



## PENGUNAAN AUTOGRADER DALAM UJIAN SUMATIF SISWA KELAS X

### THE USE OF AUTOGRADER IN THE SUMMATIVE ASSESSMENT OF TENTH GRADE VOCATIONAL STUDENTS

Ayra Zalikha Munir<sup>1</sup>, Adelia Ramadhani<sup>2</sup>, Nadifah Amaliyah<sup>3</sup>, Musakkar<sup>4</sup>, Nur Aeni<sup>5\*</sup>

<sup>1,2,3,4,5\*</sup> Universitas Negeri Makassar, Makassar, Makassar, Indonesia

\*email (<sup>1</sup>ayramunir2407@gmail.com, <sup>2</sup>adeliaramadhani664@gmail.com,

<sup>3</sup>nadiifahamaliyah@gmail.com, <sup>4</sup>anugrah.putra828@gmail.com <sup>6</sup>nur\_aeni@unm.ac.id)

**Abstrak:** Kegiatan pengabdian kepada masyarakat ini dilatarbelakangi oleh permasalahan penilaian ujian secara manual yang masih umum dilakukan di sekolah, termasuk di SMKN 10 Makassar. Tujuan dari kegiatan ini adalah untuk mengimplementasikan aplikasi *Autograder* berbasis kecerdasan buatan guna membantu guru dalam melakukan penilaian ujian sumatif secara lebih efisien dan objektif. Metode yang digunakan meliputi pendampingan saat pelaksanaan ujian, serta evaluasi melalui observasi dan survey kepada siswa. Hasil kegiatan menunjukkan bahwa penggunaan *Autograder* mampu menghemat waktu koreksi dan memberikan hasil penilaian secara otomatis, terutama untuk soal pilihan ganda. Namun, untuk soal esai, aplikasi ini masih memiliki keterbatasan dalam menilai ragam jawaban yang berbeda struktur tetapi memiliki makna serupa.

**Kata Kunci:** *Autograder*, kecerdasan buatan, penilaian otomatis, ujian sumatif, SMK.

**Abstract:** *This community service activity is motivated by the problem of manual exam grading which is still commonly done in schools, including at SMKN 10 Makassar. The purpose of this activity is to implement an artificial intelligence-based Autograder application to assist teachers in conducting summative exam grading more efficiently and objectively. The methods used include assistance during the implementation of the exam, as well as evaluation through observations and surveys to students. The results of the activity show that the use of Autograder is able to save correction time and provide assessment results automatically, especially for multiple choice questions. However, for essay questions, this application still has limitations in assessing a variety of answers that are different in structure but have similar meanings.*

**Keywords:** *Autograder, artificial intelligence, automated assessment, summative assessment, vocational school.*

#### Article History:

Received	Revised	Published
20 April 2025	10 Mei 2025	15 Mei 2025

#### Pendahuluan

Berdasarkan Permendikbud No. 23 Tahun 2016, penilaian merupakan proses yang melibatkan pengumpulan dan analisis informasi untuk mengukur pencapaian hasil belajar siswa. Proses ini tidak hanya bertujuan untuk mengetahui sejauh mana siswa memahami materi, tetapi juga sebagai bagian integral dari pembelajaran itu sendiri. Penilaian yang menyeluruh memerlukan berbagai teknik asesmen dengan instrumen yang valid dari beragam sumber, guna memberikan umpan balik yang bermakna bagi pendidik, peserta didik, dan orang tua/wali siswa (Purnawanto, 2022). Evaluasi dalam pembelajaran memiliki peran krusial, tidak

hanya dalam menilai kemampuan peserta didik, tetapi juga dalam mengukur efektivitas metode pembelajaran yang digunakan guru serta menentukan arah pengembangan selanjutnya.

Namun, dalam praktiknya, proses penilaian—khususnya ujian sumatif secara manual—masih menghadapi berbagai tantangan di lingkungan SMK. Salah satu persoalan utama adalah tidak adanya standar nilai yang jelas, yang membuat siswa kurang menghargai proses evaluasi. Perbedaan standar antar guru atau sekolah dapat memunculkan ketidakadilan dan menimbulkan pertanyaan terhadap objektivitas hasil penilaian (Hadrianingtyas et al., 2024). Di sisi lain, penilaian manual memerlukan waktu dan tenaga yang besar, mulai dari penyusunan soal, pelaksanaan ujian, hingga proses koreksi, yang pada akhirnya membebani guru, apalagi jika tidak didukung dengan infrastruktur atau SDM yang memadai (Ramesh & Sanampudi, 2022).

Seiring berkembangnya teknologi digital, penilaian otomatis menjadi salah satu solusi yang mulai banyak dikembangkan dan diadopsi, terutama untuk mengatasi tantangan penilaian manual di kelas besar. Sistem ini memungkinkan guru menilai banyak tugas dalam waktu singkat tanpa mengorbankan kualitas, objektivitas, maupun keadilan. Penilaian otomatis juga mengurangi potensi bias dan inkonsistensi penilaian yang dilakukan secara manual, serta memberikan umpan balik yang cepat kepada siswa (Acuña & Bansal, 2024).

Meskipun sistem ujian berbasis komputer awalnya hanya mampu menangani soal pilihan ganda, perkembangan teknologi kecerdasan buatan (AI) dan pemrosesan bahasa alami (Natural Language Processing/NLP) memungkinkan diterapkannya penilaian otomatis untuk soal essay. Tantangan dalam menilai aspek seperti relevansi konten, kohesi, dan pengembangan ide dalam jawaban essay kini mulai diatasi dengan kehadiran large language model (LLM) berbasis AI, meskipun masih terus disempurnakan (Ramesh & Sanampudi, 2022; Tisha et al., 2022).

Menjawab kebutuhan tersebut, mahasiswa Program Studi Pendidikan Teknik Informatika dan Komputer (PTIK) Universitas Negeri Makassar mengembangkan sebuah aplikasi berbasis AI bernama *Autograd*. Aplikasi ini memiliki dua fitur utama, yaitu:

1. Penilaian Otomatis Essay, yang menggunakan teknologi LLM berbasis DeepSeek API untuk menilai jawaban siswa secara cepat dan objektif.
2. Ujian Pilihan Ganda Otomatis, yang memungkinkan pelaksanaan ujian berbasis digital dengan sistem skor otomatis dan penyimpanan riwayat nilai siswa.

Pengembangan aplikasi ini menggunakan pendekatan Agile Development, melalui tahapan iteratif seperti perencanaan, analisis kebutuhan, desain sistem, implementasi, dan pengujian secara berkelanjutan. Dengan demikian, *Autograder* tidak hanya menjadi alat bantu teknis, tetapi juga solusi praktis dan inovatif untuk meningkatkan efisiensi, transparansi, serta objektivitas penilaian, khususnya dalam pelaksanaan ujian sumatif di tingkat SMK yang membutuhkan akurasi dan kecepatan dalam proses evaluasi.

Oleh karena itu, kegiatan pengabdian kepada masyarakat (PkM) ini bertujuan untuk mengimplementasikan dan mengevaluasi kinerja *Autograder* dalam konteks nyata melalui ujian siswa kelas X TKJ di SMKN 10 Makassar, guna melihat manfaat, respon siswa, dan tantangan teknis yang mungkin dihadapi dalam penggunaannya.

## Metode

Kegiatan ini merupakan program yang berfokus pada pemanfaatan teknologi pendidikan, khususnya dalam bentuk penerapan *Autograder* sebagai alat bantu penilaian ujian sumatif. Sasaran utama kegiatan adalah siswa kelas X dari jurusan Teknik Komputer dan Jaringan (TKJ) di Sekolah Menengah Kejuruan (SMK) Negeri 10 Makassar, yang telah memiliki dasar pengetahuan dalam bidang pemrograman.

Platform *Autograder* menggunakan large language model (LLM) berbasis deepseek API, *Autograder* dapat menilai jawaban essay siswa secara cepat dan objektif, sehingga mengurangi potensi subjektivitas penilaian oleh guru dan dipilih karena kemampuannya dalam memberikan penilaian otomatis terhadap jawaban berbentuk *multiple choice*. Kegiatan dilaksanakan melalui beberapa tahapan, yaitu: (1) implementasi *Autograder* secara langsung dalam pelaksanaan ujian sumatif; dan (2) evaluasi hasil penilaian dengan membandingkan efektivitas antara metode otomatis dan manual.

Teknik pengumpulan data meliputi observasi langsung selama proses implementasi, penyebaran survey kepada siswa untuk mengetahui respon terhadap penggunaan *Autograder*, serta dokumentasi hasil ujian siswa. Analisis data dilakukan secara kuantitatif deskriptif dengan membandingkan waktu koreksi ujian dan konsistensi nilai antara dua metode penilaian. Hasil dari analisis ini menjadi dasar dalam menilai efektivitas *Autograder* sebagai alat bantu penilaian yang efisien dan objektif.

## Hasil dan Pembahasan

### Hasil

#### 1. Perbandingan Waktu Koreksi Manual vs *Autograder*

Hasil implementasi menunjukkan bahwa penggunaan *Autograder* secara signifikan mempercepat proses koreksi dibandingkan dengan metode manual. Dalam praktiknya, koreksi manual memerlukan waktu yang cukup lama, terutama apabila guru harus menilai soal-soal esai secara satu per satu. Proses ini tidak hanya menghabiskan waktu tetapi juga tenaga, terlebih jika jumlah siswa yang mengikuti ujian cukup besar. Hal ini menjadi tantangan tersendiri bagi guru yang harus menyelesaikan penilaian dalam waktu terbatas, terutama menjelang akhir semester.

Sebaliknya, *Autograder* mampu memberikan skor secara otomatis dalam waktu yang sangat singkat, hanya beberapa detik setelah siswa menyelesaikan ujian. Kecepatan ini menjadi salah satu keunggulan utama *Autograder* dalam meningkatkan efisiensi pelaksanaan ujian. Berdasarkan hasil survei yang dilakukan terhadap siswa kelas X, sebagian besar siswa memberikan tanggapan positif terhadap kecepatan sistem dalam menampilkan nilai. Sekitar 80% responden menyatakan bahwa skor yang mereka peroleh muncul dengan cepat setelah ujian berakhir. Hal ini menunjukkan bahwa sistem berhasil mengurangi waktu tunggu dan memberikan kepastian kepada siswa mengenai hasil yang mereka peroleh, yang pada akhirnya berdampak pada peningkatan transparansi dan kepuasan siswa terhadap proses evaluasi.

Keunggulan dalam hal kecepatan ini sangat penting dalam konteks pendidikan vokasi yang cenderung menuntut efisiensi dan ketepatan waktu. Dengan berbagai kegiatan praktik yang harus dijalani siswa SMK, waktu pelaksanaan ujian dan koreksi perlu dilakukan seefisien mungkin agar tidak mengganggu proses pembelajaran lainnya. Dengan demikian, *Autograder* terbukti mampu mendukung kebutuhan praktis di lingkungan SMK, terutama dalam pengelolaan waktu koreksi ujian.

#### 2. Konsistensi Nilai antara Siswa dengan Kunci Jawaban

Meskipun keunggulan dari segi efisiensi waktu sangat menonjol, hasil implementasi juga mengungkap adanya tantangan dalam hal konsistensi nilai, terutama pada soal esai. *Autograder* yang digunakan cenderung hanya mengenali jawaban yang sesuai secara literal atau sangat mendekati kunci jawaban yang telah dimasukkan oleh guru. Akibatnya, variasi dalam jawaban siswa yang secara makna benar, namun menggunakan pilihan kata atau struktur kalimat yang berbeda dari kunci, sering kali tidak dikenali oleh sistem. Hal ini berujung

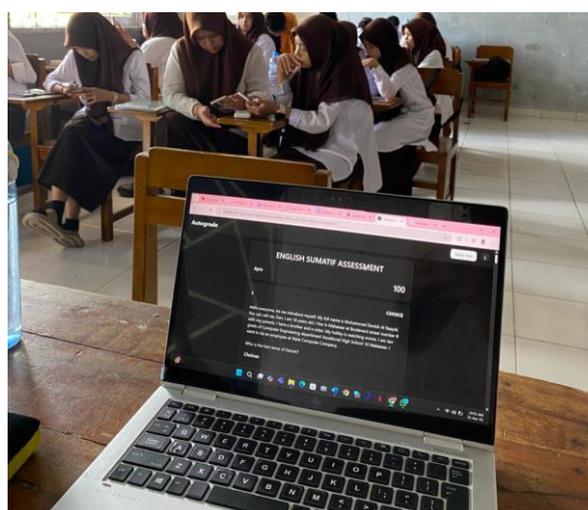
pada nilai yang rendah atau bahkan nol, meskipun secara substansi jawaban tersebut layak mendapatkan poin.

Data dari survei mendukung temuan ini. Hanya sekitar 34% siswa yang menyatakan bahwa nilai yang diberikan pada soal esai sesuai dengan jawaban yang mereka berikan. Sementara itu, sisanya merasa bahwa sistem belum sepenuhnya adil dalam menilai jawaban mereka, khususnya untuk soal esai yang bersifat lebih terbuka dan membutuhkan interpretasi. Ketidaksesuaian ini dapat menimbulkan rasa ketidakpuasan di kalangan siswa dan dapat menurunkan kepercayaan terhadap keadilan sistem evaluasi jika tidak segera diperbaiki.

Sebaliknya, pada soal pilihan ganda, *Autograder* menunjukkan performa yang sangat baik dan konsisten. Karena soal jenis ini bersifat objektif dan memiliki jawaban yang pasti, sistem dapat memberikan penilaian secara akurat tanpa banyak kendala. Dalam survei, tidak ditemukan keluhan terkait ketidaksesuaian skor pada soal pilihan ganda. Ini menunjukkan bahwa *Autograder* sangat cocok digunakan untuk soal-soal objektif yang membutuhkan ketepatan tunggal dalam jawaban. Namun, untuk soal subjektif seperti esai, keterlibatan guru dalam proses koreksi tetap diperlukan guna menjaga keadilan dan akurasi penilaian.



**Gambar 1.** Proses Pengerjaan Ujian Sumatif Menggunakan *Autograder*



**Gambar 2.** Pengawasan Ujian Sumatif Siswa Melalui *Autograder*



**Gambar 3.** Hasil Ujian Sumatif Siswa di *Autograder*

## **Pembahasan**

### **1. Efektivitas Penggunaan *Autograder* dalam Pendidikan Vokasi**

Dari hasil implementasi yang telah dijelaskan, dapat disimpulkan bahwa *Autograder* memberikan kontribusi yang signifikan terhadap peningkatan efektivitas proses evaluasi dalam pendidikan vokasi. SMK sebagai lembaga pendidikan yang menyiapkan siswa untuk dunia kerja membutuhkan sistem pembelajaran dan evaluasi yang cepat, tepat, dan efisien. *Autograder* sebagai sistem koreksi otomatis memberikan solusi terhadap permasalahan waktu dan tenaga dalam proses penilaian, terutama untuk mata pelajaran dengan jumlah siswa yang besar.

Selain dari segi efisiensi waktu, penggunaan *Autograder* juga mendapatkan tanggapan positif dari sisi pengalaman pengguna. Sebagian besar siswa menyatakan bahwa aplikasi *Autograder* mudah digunakan, dengan tampilan yang menarik dan navigasi yang intuitif. Sekitar 76% siswa setuju bahwa aplikasi mudah diakses dan digunakan, dan 82% siswa merasa bahwa tampilannya mudah dipahami. Ini merupakan indikator bahwa teknologi ini dapat diterapkan secara luas di lingkungan SMK tanpa perlu pelatihan intensif atau penyesuaian besar.

Lebih jauh lagi, penggunaan *Autograder* tidak hanya berdampak pada percepatan koreksi dan kemudahan penggunaan, tetapi juga berpotensi membawa perubahan yang lebih luas dalam ekosistem pembelajaran. Sebagaimana dinyatakan oleh Hagerer et al. (2021), penggunaan sistem *Autograder* dapat berkontribusi terhadap perubahan signifikan dalam data pembelajaran, seperti meningkatnya interaksi antara tutor dan siswa, peningkatan kualitas keseluruhan kursus, peningkatan keberhasilan belajar, peningkatan waktu yang dihabiskan siswa untuk belajar, serta menurunnya tingkat kesulitan yang dirasakan siswa dalam memahami materi. Hal ini menunjukkan bahwa teknologi koreksi otomatis seperti *Autograder* tidak hanya berdampak pada aspek teknis evaluasi, tetapi juga secara tidak langsung mendukung proses pembelajaran yang lebih menyeluruh dan berdampak positif terhadap hasil belajar siswa.

Dengan demikian, *Autograder* terbukti menjadi alat yang efektif dalam mendukung pendidikan vokasi karena mampu mengoptimalkan proses penilaian sekaligus meningkatkan kualitas pembelajaran secara keseluruhan.

## 2. Potensi Penggunaan *Autograder* di Mata Pelajaran Lain

Melihat hasil implementasi yang menunjukkan efisiensi dan efektivitas dalam pelaksanaan ujian sumatif, *Autograder* memiliki potensi besar untuk direplikasi ke berbagai mata pelajaran lain, terutama yang memiliki dominasi soal pilihan ganda atau jenis soal dengan jawaban tunggal. Mata pelajaran seperti Matematika, Bahasa Inggris, serta mata pelajaran produktif tertentu di SMK yang dapat dievaluasi dengan soal objektif merupakan kandidat yang sangat potensial untuk menerapkan *Autograder*.

Lebih lanjut, menurut Schneider et al. (2022), sistem *Autograder* memiliki potensi kuat untuk diterapkan dalam berbagai mata pelajaran, tidak terbatas pada bahasa Inggris saja, tetapi juga mencakup matematika dan bahasa lainnya. Dengan penyesuaian dan pelibatan aktif dari guru, sistem ini bahkan dapat mencapai tingkat akurasi yang mendekati penilaian manusia, serta mampu mendukung lingkungan pembelajaran yang multibahasa dan berbagai tipe soal. Ini berarti bahwa dengan pengembangan dan penerapan yang tepat, *Autograder* dapat menjadi solusi evaluasi yang adaptif dan fleksibel bagi berbagai konteks pembelajaran.

Namun demikian, penting untuk dicatat bahwa *Autograder* saat ini belum sepenuhnya optimal dalam menangani soal-soal subjektif, seperti esai atau uraian terbuka. Hal ini karena sistem masih cenderung hanya mengenali jawaban yang sesuai secara literal dengan kunci jawaban yang telah ditentukan. Untuk menjawab tantangan ini, pengembangan lebih lanjut perlu diarahkan pada peningkatan kemampuan semantik sistem, agar dapat memahami variasi jawaban yang benar meskipun berbeda secara struktur atau pilihan kata.

Dengan demikian, *Autograder* tidak hanya bermanfaat untuk satu jenis mata pelajaran atau format soal, tetapi memiliki potensi besar untuk dikembangkan dan diintegrasikan secara lebih luas dalam sistem pendidikan vokasi di Indonesia. Dengan dukungan kebijakan, pelatihan guru, dan peningkatan kualitas sistem, *Autograder* dapat menjadi bagian integral dari transformasi digital pendidikan yang lebih modern, efisien, dan berorientasi pada masa depan.

## Kesimpulan

Berdasarkan hasil implementasi, penggunaan *Autograder* memberikan dampak yang nyata terhadap efisiensi dan efektivitas evaluasi pembelajaran di SMK. Sistem ini mampu mempercepat proses koreksi, khususnya pada soal pilihan ganda, serta mengurangi beban administratif guru. Sebagian besar siswa juga merespons positif kemudahan akses dan tampilan sistem yang intuitif. Temuan ini didukung oleh Hagerer et al. (2021), yang menyatakan bahwa *Autograder* dapat meningkatkan interaksi antara guru dan siswa, meningkatkan kualitas pembelajaran, serta membantu siswa lebih fokus dalam proses belajar. Meskipun demikian, keterbatasan masih ditemukan pada penilaian soal esai, di mana sistem belum mampu menilai jawaban yang bermakna serupa namun berbeda secara struktur atau diksi.

Melihat potensi yang dimiliki, *Autograder* sangat mungkin diterapkan dalam berbagai mata pelajaran lain seperti Matematika, Bahasa Inggris, maupun mata pelajaran produktif di SMK yang menggunakan soal objektif. Schneider et al. (2022) mengungkapkan bahwa dengan pelibatan guru dan penyempurnaan sistem, *Autograder* dapat mencapai akurasi setara dengan penilaian manusia, bahkan dalam lingkungan multibahasa dan multi-format soal. Oleh karena itu, direkomendasikan agar pengembangan *Autograder* difokuskan pada peningkatan kemampuan semantik sistem dan implementasi pendekatan hybrid, yang menggabungkan koreksi otomatis dan manual. Selain itu, pelatihan guru serta dukungan kebijakan sangat diperlukan agar *Autograder* dapat menjadi bagian integral dalam digitalisasi pendidikan vokasi yang modern dan berkelanjutan.

## Ucapan Terima Kasih

Ucapan terima kasih yang tulus kami sampaikan kepada pihak SMKN 10 Makassar, khususnya guru dan siswa kelas X TKJ, atas dukungan dan partisipasinya dalam pelaksanaan kegiatan pengabdian masyarakat ini. Kami juga mengapresiasi Tim AjarMi Satuan Pendidikan Program Studi Pendidikan Teknik Informatika dan Komputer Universitas Negeri Makassar atas segala bantuan, fasilitas, serta bimbingan yang telah diberikan, sehingga kegiatan ini dapat berjalan dengan lancar dan memberikan dampak yang positif.

## Referensi

- Acuña, R., & Bansal, A. (2024). Improving Student Learning with Automated Assessment. *Proceedings of the 2024 on Innovation and Technology in Computer Science Education V. 1*. <https://doi.org/10.1145/3649217.3653603>.
- Hadrianingtyas, B., Aisyah, N., Zhulhijah, S., & Amalia, K. (2024). Evaluasi Program Pendidikan di SMK PGRI 7 Surabaya. *TSAQOFAH*. <https://doi.org/10.58578/tsaqofah.v4i3.2857>.
- Hagerer, G., Lahesoo, L., Anschütz, M., Krusche, S., & Groh, G. (2021). An Analysis of Programming Course Evaluations Before and After the Introduction of an *Autograder*. *2021 19th International Conference on Information Technology Based Higher Education and Training (ITHET)*, 01-09. <https://doi.org/10.1109/ITHET50392.2021.9759809>.
- Purnawanto, Ahmad Teguh. (2022). Perencanaan Pembelajaran Bermakna Dan Asesmen Kurikulum Merdeka. *Jurnal Ilmiah Pedagogy* 20(1):75–94.
- Ramesh, D., Sanampudi, S.K. An automated essay scoring systems: a systematic literature review. *Artif Intell Rev* 55, 2495–2527 (2022). <https://doi.org/10.1007/s10462-021-10068-2>.
- Schneider, J., Richner, R., & Riser, M. (2022). Towards Trustworthy AutoGrading of Short, Multi-lingual, Multi-type Answers. *International Journal of Artificial Intelligence in Education*, 33, 88-118. <https://doi.org/10.1007/s40593-022-00289-z>.
- Tisha, S., Oregon, R., Baumgartner, G., Alegre, F., & Moreno, J. (2022). An Automatic Grading System for a High School-level Computational Thinking Course. *2022 IEEE/ACM 4th International Workshop on Software Engineering Education for the Next Generation (SEENG)*, 20-27. <https://doi.org/10.1145/3528231.3528357>.