

SKRIPSI

**SISTEM KOREKSI JAWABAN *ESSAY*
MENGUNAKAN ALGORITMA NAZIEF & ADRIANI
DI SMA NEGERI 1 MAGELANG**



**CATUR MULYAWAN
NPM. 11.0504.0016**

**PROGRAM STUDI TEKNIK INFORMATIKA S1
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAGELANG
2018**

**SISTEM KOREKSI JAWABAN *ESSAY*
MENGUNAKAN ALGORITMA NAZIEF & ADRIANI
DI SMA NEGERI 1 MAGELANG**

Disusun Sebagai Salah Satu Syarat Memperoleh Gelar Sarjana Komputer (S.Kom)
Program Studi Teknik Informatika Jenjang Strata Satu (S-1) Fakultas Teknik
Universitas Muhammadiyah Magelang



OLEH :

**CATUR MULYAWAN
NPM. 11.0504.0016**

**PROGRAM STUDI TEKNIK INFORMATIKA S1
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAGELANG
FEBRUARI, 2018**

HALAMAN PENEGASAN

Skripsi ini adalah hasil karya saya sendiri dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Catur Mulyawan

NPM : 11.0504.0016

Magelang, 14 Februari 2018

Yang menyatakan,



CATUR MULYAWAN

NPM. 11.0504.0016

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

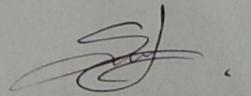
Nama : Catur Mulyawan
NPM : 11.0504.0016
Program Studi : Teknik Informatika
Fakultas : Teknik
Alamat : Krajan RT 02 RW 01, Tlogokotes, Bagelen, Purworejo
Judul Skripsi : Sistem Koreksi Jawaban *Essay* Di SMA Negeri 1
Magelang

Dengan ini menyatakan bahwa Laporan Skripsi ini merupakan hasil karya sendiri dan bukan merupakan plagiat dari hasil karya orang lain. Dan bila dikemudian hari terbukti bahwa karya tersebut merupakan plagiat, maka saya bersedia menerima sanksi administrasi maupun sanksi apapun.

Demikian surat pernyataan ini saya buat dengan penuh kesadaran dan sebenarnya serta penuh tanggung jawab.

Magelang, 14 Februari 2018

Yang menyatakan,



CATUR MULYAWAN

NPM. 11.0504.0016

HALAMAN PENGESAHAN

SKRIPSI

**SISTEM KOREKSI JAWABAN ESSAY MENGGUNAKAN
ALGORITMA NAZIEF & ADRIANI DI SMA NEGERI 1
MAGELANG**

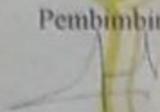
dipersiapkan dan disusun oleh

**CATUR MULYAWAN
NPM. 11.0504.0016**

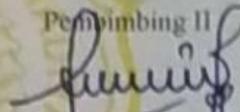
Telah dipertabangkan di depan Dewan Penguji
Pada tanggal 14 Februari 2018

Susunan Dewan Penguji

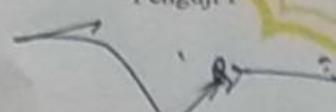
Pembimbing I


Andi Widiyanto, S.Kom., M.Kom
NIDN. 0623087901

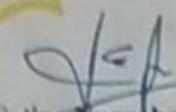
Pembimbing II


Endah Ratna Arum, M.Cs
NIDN. 0601129001

Penguji I

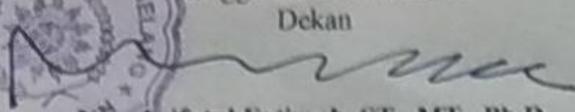

Nuryanto, ST., M.Kom
NIDN. 0605037002

Penguji II


Setiyo Nugroho, M. Eng
NIDN. 0631088203



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
tanggal 14 Februari 2018
Dekan


Yun Arifatul Fatimah, ST., MT., Ph.D.
NIK. 987408139

KATA PENGANTAR

Puji syukur dipanjatkan ke hadirat Allah SWT, karena atas berkat nikmat dan karunia-Nya, Skripsi ini dapat diselesaikan. Penyusunan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Magelang.

Penyelesaian Skripsi ini banyak memperoleh bantuan dan bimbingan dari berbagai pihak. Oleh karena itu, diucapkan terima kasih kepada :

1. Ir. Eko Muh Widodo, MT selaku Rektor yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penyusunan skripsi ini.
2. Yun Arifatul Fatimah, ST., MT selaku Dekan Fakultas Teknik yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penyusunan skripsi ini.
3. Agus Setiawan, M.Eng selaku Kepala Program Studi Teknik Informatika S1 Universitas Muhammadiyah Magelang.
4. Andi Widiyanto, M.Kom dan Endah Ratna Arumi, M.Cs selaku dosen pembimbing pendamping yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penyusunan skripsi ini.
5. Beberapa pihak yang telah banyak membantu dalam usaha memperoleh data yang diperlukan.
6. Orang tua dan keluarga yang telah memberikan bantuan dukungan material dan moral.

Akhir kata, semoga Allah SWT berkenan membalas segala kebaikan semua pihak yang telah membantu dan semoga Skripsi ini membawa manfaat bagi pengembangan ilmu.

Magelang, 14 Februari 2018


Catur Mulvawan
NPM. 11.0504.0016

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN	ii
PERNYATAAN KEASLIAN.....	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	xii
ABSTRAK	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
A. Latar Belakang Permasalahan.....	1
B. Rumusan Masalah.....	3
C. Tujuan Penelitian	3
D. Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA.....	4
A. Penelitian Relevan	4
B. Variabel Penelitian.....	5
1. Sistem	5
2. Ujian Essay	5
3. Cosine Similarity	6
4. Teks Mining.....	7
5. Algoritma Nazief & Adriani.....	7
6. PHP – MySQL	13
C. Landasan Teori.....	14

BAB III ANALISIS DAN PERANCANGAN	15
A. Analisis Sistem.....	15
1. Analisis Sistem Yang Sedang Berjalan	15
2. Koreksi Ujian Essay Baru.....	16
3. Contoh Koreksi Jawaban Essay.....	16
B. Perancangan Sistem	19
1. DFD	19
2. ERD	22
3. Perancangan Basis Data.....	23
4. Perancangan Antarmuka.....	26
BAB IV IMPLEMENTASI DAN PENGUJIAN.....	35
A. Implementasi Sistem	35
B. Implementasi Metode <i>Cosine Similarity</i>	38
C. Implementasi <i>Interface</i>	45
D. Pengujian Sistem.....	53
BAB V HASIL DAN PEMBAHASAN	65
A. Hasil Sistem Koreksi Jawaban Essay	65
B. Pembahasan	67
C. Kelebihan Sistem	75
D. Manfaat Sistem	75
BAB VI PENUTUP	76
A. Kesimpulan	76
B. Saran	76
DAFTAR PUSTAKA	77

DAFTAR GAMBAR

Gambar 3.1 Flowchart sistem lama koreksi jawaban essay.....	15
Gambar 3.2 Flowchart sistem baru koreksi jawaban essay.....	16
Gambar 3.3 Diagram Kontek	20
Gambar 3.4 DFD Level 0.....	21
Gambar 3.5 ERD.....	22
Gambar 3.6 Relasi tabel	23
Gambar 3.7 Rancangan halaman login	27
Gambar 3.8 Rancangan halaman data siswa	27
Gambar 3.9 Rancangan halaman data guru.....	28
Gambar 3.10 Rancangan halaman data admin	28
Gambar 3.11 Rancangan halaman data kelas.....	29
Gambar 3.12 Rancangan halaman data mata pelajaran	29
Gambar 3.13 Rancangan halaman data ujian	30
Gambar 3.14 Rancangan halaman data soal ujian	30
Gambar 3.15 Rancangan halaman data hasil ujian	31
Gambar 3.16 Rancangan halaman data detail ujian.....	31
Gambar 3.17 Rancangan halaman awal	32
Gambar 3.18 Rancangan halaman pilih data ujian.....	32
Gambar 3.19 Rancangan halaman detail ujian.....	33
Gambar 3.20 Rancangan halaman ujian.....	33
Gambar 3.21 Rancangan halaman hasil ujian	34
Gambar 4.1. Tabel Admin.....	36
Gambar 4.2. Tabel Guru.....	36
Gambar 4.3. Tabel Kamus	36
Gambar 4.4. Tabel Kelas.....	36
Gambar 4.5. Tabel Mapel.....	37
Gambar 4.6. Tabel Siswa	37
Gambar 4.7. Tabel Siswa Ujian	37
Gambar 4.8. Tabel Siswa Ujian Hasil.....	37

Gambar 4.9. Tabel Siswa Ujian Jawaban	38
Gambar 4.10. Tabel Ujian.....	38
Gambar 4.11. Tabel Ujian Soal.....	38
Gambar 4.12. <i>Script Case Folding</i>	39
Gambar 4.13. <i>Script Tokenizing</i>	39
Gambar 4.14. <i>Script Filtering</i>	39
Gambar 4.15. <i>Script Stemming</i>	39
Gambar 4.16. <i>Analizing</i>	40
Gambar 4.17. <i>Similarity</i>	40
Gambar 4.18.1. <i>Script Stemming</i>	41
Gambar 4.18.2. <i>Script Stemming</i>	42
Gambar 4.18.3. <i>Script Stemming</i>	43
Gambar 4.18.4. <i>Script Stemming</i>	44
Gambar 4.19. Halaman Login Siswa	45
Gambar 4.20. Halaman Login Guru.....	46
Gambar 4.21 Login Admin	46
Gambar 4.22. Halaman Utama Siswa	47
Gambar 4.23. Daftar Ujian.....	47
Gambar 4.24. Hasil Ujian.....	48
Gambar 4.25. Halaman Utama Guru.....	48
Gambar 4.26. Data Ujian	49
Gambar 4.27. Data Soal Ujian	49
Gambar 4.26. Data Ujian	50
Gambar 4.27. Data Soal Ujian	50
Gambar 4.28. Hasil Ujian.....	50
Gambar 4.29. Halaman Utama Admin.....	51
Gambar 4.30. Halaman Data Siswa	51
Gambar 4.31. Halaman Data Guru.....	51
Gambar 4.32. Data Mata Pelajaran	52
Gambar 4.33. Data Kelas	52
Gambar 4.34. Data Kamus	53

Gambar 5.1. Input Data Ujian	65
Gambar 5.2. Soal dan Kunci Jawaban Yang Sudah Di Input	65
Gambar 5.3. Daftra Ujian.....	66
Gambar 5.4. Hasil Ujian.....	66
Gambar 5.5. Hasil Ujian Pada Guru.....	66
Gambar 5.6. Detail Ujian	67

DAFTAR TABEL

Tabel 2.1 Kombinasi Awalan Akhiran Yang Tidak Diizinkan.....	8
Tabel 2.2 Aturan Pemenggalan Awalan	9
Tabel 2.3 Potongan Dan Jumlah Kata Dari Kalimat 1 Dan 2	11
Tabel 2.4 Vektor Berdasarkan Aturan Metode Cosine Similarity	11
Tabel 2.5 Kalimat 1 Dan 2 Yang Diubah Berdasarkan	12
Tabel 2.6 Kata-Kata Kalimat 1 Dan 2 Yang Telah Diubah Berdasarkan Aturan Algoritma Nazief & Adriani	12
Tabel 2.7 Vektor Berdasarkan Aturan Metode Cosine Similarity	13
Tabel 3.1 Case Folding	17
Tabel 3.2 Tokenizing	17
Tabel 3.3 Filtering.....	18
Tabel 3.4 Hasil stemming algoritma Nazief & Adriani	18
Tabel 3.5 Hasil stemming algoritma Nazief & Adriani	18
Tabel 3.6 Struktur tabel kelas.....	23
Tabel 3.7 Struktur tabel siswa.....	24
Tabel 3.8 Struktur tabel mapel	24
Tabel 3.9 Struktur tabel guru	24
Tabel 3.10 Struktur tabel ujian.....	25
Tabel 3.11 Struktur tabel soal ujian	25
Tabel 3.12 Struktur tabel siswa ujian.....	25
Tabel 3.13 Struktur tabel jawaban siswa.....	25
Tabel 3.14 Struktur tabel hasil ujian	26
Tabel 3.15 Struktur tabel kamus	26
Tabel 3.16 Struktur tabel admin.....	26
Tabel 4.1 Pengujian Hak Akses Admin, Siswa dan Guru	53
Tabel 4.2. Pengujian Koreksi Jawaban Essay.....	55
Tabel 5.1 <i>Case Folding</i> soal 1.....	67
Tabel 5.2 <i>Tokenizing</i> soal 1.....	68
Tabel 5.3 <i>Filtering</i> soal 1	68
Tabel 5.4 <i>Stemming</i> soal 1	68
Tabel 5.5 <i>Analizing</i> soal 1.....	68
Tabel 5.6 <i>Case Folding</i> soal 2.....	69
Tabel 5.7 <i>Tokenizing</i> soal 2.....	69
Tabel 5.8 <i>Filtering</i> soal 2.....	70

Tabel 5.9 <i>Stemming</i> soal 2	70
Tabel 5.10 <i>Analizing</i> soal 2.....	70
Tabel 5.11 <i>Case Folding</i> soal 3.....	71
Tabel 5.12 <i>Tokenzing</i> soal 3.....	71
Tabel 5.13 <i>Filtering</i> soal 3	71
Tabel 5.14 <i>Stemming</i> soal 3	71
Tabel 5.15 <i>Analizingg</i> soal 3.....	72
Tabel 5.16 <i>Case Folding</i> soal 4.....	72
Tabel 5.17 <i>Tokenzing</i> soal 4.....	73
Tabel 5.18 <i>Filtering</i> soal 4.....	73
Tabel 5.19 <i>Stemming</i> soal 4	73
Tabel 5.20. <i>Analizing</i> soal 4.....	73
Tabel 5.21. Tabel Nilai Rata-Rata Kemiripan Jawaban Siswa Dengan Metode <i>Cosine Similarity</i>	74

ABSTRAK

SISTEM KOREKSI JAWABAN *ESSAY* MENGGUNAKAN ALGORITMA NAZIEF & ADRIANI DI SMA NEGERI 1 MAGELANG

Oleh : Catur Mulyawan
Pembimbing : 1. Andi Widiyanto, M.Kom
2. Endah Ratna Arumi, M.Cs

Sistem koreksi jawaban *essay* merupakan sistem yang dirancang untuk memudahkan guru untuk mengoreksi jawaban *essay*. *Text mining* adalah proses menganalisis teks untuk mengekstrak informasi yang berguna untuk tujuan tertentu. *Cosine Similarity* adalah salah satu metode dalam menentukan nilai kemiripan antar dua objek. *Cosine Similarity* dapat diterapkan dalam menentukan nilai kemiripan pada dua berkas dokumen teks. Tahap-tahap pada *text mining* yang di gunakan pada metode *Cosine Similarity* terdiri dari *case folding*, *tokenizing*, *filtering*, *stemming*, dan *analyzing*. Tujuan penelitian ini adalah membangun sebuah sistem yang dapat digunakan untuk mendeteksi kesamaan atau membandingkan antara kunci jawaban dengan jawaban, yang berbentuk dokumen teks dan mampu memberikan nilai kemiripan yang dimiliki dua atau lebih dokumen teks. Hasil akhir dalam koreksi jawaban *essay* adalah akan menampilkan soal secara acak, kesamaan antara kunci jawaban dengan jawaban, serta nilai kemiripan *teks*. Kesamaan kata jawaban siswa dengan kunci jawaban memiliki deteksi kemiripan yang baik yaitu dengan nilai rata-rata kemiripan sebesar 94,291%. Berdasarkan hasil tersebut dapat disimpulkan bahwa sistem koreksi jawaban *essay* ini berguna bagi guru untuk membantu koreksi jawaban *essay*.

Kata kunci: Sistem Koreksi, *Text mining*, *Cosine Similarity*

ABSTRACT

THE ESSAY ANSWER CORRECTION SYSTEM AT SMA NEGERI 1 MAGELANG USE NAZIEF & ADRIANI ALGORITHM

By : Catur Mulyawan
Menthor : 3. Andi Widiyanto, M.Kom
4. Endah Ratna Arumi, M.Cs

The essay answer correction system is a system designed to correct essay answers have by teachers. Text mining is the process of text analyzing to extract useful information for a particular purpose. Cosine Similarity is a methode in determinibg the similarity 2 objects. The steps of text mining used in the Cosine Similarity Consist of case folding, tokenizing, filtering, stemming, dan analyzing. The purpose of this research is to build a system that can be used to detect the similarity or compare between the key answers and the answer made by the student, which is in the form of a text document and able to give the value of similarity of two or more text documents. The final result in essay's answer correction is display random questions, similarities between key answers and the answers, it as will as the value of text similarities. The similarity between the students' answersd the kay answers gas good similarity rate, that is it has the average of similarity value of 94,291%. Based on this results it can be concluded that the essay answer correction system is useful for teachers to help correct essay answers.

Key words: *correct system, Text mining, Cosine Similarity*

BAB I

PENDAHULUAN

A. Latar Belakang Permasalahan

Teknologi saat ini berkembang menjadi semakin canggih. Dapat mengubah berkas yang awalnya harus ditulis atau diketik ke dalam lembaran kertas menjadi sebuah teknologi berupa data komputer yang disebut sebagai dokumen digital yang bisa dicetak untuk keperluan tertentu atau digunakan sebagai data yang akan digunakan di masa yang akan datang.

Dengan perkembangan teknologi saat ini dapat menimbulkan dampak negatif dan positif. Salah satu dampak negatif yaitu plagiatisme, plagiat adalah pengambilan karangan (pendapat dan sebagainya) orang lain dan menjadikan seolah-olah menjadikan karangan (pendapat) sendiri (Purwani Istiana, Purwoko, 2013). Salah satu dampak positif perkembangan teknologi adalah E-learning untuk mempermudah proses pengajaran dan pembelajaran. *E-learning* adalah sebagai sembarang pengajaran dan pembelajaran yang menggunakan rangkaian elektronik (LAN, WAN, atau internet) untuk menyampaikan isi pembelajaran, interaksi, atau bimbingan (Jaya Kumar C. Koran, 2002). SMA Negeri 1 Magelang adalah salah satu sekolah yang sudah menggunakan *e-learning* dalam proses pembelajarannya.

Dalam proses pembelajaran tentunya akan ada ujian untuk mengetahui hasil dari proses pembelajaran yang telah di sampaikan, untuk mengetahui sejauh mana siswa dapat memahami materi yang di sampaikan. Banyaknya tanggung jawab dari guru, terkadang guru kekurangan waktu untuk mengoreksi hasil dari ujian yang telah di lakukan, hal itu berdampak pada keterlambatan nilai hasil ujian keluar. Dengan keterlambatan nilai hasil ujian keluar, dapat merugikan siswa atau peserta didik jika ada nilai mata pelajaran yang nilainya kurang baik akan tetapi nilai tersebut terlambat keluar, maka siswa atau peserta didik akan kekurangan waktu belajar untuk

mempersiapkan ujian ulang pada mata pelajaran yang memiliki nilai jelek. Sehingga hasil ujian ulang tidak jauh beda dengan hasil ujian. Beberapa pengajar masih menggunakan cara membandingkan dua atau lebih berkas dokumen dan menelusuri paragraf per paragraf dan menentukan kata, kalimat atau paragraf yang memiliki kemiripan untuk menentukan kemiripan suatu dokumen teks. Cara tersebut merupakan cara yang kurang efektif dan efisien. Oleh, sebab itu diperlukan adanya aplikasi yang mampu mendeteksi kemiripan pada dokumen teks.

Text *mining* adalah proses menganalisis teks untuk mengekstrak informasi yang berguna untuk tujuan tertentu . Tahap-tahap pada text mining secara umum terdiri dari *case folding, tokenizing, filtering, stemming, tagging dan analyzing*. Algoritma Nazief & Adriani adalah salah satu algoritma yang digunakan dalam tahap *stemming*. Algoritma Nazief & Adriani merupakan algoritma untuk mengubah kata yang memiliki *sufiks, prefiks dan konfiks* menjadi bentuk kata dasar. Algoritma Nazief & Adriani digunakan dalam penelitian ini sebagai algoritma pendukung dalam proses penentuan nilai kemiripan pada dokumen teks. Terdapat beberapa metode dalam menentukan kemiripan antar dua objek. *Cosine Similarity* adalah salah satu metode dalam menentukan nilai kemiripan antar dua objek. Salah satu contoh penerapan adalah penentuan kemiripan pada sidik jari manusia. *Cosine Similarity* dapat diterapkan dalam menentukan nilai kemiripan pada dua berkas dokumen teks. Parameter yang digunakan adalah jumlah kata-kata pada dua dokumen teks yang dibandingkan. *Cosine Similarity* menggunakan dua vektor yang mempresentasikan dua dokumen teks dimana nilai sudut kosinus dari kedua vektor tersebut adalah nilai kemiripan dari dua dokumen teks tersebut. Batasan nilai yang dihasilkan mulai dari 0 sampai dengan 1. Komposisi kata pada dokumen teks merupakan penentuan nilai yang diperoleh dari metode *Cosine Similarity*. Metode *Cosine Similarity* tidak dapat menentukan kesamaan makna dari setiap kata. Setiap kata yang memiliki komponen huruf yang berbeda dianggap kata yang berbeda. Algoritma Nazief & Adriani dapat

mengenal kesamaan makna dari setiap kata, dimana kata yang memiliki sufiks, prefiks dan konfiks diubah menjadi bentuk dasar.

Berdasarkan uraian diatas maka perlu dilakukan suatu penelitian dalam hal penyelesaian masalah pengajar dalam mengoreksi jawaban dalam bentuk *essay*, dengan membuat aplikasi berbasis *web* dengan memanfaatkan metode *Cosine Similarity* dan algoritma Nazief & Adriani, maka diangkatlah sebuah judul skripsi “SISTEM KOREKSI JAWABAN *ESSAY* DI SMA NEGERI 1 MAGELANG”

B. Rumusan Masalah

Dari uraian latar belakang diatas, maka dapat dirumuskan bahwa masalah yang melatarbelakangi penelitian ini adalah bagaimana membangun sebuah sistem dengan memanfaatkan algoritma Nazief dan Adriani dengan metode *Cosine Similarity* sebagai cara untuk menyelesaikan masalah pengajar dalam mengoreksi sebuah jawaban *essay* dengan membandingkan antara jawaban dengan kunci jawaban

C. Tujuan Penelitian

Tujuan penelitian ini adalah membangun sebuah sistem yang dapat digunakan untuk mendeteksi kesamaan atau membandingkan antara kunci jawaban dengan jawaban, yang berbentuk dokumen teks dan mampu memberikan nilai kemiripan yang dimiliki dua atau lebih dokumen teks.

D. Manfaat Penelitian

Manfaat dari penelitian setelah tujuan penelitian tercapai adalah untuk membantu pengajar dalam mengoreksi jawaban sehingga pengajar tidak memerlukan waktu lama untuk mengoreksi jawaban dan dapat menunjukkan letak kesamaan antara dokumen *essay* dengan jawabannya.

BAB II

TINJAUAN PUSTAKA

A. Penelitian Relevan

1. Penelitian yang dilakukan oleh Dewa Ayu Rai Ariantini, Arie S. M. Lumenta, Agustinus Jacobus (2016) Universitas Sam Ratulangi Manado dalam penelitiannya yang berjudul “*Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity*” tujuan dari penelitian ini adalah membangun sebuah sistem untuk mengukur tingkat kemiripan antar dokumen teks bahasa indonesia. Penelitian ini menggunakan metode *Cosine Similarity*. Tahap pengujian sistem dilakukan untuk evaluasi perangkat lunak yang dibangun dengan hasil yang diharapkan, dapat dilakukan dengan cara membandingkan nilai hasil prediksi dan actual, nilai prediksi diperoleh dari sistem, kemudian nilai actual diperoleh dari nilai yang diperiksa secara manual, pada sistem deteksi kemiripan dokumen sudah berhasil diterapkan dan sudah bisa digunakan untuk mengukur tingkat kemiripan dokumen.
2. Penelitian yang dilakukan oleh Meiyanto Eko Sulisty, Ristu Saptono, Adam Asshidiq (2015) Universitas Sebelas Maret, Surakarta dalam penelitiannya yang berjudul “*Penilaian Ujian Bertipe Essay Menggunakan Metode Text Similarity*” mengatakan bahwa : Penelitian ini berfokus pada pembuatan ekstensi penilaian otomatis untuk ujian bertipe uraian. Hasil penelitian ini adalah nilai akhir hasil penilaian menggunakan sistem text similarity tidak berbeda signifikan dengan hasil penilaian dari pakar. Namun, terdapat beberapa tipe soal yang hasilnya berbeda signifikan dikarenakan terdapat karakter unik yang terrekam di dalam database serta jawaban pada kunci tidak mengandung keyword yang sesuai dengan relevansi jawaban yang benar.
3. Penelitian yang dilakukan oleh Dheru Alam Perkasa, Eki Saputra, Mona Fronita (2015) Universitas Islam Negeri Sultan Syarif Riau dalam penelitiannya yang berjudul “*Sistem Ujian Online Essay Dengan Penilaian Menggunakan Metode Latent Semantic Analysis (LSA)*” mengatakan bahwa

: Ujian merupakan salah satu cara untuk mengevaluasi proses belajar. Ujian dalam bentuk essay dinilai merupakan metode yang tepat untuk mengukur kemampuan mahasiswa untuk mengingat, mengorganisasikan, dan mengekspresikan gagasan yang dimiliki. Sayangnya, penilaian essay membutuhkan waktu yang lama karena pemeriksaan jawaban dilakukan secara manual satu persatu.

Berdasarkan ketiga penelitian tersebut telah dianggap relevan sebagai acuan untuk perancangan sistem koreksi akan dibuat. Dari ketiga penelitian di atas terdapat kesamaan dalam hal metode dan objek penelitian yang digunakan dengan metode penelitian yang sedang penulis lakukan yaitu membangun sebuah sistem koreksi jawaban *essay* dengan metode cosine similarity.

B. Variabel Penelitian

1. Sistem

Sistem bisa ditafsirkan sebagai kesatuan elemen yang memiliki keterkaitan. Beberapa elemen dapat digabung menjadi suatu unit, kelompok atau komponen sistem dengan fungsi tertentu. Komponen sistem ini bisa dilihat, dianggap atau memang dirancang untuk berfungsi mandiri sebagai modul sistem (Taufiq. 2002).

2. Ujian Essay

Essay adalah uraian yang membicarakan bermacam ragam, tidak tersusun secara teratur tetapi seperti dipetik dari bermacam jalan pikiran. Dalam esai terlihat keinginan, sikap terhadap soal yang dibicarakan, kadang-kadang terhadap soal yang dibicarakan. Pengertian esai sebagai karangan yang sedang panjangnya, yang membahas persoalan secara mudah dan sepiantas lalu dalam bentuk prosa (Menurut H.B Jassin).

Ujian esai adalah tes yang disusun dalam bentuk pertanyaan terstruktur dan peserta didik menyusun, mengorganisasikan sendiri jawaban tiap pertanyaan.

3. Cosine Similarity

Cosine similarity (cosine) adalah ukuran kesamaan diantara dua vektor dari sebuah *inner product space* yang mengukur kosinus dari sudut diantara dua vektor tersebut. (Michelle Deza & Elena Deza, 2009)

Cosine similarity digunakan dalam ruang positif, dimana hasilnya dibatasi antara nilai 0 dan 1. Jika nilainya 1 maka dokumen tersebut dikatakan mirip dan jika hasilnya 0 maka nilai tersebut dikatakan tidak mirip. Perhatikan bahwa batas ini berlaku untuk sejumlah dimensi, dan *Cosine similarity* ini paling sering digunakan dalam ruang positif dimensi tinggi. Misalnya, dalam *Information Retrieval*, masing-masing kata/istilah (term) diasumsikan sebagai dimensi yang berbeda dan dokumen ditandai dengan vector dimana nilai masing-masing dimensi sesuai dengan berapa istilah muncul dalam dokumen. Berikut adalah rumus *cosine similarity*

$$\text{Similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (1)$$

Keterangan : A = vektor
 B = vektor
 A_i = bobot term i dalam blok A_i
 B_i = bobot term i dalam blok B_i
 i = jumlah term dalam kalimat
 n = jumlah vektor

Dimana A merupakan bobot setiap ciri pada vector A, dan B merupakan bobot setiap ciri pada B. jika dikaitkan dengan *information retrieval* maka A adalah bobot setiap istilah pada dokumen A, dan B merupakan bobot setiap istilah pada dokumen B. Pada penelitian ini digunakan *cosine similarity* karena pengukuran kemiripan dapat dilakukan dengan membandingkan dokumen 1 dengan dokumen 2 kemudian sistem akan menghitung nilai kemiripan. A_i.B_i adalah nilai yang diperoleh dari term A dan term B kemudian kedua nilai tersebut dijumlahkan, kemudian nilai A_i² semua nilai term dokumen A semua

nilainya dipagkatkan dua, begitu juga dengan term Bi2 semua nilai yang diperoleh dipagkatkan dua kemudian semua nilai yang diperoleh dijumlahkan

4. Teks Mining

Text mining memiliki definisi menambang data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen (Harlian,2006).

a. *Case Folding dan Tokenizing*

Case folding adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf a sampai dengan z yang diterima. Tahap *tokenizing/parsing* adalah tahap pemotongan string input berdasarkan tiap kata yang menyusunnya.

b. *Filtering*

Filtering adalah tahap mengambil kata-kata penting dari hasil *token*.

c. *Stemming*

Stemming adalah proses untuk menggabungkan atau memecahkan setiap varian-varian suatu kata menjadi kata dasar. Terdapat beberapa algoritma dalam stemming, salah satunya adalah algoritma Nazief & Adriani.

d. *Analyzing*

Tahap *Analyzing* merupakan tahap penentuan seberapa jauh kemiripan antar dokumen teks

5. Algoritma Nazief & Adriani

Konjungsi adalah Algoritma stemming Nazief dan Adriani dikembangkan berdasarkan aturan morfologi Bahasa Indonesia yang mengelompokkan imbuhan menjadi awalan (*prefix*), sisipan (*infix*), akhiran (*suffix*) dan gabungan awalan akhiran (*confixes*).

Algoritma Nazief & Adriani memiliki tahap-tahap sebagai berikut:

- a. Cari kata dalam kamus jika ditemukan maka diasumsikan bahwa kata tersebut adalah kata dasar. Algoritma berhenti. Jika tidak ditemukan maka lakukan langkah 2.
- b. Hilangkan *inflectional suffixes* bila ada. Dimulai dari *inflectional particle* (“-lah”, “-kah”, “-tah” dan “-pun”), kemudian *possessive pronoun* (“-ku”, “-mu” dan “-nya”). Cari kata pada kamus jika ditemukan algoritma berhenti, jika kata tidak ditemukan dalam kamus lakukan langkah 3.
- c. Hilangkan *derivation suffixes* (“-an”, “-i” dan “-kan”). Jika akhiran “-an” dihapus dan ditemukan akhiran “-k”, maka akhiran “-k” dihapus. Cari kata pada kamus jika ditemukan algoritma berhenti. jika kata tidak ditemukan maka lakukan langkah 4.
- d. Pada langkah 4 terdapat tiga iterasi. Iterasi berhenti jika :
Ditemukannya kombinasi awalan yang tidak diizinkan berdasarkan awalan

1) Iterasi berhenti jika :

- a) Ditemukannya kombinasi awalan yang tidak diizinkan berdasarkan awalan

Tabel 2.1 Kombinasi Awalan Akhiran Yang Tidak Diizinkan

AWALAN	AKIRAN YANG TIDAK DI IJINKAN
be-	-i
di-	-an
ke-	-i, -kan
me-	-an
se-	-i, -kan

- b) Awalan yang dideteksi saat ini sama dengan awalan yang dihilangkan sebelumnya.

c) Tiga awalan telah dihilangkan.

- 2) Identifikasikan tipe awalan dan hilangkan. Awalan terdiri dari dua tipe:

- a) Standar (“di-“, “ke-“, “se-”) yang dapat langsung dihilangkan dari kata.
- b) Kompleks (“me-“, “be-“, “pe-“, “te”) adalah tipe-tipe awalan yang dapat bermorfologi sesuai kata dasar yang mengikutinya. Oleh karena itu dibutuhkan aturan untuk mendapatkan hasil pemenggalan yang tepat.

Tabel 2.2 Aturan Pemenggalan Awalan

ATURAN	FORMAT KATA	PEMENGGALAN
1	berV...	ber-V ... be-rV
2	berCAP...	ber-CAP... dimana C != 'r' & P != 'er'
3	BerCAerV	ber-CaerV... dimana C != 'r'
4	Belajar	bel-ajar
5	berC1erC2...	be-C1erC2... dimana C1 != 'r' 'l'
6	terV...	ter-V... te-rV...
7	terCerV...	ter-CerV dimana C != 'r'
8	terCP...	Ter-CP... dimana C != 'r' dan P != 'er'
9	teC1erC2...	Te-C1erC2... dimana C1 != 'r'
10	me{l r w y}V...	me - {l r w y} V...
11	mem{b f v}...	mem-{b f v}...
12	Mempe	mem-pe...
13	mem{rV V}...	me-m{rV V}... me-p{rV V}
14	men{c d j s z}...	men-{c d j s z}...
15	menV...	me-nV... me-tV
16	meng{g h q k}...	meng-{g h q k}...
17	mengV...	meng-V... meng-kV... mengV-... jika V='e'
18	menyV...	meny-sV....
19	mempA...	mem-pA... dimana A != 'e'
20	pe{w y}V...	pe-{w y}V...
21	perV...	per-V... pe-rV...

Tabel 2.2. Aturan Pemenggalan Awalan (2)

ATURAN	FORMAT KATA	PEMENGKALAN
22	perCAP...	per-CAP... dimana C != 'r' dan P != 'er'
23	perCAerV...	per-CAerV... dimana C != 'r'
24	pem{b f V}...	pem-{b f V}...
25	pem{rV V}...	pe-m{rV V}... pe-p{rV V}...
26	pen{c d j z}...	pen-{c d j z}...
27	penV...	pe-nV... pe-tV...
28	pengC...	peng-C...
29	pengV...	peng-V... peng-kV... pengV-... jika V='e'
30	penyV...	peny-sV...
31	pelV...	pe-lV... kecuali "pelajar" yang menghasilkan "ajar"
32	peCerV...	Per-erV ... dimana C!={r w y l m n}
33	PeCP	Pe-CP... dimana C!={r w y l m n} dan P!= 'er'
34	terC1erC2...	ter-C1erC2... dimana C1!= 'r'
35	peC1erC2...	pe-C1erC2... dimana C1!={r w y l m n}

Keterangan simbol huruf

C : huruf konsonan

V : huruf vocal

A : huruf vocal atau konsonan

P : partikel atau fragmen dari setiap kata, misalnya "er"

- c) Cari kata yang telah dihilangkan awalannya. Apabila tidak ditemukan, maka langkah 4 diulang kembali. Apabila ditemukan, maka algoritma berhenti.
- d) Apabila setelah langkah 4 kata dasar masih belum ditemukan, maka proses *recording* dilakukan dengan mengacu pada aturan tabel 2.2 *recording* dilakukan dengan menambahkan karakter *recording* di awal kata yang dipenggal. Pada tabel 2.2, karakter *recording* adalah huruf kecil setelah tanda hubung ('-') dan terkadang berada sebelum tanda kurung. Sebagai

contoh, kata “menangkap” (aturan 15) pada tabel 2.2, setelah dipenggal menjadi “nangkap”. Karena tidak valid, maka recording dilakukan dan menghasilkan kata “tangkap”.

- e) Jika semua langkah gagal, maka input kata yang diuji pada algoritma ini di anggap sebagai kata dasar.

Algoritma Nazief & Adriani pada penelitian ini merupakan algoritma yang digunakan untuk mempengaruhi nilai yang dihasilkan oleh metode *Cosine Similarity*. Contoh, Kalimat 1 yaitu, “Komputer membantu pekerjaan manusia menjadi lebih mudah”. Kalimat 2 yaitu, “Pekerjaan manusia menjadi lebih mudah bila dibantu oleh komputer”.

Tabel 2.3 Potongan Dan Jumlah Kata Dari Kalimat 1 Dan 2

NO	KATA	KALIMAT PERTAMA	KALIMAT KEDUA
1	Bila	0	1
2	Dibantu	0	1
3	Computer	1	1
4	Lebih	1	1
5	Manusia	1	1
6	Membantu	1	0
7	Menjadi	1	1
8	Mudah	1	1
9	Oleh	0	1
10	Pekerjaan	1	1

Tabel 2.3 merupakan kata-kata yang menyusun kalimat pertama dan kedua Metode *Cosine Similarity* menggunakan parameter jumlah kata yang sama dan tidak sama dari kedua kalimat untuk membentuk vektor. Berdasarkan tabel 2.3, vektor yang dihasilkan antara lain:

Tabel 2.4 Vektor Berdasarkan Aturan Metode *Cosine Similarity*

NO	KALIMAT	VEKTOR YANG DIHASILKAN
1	Komputer membantu pekerjaan manusia menjadi lebih mudah	$A = (0,0,1,1,1,1,1,1,0,1)$
2	Pekerjaan manusia menjadi lebih mudah bila dibantu oleh manusia	$B = (1,1,1,1,1,0,1,1,1,1)$

Keterangan

A = Vektor untuk kalimat pertama

B = Vektor untuk kalimat kedua

Berdasarkan vektor yang dihasilkan dari tabel 2.3, maka hasil yang diperoleh dari perhitungan *Cosine Similarity* adalah 0,7559 dan persentase kemiripan dokumen adalah 75,59%. Untuk mengoptimalkan nilai dari metode *Cosine Similarity*, maka algoritma Nazief & Adriani digunakan sebelum proses perhitungan. Algoritma Nazief & Adriani adalah algoritma yang digunakan pada aplikasi untuk mengubah kata menjadi bentuk kata dasar. Berdasarkan aturan dari algoritma ini maka perubahan kalimat pertama dan kedua dapat dilihat pada tabel

Tabel 2.5 Kalimat 1 Dan 2 Yang Diubah Berdasarkan Aturan Algoritma Nazief&Adriani

NO	KALIMAT	BENTUK PERUBAHAN KALIMAT
1	Komputer membantu pekerjaan manusia menjadi lebih mudah	Komputer dapat bantu kerja manusia jadi lebih mudah
2	Pekerjaan manusia manjadi lebih mudah bila dibantu oleh komputer	Kerja manusia jadi lebih mudah bila bantu oleh komputer

Pada tabel 2.5 terdapat dua kata dari kedua kalimat yang diubah menjadi bentuk kata dasar berdasarkan aturan algoritma Nazief &

Adriani. Transformasi kata yang terjadi yaitu “membantu” menjadi “bantu”, “pekerjaan” menjadi “kerja” dan “menjadi” menjadi “jadi”.

Tabel 2.6 Kata-Kata Kalimat 1 Dan 2 Yang Telah Diubah Berdasarkan Aturan Algoritma Nazief & Adriani

NO	KATA	KALIMAT PERTAMA	KALIMAT KEDUA
1	Bantu	1	1
2	Bila	0	1
3	Komputer	1	1
4	Lebih	1	1
5	Manusia	1	1
6	Menjadi	1	0
7	Mudah	1	1
8	Oleh	0	1
9	Kerja	1	1

Tabel 2.5 merupakan kata-kata yang menyusun kalimat pertama dan kedua yang telah diubah berdasarkan aturan algoritma Nazief & Adriani. Berdasarkan tabel 2.5, vektor yang dihasilkan berdasarkan aturan metode Cosine Similarity antara lain:

Tabel 2.7 Vektor Berdasarkan Aturan Metode *Cosine Similarity*

NO	KALIMAT	VEKTOR YANG DI BENTUK
1	Komputer bantu pekerjaan manusia jadi lebih mudah	$A = (1,0,1,1,1,1,1,0,1)$
2	Pekerjaan manusia jadi lebih mudah bilabantu oleh komputer	$B = (1,1,1,1,1,0,1,1,1)$

Keterangan :

A : vektor untuk kalimat pertama

B : vektor untuk kalimat kedua

Berdasarkan vektor yang dihasilkan dari tabel 2.7, hasil yang diperoleh dari perhitungan *Cosine Similarity* adalah 0,8018 dan persentase kemiripan dokumen adalah 80,18%. Selisih dari perhitungan pertama dan kedua adalah 0,0495 dan selisih persentase adalah 4,95%.

Hipotesis yang diperoleh berdasarkan ilustrasi di atas adalah algoritma Nazief & Adriani mampu mempengaruhi nilai kemiripan yang dihasilkan oleh metode *Cosine Similarity*.

6. PHP – MySQL

PHP adalah kependekan dari *Hypertext Preprocessor*, bahasa interpreter yang mempunyai kemiripan dengan C dan Perl.PHP dapat digunakan bersama HTML sehingga memudahkan dalam membuat aplikasi dengan cepat.

MySQL adalah *Relational Database Management System* (RDBMS) yang didistribusikan secara gratis dibawah lisensi GPL (*General Public License*). Dimana setiap orang bebas untuk menggunakan MySQL, namun tidak boleh dijadikan produk turunan yang bersifat *close source* atau komersial. MySQL sebenarnya merupakan turunan salah satu konsep utama dalam *database* sejak lama, yaitu SQL (*Structured Query Language*).SQL adalah sebuah konsep pengoperasian *database*, terutama untuk pemilihan atau seleksi dan pemasukan data, yang memungkinkan pengoperasian data dikerjakan dengan mudah secara otomatis.

C. Landasan Teori

Dengan adanya sistem koreksi jawaban essay guru matapelajaran yang menggunakan bahasa indonesia SMA Negeri 1 Magelang dapat terbantu dalam mengoreksi jawaban essay, karena guru tidak perlu melakukan koreksi jawaban satu persatu atau membandingkan dokumen teks jawaban satu persatu yang memakan waktu lama. Dengan adanya aplikasi ini akan mengurangi resiko keterlambatan nilai keluar sehingga akan berpengaruh terhadap nilai ujian susulan karena siswa memiliki waktu yang lebih lama untuk mempelajari materi.

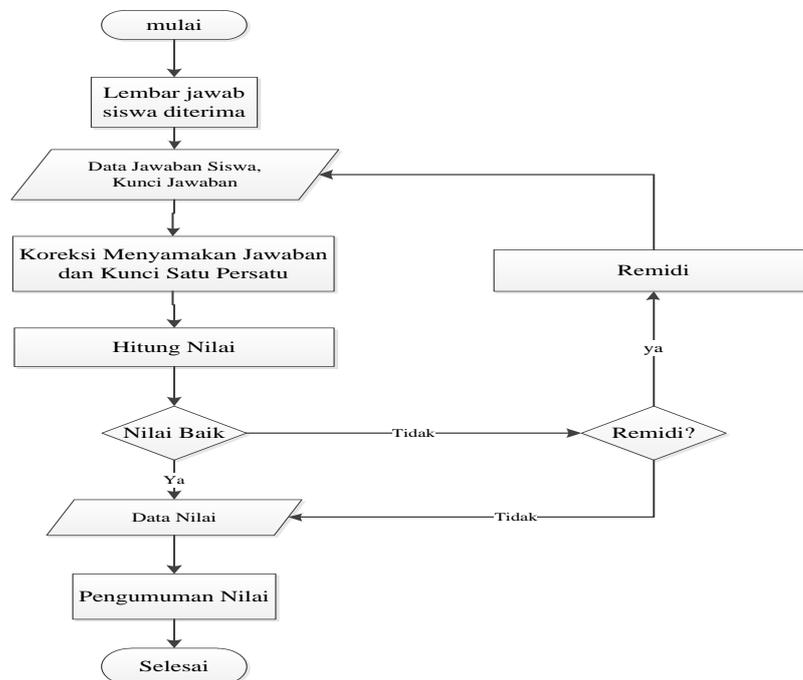
BAB III

ANALISIS DAN PERANCANGAN

A. Analisis Sistem

1. Analisis Sistem Yang Sedang Berjalan

Dari hasil pengamatan dan wawancara yang dilakukan, didapatkan alur koreksi jawaban essay di SMA Negeri 1 Magelang seperti yang digambarkan pada flowchart berikut ini.



Gambar 3.1 Flowchart sistem lama koreksi jawaban essay

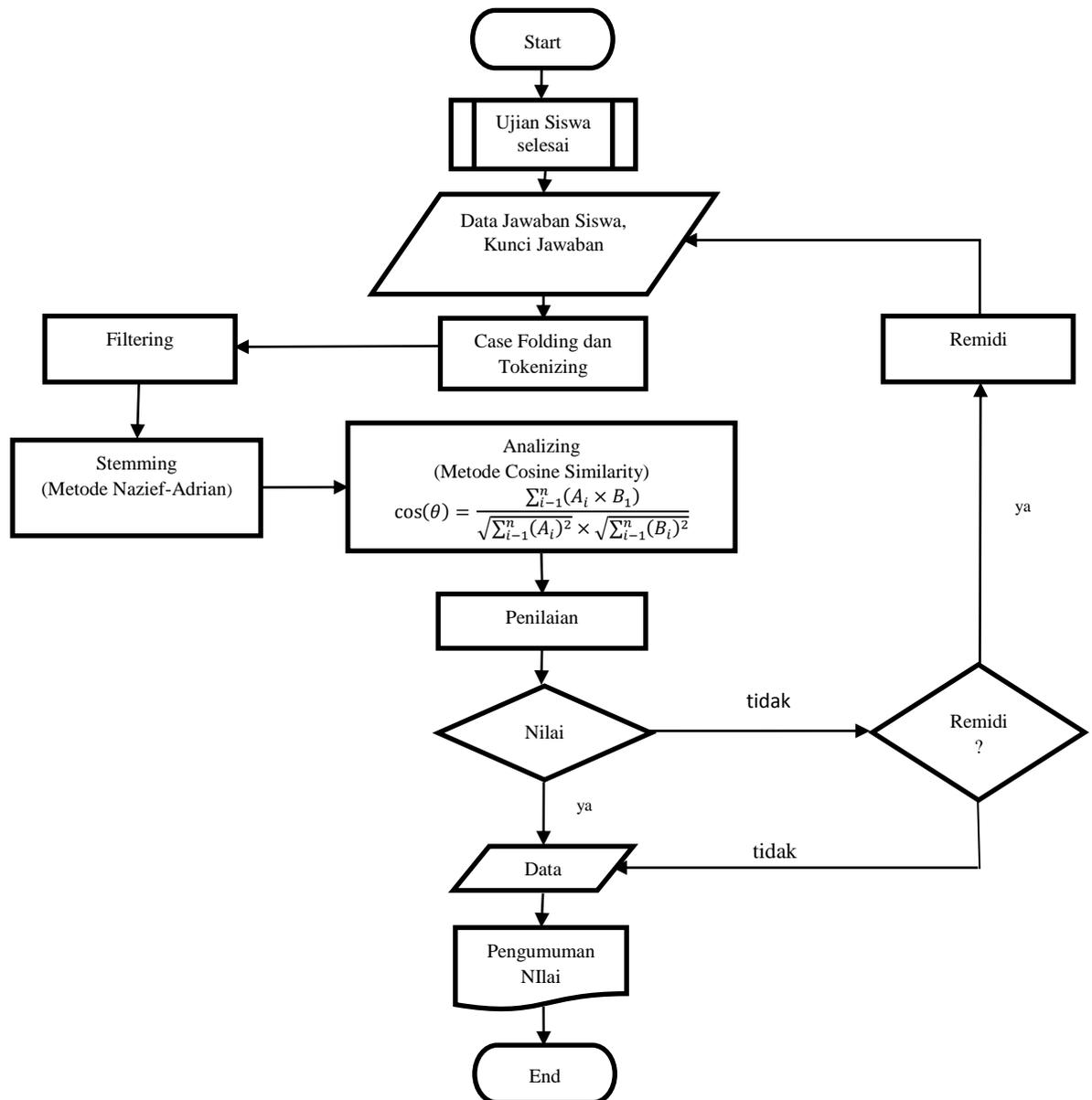
Dari hasil analisis sistem koreksi ujian essay yang menggunakan cara membandingkan dua atau lebih berkas dokumen dan menelusuri paragraph per paragraph yang memiliki kemiripan untuk menentukan kemiripan suatu dokumen dalam hal ini antara jawaban siswa dengan kunci jawaban untuk menentukan jawaban benar atau salah, cara tersebut kurang efektif dan efisien, sehingga ditemukan kendala-kendala. Kendala yang ditemukan yaitu waktu yang digunakan untuk melakukan koreksi cukup lama, dikarenakan banyaknya tanggung jawab guru, terkadang guru

kekurangan waktu untuk mengoreksi hasil ujian yang telah dilakukan sehingga berdampak kepada keterlambatan nilai hasil ujian keluar. Sehingga terkadang tidak sempat untuk dilakukan remidi atau perbaikan nilai siswa.

2. Koreksi Ujian Essay Baru

Sistem koreksi jawaban essay secara otomatis berbasis komputersasi dilakukan dalam beberapa tahap, yaitu tahap *Case Folding* dan *Tekonizing*, *Filtering*, *Stemming* dan *Analyzing*. Flowchart sistem koreksi jawaban essay dapat dilihat pada gambar berikut:

3. Contoh Koreksi Jawaban Essay



Gambar 3.2 Flowchart sistem baru koreksi jawaban essay

Semisal ada sebuah pertanyaan sejarah “Pembentukan Jawa Hokokai dimaksudkan untuk?”.Diketahui jawaban siswa adalah “Menyuruh rakyat mengabdikan diri kepada Jepang” dan diketahui kunci jawabannya adalah “Meminta kepada rakyat untuk mengabdikan diri kepada pemerintah Jepang”

Dari kalimat diatas didapatkan bahwa ada 2 kalimat yang akan dicari prosesntase kesamaannya, yaitu:

Kalimat A: Menyuruh rakyat mengabdikan diri kepada Jepang

Kalimat B :Meminta kepada rakyat untuk mengabdikan diri kepada pemerintah Jepang

a. *Case Folding* dan *Tokenizing*

Case Folding, pada tahap ini dilakukan perubahan huruf ke huruf kecil semua, menjadi:

Tabel 3.1 *Case Folding*

KALIMAT	HASIL CASE FOLDING
A	menyuruh rakyat mengabdikan diri kepada jepang
B	meminta kepada rakyat untuk mengabdikan diri kepada pemerintah jepang

Lalu dilanjutkan ke tahap tokenizing, yaitu memecah kalimat menjadi kata-kata:

Tabel 3.2 *Tokenizing*

KALIMAT	HASIL TOKENIZING
A	menyuruh, rakyat, mengabdikan, diri, kepada,jepang
B	meminta, kepada, rakyat, untuk, mengabdikan, diri, kepada, pemerintah,jepang

b. *Filtering*

Pada tahap ini digunakan untuk menghapus kata-kata yang dianggap tidak berpengaruh kepada inti kalimat, pada penelitian ini menggunakan metode word list, dimana data kata disimpan dalam daftar kata yang nantinya bisa dihapus.

Diketahui dari daftar kata tersebut yang dapat dihilangkan adalah "untuk", maka kata untuk dihilangkan dari hasil *tokenizing*.

Tabel 3.3 *Filtering*

KALIMAT	HASIL <i>FILTERING</i>
A	menyuruh, rakyat, mengabdikan, diri, kepada,jepang
B	meminta, kepada, rakyat,mengabdikan, diri, kepada, pemerintah,jepang

c. *Stemming*

Pada tahap ini adalah memecah varian kata menjadi kata dasar, pada tahap ini digunakan algoritma Nazief & Adriani.

Tabel 3.4 Hasil *stemming* algoritma Nazief & Adriani

KALIMAT	HASIL <i>STEMMING</i>
A	suruh, rakyat,abdi, diri, kepada,jepang
B	minta, kepada, rakyat,abdi, diri, kepada,perintah,jepang

d. *Analizing*

Pada tahap ini digunakan metode Cosine Similarity untuk mencari nilai persamaannya, berikut jumlah kata dari kalimat A dan B.

Tabel 3.5 Hasil *stemming* algoritma Nazief & Adriani

NO	KATA	KALIMAT A	KALIMAT B
1	Abdi	1	1
2	Diri	1	1
3	Jepang	1	1
4	Kepada	1	2
5	Minta	0	1
6	Perintah	0	1
7	Rakyat	1	1
8	Suruh	1	0

Kemudian dicari persamaannya dengan persamaan 2.1

$$Similarity = \frac{(1x1) + (1x1) + (1x1) + (1x2) + (0x1) + (0x1) + (1x1) + (0x1)}{\sqrt{(1^2 + 1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 1^2 + 1^2)} \times \sqrt{(1^2 + 1^2 + 1^2 + 2^2 + 1^2 + 1^2 + 0^2)}}$$

$$\begin{aligned} \text{Similarity} &= \frac{1 + 1 + 1 + 2 + 0 + 0 + 1 + 0}{\sqrt{(1 + 1 + 1 + 1 + 0 + 0 + 1 + 1)} \times \sqrt{(1 + 1 + 1 + 4 + 1 + 1 + 1 + 0)}} \\ \text{Similarity} &= \frac{6}{2.45 \times 3.16} \\ \text{Similarity} &= \frac{6}{7.742} \\ \text{Similarity} &= 0.775 \end{aligned}$$

Hasil yang diperoleh dari perhitungan *Cosine Similarity* adalah 0.775 dan prosentase kemiripan dokumen adalah 77.5%.

e. Penilaian

Untuk mendapatkan hasil akhir nilai siswa menggunakan cara berikut:

Semisal ada 5 soal, dari perhitungan *Cosine Similarity* didapatkan prosentase 77.5%, 80.5%, 90%, 65% dan 78%

$$\text{Nilai} = \frac{77.5 + 80.5 + 90 + 65 + 78}{5} = 78.2.$$

Jadi nilai akhir hasil ujian essay siswa adalah 78.2.

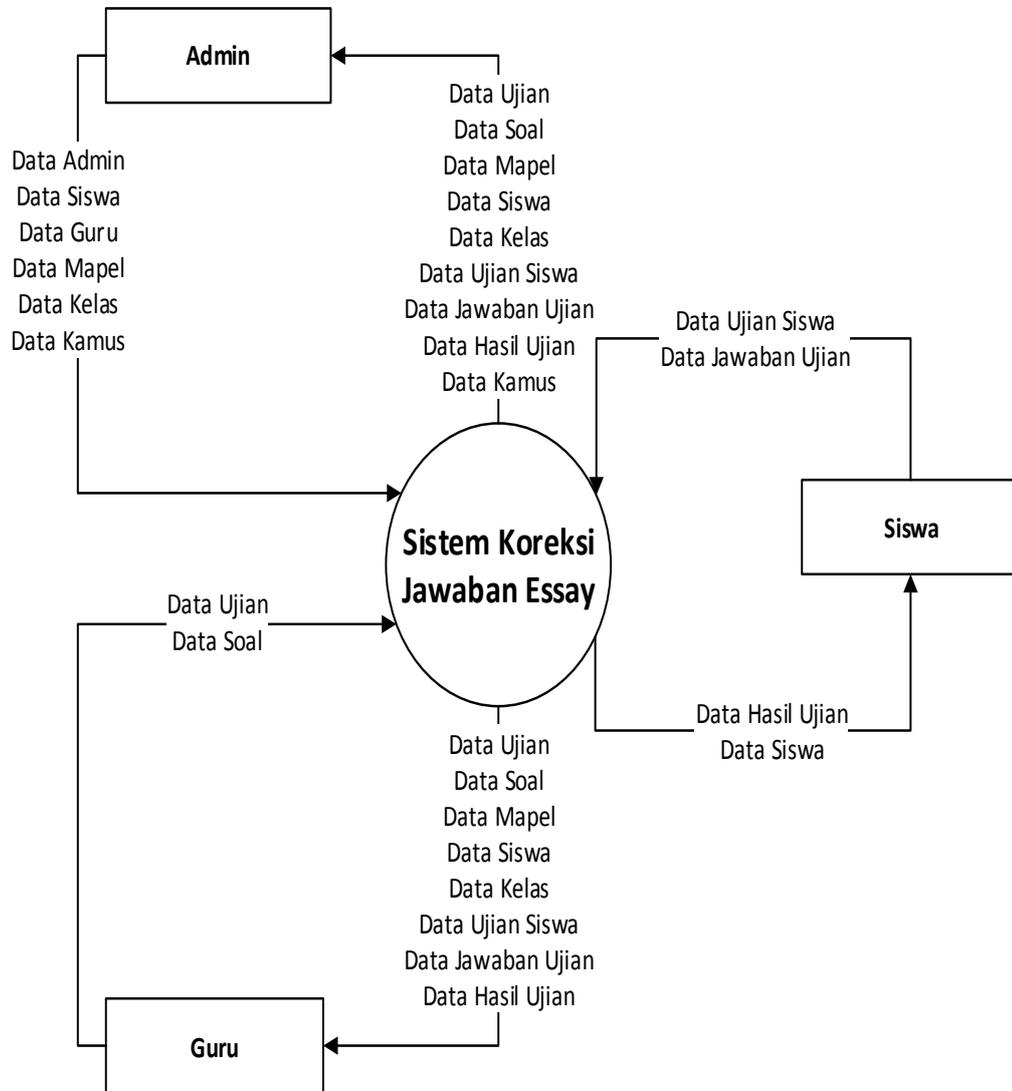
B. Perancangan Sistem

1. DFD

Perancangan sistem pada sistem koreksi jawaban *essay* ini menggunakan DFD. Pembuatan DFD ini meliputi Diagram Kontek dan DFD Level, pada sistem ini akan dipergunakan Diagram Kontek dan DFD Level 0.

a. Diagram Kontek

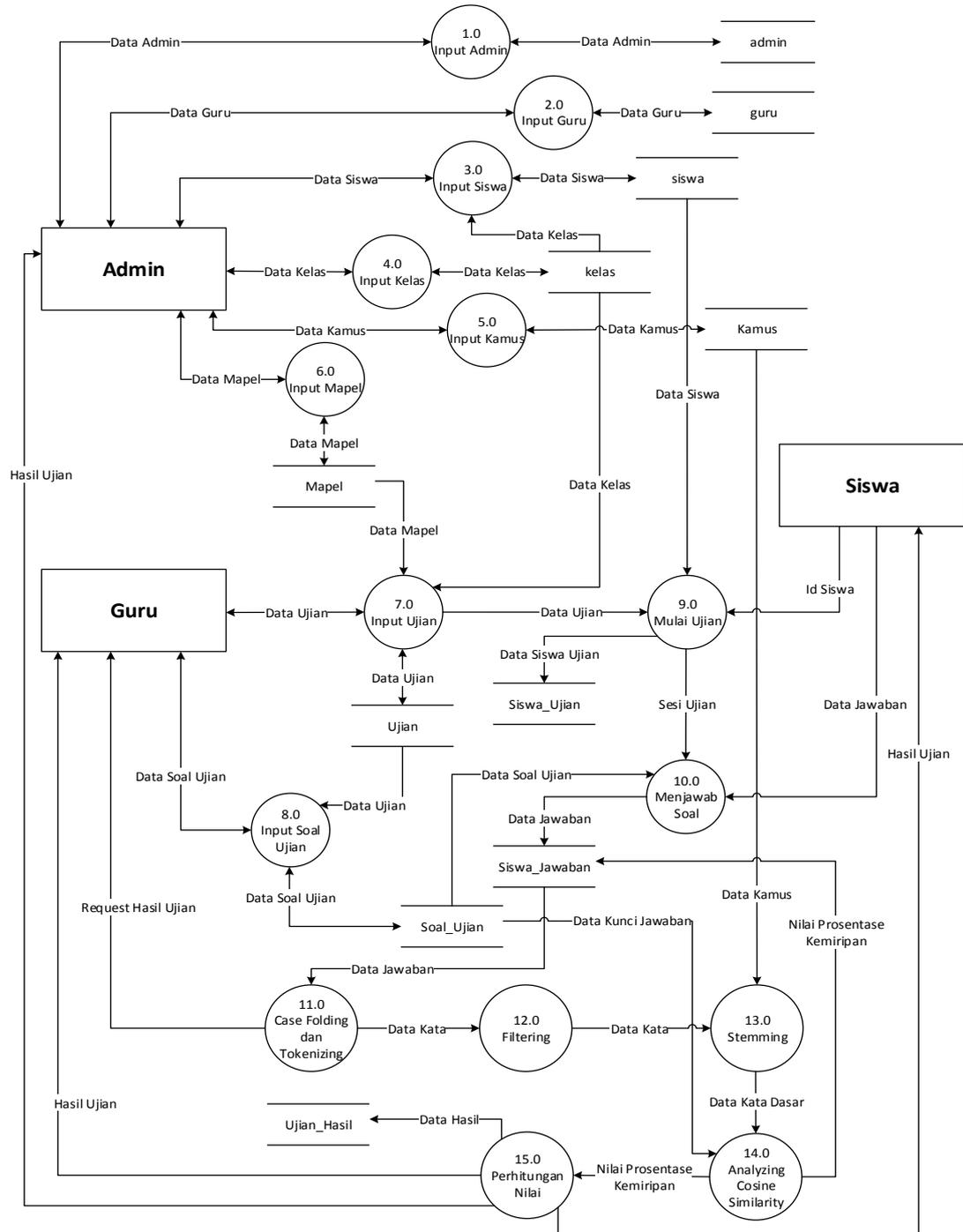
Diagram Kontek pada sistem koreksi jawaban *essay* dapat dilihat pada gambar berikut.



Gambar 3.3 Diagram Kontek

b. DFD Level 0

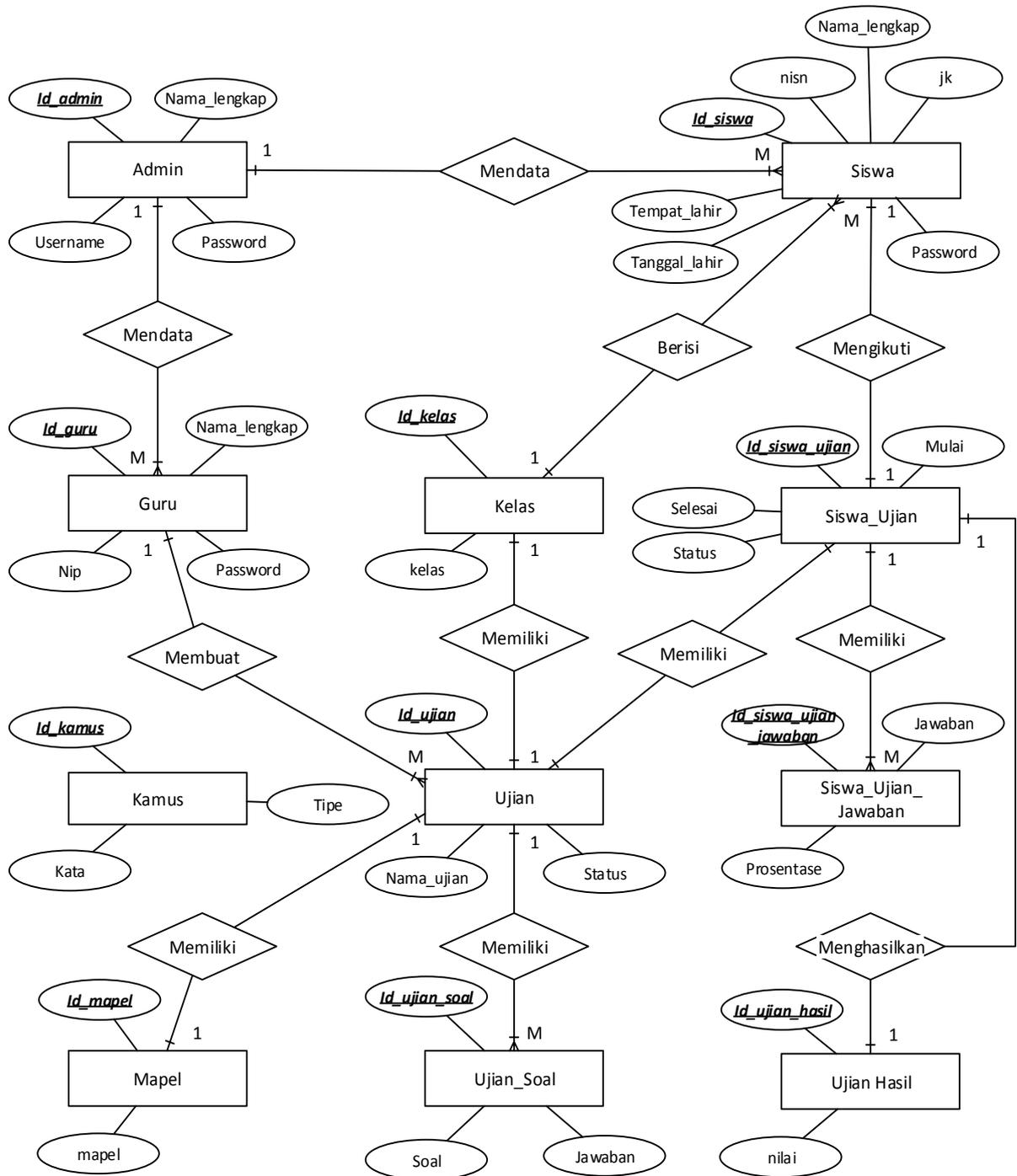
DFD level 0 merupakan diagram yang menggambarkan proses yang terjadi pada sebuah sistem berdasar DFD Kontek.



Gambar 3.4 DFD Level 0

2. ERD

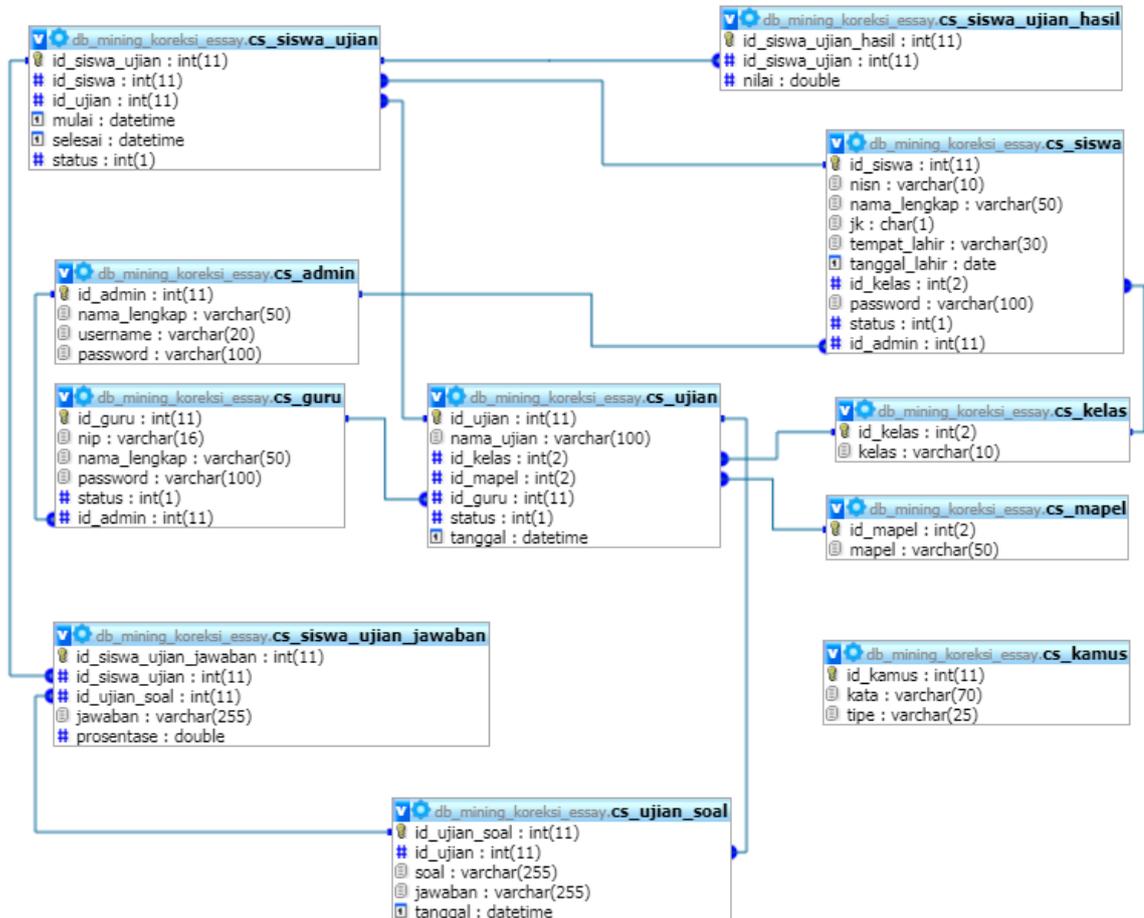
ERD pada sistem koreksi jawaban essay dapat dilihat pada gambar berikut:



Gambar 3.5ERD

3. Perancangan Basis Data

Relasi antar tabel pada basis data sistem koreksi jawaban *essay* dapat dilihat pada gambar dibawah.



Gambar 3.6 Relasi tabel

Berikut adalah struktur tabel-tabel yang terdapat dalam basis data:

a. Struktur tabel jenis

Tabel ini digunakan untuk menyimpan data kelas.

Tabel 3.6 Struktur Tabel Kelas

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_kelas</i>	Int	2	Kunci utama
Kelas	Varchar	10	-

b. Struktur tabel siswa

Tabel ini digunakan untuk menyimpan data siswa.

Tabel 3.7 Struktur Tabel Siswa

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_siswa</i>	Int	11	Kunci utama
Nisn	Varchar	10	-
nama_lengkap	Varchar	30	-
Jk	Char	1	Jenis kelamin (L/P)
tempat_lahir	varchar	30	-
tanggal_lahir	Date	-	-
id_kelas	Int	2	Kelas siswa
Password	varchar	15	-
Status	Int	1	1=aktif, 2=tidak
id_admin	Int	11	-

c. Struktur tabel mapel

Tabel ini digunakan untuk menyimpan data mata pelajaran.

Tabel 3.8 Struktur Tabel Mapel

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_mapel</i>	Int	2	Kunci utama
Mapel	varchar	50	-

d. Struktur tabel guru

Tabel ini digunakan untuk menyimpan data guru.

Tabel 3.9 Struktur Tabel Guru

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_guru</i>	int	11	Kunci utama
Nip	varchar	16	-
nama_lengkap	varchar	30	-
Password	varchar	15	-
Status	int	1	1=aktif, 2=tidak
id_admin	int	11	-

e. Struktur tabel ujian

Tabel ini digunakan untuk data ujian seperti nama ujian, mata pelajaran dan kelas ujian.

Tabel 3.10 Struktur Tabel Ujian

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_ujian</i>	Int	11	Kunci utama
nama_ujian	Varchar	100	-
id_kelas	Int	2	-
id_mapel	Int	2	-
id_guru	Int	11	-
Status	Int	1	1=aktif, 2=tidak
Tanggal	Datetime	-	-

f. Struktur tabel soal ujian

Tabel ini digunakan untuk menyimpan data soal dalam setiap ujian.

Tabel 3.11 Struktur Tabel Soal Ujian

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_ujian_soal</i>	Int	11	Kunci utama
id_ujian	Int	11	-
Soal	Varchar	255	-
Jawaban	Varchar	255	Kunci jawaban
Tanggal	Datetime	-	-

g. Struktur tabel siswa ujian

Tabel ini digunakan untuk menyimpan data siswa saat ujian.

Tabel 3.12 Struktur Tabel Siswa Ujian

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_siswa_ujian</i>	Int	11	Kunci utama
id_siswa	Int	11	-
id_ujian	Int	11	-
Mulai	Datetime	-	Waktu mulai ujian
Selesai	Datetime	-	Waktu selesai ujian
Status	Int	1	1=selesai, 0=belum

h. Struktur tabel jawaban siswa

Tabel ini digunakan untuk menyimpan data jawaban siswa untuk setiap soal.

Tabel 3.13 Struktur Tabel Jawaban Siswa

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_siswa_ujian_jawaban</i>	Int	11	Kunci utama
id_siswa_ujian	int	11	-
id_soal	Int	11	-
Jawaban	Varchar	255	-
Prosentase	double	-	Prosentase kesamaan

i. Struktur tabel hasil ujian

Tabel ini digunakan untuk menyimpan data hasil nilai ujian siswa.

Tabel 3.14 Struktur Tabel Hasil Ujian

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_siswa_ujian_hasil</i>	int	11	Kunci utama
id_siswa_ujian	int	11	-
Nilai	double	-	Hasil akhir

j. Struktur tabel kamus

Tabel ini digunakan untuk menyimpan data kamus kata dasar Bahasa Indonesia yang nantinya digunakan pada saat *stemming*.

Tabel 3.15 Struktur Tabel Kamus

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_kamus</i>	Int	11	Kunci utama
Kata	varchar	70	Kata dasar
Tipe	varchar	25	Tipe kata dasar

k. Struktur tabel admin

Tabel ini digunakan untuk menyimpan data admin.

Tabel 3.16 Struktur Tabel Admin

NAMA KOLOM	TIPE DATA	UKURAN	KETERANGAN
<i>id_admin</i>	Int	11	Kunci utama
nama_lengkap	varchar	30	-
Username	varchar	20	-
Password	varchar	15	-

4. Perancangan Antarmuka

Perancangan antarmuka digunakan untuk mempermudah mendesain halaman pada tahap implementasi.

a. Rancangan halaman admin

Berikut merupakan rancangan halaman pada sistem admin.

1) Rancangan halaman login

Halaman ini ditampilkan pertama kali ketika admin mengakses halaman admin untuk verifikasi dan keamanan sistem, halaman ini juga digunakan oleh guru untuk login.

Login Sistem (Guru | Admin)

Sistem Koreksi Jawaban Essay

Username

Password

Gambar 3.7 Rancangan halaman login

2) Rancangan halaman data siswa

Halaman ini digunakan admin untuk melihat, menambah, merubah dan menghapus data siswa.

Koreksi Jawaban Essay | SMA Negeri 1 Magelang
Admin : Catur Mulya | Logout

[Data Siswa](#) | [Data Guru](#) | [Pengaturan](#)

Data Siswa

NISN Tanggal Lahir

Nama Lengkap Kelas

Jenis Kelamin Status

Tempat Lahir

Kelas

No	NISN	Nama Lengkap	JK	Tempat, Tgl Lahir	Status	Tambah
						Ubah Hapus
						Ubah Hapus
						Ubah Hapus

Gambar 3.8 Rancangan halaman data siswa

3) Rancangan halaman data guru

Halaman ini digunakan admin untuk melihat, menambah, merubah dan menghapus data guru.

Koreksi Jawaban Essay | SMA Negeri 1 Magelang
Admin : Catur Mulya | Logout

Data Siswa | Data Guru | Pengaturan

Data Guru

NIP

Nama

Status ▼

Pencarian

No	NIP	Nama Lengkap	Status	Tambah
				Ubah Hapus
				Ubah Hapus
				Ubah Hapus

Gambar 3.9 Rancangan halaman data guru

4) Rancangan halaman data admin

Halaman ini untuk melihat, menambah, merubah dan menghapus data admin.

Koreksi Jawaban Essay | SMA Negeri 1 Magelang
Admin : Catur Mulya | Logout

Data Siswa | Data Guru | Pengaturan

Data Admin

Nama Lengkap

Username

Password

Ulangi Password

Pencarian

No	Nama Lengkap	Username	Tambah
			Ubah Hapus
			Ubah Hapus
			Ubah Hapus

Gambar 3.10 Rancangan halaman data admin

5) Rancangan halaman data kelas

Halaman ini digunakan untuk mengelola data kelas oleh admin.

Koreksi Jawaban Essay SMA Negeri 1 Magelang		Admin : Catur Mulya Logout
Data Siswa Data Guru Pengaturan		
Data Kelas		
Kelas	<input type="text"/>	<input type="button" value="Simpan"/>
Pencarian	<input type="text"/>	<input type="button" value="Cari"/>
No	Kelas	Tambah
		Ubah Hapus
		Ubah Hapus
		Ubah Hapus

Gambar 3.11 Rancangan halaman data kelas

6) Rancangan halaman data mata pelajaran

Halaman ini digunakan untuk mengelola data mata pelajaran.

Koreksi Jawaban Essay SMA Negeri 1 Magelang		Admin : Catur Mulya Logout
Data Siswa Data Guru Pengaturan		
Data Mata Pelajaran		
Mata Pelajaran	<input type="text"/>	<input type="button" value="Simpan"/>
Pencarian	<input type="text"/>	<input type="button" value="Cari"/>
No	Mata Pelajaran	Tambah
		Ubah Hapus
		Ubah Hapus
		Ubah Hapus

Gambar 3.12 Rancangan halaman data mata pelajaran

b. Rancangan halaman guru

Berikut merupakan rancangan halaman pada guru.

1) Rancangan halaman data ujian

Halaman ini dapat digunakan oleh guru untuk mengelola data ujian (menambah, merubah maupun menghapus dan mengaktifkan atau menonaktifkan ujian).

Koreksi Jawaban Essay | SMA Negeri 1 Magelang
Guru : Catur | Logout

[Data Ujian](#) | [Data Soal Ujian](#) | [Hasil Ujian](#)

Data Ujian

Nama Ujian

Kelas ▼

Mata Pelajaran ▼

Status ▼

Pencarian

No	Nama Ujian	Kelas	Mata Pelajaran	Status	Tambah
					Ubab Hapus
					Ubab Hapus
					Ubab Hapus

Gambar 3.13 Rancangan halaman data ujian

2) Rancangan halaman data soal ujian

Halaman ini digunakan guru untuk mengelola data soal untuk setiap ujian dan memasukkan kunci jawabannya.

Koreksi Jawaban Essay | SMA Negeri 1 Magelang
Guru : Catur | Logout

[Data Ujian](#) | [Data Soal Ujian](#) | [Hasil Ujian](#)

Data Soal Ujian

Nama Ujian

Soal

Kunci Jawaban

Pilih Ujian ▼ Kata Pencarian

No	Soal	Kunci Jawaban	Tambah
			Ubab Hapus
			Ubab Hapus
			Ubab Hapus

Gambar 3.14 Rancangan halaman data soal ujian

3) Rancangan halaman data hasil ujian

Halaman ini digunakan untuk melihat data nilai hasil ujian untuk semua siswa yang mengikuti ujian.

Koreksi Jawaban Essay SMA Negeri 1 Magelang					Guru : Catur Logout	
Data Ujian Data Soal Ujian Hasil Ujian						
Data Hasil Ujian						
Pilih Ujian	<input type="text" value="Nama Ujian"/>	<input type="text" value="Kata Pencarian"/>	<input type="button" value="Tampilkan"/>			
No	NISN	Nama Lengkap	Kelas	Nilai		
					Detail	Hapus
					Detail	Hapus
					Detail	Hapus

Gambar 3.15 Rancangan halaman data hasil ujian

4) Rancangan halaman data detail ujian

Halaman ini untuk melihat data detail jawaban siswa untuk setiap ujian siswa.

Koreksi Jawaban Essay SMA Negeri 1 Magelang					Guru : Catur Logout	
Data Ujian Data Soal Ujian Hasil Ujian						
Data Hasil Ujian Detail						
Nama Ujian	: Ujian Bhs Indo 1					
Kelas	: Kelas 1					
Mata Pelajaran	: Bahasa Indonesia					
NISN	: 1300000221					
Nama Siswa	: Mulyawan					
No	Soal	Kunci Jawaban	Jawaban	Prosentase Kesamaan		
<input type="button" value="Kembali"/>						

Gambar 3.16 Rancangan halaman data detail ujian

c. Rancangan halaman siswa

Berikut merupakan rancangan halaman untuk siswa.

1) Rancangan halaman awal

Halaman ini ditampilkan sebelum siswa melakukan login.

SMA Negeri 1 Magelang | Aplikasi Ujian Essay

Data Ujian | Hasil Ujian

Login Siswa

Aplikasi Ujian Essay

Selamat data di aplikasi ujian essay, pada saat ujian Anda akan dihadapkan pada beberapa Pertanyaan yang akan ditampilkan 1 per satu, disediakan kotak untuk menuliskan jawaban Anda, huruf besar atau kecil tidak akan mempengaruhi jawaban Anda, penilaian sangat berpengaruh pada penggunaan Bahasa Indonesia yang sesuai dengan EYD, silahkan gunakan penulisan sesuai Bahasa Indonesia yang baik dan benar.

Gambar 3.17 Rancangan halaman awal

2) Rancangan halaman pilih data ujian

Halaman ini menampilkan daftar ujian yang dapat diikuti oleh siswa yang kemudian dapat dipilih untuk melakukan ujian.

SMA Negeri 1 Magelang | Aplikasi Ujian Essay

Data Ujian | Hasil Ujian

Data Siswa

1300000321
Mulyawan

[Ganti Password](#)

[Logout](#)

Pilih Data Ujian

No	Nama Ujian	Kelas	Mata Pelajaran	Aksi
				Mulai Ujian
				Mulai Ujian
				Mulai Ujian

Gambar 3.18 Rancangan halaman pilih data ujian

3) Rancangan halaman detail ujian

Halaman ini menampilkan data detail ujian setelah siswa memilih ujian.

SMA Negeri 1 Magelang Aplikasi Ujian Essay	
Data Ujian Hasil Ujian	
Data Siswa 1300000321 Mulyawan Ganti Password Logout	Data Ujian - Mulai Nama Ujian : Ujian Bhs Indo 1 Kelas : Kelas 1 Mata Pelajaran : Bahasa Indonesia Guru Pembuat : Ujian Bhs Indo 1 Jumlah Soal : 5 Waktu : 60 Menit <div style="text-align: center;">Mulai Ujian</div>

Gambar 3.19 Rancangan halaman detail ujian

4) Rancangan halaman ujian

Halaman ini digunakan untuk melakukan ujian bagi siswa.

SMA Negeri 1 Magelang Aplikasi Ujian Essay	
Data Ujian Hasil Ujian	
Data Siswa 1300000321 Mulyawan Ganti Password Logout	Ujian (Nama Ujian) Kelas 1 Bahasa Indonesia 5 Soal Soal 1 Pertanyaan : <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;">Isi pertanyaan disini</div> Jawaban Anda <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;">Tuliskan jawaban disini</div> <div style="text-align: center;">Lanjutkan</div>
Timer Waktu Pengerjaan 00:50	

Gambar 3.20 Rancangan halaman ujian

5) Rancangan halaman hasil ujian

Halaman ini digunakan untuk menampilkan hasil nilai ujian bagi siswa.

SMA Negeri 1 Magelang Aplikasi Ujian Essay				
Data Ujian Hasil Ujian				
Data Siswa 1300000321 Mulyawan Ganti Password Logout		Nilai Hasil Ujian		
No	Nama Ujian	Kelas	Mata Pelajaran	Nilai

Gambar 3.21 Rancangan halaman hasil ujian

BAB VI

PENUTUP

A. Kesimpulan

Berdasarkan hasil pengujian, implementasi serta pembahasan mengenai sistem koreksi jawaban *essay* menggunakan algoritma *Nazief & Adriani* dan metode *Cosine Similarity* yang diterapkan pada sistem koreksi jawaban *essay* yang dilakukan di SMAN 1 kota magelang sistem koreksi ini memiliki tingkat koreksi kesamaan yang baik, kesamaan kata jawaban siswa dengan kunci jawaban memiliki deteksi kemiripan yang baik yaitu dengan nilai rata-rata kemiripan sebesar 94,291%.

B. Saran

Pengembangan sistem koreksi jawaban *esaay* ini diharapkan mampu memberikan manfaat bagi guru dalam mengoreksi jawaban *essay*. Proses update pada kamus perlu dilakukan secara berkala untuk memelihara keakuratan sistem. Adapun saran untuk pengembangan selanjutnya adalah

1. perlu adanya inovasi pengembangan sistem koreksi ini agar sistem koreksi dapat membaca angka, tanda baca dan gambar.
2. Diharapkan sistem yang akan di kembangkan dapat mendeteksi sinonim kata.
3. Diharapkan sistem yang akan di kembangkan dapat mendeteksi negasi.

DAFTAR PUSTAKA

- Arianita, D.A., Lumenta, A.S., & Jacobus, A. 2016. Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity. Teknik Informatika, Universitas Sam Ratulangi, Manado.
- Firdaus, A., Ernawati, & Vatesia, A. 2014. Aplikasi Pendeteksi Kemiripan Pada Dokumen Teks Menggunakan Algoritma Nazief & Adriyani Dan Metode Cosine Similarity. *Skripsi*. Program Studi Teknik Informatika, Fakultas Teknik, Universitas Bengkulu.
- Fitri, Rahim., Asyikin, Arifin Noor. (2015). Aplikasi Penilaian Ujian Essay Otomatis Menggunakan Metode Cosine Similarity, *Poros Teknik*, Volume 7 No.2, ISSN 2442-7764.
- Ariyanti, Dewa A.R., Lumenta, Arie S.M., Jacobus Lumenta. (2016). Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity. *E-Journal Teknik Informatika*. Fitri, Rahim., Asyikin, Arifin Noor. (2015). Aplikasi Penilaian Ujian Essay Otomatis Menggunakan Metode Cosine Similarity, *Poros Teknik*, Volume 7 No.2, ISSN 2442-7764. Volume 9, No 1. ISSN : 2301-8364
- Hamzah, S., Saroso, M., & Santoso, P.B. 2013. Sistem Koreksi Soal Essay Otomatis Dengan Menggunakan Metode Rabin Krap.
- Imbar, R.V., Adelia, Ayub, M., Rehatta, A. 2014. Implementasi Cosine Similarity dan Algoritma Smith-WATERMAN untuk mendeteksi kemiripan teks. Jurusan S1 Sistem Informasi, Jurusan S1 Teknik Informatika, Fakultas Teknik Informatika, Universitas Kristen Maranatha, Bandung.
- Putra, D.A., Sujaini, H., & Pratiwi, H.S. 2015. Implementasi Algoritma Rabin-Krab Untuk Membantu Pendeteksi Plagiat Pada Karya Ilmiah. Program Studi Teknik Informatika, Fakultas Teknik, Universitas Tanjung Pura.
- Syahputa, A.R. 2015. Implementasi Algoritma Winnowing Untuk Deteksi Kemiripan Text. *Skripsi*. Program Studi Teknik Informatika, STMIK Budidarma, Medan.