



**RANCANGAN MODEL PENILAIAN UJIAN ESAI OTOMATIS
MENGUNAKAN METODE *COSINE SIMILARITY* DAN
*SYNONYM RECOGNITION***

SKRIPSI

PUTRI APRICANIA

1910511012

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN

JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2023



**RANCANGAN MODEL PENILAIAN UJIAN ESAI OTOMATIS
MENGUNAKAN METODE *COSINE SIMILARITY* DAN
*SYNONYM RECOGNITION***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Komputer**

PUTRI APRICANIA

1910511012

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA**

2023

PERNYATAAN ORISINALITAS

Tugas skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Putri Apricania

NIM : 1910511012

Tanggal : 16 Juli 2023

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 16 Juli 2023

Yang menyatakan,



Putri
Putri Apricania

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK
KEPENTINGAN AKADEMIS**

Saya civitas akademis Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Putri Apricania

NIM : 1910511012

Fakultas : Ilmu Komputer

Program Studi : S1 Informatika

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan pelayanan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**Rancangan Model Penilaian Ujian Esai Otomatis Menggunakan Metode
*Cosine Similarity Dan Synonym Recognition.***

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 16 Juli 2023

Yang menyatakan,



Putri Apricania

LEMBAR PENGESAHAN

Yang bertanda tangan di bawah ini:

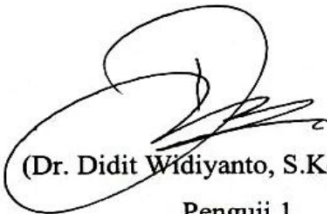
Nama : Putri Apricania

NIM : 1910511012


Program Studi : S-1 Informatika

Judul Skripsi/TA : Rancangan Model Penilaian Ujian Esai Otomatis Menggunakan Metode *Cosine Similarity* Dan *Synonym Recognition*.

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



(Dr. Didit Widiyanto, S.Kom, M.Si.)
Penguji 1



(Bayu Hananto, S.Kom., M.Kom.)
Penguji 2



(Jayanta, S.Kom., M.Si.)
Pembimbing



(Dr. Ermatita, M.Kom.)
Dekan



(Dr. Widya Cholil, M.I.T.)
Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 6 Juli 2023



RANCANGAN MODEL PENILAIAN UJIAN ESAI OTOMATIS MENGUNAKAN METODE *COSINE SIMILARITY* DAN *SYNONYM RECOGNITION*

Putri Apricania

ABSTRAK

Penilaian esai oleh guru membutuhkan waktu yang lama karena harus membaca jawaban siswa dan mencocokkannya dengan kunci jawaban. Untuk mengatasi hal ini, sistem penilaian otomatis dengan menggunakan metode cosine similarity dan synonym recognition dikembangkan. Metode cosine similarity mengukur kesamaan kosinus antara jawaban siswa dan jawaban yang benar dengan menggunakan vektor, sedangkan synonym recognition mengidentifikasi kata-kata dengan makna yang sama dengan kata kunci dari sinonim. Metode ini efektif dalam memberikan penilaian yang akurat dan konsisten serta meningkatkan akurasi dari penilaian. Dengan menggunakan term frequency yaitu menghitung bobot kata atau jumlah kata yang muncul dalam sebuah dokumen. Berdasarkan hasil penelitian yang didapatkan bahwa sistem penilaian ujian esai otomatis menggunakan metode cosine similarity dan synonym recognition bekerja dengan baik. Namun, terdapat tingkat kesalahan dalam memprediksi nilai siswa dengan Mean Absolute Error (MAE) sebesar 8.4 untuk 20 data uji dan 11 untuk 10 data uji yang bisa dikatakan memiliki tingkat kesalahan sistem rendah. Penelitian ini memberikan kontribusi dalam pengembangan sistem penilaian ujian esai otomatis yang dapat mempercepat proses penilaian dan memberikan penilaian yang konsisten kepada siswa.

Kata Kunci: Ujian Esai, *Synonym Recognition*, *Cosine Similarity*, *Term Frequency*

**DESIGN OF AUTOMATIC ESSAY EXAM ASSESSMENT
MODELS USING COSINE SIMILARITY AND SYNONYM
RECOGNITION METHODS**

Putri Apricania

ABSTRACT

Essay assessment by the teacher takes a long time because they have to read student answers and match them with the answer key. To overcome this, an automatic scoring system using cosine similarity and synonym recognition methods was developed. The cosine similarity method measures the cosine similarity between students' answers and the correct answers using vectors, while synonym recognition identifies words with the same meaning as synonym keywords. This method is effective in providing accurate and consistent assessments and increasing the accuracy of the assessment. By using the term frequency, namely calculating the weight of words or the number of words that appear in a document. Based on the research results, it was found that the automatic essay examination scoring system using the cosine similarity and synonym recognition methods worked well. However, there is an error rate in predicting student scores with a Mean Absolute Error (MAE) of 8.4 for 20 test data and 11 for 10 test data which can be said to have a low system error rate. This research contributes to the development of an automatic essay exam scoring system that can speed up the assessment process and provide consistent assessments to students.

Keywords: Essay exam, Synonym Recognition, Cosine Similarity, Term Frequency

KATA PENGANTAR

Dengan rasa syukur dan pujian kepada Allah SWT atas berlimpahnya rahmat, kesehatan, bimbingan, dukungan, dan semua anugerah-Nya, penulis berhasil menyelesaikan tugas akhir ini yang berjudul: “Rancangan Model Penilaian Ujian Esai Otomatis Menggunakan Metode *Cosine Similarity* Dan *Synonym Recognition*”.

Penelitian ini disusun peneliti sebagai syarat untuk memperoleh Gelar Sarjana di Fakultas Ilmu Komputer, Program Studi Informatika, Universitas Pembangunan Nasional "Veteran" Jakarta. Dalam penyusunan penelitian skripsi ini, peneliti juga mendapatkan dukungan yang sangat banyak tidak lupa bimbingan, bantuan, serta kerja sama dari berbagai pihak, baik dalam hal dukungan moral, materi, maupun spiritual, yang secara keseluruhan telah memungkinkan peneliti menyelesaikan skripsi ini dengan lancar dan baik. Dengan penuh rasa hormat, peneliti ingin menyampaikan ungkapan terima kasih yang tulus kepada semua pihak yang telah memberikan dukungan dan kontribusi berarti dalam penyusunan skripsi ini:

1. Ibu, Ayah, Kakak, Adik dan seluruh keluarga peneliti yang selalu mendukung peneliti melalui doa, serta semangat kepada peneliti sehingga dapat menyelesaikan penelitian dan skripsi.
2. Bapak Jayanta, S. Kom., M. Si sebagai dosen pembimbing peneliti yang senantiasa membimbing serta memberikan masukan, doa, dan semangat kepada peneliti dalam menyelesaikan penelitian ini.
3. Ibu Dr. Widya Cholil, M.I.T. sebagai Ketua Program Studi Informatika.
4. Ibu Dr. Ermatita, M. Kom. sebagai Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
5. Bapak/Ibu Dosen Fakultas Ilmu Komputer yang berdedikasi mengajar dan membagikan ilmu sehingga peneliti dapat memperluas wawasan serta ilmu pengetahuan selama menjadi mahasiswa di Fakultas Ilmu Komputer.

6. Bapak Zaky Syahri Ramadhan sebagai Guru Sejarah Indonesia di SMA Negeri 44 Jakarta yang telah memberikan bantuan dan dukungan kepada peneliti sehingga dapat menjadi penunjang penelitian ini.
7. Nur Frita S.E atas bantuan bimbingan dalam pengerjaan dokumen skripsi ini serta penyemangat dan pemberi motivasi dalam penyusunan skripsi peneliti.
8. Teman-teman peneliti dan seluruh pihak lain yang tidak dapat disebutkan satu persatu atas kontribusi dalam penelitian serta penyusunan skripsi baik secara langsung maupun tidak langsung.

Peneliti menyadari bahwa masih banyak kekurangan dan kesalahan pada penelitian serta skripsi ini karena dengan keterbatasan pengetahuan dan kemampuan yang dimiliki oleh peneliti. Oleh karena itu, peneliti sangat mengharapkan skripsi ini dapat memberikan informasi dan pengetahuan yang berharga untuk pembaca.

Jakarta, 29 Mei 2023

Peneliti,

DAFTAR ISI

	Halaman
RANCANGAN MODEL PENILAIAN UJIAN ESAI OTOMATIS MENGUNAKAN METODE <i>COSINE SIMILARITY</i> DAN <i>SYNONYM RECOGNITION</i>	i
.....	ii
PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xiv
BAB 1	1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan Penelitian.....	4
1.4. Manfaat Penelitian.....	4
1.5. Ruang Lingkup.....	5
1.6. Luaran.....	5

1.7. Sistematika Penulisan.....	5
BAB 2	7
TINJAUAN PUSTAKA	7
2.1. Ujian Esai	7
2.2. Text Mining.....	8
2.2.1. <i>Case Folding</i>	8
2.2.2. <i>Tokenizing</i>	9
2.2.3. <i>Filtering</i>	9
2.2.4. <i>Stemming</i>	9
2.2.5. <i>Analyzing</i>	11
2.3. <i>Synonym Recognition</i>	12
2.4. <i>Term Frequency</i>	13
2.5. <i>Cosine Similarity</i>	13
2.6. <i>Mean Absolute Error (MAE)</i>	14
2.7. Review Jurnal Terdahulu.....	15
BAB 3	18
METODOLOGI PENELITIAN.....	18
3.1. Tahapan Penelitian	18
3.2. Identifikasi Masalah	19
3.3. Studi Pustaka	19
3.4. Uraian Penelitian	19
3.4.1. Pengumpulan Data	20
3.4.2. <i>Preprocessing Data</i>	21
3.4.3. <i>Synonym Recognition</i>	26

3.4.4.	<i>Term Frequency – Inverse Document Frequency</i>	28
3.4.5.	<i>Cosine Similarity</i>	30
3.4.6.	Perancangan Sistem	33
3.4.7.	Pengujian Sistem.....	33
3.4.8.	Implementasi Sistem	34
3.5.	Perangkat Penelitian	34
3.5.1.	Perangkat Keras (<i>Hardware</i>)	34
3.5.2.	Perangkat Lunak (<i>Software</i>).....	35
3.6.	Jadwal Penelitian	35
BAB 4	37
HASIL DAN PEMBAHASAN	37
4.1.	Pengumpulan Data	37
4.2.	Skenario Pengujian.....	39
4.2.1.	<i>Preprocessing</i>	39
4.2.2.	<i>Synonym Recognition</i>	42
4.2.3.	<i>Term Frequency – Inverse Document Frequency</i>	44
4.2.4.	<i>Cosine Similarity</i>	46
4.3.	Pengujian Sistem	46
4.4.	<i>Unified Modelling Language (UML)</i>	49
4.4.1.	<i>Activity Diagram</i>	49
4.4.2.	<i>Use Case Diagram</i>	51
4.4.3.	Rancangan Basis Data.....	52
4.5.	Implementasi Antar Muka.....	52
4.5.1.	Halaman Registrasi	53

4.5.2.	Halaman Login.....	54
4.5.3.	Halaman Dashboard.....	54
4.5.4.	Halaman Penilaian Jawaban.....	55
4.5.5.	Formulir Input Kunci Jawaban dan Jawaban Siswa	56
4.5.6.	Halaman Tabel Proses Jawaban.....	57
4.5.7.	Halaman Input Mata Pelajaran.....	58
4.5.8.	Halaman Profil Admin	59
4.5.9.	Halaman Daftar Nilai Siswa	60
4.5.10.	Halaman Data Sinonim	61
4.6.	Pengujian Fungsional	62
BAB 5	65
KESIMPULAN DAN SARAN.....	65
5.1.	Kesimpulan.....	65
5.2.	Saran.....	65
DAFTAR PUSTAKA	67
RIWAYAT HIDUP.....	69
LAMPIRAN.....	70

DAFTAR GAMBAR

Gambar 1. 1 Tingkat Similaritas Jawaban Esai Siswa SMAN 44 Jakarta	2
Gambar 3. 1 Tahapan Penelitian	18
Gambar 3. 2 Rancangan Sistem Penilaian Ujian Esai.....	20
Gambar 3. 3 Tahapan <i>Preprocessing</i>	21
Gambar 3. 4 Tahapan Proses <i>Case Folding</i>	22
Gambar 3. 5 Tahapan Proses <i>Tokenizing</i>	23
Gambar 3. 6 Proses Tahapan <i>Filtering</i>	24
Gambar 3. 7 Proses Tahapan <i>Stemming</i>	25
Gambar 3. 8 Tahapan Proses <i>Synonym Recognition</i>	27
Gambar 3. 9 Diagram <i>Cosine Similarity</i>	30
Gambar 3. 10 Tahapan Proses <i>Cosine Similarity</i>	32
Gambar 4. 1 <i>Activity Diagram</i> Sistem	50
Gambar 4. 2 <i>Use Case Diagram</i> Sistem	51
Gambar 4. 3 Rancangan Basis Data.....	52
Gambar 4. 4 Halaman Registrasi	53
Gambar 4. 5 Halaman Login.....	54
Gambar 4. 6 Halaman Dashboard	55
Gambar 4. 7 Halaman Penilaian Jawaban.....	56
Gambar 4. 8 Formulir Input Kunci Jawaban dan Jawaban Siswa.....	57
Gambar 4. 9 Halaman Tabel Proses Jawaban	58
Gambar 4. 10 Halaman Input Mata Pelajaran.....	59
Gambar 4. 11 Halaman Profil Admin	60
Gambar 4. 12 Halaman Daftar Nilai Siswa.....	61
Gambar 4. 13 Halaman Data Sinonim	62

DAFTAR TABEL

Tabel 3. 1 Data Kunci Jawaban dan Jawaban Siswa	21
Tabel 3. 2 Hasil Proses <i>Case Folding</i>	22
Tabel 3. 3 Hasil Proses <i>Tokenizing</i>	23
Tabel 3. 4 Proses Hasil <i>Filtering</i>	25
Tabel 3. 5 Proses Hasil <i>Stemming</i>	26
Tabel 3. 6 Tabel Proses <i>Synonym Recognition</i>	27
Tabel 3. 7 Pembobotan Kata	28
Tabel 3. 8 <i>Inverse Document Frequency</i>	29
Tabel 3. 9 Jadwal Penelitian.....	35
Tabel 4. 1 Data Kunci Jawaban dan Jawaban Siswa	38
Tabel 4. 2 Kunci Jawaban dan Jawaban Siswa	39
Tabel 4. 3 Hasil Proses <i>Case Folding</i>	40
Tabel 4. 4 Hasil Proses <i>Tokenizing</i>	41
Tabel 4. 5 Hasil Proses <i>Filtering</i>	41
Tabel 4. 6 Hasil Proses <i>Stemming</i>	42
Tabel 4. 7 Hasil Proses <i>Synonym Recognition</i>	43
Tabel 4. 8 Hasil Proses Pembobotan Kata	44
Tabel 4. 9 Proses Hasil <i>Inverse Document Frequency</i>	45
Tabel 4. 10 Data Nilai Guru dan Nilai Sistem	47
Tabel 4. 11 Pengujian Fungsional	62