

**PENERAPAN ALGORITMA COSINE SIMILARITY
UNTUK OTOMATISASI PENILAIAN
JAWABAN ESSAY**

SKRIPSI



disusun oleh

Frendy A. Nancy

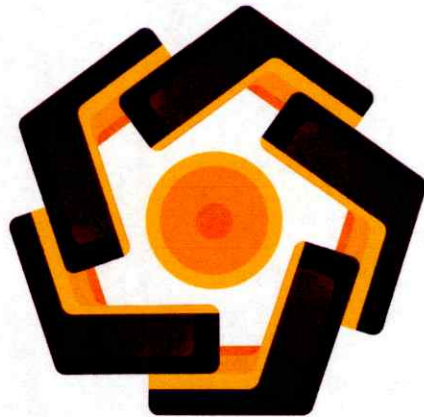
16.11.0870

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**

**PENERAPAN ALGORITMA COSINE SIMILARITY
UNTUK OTOMATISASI PENILAIAN
JAWABAN ESSAY**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Frendy A. Nancy

16.11.0870

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**



PERSETUJUAN

SKRIPSI

PENERAPAN ALGORITMA COSINE SIMILARITY UNTUK OTOMATISASI PENILAIAN JAWABAN ESSAY

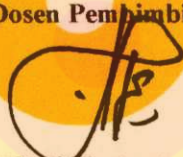
yang dipersiapkan dan disusun oleh

Frendy A. Nancy

16.11.0870

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Mei 2019

Dosen Pembimbing,



Anggit Dwi Hartanto, M.Kom

NIK. 190302163

PENGESAHAN

SKRIPSI

PENERAPAN ALGORITMA COSINE SIMILARITY UNTUK OTOMATISASI PENILAIAN JAWABAN ESSAY

yang dipersiapkan dan disusun oleh

Frendy A. Nancy

16.11.0870

telah dipertahankan di depan Dewan Penguji
pada tanggal 19 November 2019

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ferry Wahyu Wibowo, S.Si, M.Cs
NIK. 190302235



Yuli Astuti, M.Kom
NIK. 190302146



Anggit Dwi Hartanto, M.Kom
NIK. 190302163

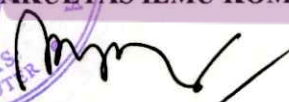


Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 28 November 2019

DEKAN FAKULTAS ILMU KOMPUTER



Krisnawati, S.Si, M.T.
NIK. 190302038



PERNYATAAN

Saya yang bertandatangan di bawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 28 November 2019



Frendy A. Nancy

NIM. 16.11.0870

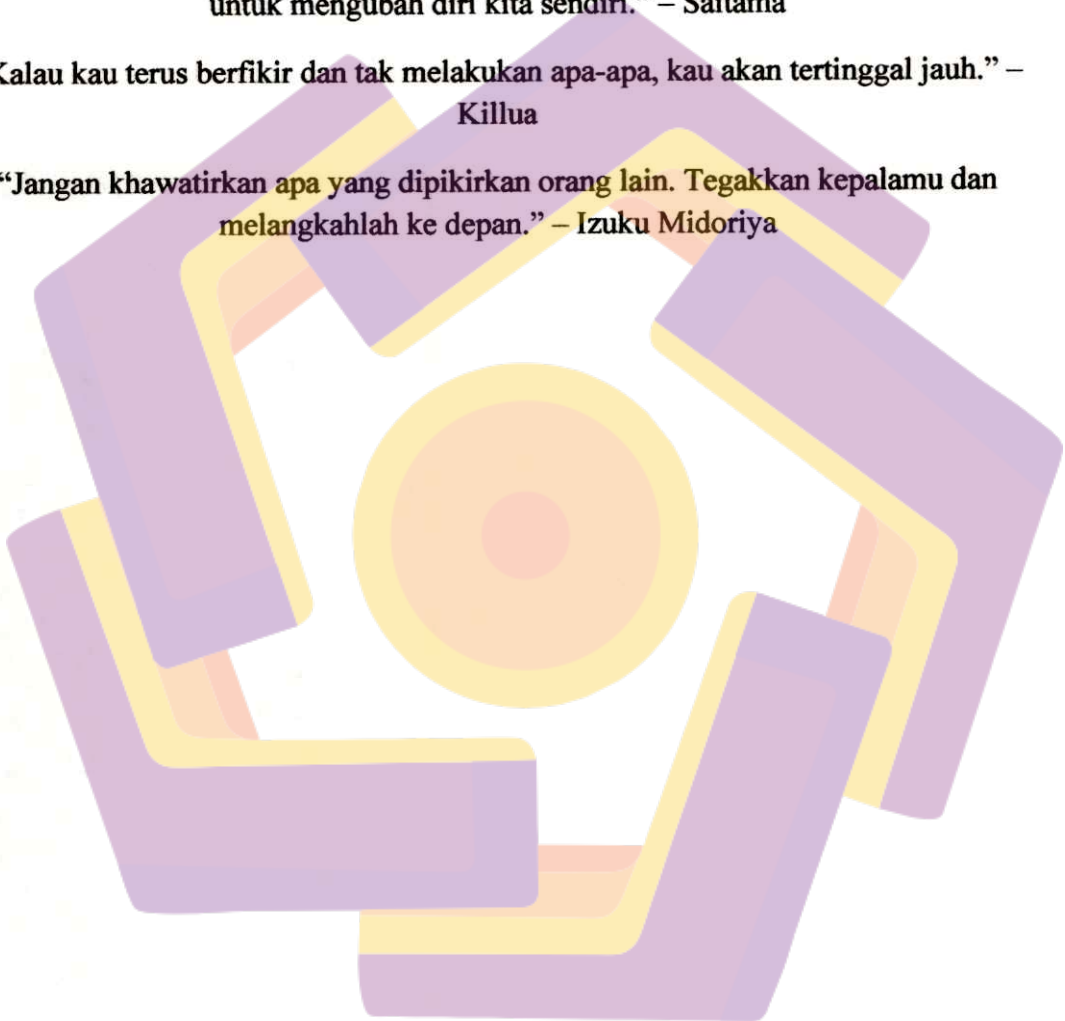
MOTTO

“Setiap orang itu jenius, tapi jika Anda mengecap sebuah ikan akan kemampuannya dalam memanjat pohon, maka seumur hidupnya ia akan mempercayai bahwa dirinya itu bodoh.” – Albert Einstein

“Kekuatan sejati dari umat manusia adalah bahwa kita memiliki kuasa penuh untuk mengubah diri kita sendiri.” – Saitama

“Kalau kau terus berfikir dan tak melakukan apa-apa, kau akan tertinggal jauh.” – Killua

“Jangan khawatirkan apa yang dipikirkan orang lain. Tegakkan kepalamu dan melangkahlah ke depan.” – Izuku Midoriya



PERSEMBAHAN

Saya mempersembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung dalam proses pembuatan skripsi.

1. Tuhan Yang Maha Esa yang memberikan segala nikmat dan kasih sayangnya sampai sejauh ini.
2. Kedua orangtua saya, yang selalu mendoakan, menyemangati dan menjadi pendorong saya untuk segera menyelesaikan skripsi.
3. Bapak Anggit Dwi Hartanto, M.Kom. yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
4. Dosen-dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama perkuliahan.
5. Teman-teman kelas IF14 2016 yang selalu menemani perkuliahan, mendukung dan memberikan semangat sampai saat ini. Semoga selalu bahagia dan menjadi pribadi yang lebih baik lagi.

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat, hidayah dan kekuatan sehingga saya dapat menyelesaikan skripsi yang berjudul Penerapan Algoritma Cosin Similarity untuk Otomatisasi Penilaian Jawaban Essay.

Skripsi ini saya buat guna menyelesaikan studi jenjang Strata Satu (S1) pada program studi Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata satu dan untuk memperoleh gelar Sarjana Komputer. Dengan selesainya skripsi ini maka pada kesempatan ini saya mengucapkan terima kasih pada:

1. Bapak Prof. Dr. M. Suyanto, MM. Selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Bapak Anggit Dwi Hartanto, M.Kom. selaku dosen pembimbing yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini.
4. Dosen Penguji (Yuli Astuti, M.Kom., Ferry Wahyu Wibowo, S.Si, M.Cs) dan segenap Dosen dan Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.
5. Tenaga pengajar yang bersedia menjadi penyedia sumber data penelitian ini.
6. Kedua orang tua saya yang telah mendoakan, mendukung dan memberikan semangat.
7. Saudara-saudara yang senantiasa mendoakan, memberi semangat dan bantuan.
8. Teman-teman IF16 angkatan 2016 yang telah menemani selama proses perkuliahan.

9. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.

Semoga Tuhan memberikan balasan yang lebih kepada semua pihak yang telah ikut membantu saya menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun diterima dengan senang hati dan rasa terima kasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

Yogyakarta, 28 November 2019

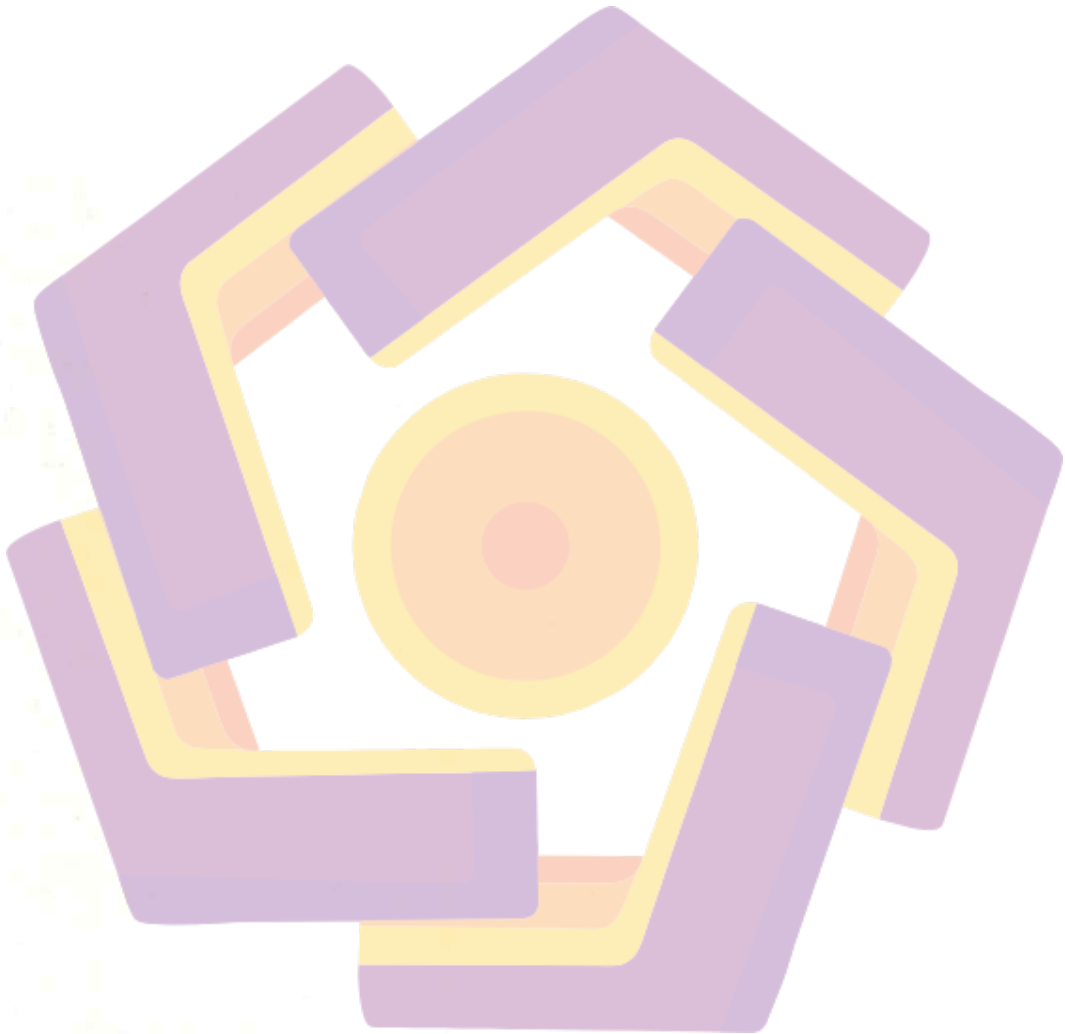
Frendy A. Nancy

DAFTAR ISI

JUDUL	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN.....	iv
MOTTO	v
PERSEMBAHAN.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xv
INTISARI.....	xix
ABSTRACT.....	xx
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Penelitian	3
1.4 Maksud dan Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Metode Penelitian.....	4
1.6.1 Metode Pengumpulan Data.....	5
1.6.2 Metode Analisis	6
1.6.3 Metode Perancangan.....	6
1.6.4 Metode Coding	7
1.6.5 Metode Testing	7
1.7 Sistematika Penulisan.....	8

BAB II LANDASAN TEORI	10
2.1 Tinjauan Pustaka	10
2.2 Dasar Teori	12
2.2.1 Esai	12
2.2.2 Otomatisasi	13
2.2.3 Otomatisasi Esai	13
2.2.4 Text Mining	13
2.2.5 Text Preprocessing.....	14
2.2.5.1 Text Transformation.....	16
2.2.6 Cosine Similarity	16
2.2.7 TF (Term Frequency)	17
2.2.8 DF(Document Frequency).....	18
2.2.9 Confusion Matrix.....	19
2.2.10 Unified Modelling Language (UML)	20
2.2.10.1 Use Case Diagram	20
2.2.10.2 Class Diagram	21
2.2.10.3 Sequence Diagram.....	23
2.2.10.4 Activity Diagram.....	25
2.2.11 Bahasa Pengkodingan	27
2.2.11.1 HTML.....	27
2.2.11.2 CSS.....	27
2.2.11.3 PHP.....	28
2.2.11.4 Javascript.....	28
BAB III ANALISIS DAN PERANCANGAN	29
3.1. Identifikasi Masalah	29
3.2. Analisis Masalah	29
3.3. Hasil Analisis	30
3.4. Deskripsi Sistem.....	30
3.5. Perhitungan Manual Cosine Similarity	32
3.5.1 Teks jawaban dan kunci.....	32

3.5.2	Text processing.....	32
3.5.3	Menghitung TF-IDF dan pembobotan kata	34
3.5.4	Perhitungan kemiripan menggunakan Cosine similarity	35
3.5.5	Konversi Nilai Kemiripan menjadi Nilai Ujian Esai	35
3.6.	Analisis Kebutuhan	36
3.6.1	Analisis Kebutuhan Fungsional.....	36
3.6.1.1	Definisi Fitur	36
3.6.1.2	Aktor.....	38
3.6.1.3	Use Case	38
3.6.1.4	Skenario Use Case.....	40
3.6.2	Analisis Kebutuhan Non-Fungsional.....	52
3.7.	Perancangan Perangkat Lunak	54
3.7.1.	Class Diagram	54
3.7.2.	Activity Diagram.....	54
3.7.3.	Sequence Diagram	67
3.8.	Rancangan Antarmuka Pengguna (User Interface).....	78
BAB IV HASIL DAN PEMBAHASAN		85
4.1.	Database dan Tabel	85
4.2.	Implementasi Antarmuka Pengguna	89
4.3.	Pembahasan Source Code	99
4.4.	Pengujian Sistem	106
4.4.1.	White Box Testing	106
4.4.2.	Black Box Testing.....	107
4.5.	Hasil dan Analisis Pengujian.....	108
4.5.1.	Pengujian Sistem Dengan Jawaban Siswa	108
4.5.2.	Perhitungan Dalam Sistem.....	112
4.5.3.	Kelebihan dan Kekurangan Sistem.....	115
BAB V PENUTUP.....		116
5.1.	Kesimpulan.....	116
5.2.	Saran.....	116



DAFTAR TABEL

Tabel 2.1 Perbedaan Penelitian.....	11
Tabel 2.2 Tabel Confusion Matrix.....	19
Tabel 2.3 Simbol <i>Use Case Diagram</i>	20
Tabel 2.4 Simbol <i>Class Diagram</i>	21
Tabel 2.5 Simbol <i>Sequence Diagram</i>	23
Tabel 2.6 Simbol <i>Activity Diagram</i>	25
Tabel 3.1 Tabel Jawaban Dan Kunci	32
Tabel 3.2 Tabel Text Processing.....	32
Tabel 3.3 Tabel Perhitungan Tf-Idf	34
Tabel 3.4 Tabel Hasil Tf-Idf	34
Tabel 3.5 Tabel Rentang Nilai	35
Tabel 3.6 Tabel Definisi Fitur.....	36
Tabel 3.7 Tabel Aktor	38
Tabel 3.8 Skenario <i>Use Case Login</i>	40
Tabel 3.9 Skenario <i>Use Case Mengerjakan Soal</i>	41
Tabel 3.10 Skenario <i>Use Case Lihat Riwayat Ujian</i>	41
Tabel 3.11 Skenario <i>Use Case Ubah Profil Siswa</i>	42
Tabel 3.12 Skenario <i>Use Case Ubah Profil Admin</i>	43
Tabel 3.13 Skenario <i>Use Case Manipulasi Data Mata Pelajaran</i>	43
Tabel 3.14 Skenario <i>Use Case Manipulasi Data Siswa</i>	44
Tabel 3.15 Skenario <i>Use Case Manipulasi Data Guru</i>	45
Tabel 3.16 Skenario <i>Use Case Manipulasi Data Admin</i>	46
Tabel 3.17 Skenario <i>Use Case Manipulasi Data Kelas</i>	48

Tabel 3.18 Skenario *Use Case* Manipulasi Data Ujian.....49

Tabel 3.19 Skenario *Use Case* Manipulasi Data Soal50

Tabel 3.20 Skenario *Use Case* Ubah Profil Guru51

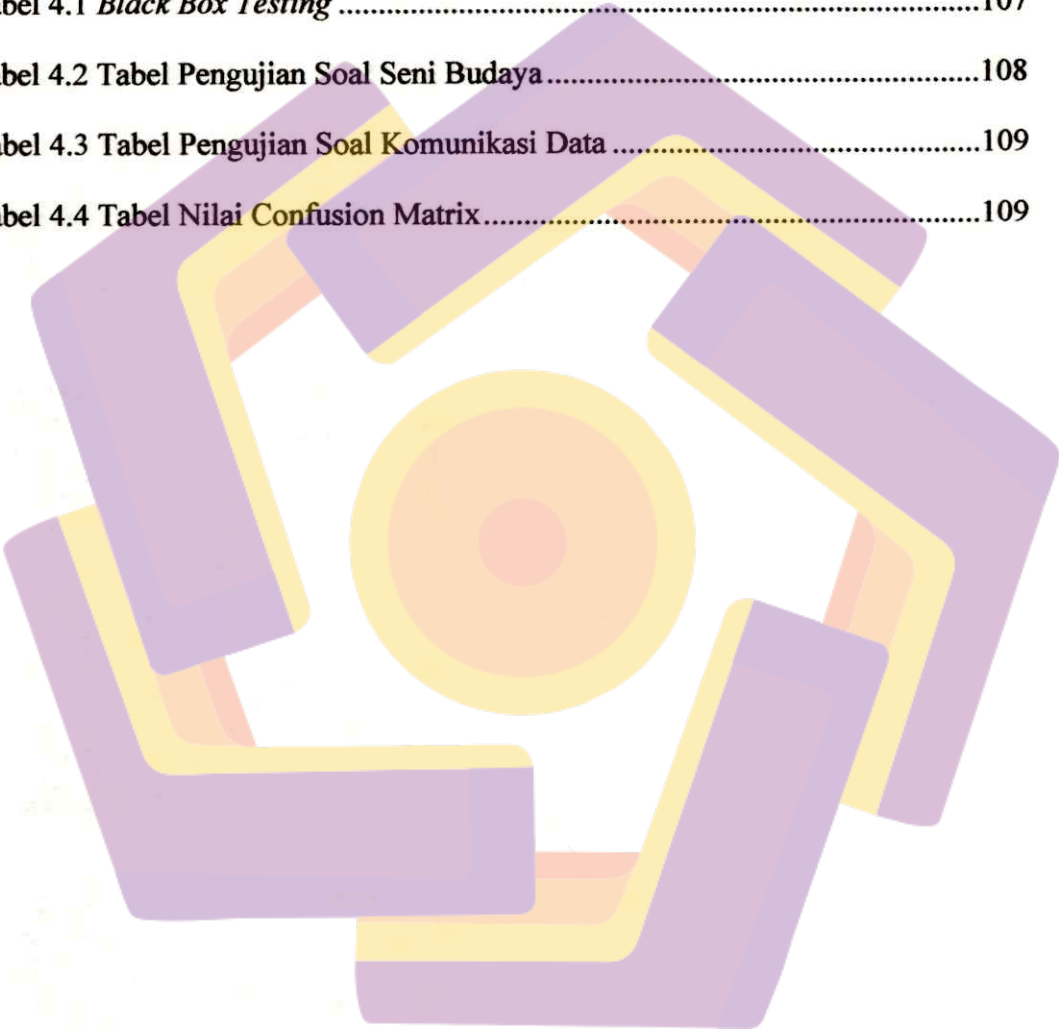
Tabel 3.21 Skenario *Use Case* Lihat Nilai Siswa52

Tabel 4.1 *Black Box Testing*107

Tabel 4.2 Tabel Pengujian Soal Seni Budaya.....108

Tabel 4.3 Tabel Pengujian Soal Komunikasi Data109

Tabel 4.4 Tabel Nilai Confusion Matrix.....109



DAFTAR GAMBAR

Gambar 3.2 Use Case Sistem Aktor Siswa	38
Gambar 3.3 Use Case Sistem Aktor Guru	39
Gambar 3.4 Use Case Sistem Aktor Admin	39
Gambar 3.5 <i>Class Diagram</i> Sistem.....	54
Gambar 3.6 <i>Activity Diagram</i> Login.....	55
Gambar 3.7 <i>Activity Diagram</i> Ubah Profil Siswa	56
Gambar 3.8 <i>Activity Diagram</i> Mengerjakan Ujian.....	57
Gambar 3.9 <i>Activity Diagram</i> Lihat Riwayat Ujian.....	58
Gambar 3.10 <i>Activity Diagram</i> Ubah Profil Admin.....	58
Gambar 3.11 <i>Activity Diagram</i> Manipulasi Data Pelajaran	59
Gambar 3.12 <i>Activity Diagram</i> Manipulasi Data Siswa	60
Gambar 3.13 <i>Activity Diagram</i> Manipulasi Data Guru.....	61
Gambar 3.14 <i>Activity Diagram</i> Manipulasi Data Admin.....	62
Gambar 3.15 <i>Activity Diagram</i> Manipulasi Data Kelas.....	63
Gambar 3.16 <i>Activity Diagram</i> Manipulasi Data Ujian	64
Gambar 3.17 <i>Activity Diagram</i> Manipulasi Data Soal Ujian.....	65
Gambar 3.18 <i>Activity Diagram</i> Ubah Profil Guru.....	66
Gambar 3.19 <i>Activity Diagram</i> Lihat Nilai Siswa.....	66
Gambar 3.20 <i>Sequence Diagram</i> Login.....	67
Gambar 3.21 <i>Sequence Diagram</i> Ubah Profil Siswa	67
Gambar 3.22 <i>Sequence Diagram</i> Mengerjakan Ujian	68
Gambar 3.23 <i>Sequence Diagram</i> Lihat Riwayat Ujian.....	68
Gambar 3.24 <i>Sequence Diagram</i> Ubah Profil Admin.....	69

Gambar 3.25 Sequence Diagram Manipulasi Data Pelajaran	70
Gambar 3.26 Sequence Diagram Manipulasi Data Siswa	71
Gambar 3.27 Sequence Diagram Manipulasi Data Guru.....	72
Gambar 3.28 Sequence Diagram Manipulasi Data Admin.....	73
Gambar 3.29 Sequence Diagram Manipulasi Data Kelas.....	74
Gambar 3.30 Sequence Diagram Manipulasi Data Ujian.....	75
Gambar 3.31 Sequence Diagram Manipulasi Data Soal.....	76
Gambar 3.32 Sequence Diagram Ubah Profil Guru	77
Gambar 3.33 Sequence Diagram Lihat Nilai Siswa	77
Gambar 3.34 User Interface Login	78
Gambar 3.35 User Interface Mengerjakan Ujian.....	78
Gambar 3.36 User Interface Riwayat Ujian.....	79
Gambar 3.37 User Interface Ubah Profil Siswa.....	79
Gambar 3.38 User Interface Ubah Profil Admin	80
Gambar 3.39 User Interface Manipulasi Data Pelajaran.....	80
Gambar 3.40 User Interface Manipulasi Data Siswa.....	81
Gambar 3.41 User Interface Manipulasi Data Guru	81
Gambar 3.42 User Interface Manipulasi Data Admin	82
Gambar 3.43 User Interface Manipulasi Data Kelas	82
Gambar 3.44 User Interface Manipulasi Data Ujian.....	83
Gambar 3.45 User Interface Manipulasi Data Soal	83
Gambar 3.46 User Interface Ubah Profil Guru	84
Gambar 3.47 User Interface Lihat Nilai Siswa	84
Gambar 4.1 Struktur Tabel Admin	85

Gambar 4.2 Struktur Tabel Guru	86
Gambar 4.3 Struktur Tabel Siswa	86
Gambar 4.4 Struktur Tabel Kelas	87
Gambar 4.5 Struktur Tabel Pelajaran.....	87
Gambar 4.6 Struktur Tabel Ujian.....	87
Gambar 4.7 Struktur Tabel Soal	88
Gambar 4.8 Struktur Tabel Hasil Ujian	88
Gambar 4.9 Struktur Tabel <i>Dictionary</i>	89
Gambar 4.10 Struktur Tabel Kata Hubung	89
Gambar 4.11 Form Login.....	90
Gambar 4.12 Halaman Mengerjakan Ujian	90
Gambar 4.13 Halaman Riwayat Ujian	91
Gambar 4.14 Halaman Edit Profil Siswa.....	92
Gambar 4.15 Halaman Edit Profil Admin	92
Gambar 4.16 Halaman Manipulasi Data Pelajaran.....	93
Gambar 4.17 Halaman Manipulasi Data Siswa	94
Gambar 4.18 Halaman Manipulasi Data Guru.....	94
Gambar 4.19 Halaman Manipulasi Data Admin.....	95
Gambar 4.20 Halaman Manipulasi Data Kelas.....	96
Gambar 4.21 Halaman Manipulasi Data Ujian.....	96
Gambar 4.22 Halaman Manipulasi Data Soal.....	97
Gambar 4.23 Halaman Edit Profil Guru	98
Gambar 4.24 Halaman Lihat Nilai Siswa 1	98
Gambar 4.25 Halaman Lihat Nilai Siswa 2	99

Gambar 4.26 Kode Fungsi Preproses.....	99
Gambar 4.27 Kode Proses Filtering.....	100
Gambar 4.28 Kode Fungsi Hapus_Symbol.....	100
Gambar 4.29 Kode Fungsi Pecah_Kata.....	101
Gambar 4.30 Kode Fungsi Cari_Kata_Dasar.....	101
Gambar 4.31 Kode File Proses_Ujian.Php.....	103
Gambar 4.32 Kode Fungsi Hapus_Symbol.....	104
Gambar 4.33 Kode Tf-Idf.....	104
Gambar 4.34 Kode Cosine Similarity.....	105
Gambar 4.28 Kode Konversi Nilai.....	106
Gambar 4.29 Memasukkan Teks Uji Jawaban Dan Kunci Jawaban.....	112
Gambar 4.30 Hasil Case Folding Dan Hapus Symbol.....	113
Gambar 4.31 Hasil Tokenizing.....	113
Gambar 4.32 Hasil Filtering.....	113
Gambar 4.33 Hasil Stemming.....	113
Gambar 4.34 Hasil Merging Term Yang Mewakili Tiap Dokumen.....	114
Gambar 4.35 Hasil Perhitungan Tf-Idf.....	114
Gambar 4.36 Hasil Cosine Similarity.....	114

INTISARI

Penelitian ini bertujuan untuk mengimplementasikan text mining dengan algoritma cosine similarity untuk menganalisis jawaban esai serta menganalisis tingkat keakuratan kemiripan pada jawaban soal esai. Penelitian ini akan meneliti sistem pengoreksian jawaban esai berbasis web menggunakan bahasa pemrograman PHP dengan algoritma cosine similarity pada dokumen yang telah dikumpulkan dari tenaga pengajar.

Metode penelitian pada penelitian ini menggunakan *research and development*. Data penelitian ini diperoleh dengan cara observasi, studi pustaka, dan analisis dokumen. data dianalisis dengan *text preprocessing* dengan tahap *case folding*, *tokenizing*, *stopword removal/filtering*, *stemming* dan *analyzing*, serta pembobotan menggunakan TF-IDF.

Hasil penelitian ini adalah bahwa sistem bisa digunakan untuk mengoreksi jawaban esai berbahasa indonesia. Algoritma cosine similarity pada sistem memiliki performa yang cukup baik dan memiliki tingkat akurasi sebesar sebesar 69,10%, presisi sebesar 93,95%, *recall* sebesar 67,56%, dan *error rate* sebesar 30,90%.

Kata kunci : ujian, esai, cosine similarity, sistem pengoreksian jawaban.

ABSTRACT

This study aims to implement text mining with the cosine similarity algorithm to analyze the essay answers and analyze the level of accuracy of the similarities in the answers to essay questions. This research will examine the web-based essay answer correction system using PHP programming language with cosine similarity algorithm on documents that have been collected from the teaching staff.

The research method in this study is research and development method. The research data was obtained by observation, literature study, and document analysis. Data were analyzed by text preprocessing with case folding, tokenizing, stopword removal / filtering, stemming and analyzing, and weighting using TF-IDF.

The results of this study are that the system can be used to correct Indonesian essay answers. Cosine similarity algorithm on the system has a pretty good performance and has an accuracy rate of 69.10%, a precision of 93.95%, a recall of 67.56%, and an error rate of 30.90%.

Keyword : *cosine similarity, exam, essay, answer correction system.*