

**IMPLEMENTASI ALGORITMA JACCARD UNTUK  
OTOMATISASI KOREKSI JAWABAN ESAI  
DI SMP MUHAMMADIYAH KALIABU**

**SKRIPSI**



**disusun oleh**

**Heryan Cahya Febriansyah**

**16.11.0867**

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

**IMPLEMENTASI ALGORITMA JACCARD UNTUK  
OTOMATISASI KOREKSI JAWABAN ESAI  
DI SMP MUHAMMADIYAH KALIABU**

SKRIPSI

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



**disusun oleh**  
**Heryan Cahya Febriansyah**  
**16.11.0867**

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

# PERSETUJUAN

## SKRIPSI

### IMPLEMENTASI ALGORITMA JACCARD UNTUK OTOMATISASI KOREKSI JAWABAN ESAI DI SMP MUHAMMADIYAH KALIABU

yang dipersiapkan dan disusun oleh

**Heryan Cahya Febriansyah**

16.11.0867

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 28 November 2019

Dosen Pembimbing,



**Anggit Dwi Hartanto, M.Kom**

NIK. 190302163

**PENGESAHAN**

**SKRIPSI**

**IMPLEMENTASI ALGORITMA JACCARD UNTUK  
OTOMATISASI KOREKSI JAWABAN ESAI  
DI SMP MUHAMMADIYAH KALIABU**

yang dipersiapkan dan disusun oleh

**Heryan Cahya Febriansyah**

**16.11.0867**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 19 Desember 2019

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Dina Maulina, M.kom**  
NIK. 190302250



**Windha Mega Pradnya D, M.Kom**  
NIK. 190302185



**Anggit Dwi Hartanto, M.Kom**  
NIK. 190302163



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 28 Desember 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 tanggung jawab saya pribadi.

Yogyakarta, 28 Desember 2019



Heryan Cahya Febriansyah

NIM. 16.11.0867

## MOTTO

“Aku akan bersantai meskipun itu akan membunuhku” – Squidward Tentacles

“Aku jelek dan aku bangga” – Spongebob Squarepants

“Hidup memang tidak adil, jadi biasakanlah” – Patrick Star

“Uang selalu benar” – Eugene krab

“Rose are red, Violets are blue, omae wa mou shindeiru” – Kenshiro

“NANI!” – RIP



## PERSEMBAHAN

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

1. Ibu dan keluarga saya, yang selalu mendoakan, menyemangati dan menjadi pendorong saya untuk segera menyelesaikan skripsi.
2. Bapak Anggit Dwi Hartanto, M.Kom. yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
3. Dosen-dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama perkuliahan.
4. Teman-teman kelas IF14 2016 yang selalu menemani perkuliahan, mendukung dan memberikan semangat sampai saat ini.



## 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 Implementasi Algoritma Jaccard untuk Otomatisasi Koreksi Jawaban Esai di SMP Muhammadiyah Kaliabu.

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:

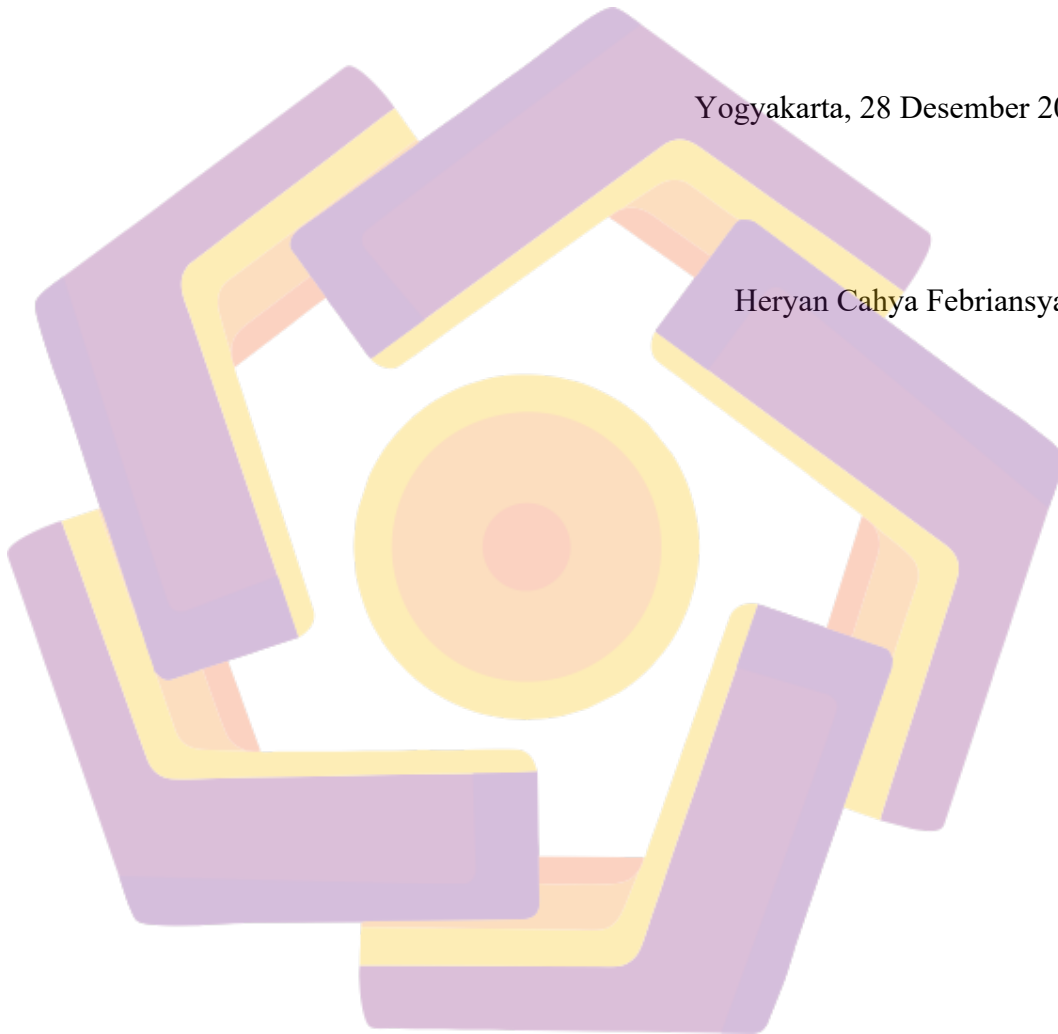
1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Bapak Anggit Dwi Hartanto, M.Kom. selaku dosen pembimbing yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini.
4. Dosen Penguji dan segenap Dosen dan Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.
5. Smp Muhammadiyah Kaliabu yang bersedia menjadi penyedia sumber data penelitian ini.
6. 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 Desember 2019

Heryan Cahya Febriansyah



## DAFTAR ISI

JUDUL .....	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR_PERSETUJUAN.....	<b>Error! Bookmark not defined.</b>
PERNYATAAN.....	<b>Error! Bookmark not defined.</b>
MOTTO.....	v
PERSEMBAHAN .....	ivi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiv
INTISARI.....	xvii
<i>ABSTRACT</i> .....	xviii
BAB I.....	1
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Penelitian .....	2
1.4 Maksud dan Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Metode Penelitian.....	4
1.6.1 Metode Pengumpulan Data .....	4
1.6.2 Metode Analisis.....	5
1.6.3 Metode Perancangan .....	5
1.6.4 Metode Pengembangan Sistem .....	6
1.6.5 Metode Testing.....	6
1.7 Sistematika Penulisan.....	7
BAB II.....	9

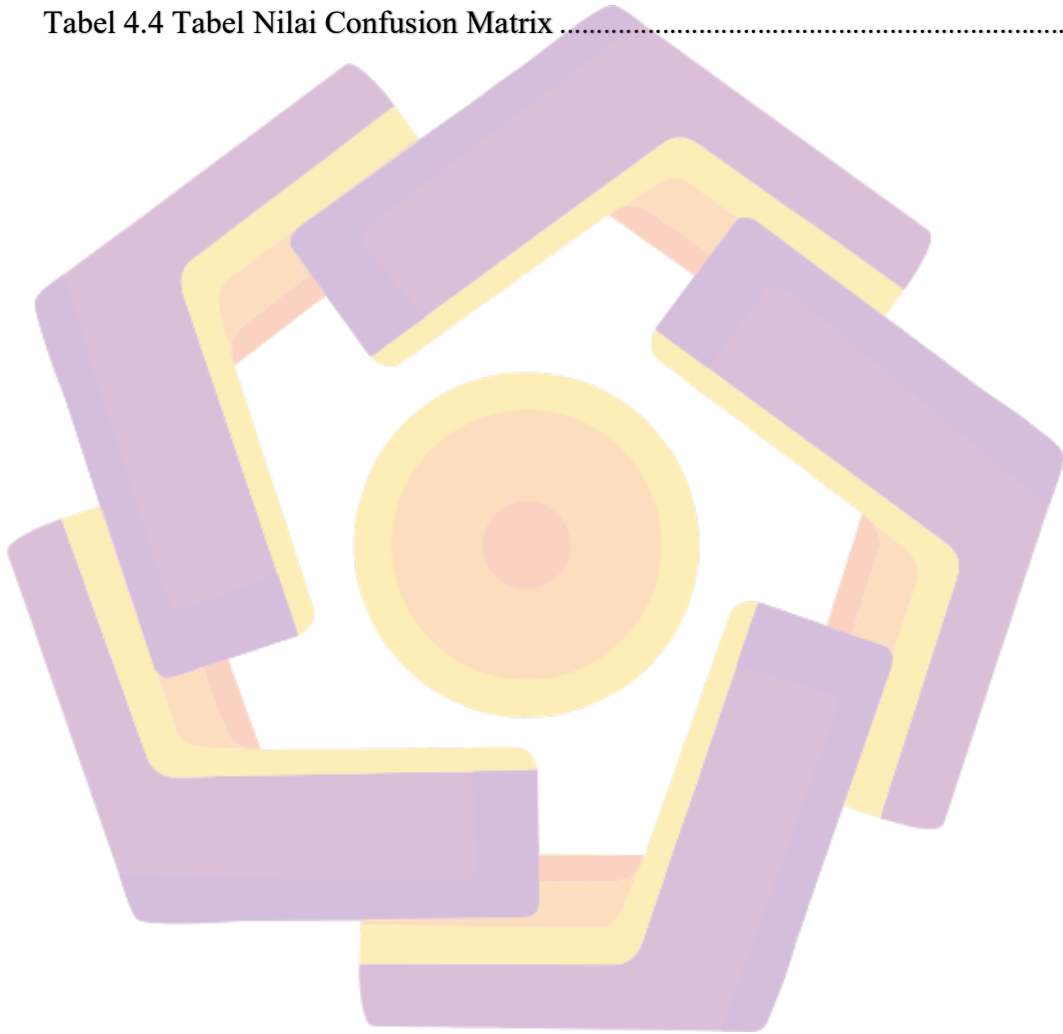
2.1	Tinjauan Pustaka .....	9
2.2	Dasar Teori.....	11
2.2.1	Esai.....	11
2.2.2	Otomatisasi.....	11
2.2.3	Otomatisasi Esai.....	12
2.2.4	Text Mining.....	12
2.2.5	Text Preprocessing .....	13
2.2.5.1	Text Transformation.....	14
2.2.6	N-Gram.....	15
2.2.7	Jaccard Similarity.....	15
2.2.8	<i>Confusion Matrix</i> .....	16
2.2.9	Unified Modelling Language (UML).....	17
2.2.9.1	Use Case Diagram.....	17
2.2.9.2	Class Diagram .....	19
2.2.9.3	Sequence Diagram.....	21
2.2.9.4	Activity Diagram.....	23
2.2.10	Bahasa Pengkodean .....	25
2.2.10.1	XML.....	25
2.2.10.2	Java.....	26
<b>BAB III</b>	.....	<b>27</b>
3.1	Identifikasi Masalah .....	27
3.2	Analisis Masalah .....	27
3.3	Hasil Analisis .....	28
3.4	Deskripsi Sistem.....	28
3.5	Perhitungan Manual Jaccard Similarity .....	30
3.5.1	Teks jawaban dan kunci .....	30
3.5.2	<i>Text processing</i> .....	31
3.5.3	N-Gram.....	31
3.5.4	Perhitungan kemiripan menggunakan <i>Jaccard similarity</i> .....	32
3.5.5	Konversi Nilai Kemiripan menjadi Nilai Ujian Esai .....	33
3.6	Analisis Kebutuhan .....	33
3.6.1	Analisis Kebutuhan Fungsional .....	33

3.6.1.1	Definisi Fitur .....	34
3.6.1.2	Aktor.....	35
3.6.1.3	Use Case.....	35
3.6.1.4	Skenario Use Case.....	36
3.6.2	Analisis Kebutuhan Non-Fungsional .....	43
3.7	Perancangan Perangkat Lunak .....	44
3.7.1	<i>Class Diagram</i> .....	44
3.7.2	<i>Activity Diagram</i> .....	45
3.7.3	<i>Sequence Diagram</i> .....	50
3.8	Rancangan Antarmuka Pengguna ( <i>User Interface</i> ) .....	54
<b>BAB IV</b>	.....	<b>58</b>
4.1	Database .....	58
4.2	Implementasi Antarmuka Pengguna .....	61
4.3	Pembahasan Source Code .....	66
4.3.1	Kode Preproses.....	66
4.3.2	Kode N-Gram dan Jaccard Similarity .....	68
4.4	Pengujian Sistem .....	69
4.4.1	<i>White Box Testing</i> .....	70
4.4.2	<i>Black Box Testing</i> .....	70
4.5	Hasil dan Analisis Pengujian.....	71
4.5.1	Pengujian Sistem Dengan Jawaban Siswa .....	71
4.5.2	Perhitungan Dalam Sistem .....	76
4.5.3	Kelebihan dan Kekurangan Sistem .....	80
4.5.3.1	Kelebihan .....	80
4.5.3.2	Kekurangan .....	81
<b>BAB V</b>	.....	<b>82</b>
5.1	Kesimpulan.....	82
5.2	Saran.....	82

## DAFTAR TABEL

Tabel 2.1 Perbedaan Penelitian.....	10
Tabel 2.2 Table Confusion Matrix .....	17
Tabel 2.3 Simbol <i>Use Case Diagram</i> .....	18
Tabel 2.4 Simbol <i>Class Diagram</i> .....	19
Tabel 2.5 Simbol <i>Sequence Diagram</i> .....	21
Tabel 2.6 Simbol <i>Activity Diagram</i> .....	23
Tabel 3.1 Tabel Jawaban dan Kunci .....	30
Tabel 3.2 Tabel Text Processing .....	31
Tabel 3.3 Tabel proses n-Gram .....	31
Tabel 3.4 Tabel irisan dan gabungan dataset.....	32
Tabel 3.5 Tabel Rentang Nilai.....	33
Tabel 3.6 Tabel Definisi Fitur.....	34
Tabel 3.7 Tabel Aktor .....	35
Tabel 3.8 Skenario <i>Use Case Register</i> .....	36
Tabel 3.9 Skenario <i>Use Case Login</i> .....	37
Tabel 3.10 Skenario <i>Use Case Mengerjakan Soal Ujian</i> .....	38
Tabel 3.11 Skenario <i>Use Case Lihat Riwayat Ujian</i> .....	39
Tabel 3.12 Skenario <i>Use Case Membuat Mata Pelajaran</i> .....	39
Tabel 3.13 Skenario <i>Use Case Membuat Soal Ujian</i> .....	40
Tabel 3.14 Skenario <i>Use Case Masuk ke Mata Pelajaran</i> .....	42
Tabel 3.15 Skenario <i>Use Case Melihat Nilai Siswa</i> .....	42

Tabel 3.16 Kebutuhan Perangkat Keras Smartphone .....	43
Tabel 4.1 <i>Black Box Testing</i> .....	70
Tabel 4.2 Tabel Pengujian Soal Bahasa Indonesia .....	71
Tabel 4.3 Tabel Pengujian Soal Pendidikan Kewarganegaraan.....	73
Tabel 4.4 Tabel Nilai Confusion Matrix .....	74



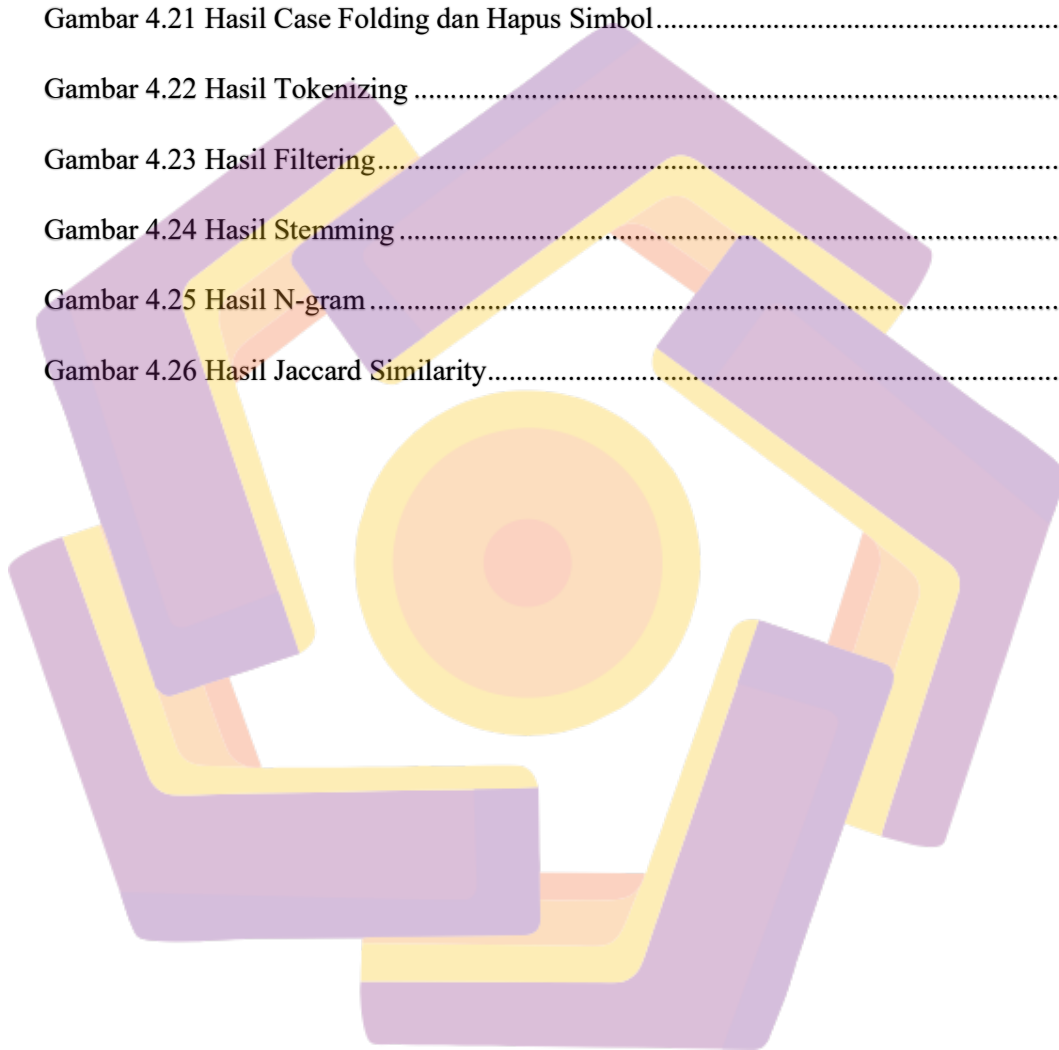
## DAFTAR GAMBAR

Gambar 2.1 Proses Case Folding .....	14
Gambar 2.2 Proses Tokenizing .....	14
Gambar 2.3 Rumus Jaccard Similarity.....	16
Gambar 3.1 Langkah Jaccard Similarity .....	29
Gambar 3.2 Use Case Sistem Aktor Siswa .....	36
Gambar 3.3 Use Case Sistem Aktor Guru.....	36
Gambar 3.4 <i>Class Diagram</i> Sistem .....	45
Gambar 3.5 <i>Activity Diagram</i> Register .....	46
Gambar 3.6 <i>Activity Diagram</i> Login .....	46
Gambar 3.7 <i>Activity Diagram</i> Masuk Ke Mata Pelajaran .....	47
Gambar 3.8 <i>Activity Diagram</i> Mengerjakan Ujian.....	47
Gambar 3.9 <i>Activity Diagram</i> Melihat Riwayat Pengerjaan Ujian .....	48
Gambar 3.10 <i>Activity Diagram</i> Membuat Soal .....	48
Gambar 3.11 <i>Activity Diagram</i> Membuat Mata Pelajaran .....	49
Gambar 3.12 <i>Activity Diagram</i> Menampilkan Nilai Murid.....	49
Gambar 3.13 <i>Sequence Diagram</i> Registration .....	50
Gambar 3.14 <i>Sequence Diagram</i> Login .....	50
Gambar 3.15 <i>Sequence Diagram</i> Mengerjakan Ujian.....	51
Gambar 3.16 <i>Sequence Diagram</i> Lihat Riwayat Ujian .....	52
Gambar 3.17 <i>Sequence Diagram</i> Membuat Mata Pelajaran .....	52
Gambar 3.18 <i>Sequence Diagram</i> Membuat Soal Ujian .....	52
Gambar 3.19 <i>Sequence Diagram</i> Masuk ke Mata Pelajaran.....	53
Gambar 3.20 <i>Sequence Diagram</i> Melihat Nilai Siswa.....	53

Gambar 3.21 User Interface Register .....	54
Gambar 3.22 User Interface Login.....	54
Gambar 3.23 User Interface Mengerjakan Ujian .....	55
Gambar 3.24 User Interface Nilai Riwayat Ujian .....	55
Gambar 3.25 User Interface Masuk Ke Mata Pelajaran.....	56
Gambar 3.26 User Interface Membuat Soal Ujian.....	56
Gambar 3.27 User Interface Menambah Mata Pelajaran .....	57
Gambar 3.28 User Interface Melihat Nilai Siswa .....	57
Gambar 4.1 Collection dan Document User .....	58
Gambar 4.2 Collection dan Document Kelas.....	59
Gambar 4.3 Collection dan Document Soal.....	60
Gambar 4.4 Collection dan Document Hasil .....	60
Gambar 4.5 Form Register.....	61
Gambar 4.6 Form Login.....	62
Gambar 4.7 Halaman Mengerjakan Ujian.....	62
Gambar 4.8 Halaman Riwayat Ujian .....	63
Gambar 4.9 Halaman Masukkan Kode Mata Pelajaran .....	63
Gambar 4.10 Halaman Membuat Ujian .....	64
Gambar 4.11 Halaman Membuat Mata Pelajaran .....	65
Gambar 4.12 Halaman Lihat Nilai Siswa.....	65
Gambar 4.13 Kode Fungsi Preproses .....	66
Gambar 4.14 Kode Proses Case Folding.....	66
Gambar 4.15 Kode Tokenisasi .....	67
Gambar 4.16 Kode Filltering .....	67



Gambar 4.17 Kode Stemming.....	68
Gambar 4.18 Kode N-Gram dan Jaccard Similarity .....	68
Gambar 4.19 Kode Konversi Nilai.....	69
Gambar 4.20 Memasukkan Teks Uji Jawaban dan Kunci Jawaban .....	77
Gambar 4.21 Hasil Case Folding dan Hapus Simbol.....	78
Gambar 4.22 Hasil Tokenizing .....	78
Gambar 4.23 Hasil Filtering.....	79
Gambar 4.24 Hasil Stemming .....	79
Gambar 4.25 Hasil N-gram .....	80
Gambar 4.26 Hasil Jaccard Similarity.....	80



## INTISARI

Penelitian ini bertujuan untuk mengimplementasikan text mining dengan algoritma Jaccard similarity untuk menganalisis jawaban esai serta menganalisis tingkat keakuratan kemiripan pada jawaban soal esai. Penelitian ini akan meneliti sistem pengoreksian jawaban esai berbasis android menggunakan bahasa pemrograman Java dengan algoritma jaccard similarity pada dokumen yang telah dikumpulkan dari Smp Muhammadiyah Kaliabu.

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 menggunakan N-Gram.

Hasil penelitian ini adalah bahwa sistem bisa digunakan untuk mengoreksi jawaban esai berbahasa Indonesia. Algoritma Jaccard similarity pada sistem memiliki performa yang cukup baik dan memiliki tingkat nilai akurasi sebesar 65,11%, presisi 92,92%, recall 57,14%, dan error rate 35,34%

**Kata kunci :** ujian, esai, Jaccard similarity, sistem pengoreksian jawaban.

## **ABSTRACT**

*This study aims to implement text mining with the Jaccard 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 Android-based essay answer correction system using Java programming language with Jaccard similarity algorithm on documents that have been collected from SMP Muhammadiyah Kaliabu.*

*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 using N-Gram.*

*The results of this study are that the system can be used to correct Indonesian essay answers. Jaccard similarity algorithm on the system has a pretty good performance and has an accuracy rate of 65.11%, a precision of 92.92%, a recall of 57.14%, and an error rate of 35.34%.*

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

