisasi jenis pengetahuan ini sangat berpengaruh pada tingkat akurasi ChatGPT dalam mengerjakan soal evaluasi. Berdasarkan penelitian sebelumnya, materi yang bersifat faktual dan prosedural memiliki tingkat keberhasilan yang tinggi. Sedangkan untuk materi yang bersifat konseptual dan meta-kognitif memerlukan *prompting* dan referensi yang harus ditambahkan secara manual (Bang et al., 2023)

Materi pengujian adalah *Core Values ASN* yang terdiri dari 8 kursus. Materi ini terdiri dari materi yang bersifat konseptual, yaitu melibatkan pemahaman tentang hubungan antara konsep dan teori, serta bagaimana konsep-konsep tersebut saling berkaitan. Berbeda dengan materi faktual dan prosedural, materi konseptual manajerial membutuhkan data tambahan agar ChatGPT mampu menginterpretasikan soal dengan benar. Pengujian dilakukan menggunakan tiga akun ChatGPT versi gratis, setiap akun mengerjakan masing-masing kursus dengan maksimal 1 kali percobaan. Sebelum melakukan pengujian, peneliti melakukan *cross check* kunci jawaban secara manual, kemudian menambahkan poin jika terdapat kesalahan. Berdasarkan hasil koreksi, terdapat 1 soal pada kursus BerAkhlak – Adaptif (3 JP) yang tidak sesuai dengan kunci jawaban.

Analisis tingkat akurasi sebuah model AI untuk evaluasi pembelajaran, secara detail dapat dilakukan melalui kriteria sebagai berikut:

1. Tidak menggunakan *prompting* tambahan yang mempengaruhi analisis soal.
2. Menggunakan akun yang berbeda untuk menghindari perbaikan ataupun redudansi data.
3. Materi yang diuji memiliki kesamaan jenis soal, subjek atau topik materi serta memiliki soal yang sama. Urutan jawaban maupun soal tidak menjadi kendala.
4. Pelaksanaan dilakukan pada rentan waktu yang sama
5. Tidak menginputkan data referensi tambahan untuk melatih AI
6. Prosedur uji coba dilakukan dengan menyalin seluruh soal dalam 1 inputan secara serentak, dilanjutkan dengan mengolah soal menggunakan ChatGPT dan terakhir menyelesaikan soal pada MOOC.

Berdasarkan hasil eksperimen yang telah dilakukan, diperoleh data keberhasilan setiap akun sebagaimana Tabel 1 berikut:

Tabel 1 Hasil Uji coba pengerjaan Evaluasi Pembelajaran menggunakan ChatGPT

No	Materi	Nilai Akun			Tingkat Kelulusan
		Akun 1	Akun 2	Akun 3	
1	Core Values ASN Berakhlak (7 JP)	80	80	80	100%
2	Berakhlak - Berorientasi Pelayanan (3 JP)	90	100	90	100%
3	Berakhlak - Akuntabel (3 JP)	90	80	80	100%
4	Berakhlak - Kompeten (3 JP)	80	80	80	100%
5	Berakhlak - Harmonis (3 JP)	90	90	90	100%
6	Berakhlak - Loyal (3 JP)	50	50	50	0%
7	Berakhlak - Adaptif (3 JP)	80	80	80	100%
8	Berakhlak - Kolaboratif (3 JP)	90	90	90	100%
Rata – rata		81.25	81.25	80	87,5%.

Berdasarkan hasil eksperimen diatas, analisis tingkat akurasi ChatGPT dalam menyelesaikan evaluasi pembelajaran pada MOOC dengan materi konseptual manajerial dapat dijabarkan berikut :

1. ChatGPT mampu menyelesaikan 7 dari 8 kursus dengan benar. Sehingga dapat disimpulkan bahwa tingkat akurasi ChatGPT untuk menjawab soal pilihan ganda yaitu sebesar 87,5%.
2. Setiap kelulusan dilakukan hanya dengan satu kali percobaan pada masing-masing akun. Artinya tanpa *prompting* lanjutan untuk melatih dalam menjawab soal, ChatGPT sudah

memiliki kemampuan menjawab dengan tingkat akurasi tinggi secara *default*. Hal ini didukung oleh dataset yang dimiliki oleh OpenAI dalam melatih model AI ChatGPT. Model ini telah mengalami perkembangan sejak pertama kali rilis.

3. Terdapat pola jawaban yang dihasilkan oleh ChatGPT, yaitu setiap jawaban selalu konsisten baik benar maupun salahnya. Terbukti dari setiap akun yang tidak menunjukkan perbedaan hasil evaluasi belajar yang signifikan. Akun 1 dan akun 2 memiliki rata-rata yang sama dan hanya 2 dari 8 materi yang memiliki perbedaan nilai rata-rata.

Selain analisis diatas, ChatGPT juga memiliki kegagalan dalam mengerjakan soal pilihan ganda. Pada kursus Berakhlak - Loyal (3 JP), tidak ada satu akun pun yang lulus. Hasil tersebut menunjukkan bahwa ChatGPT hanya mampu menjawab 5 soal di setiap akun. Berdasarkan hasil pengamatan dengan membandingkan jawaban dan kunci, kegagalan ChatGPT disebabkan oleh hal diantaranya

1. Kemiripan konseptual dalam opsi jawaban, contohnya seperti makna kata komitmen, dedikasi dan pengabdian yang menjadi opsi jawaban;
2. ChatGPT lebih cenderung overfitting ke konteks umum dalam menjawab soal penalaran yang humanis, ChatGPT membutuhkan pendekatan yang normatif dan berdasarkan regulasi panduan pemerintah;
3. Pemilihan jawaban yang memiliki makna semantik lebih cenderung menjadi pilihan daripada kalimat eksplisit di pertanyaan; dan
4. Dalam aspek bahasa, ChatGPT lebih memilih jawaban yang paling menjelaskan atau masuk akal secara normatif, padahal soal menguji pemahaman terkait terminologi resmi atau frasa khusus. Contohnya, untuk memaknai kata loyal, ChatGPT lebih memilih kata “taat” karena lebih sering muncul dalam asosiasi praktis, padahal maknanya lebih bermakna “setia”.

Selain meninjau dari aspek keberhasilan dan kegagalan ChatGPT dalam menyelesaikan evaluasi pembelajaran, perlu juga untuk menganalisis implementasi berdasarkan prosedur teknis sebagai berikut,

Tabel 2 Analisis Kriteria Persyaratan Keberhasilan ChatGPT dalam Menjawab Soal Pilihan Ganda

No	Kriteria	Hasil Pengujian
1	Wajib menggunakan akun yang berbeda untuk mendapatkan hasil yang optimal	Tidak terbukti, setiap akun baik lama maupun baru dapat menjawab soal dengan hasil yang sama
2	Menambahkan <i>prompt</i> khusus untuk mendapatkan hasil yang maksimal	Tidak terbukti, dengan menyalin dan mengolahnya di ChatGPT terbukti dapat menghasilkan jawaban yang sesuai. Namun, <i>prompt</i> khusus diperlukan untuk mengkoreksi jika ada kesalahan jawaban.
3	Menggunakan ChatGPT versi Berbayar	Tidak terbukti, penelitian ini menggunakan akun ChatGPT versi gratis. Tingkat keberhasilan bukan terletak pada lisensi yang digunakan, melainkan model AI yang digunakan. Saat penelitian ini dilakukan menggunakan model AI ChatGPT-4o.
4	Memerlukan dokumen pendukung untuk melatih berdasarkan topik	Tidak terbukti, secara <i>default</i> ChatGPT telah memiliki dataset yang cukup lengkap untuk menjawab pertanyaan soal pilihan ganda. Namun, penambahan dokumen pendukung tentu dapat meningkatkan tingkat akurasi jawaban
5	Tidak bisa menjawab soal yang memiliki makna kata yang mirip	Terbukti, ChatGPT tidak mampu menyelesaikan kursus Berakhlak - Loyal (3 JP) karena soal yang disajikan lebih banyak menggunakan kata dengan terminologi atau frasa khusus.

Potensi, Ancaman dan Rekomendasi Etis Penggunaan ChatGPT pada Evaluasi Pembelajaran Massive Open Online Course

Berdasarkan hasil ujicoba yang telah dilakukan, maka diperlukan analisis potensi, ancaman dan rekomendasi lebih lanjut terkait penggunaan ChatGPT dalam pembelajaran, khususnya dalam pembelajaran berbasis web menggunakan MOOC. Hal ini penting seiring perkembangan AI dan implikasinya pada aspek pendidikan dan pembelajaran. Adapun analisisnya sebagaimana Tabel 3.

Tabel 3 Potensi dan Ancaman Penggunaan ChatGPT pada Evaluasi Pembelajaran MOOC

No	Potensi	Ancaman
1	Peserta dapat menyelesaikan evaluasi pembelajaran MOOC tanpa mempelajari materi terlebih dahulu	Menurunnya kualitas pembelajaran khususnya pembelajaran yang menggunakan platform MOOC.
2	Tidak ada pembatasan dalam pengerjaan evaluasi pembelajaran	Peserta melakukan evaluasi secara berulang. Hal ini dapat memantik penggunaan <i>prompt</i> lanjutan untuk mengoreksi jawaban. Sehingga, bukan tidak mungkin setiap peserta mendapatkan nilai maksimal tanpa melalui proses pembelajaran yang bermakna.
3	Tidak ada pembatasan navigasi browser saat mengakses MOOC	Potensi kecurangan dengan mengerjakan soal evaluasi menggunakan bantuan AI, khususnya ChatGPT yang hanya perlu memasukkan soal tanpa ada <i>prompt</i> khusus.
4	Pembuatan akun tanpa verifikasi	Hal ini berpotensi meningkatkan kecurangan menggunakan metode <i>Copying Answers using Multiple Existences Online</i> (CAMEO)

Implementasi MOOC yang baik tidak hanya memperhatikan aspek pembelajaran saja, melainkan juga mempertimbangkan aspek teknis serta kualitas materi yang disediakan. Dalam studi oleh Margaryan, Bianco, dan Littlejohn (2015) menemukan bahwa banyak MOOC gagal memenuhi prinsip dasar desain instruksional yang mencakup kejelasan tujuan, aktivitas belajar bermakna, dan asesmen autentik. Hal ini menunjukkan bahwa keberhasilan MOOC sangat bergantung pada integrasi antara konten, desain pedagogis, dan teknologi pendukung. Konten yang tidak relevan atau disampaikan dengan cara yang monoton dapat menurunkan keterlibatan peserta dan efektivitas pembelajaran secara keseluruhan.

Aspek teknis dari MOOC mencakup aksesibilitas platform, kemudahan navigasi, dan ketersediaan dukungan teknis selama proses belajar. Albelbisi dan Yusop (2019) menekankan bahwa faktor seperti keandalan sistem, tampilan antarmuka, dan responsivitas berperan penting dalam meningkatkan kepuasan pengguna terhadap MOOC. Jika pengguna mengalami kesulitan teknis yang berulang, hal tersebut dapat menurunkan motivasi belajar. Oleh karena itu, penyelenggara MOOC perlu memastikan bahwa platform yang digunakan mampu mendukung pembelajaran dalam skala besar dengan performa teknis yang stabil.

Selain teknis dan pedagogis, kualitas materi ajar juga menjadi faktor penentu keberhasilan MOOC. Menurut Stracke et al. (2021), konten MOOC harus memenuhi standar kualitas yang mencakup relevansi topik, kejelasan penyampaian, dan keberagaman format (video, bacaan, kuis interaktif) untuk mendukung berbagai gaya belajar peserta. Tanpa perhatian terhadap kualitas konten, peserta akan kesulitan mencapai kompetensi yang ditargetkan. Oleh karena itu, kurasi dan perancangan materi MOOC harus melibatkan kolaborasi antara ahli materi, desainer instruksional, dan teknolog pendidikan.

Adapun detail rekomendasi yang perlu dilakukan untuk mengembangkan platform MOOC dapat dijabarkan sebagai berikut

Tabel 4 Rekomendasi Desain Sistem dan Pembelajaran MOOC untuk Menghindari Penyalahgunaan AI

No	Aspek	Rekomendasi

1	Teknis	1. Meskipun bersifat “Massive”, juga diperlukan sistem validasi akun yang kuat untuk membatasi pengguna membuat lebih dari satu akun. Contohnya, peserta wajib menggunakan email kampus, perusahaan atau identitas unik lainnya; 2. Membatasi navigasi browser saat mengakses halaman evaluasi pembelajaran seperti membatasi copy-paste, tidak boleh beralih tab browser dan mengunci sesi.
2	Pembelajaran	1. Menyajikan materi sesuai dengan jenis pengetahuannya, baik materi untuk jenis pengetahuan fakta, konseptual, prosedur maupun meta-kognitif; 2. Terdapat penyelenggara yang bertugas memantau ataupun memonitoring pelaksanaan MOOC; 3. Menggunakan batasan waktu penyelenggaraan MOOC seperti dapat diakses pada jam kerja, tanggal tertentu atau durasi setelah pertama kali <i>enroll</i>.
3	Kualitas Materi	1. Mengunci file materi pembelajaran menggunakan kata sandi atau membatasi akses pembacaan meta data; 2. Menggunakan video pembelajaran yang tidak disimpan pada server <i>private</i>; 3. Menyediakan media pembelajaran berupa simulasi untuk materi yang bersifat prosedural; 4. Menjadikan materi sebagai dasar utama pembuatan soal evaluasi pembelajaran.
4	Evaluasi Pembelajaran	1. Menyajikan soal dengan bentuk gambar sehingga tidak mudah disalin atau membutuhkan usaha lebih; 2. Menggunakan kosa kata dengan makna eksplisit; 3. Jika materi bersifat fakta dan prosedural, alangkah lebih baik menggunakan model evaluasi berupa penugasan berbasis project; 4. Membatasi jumlah pengulangan evaluasi, bahkan jika diperlukan hanya menyediakan satu kali pengujian untuk setiap evaluasi pembelajaran. Jika gagal, bisa menggunakan paket yang berbeda.

KESIMPULAN

ChatGPT terbukti memiliki potensi yang cukup signifikan untuk menjawab soal evaluasi. Berdasarkan hasil eksperimen, ChatGPT mampu menyelesaikan 7 dari 8 kursus dengan benar. Sehingga akurasi ChatGPT menjawab soal pilihan ganda materi konseptual manajerial sebesar 87,5%. Faktor penentu tingkat akurasi ChatGPT terletak pada soal dengan makna konseptual yang sama. Berdasarkan hasil penelitian diatas, dapat disimpulkan bahwa ChatGPT terbukti dapat membantu peserta mengerjakan soal evaluasi pembelajaran. Hal ini membuka potensi kecurangan bagi peserta. Harapannya, untuk meningkatkan kualitas pembelajaran MOOC, peneliti memberikan rekomendasi dengan melakukan tindakan diantaranya 1) menggunakan soal berupa gambar, 2) membatasi navigasi serta mencegah perpindahan tab browser laman MOOC untuk menghindari kecurangan, dan 3) gunakan soal studi kasus yang dilengkapi dengan materi konseptual. Selain itu, peneliti menyarankan bahwa perlu dilakukan penelitian lanjutan terkait desain evaluasi pembelajaran yang mampu meminimalisir kecurangan pemanfaatan model AI ChatGPT dalam pembelajaran MOOC.

Deklarasi

Penelitian ini dilakukan untuk meningkatkan kualitas pembelajaran mandiri dengan memanfaatkan platform MOOC. Harapannya, hasil penelitian ini dapat mendukung pembelajaran yang efektif dan efisien ditengah gempuran teknologi kecerdasan buatan. Semoga dapat menjadi kontribusi positif khazanah ilmu pendidikan di masa mendatang.

REFERENSI

- Adedoyin, O., & Famolu, M. (2020). *Massive Open Online Courses: Pros & Cons*.
- Albelbisi Dan, Y. (2019). Factors Influencing Learners’ Self-Regulated Learning Skills in a Massive Open Online Course (MOOC) Environment. *Turkish Online Journal of Distance Education*, 20, 1–16.

- Alseddiqi, M., AL-Mofleh, A., Albalooshi, L., & Najam, O. (2023). Revolutionizing Online Learning: The Potential of ChatGPT in Massive Open Online Courses. *European Journal of Education and Pedagogy*, 4(4), 1–5. <https://doi.org/10.24018/ejedu.2023.4.4.686>
- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., Do, Q. V, Xu, Y., & Fung, P. (2023). A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity. In *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics* (Vol. 1, pp. 675–718). Association for Computational Linguistics.
- Costello, E., Holland, J. C., & Kirwan, C. (2018). Evaluation of MCQs from MOOCs for common item writing flaws. *BMC Research Notes*, 11(1), 849. <https://doi.org/10.1186/s13104-018-3959-4>
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Sage.
- Despujol, I., Castañeda, L., Marín, V. I., & Turró, C. (2022). What do we want to know about MOOCs? Results from a machine learning approach to a systematic literature mapping review. *International Journal of Educational Technology in Higher Education*, 19(1), 53. <https://doi.org/10.1186/s41239-022-00359-1>
- Dewi Wijayanti, R., Meliala, A., Handono Sulisty Departemen Kebijakan dan Manajemen Kesehatan, D., Kedokteran, F., Masyarakat, K., & Keperawatan, dan. (2023). Implementasi Massive Open Online Course (MOOC) pada Pelatihan Jabatan Fungsional Bidan (Pangkatan Pertama) di Bapelkes Semarang. In *Jurnal Kebijakan Kesehatan Indonesia : JKKI* (Vol. 12, Issue 03).
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research In education*. Mc Graw Hill.
- Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Medical Education*, 9, e45312. <https://doi.org/10.2196/45312>
- Grassini, S. (2023). Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Education Sciences*, 13(7), 692. <https://doi.org/10.3390/educsci13070692>
- Irianto, J., Rodiyah, I., & Omar, K. (2023). The Use of the Unified Theory of Acceptance and The Use of Technology (UTAUT) to Analyze the Implementation of the Massive Open Online Course (MOOC) at the Indonesian Financial and Development Supervisory Agency. *Jurnal Borneo Administrator*, 19(2), 175–190. <https://doi.org/10.24258/jba.v19i2.1287>
- Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, 15(1). <https://doi.org/10.19173/irrodl.v15i1.1651>
- Kochanek, K., Skarzynski, H., & Jedrzejczak, W. W. (2024). Accuracy and Repeatability of ChatGPT Based on a Set of Multiple-Choice Questions on Objective Tests of Hearing. *Cureus*. <https://doi.org/10.7759/cureus.59857>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence Unleashed: An Argument for AI in Education*. Pearson.

- Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education*, 80, 77–83. <https://doi.org/10.1016/j.compedu.2014.08.005>
- N., A., Kulal, A., M.S., D., & Dinesh, S. (2025). Effectiveness of MOOCs on learning efficiency of students: a perception study. *Journal of Research in Innovative Teaching & Learning*, 18(1), 145–164. <https://doi.org/10.1108/JRIT-12-2022-0091>
- Newton, P., & Xiromeriti, M. (2024). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. *Assessment & Evaluation in Higher Education*, 49(6), 781–798. <https://doi.org/10.1080/02602938.2023.2299059>
- Northcutt, C. G., Ho, A. D., & Chuang, I. L. (n.d.). *Detecting and Preventing “Multiple-Account” Cheating in Massive Open Online Courses*.
- OpenAI Help Center. (2025, April 21). *Prompt engineering best practices for ChatGPT*. <https://help.openai.com/en/articles/10032626-prompt-engineering-best-practices-for-chatgpt>.
- Oravec, J. A. (2023). Artificial Intelligence Implications for Academic Cheating: Expanding the Dimensions of Responsible Human-AI Collaboration with ChatGPT and Bard. In *Jl. of Interactive Learning Research* (Vol. 34, Issue 2).
- Rokhim, A. A., & Rusydiyah, E. F. (2021). PEMANFAATAN MASSIVE OPEN ONLINE COURSES DALAM PEMBELAJARAN AL QUR’AN DI ERA DIGITAL. *EDURELIGIA: Jurnal Pendidikan Agama Islam*, 5(2), 83–96. <https://doi.org/10.33650/edureligia.v5i2.2518>
- Santoso, H. B. (2018). Indonesian Perspective on Massive Open Online Courses: Opportunities and Challenges. *Journal of Educators Online*, 15(1). <https://doi.org/10.9743/JEO2018.15.1.11>
- Stracke, C. M., & Trisolini, G. (2021). A systematic literature review on the quality of MOOCs. *Sustainability*, 13(11), 5817. <https://doi.org/10.3390/su13115817>
- Utami, I. S., Budi, S., Arnez, G., & Hafid, A. (2024). Needs Analysis Study of Blended Learning Model to Improve Communication for Students with Hearing Impairment. *Pedagogia : Jurnal Pendidikan*, 13(2), 216–224. <https://doi.org/10.21070/pedagogia.v13i2.1689>
- Wang, X., Lee, Y., Lin, L., Mi, Y., & Yang, T. (2021). Analyzing instructional design quality and students’ reviews of 18 courses out of the Class Central Top 20 MOOCs through systematic and sentiment analyses. *The Internet and Higher Education*, 50, 100810. <https://doi.org/10.1016/j.iheduc.2021.100810>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education-Where Are the Educators? *International Journal of Educational Technology in Higher Education*, 39.
-