

RANCANG BANGUN SISTEM PENILAIAN ESAI OTOMATIS MENGUNAKAN METODE COSINE SIMILARITY DAN MODEL BERT BERBASIS WEB

Elang Mochamad Riefki Ardian¹

¹Institut Prima Bangsa

E-mail: *¹elangmra@gmail.com

Article Info

ABSTRACT

Article history:

Received 30 09, 2025

Revised 07 10, 2025

Accepted 14 10, 2025

Keywords:

Penilaian Esai Otomatis

Cosine Similarity

Aplikasi Berbasis Web

Natural Language

Processing

Pendidikan

Penilaian esai secara manual oleh pendidik seringkali memerlukan waktu yang cukup lama dan rentan terhadap subjektivitas. Untuk mengatasi permasalahan ini, penelitian ini bertujuan untuk merancang dan membangun sebuah sistem penilaian esai otomatis yang dapat membantu meringankan beban pendidik dalam menilai soal esai. Sistem ini menggunakan metode *Cosine Similarity* dan model BERT (*Bidirectional Encoder Representations from Transformers*) yang diimplementasikan dalam sebuah aplikasi berbasis web. Metode *Cosine Similarity* digunakan untuk mengukur kesamaan antara esai siswa dengan esai kunci jawaban, sedangkan model BERT digunakan untuk menangkap konteks dan makna dari teks esai secara lebih mendalam. Kombinasi kedua metode ini diharapkan dapat memberikan penilaian yang lebih akurat dan objektif. Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model pengembangan *software waterfall*. Sistem yang dibangun telah diuji dengan menggunakan data esai dari siswa, dan hasil pengujian menunjukkan bahwa sistem ini mampu memberikan penilaian yang akurat dengan tingkat korelasi yang tinggi terhadap penilaian manual oleh pendidik. Hasil penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam bidang pendidikan, khususnya dalam proses penilaian esai, serta dapat diterapkan secara luas di berbagai institusi pendidikan.

This is an open access article under the [CC BY-SA](#) license



1. PENDAHULUAN

Dalam dunia pendidikan, evaluasi pembelajaran merupakan aspek penting yang digunakan oleh guru untuk menilai sejauh mana pemahaman dan kemampuan siswa terhadap materi yang telah diajarkan (Wulan & Aristia, 2018). Tes dibagi menjadi tiga kategori berdasarkan bentuk jawaban yaitu tes lisan, tes praktikum atau perbuatan, dan tes tertulis (Rahman et al., 2019). Salah satu bentuk evaluasi yang sering digunakan adalah tes tertulis, yang dapat dibagi menjadi dua jenis utama: tes objektif atau pilihan ganda dan tes esai atau uraian. Tes objektif memiliki jawaban yang telah ditentukan,

sehingga proses penilaiannya dapat dilakukan secara otomatis atau semi-otomatis, dengan hasil yang lebih konsisten dan cepat. Sebaliknya, tes esai memerlukan penilaian manual yang lebih kompleks karena jawaban yang diberikan siswa bersifat bebas dan beragam, dapat berupa teori maupun pendapat.

Meskipun tes esai efektif dalam mengukur kemampuan berpikir dan pemahaman siswa, penilaiannya memerlukan waktu dan tenaga yang cukup besar dari pendidik. Di SMK Wahidin Cirebon, guru - guru masih menggunakan metode penilaian secara manual untuk tes esai, yang sering kali menjadi beban karena jumlah siswa yang banyak dan jumlah soal esai yang cukup banyak. Hal ini menyebabkan proses penilaian menjadi lambat, dan kerap terjadi inkonsistensi serta kurangnya objektivitas dalam penilaian. Keadaan ini menimbulkan tantangan bagi pendidik dalam memberikan feedback yang cepat dan adil kepada siswa.

Dengan perkembangan teknologi informasi, khususnya dalam bidang pemrosesan bahasa alami (Natural Language Processing/NLP) dan kecerdasan buatan (Artificial Intelligence/AI). Teknologi tersebut memiliki potensi untuk digunakan sebagai alat otomatisasi penilaian esai. Teknologi ini memungkinkan analisis teks secara mendalam, yang dapat digunakan untuk menilai esai secara objektif dan konsisten. Salah satu pendekatan yang digunakan adalah Cosine Similarity, sebuah metode matematika untuk mengukur kesamaan antara dua teks dengan menghitung sudut kosinus dari vektor yang mewakili teks tersebut (Arbiantono & Ekohariadi, 2021). Secara umum, NLP terdiri dari tiga komponen utama: knowledge base, inference engine, dan user interface. Knowledge base mencakup fakta heuristik dan domain. Seluruh informasi yang diperoleh akan disimpan dalam basis pengetahuan dalam bentuk pola pencocokan (Suryanarayana et al., 2016). Namun, metode ini memiliki keterbatasan dalam memahami konteks semantik yang lebih kompleks.

Untuk mengatasi keterbatasan ini, penggunaan model BERT (Bidirectional Encoder Representations from Transformers) menjadi relevan. BERT adalah model pembelajaran mendalam yang dapat memahami konteks dari kedua arah dalam sebuah teks, sehingga mampu menangkap makna yang lebih kaya dan mendalam dari esai yang dianalisis (Devlin et al., 2019). Dengan menggabungkan metode Cosine Similarity dan model BERT, diharapkan sistem penilaian esai otomatis yang dibangun dapat memberikan penilaian yang lebih akurat, konsisten, dan efisien.

Penelitian ini bertujuan untuk merancang dan membangun sistem penilaian esai otomatis berbasis web menggunakan metode Cosine Similarity dan model BERT.

Sistem ini diharapkan dapat menjadi solusi untuk tantangan penilaian esai yang dihadapi pendidik, khususnya di SMK Wahidin Cirebon. Dengan adanya sistem ini, diharapkan dapat meringankan beban pendidik, meningkatkan konsistensi penilaian, serta mempercepat proses penilaian esai.

2. METODELOGI

Penelitian ini menggunakan metode Research and Development (R&D) dengan model pengembangan waterfall untuk mengembangkan sistem penilaian esai otomatis berbasis web. Metode Research and Development (R&D), adalah pendekatan penelitian yang bertujuan menciptakan produk baru yang dirancang khusus untuk memenuhi kebutuhan spesifik. Metode ini penting dalam menilai dan memastikan bahwa produk yang telah dikembangkan sesuai dan layak untuk digunakan. Inti dari metode ini adalah untuk tidak hanya menciptakan produk baru tetapi juga untuk menguji dan memvalidasi efektivitas produk yang telah dibuat (Sugiyono, 2019).

Model pengembangan waterfall dipilih karena cocok untuk pengembangan sistem yang memerlukan tahapan yang jelas dan berurutan, dimulai dari analisis kebutuhan, desain sistem, implementasi, pengujian, hingga evaluasi (Nielsen, 2019). Penelitian ini dilaksanakan di SMK Wahidin Cirebon, sebuah sekolah menengah kejuruan yang menjadi objek penelitian karena memiliki jumlah siswa yang banyak dan sistem evaluasi pembelajaran berbasis e-learning yang sudah ada. Populasi dalam penelitian ini adalah guru-guru yang terlibat dalam penilaian esai. Sampel penelitian dipilih menggunakan teknik purposive sampling, dengan kriteria guru yang telah memiliki pengalaman dalam menilai esai di sekolah tersebut.

Tahap awal dalam penelitian ini adalah analisis kebutuhan. Pada fase ini, peneliti melakukan wawancara mendalam dengan para guru di SMK Wahidin Cirebon untuk mengidentifikasi tantangan dan kebutuhan mereka terkait penilaian esai. Selain wawancara, peneliti juga melakukan observasi terkait sistem pembelajaran dan evaluasi yang digunakan pada SMK Wahidin Cirebon. Hasil dari analisis kebutuhan ini memberikan dasar bagi desain sistem yang akan dikembangkan. Desain sistem melibatkan perancangan arsitektur sistem, alur kerja, dan user interface. Penelitian ini mencakup pembuatan flowchart, entity relationship diagram, serta desain antarmuka yang akan memudahkan pengguna dalam berinteraksi dengan sistem.

Pengembangan sistem dilakukan dengan mengimplementasikan desain yang telah dirancang. Sistem ini dibangun menggunakan perangkat lunak berbasis web dengan bahasa pemrograman PHP dengan framework Laravel dan Python dengan framework Flask. Laravel memiliki kemampuan routing dan dapat digunakan untuk

mengembangkan halaman web yang komprehensif dengan efisiensi yang lebih tinggi (Khanna, 2020). Sedangkan Flask lebih mengedepankan kesederhanaan dan fleksibilitas ditambah dengan library Python yang banyak makin membuat Flask makin baik (Relan, 2019). Model BERT dan metode Cosine Similarity diintegrasikan untuk melakukan analisis teks dan penilaian esai.

Setelah sistem dikembangkan, tahap berikutnya adalah pengujian. Pengujian dilakukan dalam dua skala, yaitu skala kecil dan skala besar, dengan mengikuti standar ISO untuk memastikan kualitas dan efektivitas sistem.

Pada skala kecil, pengujian dilakukan terhadap ahli media dan materi. Pengujian ini bertujuan untuk mengevaluasi fungsionalitas sistem dari perspektif ahli yang memiliki pengetahuan mengenai media dan materi yang digunakan pada sistem. Ahli media dan materi diminta untuk menilai aspek-aspek tertentu dari sistem, seperti user interface, kemudahan penggunaan, dan kecocokan sistem dengan kebutuhan sekolah. Penilaian ini dilakukan dengan menggunakan standar ISO untuk memastikan bahwa sistem memenuhi kriteria yang ditetapkan.

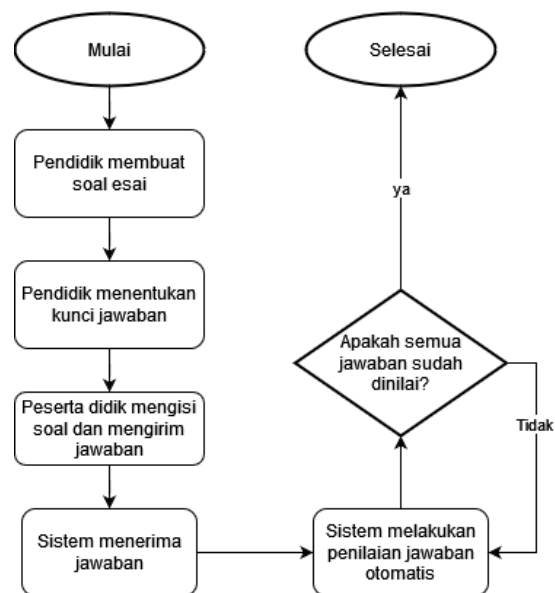
Setelah pengujian skala kecil, pengujian dilanjutkan pada skala besar, yaitu kepada pendidik di SMK Wahidin Cirebon. Pengujian ini dilakukan untuk mengevaluasi kinerja sistem secara langsung dan melihat bagaimana sistem berfungsi saat digunakan. Pendidik yang terlibat dalam pengujian ini akan menggunakan sistem untuk melakukan penilaian esai, dan hasil penilaian otomatis akan dibandingkan dengan penilaian manual yang dilakukan oleh pendidik untuk menilai akurasi dan konsistensi sistem.

Hasil dari pengujian akan dievaluasi dan feedback dari pengguna dikumpulkan melalui angket. Proses evaluasi ini memungkinkan peneliti untuk mengetahui bagian yang perlu diperbaiki dan membuat revisi pada sistem. Sistem yang telah diperbaiki kemudian diserahkan kembali kepada para ahli untuk dicek kembali sehingga nantinya dapat diimplementasikan di SMK Wahidin Cirebon.

3. HASIL DAN PEMBAHASAN

A. *Flowchart*

Flowchart merupakan gambaran simbolis dari algoritma atau metode yang diterapkan untuk menyelesaikan masalah. Dengan *flowchart*, pengguna dapat dengan mudah memeriksa bagian yang mungkin terlewat dalam analisis, serta menjadi alat komunikasi efektif antara pemrogram yang bekerja dalam sebuah tim proyek (Nurmalina et al., 2017).



Gambar 1 Flowchart

B. Hasil Pengujian Aspek Functionality

Pengujian *functionality* dilakukan untuk memastikan bahwa semua fitur utama dari sistem penilaian esai otomatis bekerja sesuai dengan spesifikasi yang telah ditentukan. Proses ini melibatkan ahli media yang mengevaluasi berbagai prosedur operasional, termasuk navigasi, proses *login* dan *logout*, akses data, serta fitur untuk membuat dan mengelola soal.

Tabel 1 Hasil Uji Aspek Functionality

No.	Prosedur	Hasil	
		Sukses	Gagal
1.	Navigasi berfungsi dengan lancar tanpa adanya tautan yang rusak.	✓	
2.	Melakukan <i>login</i> dengan <i>username</i> dan <i>password</i> yang benar.	✓	
3.	Melakukan <i>logout</i> .	✓	
Peserta Didik			
4.	Melihat data mata pelajaran	✓	
5.	Melihat data soal	✓	
6.	Mengerjakan tes yang ada	✓	
7.	Melihat hasil penilaian	✓	
Pendidik			
8.	Melihat data peserta didik	✓	
9.	Melihat data kelas	✓	
10.	Melihat data soal	✓	
11.	Membuat soal	✓	

12.	Mengubah data soal	✓
13.	Melihat hasil pengerjaan peserta didik	✓
Administrator		
14.	Melihat data kelas	✓
15.	Mengubah data kelas	✓
16.	Menambah data kelas	✓
17.	Melihat data mata pelajaran	✓
18.	Mengubah data mata pelajaran	✓
19.	Menambah data mata pelajaran	✓
20.	Melihat data pendidik	✓
21.	Menambah data pendidik	✓
22.	Mengubah data pendidik	✓
23.	Melihat data peserta didik	✓
24.	Menambah data peserta didik	✓
25.	Mengubah data peserta didik	✓
TOTAL		25

Dari hasil pengujian yang tercantum dalam Tabel 1, seluruh 25 prosedur berhasil dijalankan tanpa adanya kegagalan, yang berarti sistem telah memenuhi seluruh persyaratan *functionality*. Setiap fitur yang diuji menunjukkan performa yang stabil dan sesuai dengan kebutuhan pengguna.

C. Hasil Pengujian Aspek Usability

Pengujian *usability* bertujuan untuk mengukur sejauh mana sistem mudah digunakan dan dipahami oleh pengguna. Pengujian ini dilakukan melalui angket yang melibatkan para ahli sebagai validator dan juga pendidik. Skala penilaian yang digunakan berkisar dari 1 hingga 5, dengan 5 menunjukkan tingkat kepuasan tertinggi. Hasil pengujian menunjukkan bahwa pengguna memberikan skor yang tinggi pada aspek kemudahan navigasi, kemudahan pemahaman fitur, dan kenyamanan dalam penggunaan sehari-hari. Hal ini menandakan bahwa sistem telah berhasil dirancang dengan mempertimbangkan kebutuhan dan ekspektasi pengguna, sehingga dapat diadopsi dengan mudah di lingkungan pendidikan. Berikut adalah hasil uji *usability*.

Tabel 2 Hasil Validasi Ahli Materi Aspek *Usability*

Aspek Penilaian	Pernyataan	Skala Penilaian				
		1	2	3	4	5
Kelengkapan Fitur	1.Sistem memiliki fitur penilaian esai otomatis					✓
	2.Fitur penyimpanan data hasil penilaian					✓
	3.Fitur laporan atau <i>detail</i> hasil penilaian					✓
Akurasi	4.Penilaian esai konsisten dengan penilaian manusia				✓	
	5.Sistem mengidentifikasi kesalahan umum dalam esai					✓
	6.Sistem mampu menilai tiap bagian esai dengan benar				✓	
<i>Understandability</i>	7. <i>Interface</i> sistem mudah dipahami oleh pengguna					✓
<i>Operability</i>	8.Navigasi sistem mudah digunakan					✓
<i>Learnability</i>	9.Alur penggunaan sistem penilaian esai otomatis jelas dan mudah dipahami					✓
Keandalan Sistem	10.Sistem bekerja tanpa kendala teknis					✓
Total Skor					48	

Berdasarkan hasil validasi ahli materi pada tabel 2, maka dapat dihitung rata – rata penilaiannya adalah sebagai berikut:

$$M = \frac{\sum x}{N} = \frac{48}{10} = 4,8$$

Berdasarkan hasil perhitungan di atas diperoleh nilai rata – rata penilaian validasi oleh ahli materi sebesar 4,8. Menurut tabel konversi skala 5, nilai ini berada dalam rentang $4,24 < M \leq 5,00$ yang termasuk dalam kategori sangat baik.

Tabel 3 Hasil Validasi Ahli Media Aspek *Usability*

Aspek Penilaian	Pernyataan	Skala Penilaian				
		1	2	3	4	5
<i>User Interface</i>	1. Antarmuka pengguna mudah dipahami dan intuitif					✓
	2. Navigasi antara fitur-fitur sistem mudah dilakukan					✓
	3. Tata letak elemen visual konsisten dan rapi				✓	
	4. Teks dalam antarmuka jelas dan terbaca					✓
Kemudahan Penggunaan	5. Fitur yang ditawarkan dapat diakses dengan mudah					✓
	6. Proses penilaian esai dapat dilakukan dengan langkah sederhana					✓
	7. Sistem merespons perintah pengguna dengan cepat					✓
	8. Pengguna dapat memahami seluruh fungsionalitas sistem					✓
Desain Visual	9. Desain keseluruhan sistem menarik secara <i>visual</i>					✓
	10. Navigasi sistem mudah digunakan					✓
	11. Desain sistem responsif untuk perangkat berbeda					✓
Kualitas Teknis	12. Sistem bekerja tanpa kesalahan atau <i>bug</i>					✓
	13. Waktu muat halaman dan fitur cepat				✓	
	14. Tidak ada kendala teknis yang signifikan saat penggunaan					✓
Total Skor					68	

Berdasarkan hasil validasi ahli materi pada tabel 3, maka dapat dihitung rata – rata penilaiannya adalah sebagai berikut:

$$M = \frac{\sum x}{N} = \frac{68}{14} = 4,86$$

Berdasarkan hasil perhitungan di atas diperoleh nilai rata – rata penilaian validasi oleh ahli media sebesar 4,86. Menurut tabel konversi skala 5, nilai ini berada dalam rentang $4,24 < M \leq 5,00$ yang termasuk dalam kategori sangat baik.

Tabel 4 Hasil Validasi Sistem oleh Pendidik

No.	Pernyataan	Skala Penilaian				
		1	2	3	4	5
1.	Sistem penilaian esai otomatis memberikan hasil penilaian yang sesuai dengan penilaian pendidik.					✓
2.	Tidak terdapat perbedaan yang signifikan antara nilai yang dihasilkan sistem dengan penilaian pendidik.					✓
3.	Sistem penilaian otomatis mudah digunakan untuk memberikan penilaian esai.					✓
4.	Proses penilaian esai otomatis lebih efisien dibandingkan penilaian manual.					✓
5.	Penggunaan sistem mudah dipahami.					✓
6.	Antarmuka sistem penilaian esai otomatis dirancang dengan baik dan mudah dipahami.					✓
7.	Waktu yang dibutuhkan oleh sistem untuk menilai esai cukup cepat.					✓
8.	Sistem ini mengurangi beban kerja dalam proses penilaian esai.					✓
9.	Sistem memberikan laporan hasil penilaian yang lengkap dan mudah dipahami.					✓
10	Sistem efektif dalam mengidentifikasi dan menangani berbagai jenis kesalahan atau kelemahan dalam esai					✓
Total						50

Berdasarkan hasil validasi pendidik terhadap sistem pada tabel 4, maka dapat dihitung rata – rata penilaiannya adalah sebagai berikut:

$$M = \frac{\sum x}{N} = \frac{50}{10} = 5$$

Berdasarkan hasil perhitungan di atas diperoleh nilai rata – rata penilaian validasi oleh ahli media sebesar 5. Menurut tabel konversi skala 5, nilai ini berada dalam rentang $4,24 < M \leq 5,00$ yang termasuk dalam kategori sangat baik.

Tabel 5 Perbandingan Penilaian Pendidik dengan Sistem

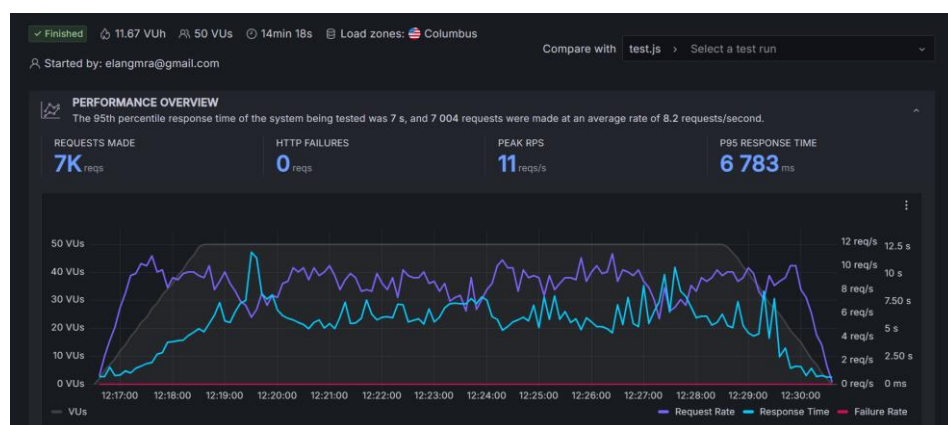
No.	Nama	Penilaian Pendidik	Penilaian Sistem
1	User1	100	95
2	User2	100	96
3	User3	95	94
4	User4	90	92
5	User5	90	89
6	User6	95	94
7	User7	100	98
8	User8	100	96
9	User9	95	91
10	User10	95	92
11	User11	100	96
12	User12	100	94
13	User13	90	92
14	User14	95	92
15	User15	90	91
16	User16	100	97
17	User17	95	93
18	User18	100	97
19	User19	100	97
20	User20	90	90
21	User21	95	92
22	User22	95	92
23	User23	100	92

24	User24	95	95
25	User25	100	98
26	User26	95	94
27	User27	95	89
28	User28	90	90
29	User29	95	95
30	User30	95	93
31	User31	90	88
32	User32	100	96

Tabel 5 menunjukkan perbandingan penilaian yang dilakukan oleh sistem dan pendidik. Berdasarkan tabel tersebut, perbedaan atau selisih nilai antara sistem dan pendidik hanya sedikit.

D. Hasil Pengujian Aspek Reliability

Pengujian aspek keandalan (*reliability*) dilaksanakan dengan alat dari k6 (www.k6.io). Alat ini mampu melakukan stress test terhadap situs web dan menilai aspek keandalan dari *website*. Dalam pengujian ini, *website* sistem penilaian esai otomatis akan diuji dengan skenario 50 pengguna yang mengakses situs selama 14 menit. Pengujian ini menggunakan metode uji tekanan atau *stress test*, dan hasilnya akan divisualisasikan dalam bentuk grafik pada gambar di bawah.



Gambar 2 Hasil Uji *Reliability* dengan K6

Gambar di atas menunjukkan hasil pengujian di mana tidak ada *request* gagal dari total 7.000 request yang diajukan. Pengujian ini dilakukan dengan total 50 *virtual users* (VUs), dengan waktu respons rata-rata 6,7 detik dan tingkat request tertinggi mencapai 11 request per detik.

Berdasarkan hasil tersebut, maka perhitungan *reliability* nya dengan model Nelson yaitu:

$$R1 = \left(1 - \frac{ne}{n}\right) \times 100\%$$

Keterangan:

R1 = nilai reliabilitas (*reliability*)

ne = total *request* gagal

n = total *request*

Maka perhitungannya sebagai berikut:

$$R1 = \left(1 - \frac{0}{7000}\right) \times 100\% = 100\%$$

Berdasarkan perhitungan tersebut, dapat disimpulkan bahwa sistem penilaian esai otomatis mencapai tingkat keandalan sebesar 100%, yang menunjukkan keandalan yang sangat tinggi.

E. Hasil Pengujian Aspek Efficiency

Pengujian *efficiency* mengukur seberapa baik sistem memanfaatkan sumber daya yang tersedia, termasuk waktu pemrosesan dan penggunaan memori. Pengujian ini dilakukan dengan memantau kinerja sistem dalam menjalankan berbagai tugas, seperti memproses penilaian esai dan mengelola data peserta didik, untuk memastikan bahwa sistem bekerja secara efisien tanpa menimbulkan beban berlebih pada perangkat keras.

Pengujian aspek efisiensi pada sistem penilaian esai otomatis dilakukan menggunakan tools *lighthouse*. Pengujian menggunakan *lighthouse* akan memberikan skor untuk situs web ketika diakses, dengan rentang skor dari 0 hingga 100. Tabel 6 di bawah ini menunjukkan hasil pengujian dengan *lighthouse*.

Tabel 6 Hasil Pengujian *Efficiency* dengan *Lighthouse*

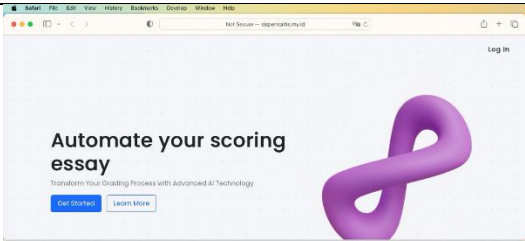
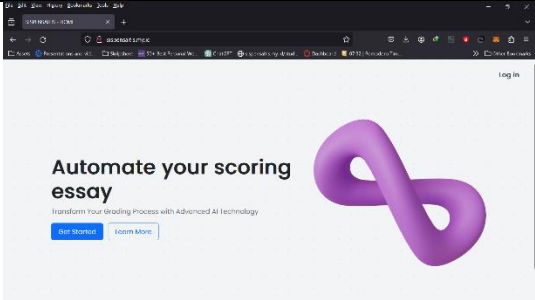
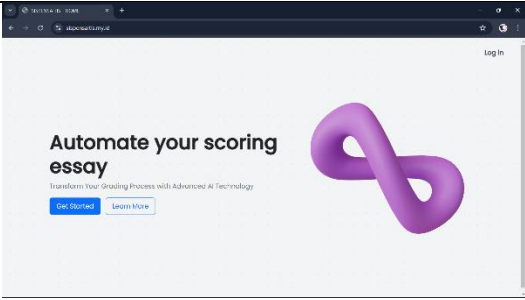
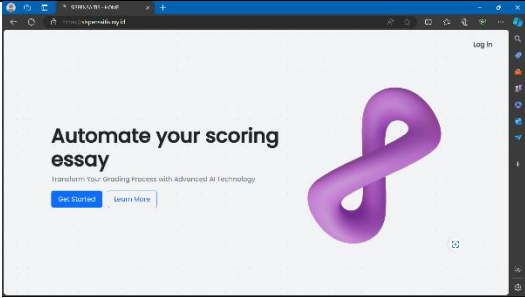
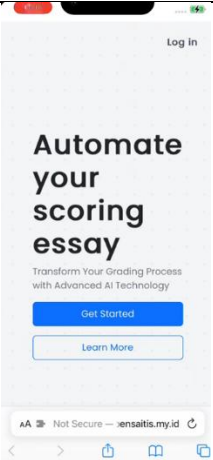
Halaman	Skor Mobile	Skor Desktop
Landing	98	99
Login	99	100
Lihat Data Mata Pelajaran	91	100
Lihat Data Soal	88	100

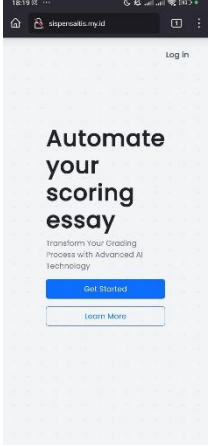
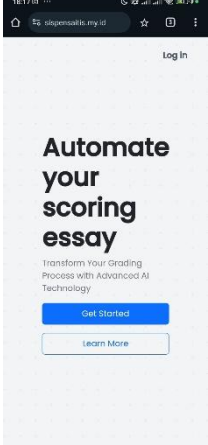
Pengerjaan Soal	97	100
Lihat Hasil Penilaian	88	100
Lihat Data Peserta Didik	96	100
Lihat Data Kelas	90	100
Lihat Data Soal Pendidik	96	100
Tambah Data Soal	97	100
Ubah Data Soal	86	100
Lihat Hasil Pengerjaan Soal	93	100
Lihat Data Kelas Admin	84	98
Tambah Data Kelas Admin	84	87
Ubah Data Kelas Admin	92	99
Lihat Data Mata Pelajaran	85	99
Tambah Data Mata Pelajaran	89	98
Ubah Data Mata Pelajaran	91	99
Lihat Data Pendidik	92	98
Tambah Data Pendidik	81	99
Ubah Data Pendidik	85	96
Lihat Data Peserta Didik	87	94
Tambah Data Peserta Didik	91	97
Ubah Data Peserta Didik	83	98
Rata – rata	89,54	99,71

F. Hasil Pengujian Aspek Portability

Pengujian *portability* bertujuan untuk memastikan bahwa sistem dapat dioperasikan pada berbagai platform dan perangkat dengan performa yang konsisten. Pengujian ini melibatkan uji coba sistem pada berbagai sistem operasi dan perangkat keras, termasuk komputer, tablet, dan *smartphone*, untuk menilai seberapa baik sistem beradaptasi dengan lingkungan yang berbeda. Uji *portabilitas* dilakukan dengan mengakses sistem penilaian esai otomatis melalui *browser* pada perangkat desktop dan mobile. Berikut ini adalah hasil pengujiannya.

Tabel 7 Hasil Pengujian *Portability*

No.	Browser	Tampilan	Keterangan
1.	Safari		Sukses
2.	Mozilla Firefox		Sukses
3.	Google Chrome		Sukses
4.	Microsoft Edge		Sukses
5.	Safari Mobile		Sukses

6. <i>Mi Browser</i>		Sukses
7. <i>Mozilla Firefox Mobile</i>		Sukses
8. <i>Google Chrome Mobile</i>		Sukses

Hasil pengujian menunjukkan bahwa sistem penilaian esai otomatis memiliki tingkat portability yang tinggi, dengan kemampuan untuk beroperasi secara lancar di berbagai platform tanpa kehilangan fungsionalitas atau kinerja. Sistem ini dapat digunakan oleh pendidik dan peserta didik di berbagai perangkat, yang sangat penting untuk fleksibilitas penggunaan di lingkungan pendidikan modern.

G. Pembahasan

Pembahasan hasil dari berbagai pengujian yang dilakukan menunjukkan bahwa sistem penilaian esai otomatis telah memenuhi standar kualitas yang tinggi di semua aspek yang diuji, termasuk *functionality*, *usability*, *reliability*, *efficiency*, dan *portability*. Pengujian *functionality* yang berhasil menunjukkan bahwa sistem ini stabil dan dapat dioperasikan sesuai spesifikasi. *Usability* yang tinggi mengindikasikan bahwa pengguna merasa nyaman dan mudah dalam menggunakan sistem, yang penting untuk adopsi dan penggunaan jangka panjang. Pengujian *reliability* mengkonfirmasi bahwa sistem dapat diandalkan dalam memberikan hasil yang konsisten, sementara *efficiency* menunjukkan bahwa sistem memanfaatkan sumber daya secara optimal. *Portability* yang tinggi memastikan bahwa sistem dapat digunakan di berbagai platform, memberikan fleksibilitas yang dibutuhkan dalam konteks pendidikan yang beragam. Dengan hasil pengujian ini, dapat disimpulkan bahwa sistem penilaian esai otomatis ini tidak hanya memenuhi ekspektasi pengguna, tetapi juga mematuhi standar ISO yang relevan, menjadikannya alat yang efektif dan terpercaya dalam mendukung proses pendidikan.

4. KESIMPULAN

Berikut adalah kesimpulan berdasarkan hasil penelitian dan analisis yang telah dilakukan:

1. Penelitian ini berhasil menciptakan sebuah sistem penilaian esai otomatis yang diintegrasikan dalam sebuah platform *website*. Sistem ini terbukti dapat mengurangi beban kerja pendidik serta menunjukkan efektivitas dan efisiensi yang signifikan.
2. Melalui serangkaian pengujian kualitas perangkat lunak berstandar ISO, sistem ini memperoleh skor dalam aspek *functionality* sebesar 100%. Dalam aspek *reliability*, sistem juga mencapai skor 100%, dengan pengujian dilakukan menggunakan *tools* dari k6. Adapun pengujian *efficiency* yang dilakukan dengan *Lighthouse* memberikan skor kinerja 89,54 untuk perangkat *mobile* dan 99,71 untuk *desktop*. Pada pengujian *portability*, sistem telah diuji coba di berbagai web *browser* dan berhasil diakses tanpa kendala, yang menandakan bahwa sistem memenuhi kriteria *portability* yang diharapkan.

Sistem juga telah diuji oleh pendidik kepada peserta didik dan sistem menghasilkan penilaian yang mendekati dengan penilaian pendidik atau penilaian manual.

DAFTAR PUSTAKA

- Arbiantono, M. W., & Ekohariadi. (2021). *Pengembangan Aplikasi Assessment Tool Menggunakan Metode Cosine Semantic Similarity untuk Automatic Scoring Jawaban Tes Uraian pada Mata Pelajaran Basis Data di SMKN 1 Surabaya*.
- Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*.
<https://github.com/tensorflow/tensor2tensor>
- Khanna, L. (2020). Laravel – A Trending PHP Framework. *International Journal of Trend in Scientific Research and Development*. <https://api.semanticscholar.org/CorpusID:220054966>
- Nielsen, L. (2019). Personas Go Agile. In L. Nielsen (Ed.), *Personas - User Focused Design* (pp. 117–121). Springer London. https://doi.org/10.1007/978-1-4471-7427-1_6
- Nurmalina, R., Yani Km, J. A., Laut, T., & Selatan, K. (2017). Perencanaan dan Pengembangan Aplikasi Absensi Mahasiswa Menggunakan Smart Card Guna Pengembangan Kampus Cerdas (Studi Kasus Politeknik Negeri Tanah Laut). In *Jurnal Integrasi* (Vol. 9, Issue 1).
- Rahman, A. A., Pd Cut, M., Nasryah, E., & Pd, M. (2019). *EVALUASI PEMBELAJARAN*. www.penerbituwais.com
- Relan, K. (2019). Beginning with Flask. *Building REST APIs with Flask*.
<https://api.semanticscholar.org/CorpusID:203527520>
- Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif dan R & D*. Alfabeta.
- Suryanarayana, D., Kanakam, P., Hussain, S. M., & Gupta, S. (2016). Knowledge representation of natural language in high performance linguistics scheme. *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 1–5.
<https://doi.org/10.1109/ICCIC.2016.7919548>
- Wulan, A., & Aristia, R. (2018). *Jenis-Jenis Instrumen dalam Evaluasi Pembelajaran*.