

Sistem Penilaian Otomatis Jawaban Esai Menggunakan Metode GLSA

Ruslan¹, Gunawan², Suhatati Tjandra³

Program Studi Teknologi Informasi
Sekolah Tinggi Teknik Surabaya
Surabaya

¹alansanmar@gmail.com, ²gunawan@stts.edu, ³tati@stts.edu

Abstrak—Sistem evaluasi merupakan salah satu cara untuk dapat mengukur kemampuan berpikir siswa terkhususnya jawaban esai. Pada umumnya bentuk evaluasi yang digunakan untuk mengukur kemampuan siswa dibagi menjadi 2 (dua) kategori tes yang meliputi: 1) Multiple Choice (pilihan ganda) dan 2) Essay (jawaban uraian). Komputer sebagai sarana penting dalam proses pembelajaran tentunya akan lebih mudah dan akurat dalam melakukan penilaian pada bentuk evaluasi pilihan ganda dan jawaban singkat dibandingkan dengan evaluasi esai dikarenakan keberagaman jawaban siswa. Hal ini pulalah yang mengakibatkan penilaian dan koreksi otomatis pada jawaban esai siswa telah menjadi kajian penelitian di bidang pendidikan sejak dulu hingga kini. Pada dasarnya penelitian ini bertujuan untuk membuat aplikasi yang dapat membantu dalam melakukan penilaian jawaban soal essay secara otomatis melalui pendekatan Algoritma Generalized Latent Semantic Analysis (GLSA). Pada Algoritma ini terdapat 3 (tiga) proses utama yang meliputi preprocessing, GLSA, dan penilaian. Pada preprocessing terdapat beberapa tahapan, seperti: tokenisasi, stopword, stemming, dan term weigh. Sementara pada proses GLSA memiliki metode reduksi dimensi untuk mengambil kata-kata kunci yang dibandingkan dengan jawaban-jawaban kunci yang tersedia. Proses ini menggunakan n-gram yang berupa unigram, bigram, dan trigram sebagai tahapan kombinasi kata pada pembentukan matriks jawaban siswa dan jawaban guru. Dari proses ini, selanjutnya dibandingkan dengan tahapan penilaian melalui cosine similarity untuk memperoleh nilai siswa berdasarkan kesamaan n-gram dari jawaban siswa dan jawaban guru. Selain itu kami juga mengembangkan penelitian kami sebelumnya dengan menambahkan training esai dan melakukan evaluasi dari jawaban siswa berdasarkan data training esai. Simpulan hasil akurasi yang diperoleh adalah rata-rata untuk mata pelajaran Sejarah, Sosiologi dan Pendidikan Kewarganegaran pada jenjang Sekolah Menengah Atas (SMA) mencapai 90,39%.

Kata kunci—Penilaian Otomatis, Latent Semantic Analysis, Generalized Latent Semantic Analysis, Singular Value Decomposition, N-gram.

I. PENDAHULUAN

Sistem penilaian dan koreksi otomatis pada jawaban tes esai (essay test) untuk penilaian hasil belajar siswa telah menjadi kajian penelitian menarik di bidang pendidikan sejak dulu hingga kini [1], [2]. Namun, penelitian penilaian esai secara otomatis untuk bahasa Indonesia masih kurang dan belum terdapat implementasi secara menyeluruh pada semua jenjang pendidikan di Indonesia. Dengan adanya kemajuan teknologi informasi, proses penilaian menjadi bagian yang penting selama

proses pembelajaran dengan tujuan untuk mengevaluasi tingkat pengetahuan dan pemahaman siswa.

Proses evaluasi yang dilakukan untuk menganalisis pengetahuan siswa melibatkan evaluasi dalam bentuk tes esai, dimana tes esai dianggap dapat memberikan wawasan tentang pembelajaran yang lebih dalam dan ketrampilan tingkat tinggi seperti berpikir kritis dan inovatif [3]. Namun, proses penilaian esai tidaklah mudah hal ini dikarenakan guru seringkali masih merasa kesulitan saat melakukan evaluasi pembelajaran khususnya saat melakukan koreksi pada jawaban tes esai.

Proses yang dilakukan guru saat melakukan evaluasi adalah dengan mengoreksi secara manual dan saat pekerjaan koreksi itu tidak selesai maka pekerjaan akan dibawa ke rumah. Bila ini terus dilakukan dan menjadi sebuah rutinitas, tentu saja akan membebani guru dari sisi manajemen waktu. Pekerjaan guru menjadi tidak efektif dan efisien karena waktu yang diperlukan menjadi relatif lama.

Untuk meningkatkan objektivitas dan efisiensi waktu dalam penilaian esai maka sistem penilaian esai langsung terhadap jawaban siswa sangat dibutuhkan. Sistem penilaian esai otomatis telah dikembangkan dan kebanyakan digunakan untuk menilai esai bahasa Inggris dengan menggunakan berbagai algoritma komputer yang dirancang untuk meniru penilaian manusia serta memiliki akurasi 80-93% [1].

Dengan demikian, muncul berbagai teknik penilaian esai otomatis yang menggunakan komputasi linguistik sebagai acuan untuk melakukan penilaian [2]. Teknik yang digunakan dalam melakukan penilaian otomatis antara lain: Latent Semantic Analysis (LSA), Probabilistic Latent Semantic Analysis (PLSA), dan Generalized Latent Semantic Analysis (GLSA) [4], [5], [6]. Pada LSA, kata direpresentasikan dalam sebuah matriks semantik dan kemudian diolah secara matematis menggunakan teknik aljabar linier Singular Value Decomposition (SVD) [7]. Metode ini relatif sederhana, namun memiliki tingkat korelasi yang cukup tinggi bila dibandingkan dengan penilaian yang dilakukan manusia secara manual. Metode LSA hanya menilai kemiripan antar dokumen text melalui frekuensi term yang ada pada masing-masing dokumen text sehingga mempunyai kelemahan yaitu tidak memperhatikan urutan kata atau tata letak term tersebut yang secara tidak langsung berpengaruh terhadap makna yang terkandung pada masing-masing dokumen. Sedangkan pada PLSA merupakan metode yang menggunakan teknik Expectation Maximization yang menilai esai dengan melakukan prosedur berulang secara efisien untuk menghitung

kemungkinan/probabilitas maksimum dalam memperkirakan ada tidaknya kata dari kelompok kata pada matriks dokumen [4]. Teknik berikutnya adalah GLSA [6], [8] yang merupakan pengembangan lebih lanjut dari algoritma LSA dengan membuat n-gram berdasarkan matriks dokumen. GLSA mempertimbangkan urutan kata kalimat dalam dokumen dan menjaga kedekatan kata dalam kalimat. Secara umum GLSA memperbaiki keterbatasan LSA dalam hal pendeteksian kalimat yang memiliki kesalahan sintaksis ataupun kata-kata umum yang hilang [9].

Sistem penilaian esai pada umumnya bekerja dengan menganalisis dokumen jawaban siswa dengan dokumen jawaban guru [10]. Kemudian dokumen jawaban setiap siswa dan guru akan dirubah kedalam bentuk vektor yang direpresentasikan melalui matriks U, matriks S dan matriks V yang mana semakin dekat nilai kedua vektor yang dihasilkan, semakin tinggi nilai yang didapatkan. Kami melakukan penelitian terlebih dahulu menggunakan GLSA dan menghasilkan sebuah sistem penilaian untuk esai berbahasa Indonesia dengan tingkat akurasi yang paling tinggi mencapai 88%. Pada penelitian kami sebelumnya proses GLSA tidak melibatkan training data sehingga hanya menggunakan 1 atau 2 jawaban guru sebagai pedoman evaluasi jawaban siswa.

Pada penelitian ini kami mencoba menggunakan data training dan menggunakan TF-IDF dalam perhitungan term-weightnya serta kami menambahkan perhitungan perbedaan panjang vektor antar dokumen.

II. KAJIAN PUSTAKA

Pada bagian ini kami akan membahas metode yang sudah ada yaitu metode LSA yang kami jadikan sebagai acuan dalam melakukan penelitian selanjutnya. Tahapan dari alur metode LSA secara sederhana dapat dilihat pada gambar 1. Selain itu pada bagian ini kami juga akan membahas mengenai penelitian kami sebelumnya yaitu metode GLSA yang merupakan pengembangan dari metode LSA. Pada metode GLSA secara lengkap dapat ditunjukkan pada gambar 2.

A. Latent Semantic Analysis

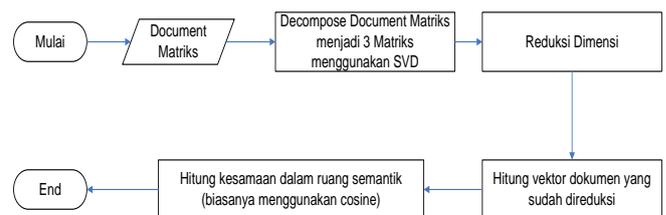
LSA merupakan metode untuk mengekstraksi dan merepresentasikan makna konseptual kata dengan menggunakan perhitungan statistik yang digunakan untuk menganalisa struktur semantik suatu teks [11]. Metode ini mempunyai ciri khas untuk mengekstrak dan merepresentasikan kalimat dengan perhitungan matematis dan mementingkan kata-kata kunci yang terkandung dalam sebuah kalimat tanpa memperhatikan karakteristik linguistiknya. Perhitungan matematis dilakukan dengan memetakan ada atau tidak adanya kata dari kelompok kata pada matriks semantik dan kemudian diolah menggunakan teknik aljabar linier Singular Value Decomposition (SVD). LSA digunakan untuk menilai esai dengan mengkonversi esai menjadi matriks-matriks yang diberi nilai pada masing-masing term untuk dicari kesamaan dengan term referensi. Sehingga dapat dikatakan LSA merupakan sebuah teknik statisitik yang digunakan untuk memperoleh hubungan makna kontekstual dari sebuah text.

Langkah-langkah LSA dalam menilai esai meliputi: Term document matriks yang akan merepresentasikan text jawaban

guru dan jawaban siswa sebagai matriks dokumen, matriks ini terdiri dari baris dan kolom dimana baris mewakili sebuah kata yang unik dari jawaban guru secara keseluruhan sementara kolom mewakili dokumen jawaban siswa yang mana nilai setiap baris dan kolom adalah frekuensi kemunculan istilah dalam dokumen.

Selanjutnya proses dekomposisi matriks menggunakan SVD serta proses reduksi dimensi yang ditujukan untuk mengurangi besarnya matriks dokumen yang dihasilkan dan membuang data yang tidak memiliki korelasi yang kuat dengan data dokumen jawaban guru. Setiap kata dalam paragraf kalimat direpresentasikan sebagai baris dan kolom matriks. Dengan menggunakan teknik matrik aljabar SVD (Singular Value Decomposition), matrik didekomposisi menjadi tiga komponen matrik, yaitu dua matriks orthogonal dan satu matriks diagonal singular

Proses yang menggambarkan tahapan LSA dapat dilihat pada gambar 1.



Gambar 1. Flowchart tahapan LSA

Pada bagan LSA terdapat *Singular Value Decomposition* (SVD) yang merupakan teorema aljabar linear yang mengatakan bahwa persegi panjang matriks A dapat dipecah menjadi tiga matriks, yaitu:

- a.) Matriks orthogonal U
- b.) Diagonal matriks S
- c.) *Transpose* dari matriks V orthogonal.

Yang dapat dirumuskan seperti pada persamaan 1 berikut ini:

$$A_{mn} = U_{mm} \times S_{mn} \times V_{nn}^T \quad (1)$$

A_{mn} = matriks awal

U_{mm} = matriks orthogonal U

S_{mn} = diagonal matriks S

V_{nn}^T = *transpose* matriks V orthogonal

Tahap akhir dari proses LSA adalah menghitung kesamaan dalam ruang semantik yang biasanya menggunakan *Cosine Similarity*. *Cosine Similarity* digunakan untuk menghitung nilai kosinus sudut antara dua vektor (vektor dokumen dan vektor *query*, semakin kecil sudut yang dihasilkan, maka tingkat kemiripan esai semakin tinggi. Vektor merupakan besaran yang memiliki nilai dan arah.

Cosine Similarity dapat dirumuskan seperti pada persamaan 2 berikut ini:

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

Keterangan :

- A** = **Vektor** dokumen
- B** = **Vektor query**
- A • B** = perkalian *dot* **Vektor A** dan **Vektor B**
- |A|** = panjang **Vektor A**
- |B|** = panjang **Vektor B**
- |A||B|** = *cross product* antara **|A|** dan **|B|**
- α = sudut yang terbentuk antar **Vektor A** dan **B**

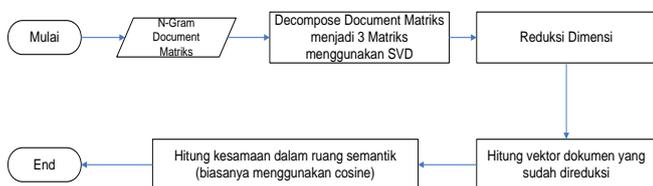
B. Generalized Latent Semantic Analysis

Generalized Latent Semantic Analysis (GLSA) merupakan pengembangan lebih lanjut dari algoritma LSA dengan membuat n-gram berdasarkan matriks dokumen. GLSA mempertimbangkan urutan kata kalimat dalam dokumen dan menjaga kedekatan kata dalam kalimat. Seperti contoh misalnya dalam LSA untuk formasi kata “carbon dioxide” menghasilkan arti yang sama dengan “dioxide carbon”. Menurut GLSA, bi-gram vektor untuk “carbon dioxide” itu mempunyai arti yang satu daripada kombinasi “carbon” dan “dioxide”.

GLSA ini tidak berbasis pada vektor dokumen dengan pendekatan *bag of words*. Melainkan proses dimulai dengan kesamaan-kesamaan antar kata yang berpasangan secara semantik untuk melakukan komputasi representasi untuk kata-kata [9]. Hal ini berbeda dari pendekatan yang umumnya representasi dokumen-kata menjadi representasi kata yang mempunyai maksud sebagai berikut:

- Kata-kata mempunyai tingkat fleksibilitas yang lebih tinggi dalam hal melakukan pencarian hubungan kesamaan yang ada daripada dokumen-dokumen.
- Ketersediaan koleksi yang banyak seperti web menawarkan sumber daya yang luar biasa untuk pendekatan statistik.

Kata-kata yang mengandung makna, contohnya kata-kata yang menyampaikan informasi semantik paling banyak, sering kali dikombinasikan menjadi kelas-kelas semantik yang berhubungan kepada aktifitas atau relasi tertentu dan mengandung sinonim dan kata yang berhubungan secara semantik. Oleh karena itu, sudah menjadi hal yang alami untuk menyajikan kata-kata sebagai vektor-vektor yang berdimensi rendah pada konsep-konsep semantik yang ada [9]. Proses yang menggambarkan GLSA secara sederhana dapat ditunjukkan pada gambar 2.



Gambar 2. Flowchart tahapan GLSA

Proses pada GLSA memiliki kemiripan dengan LSA, yang membedakannya hanya pada tahapan pembentukan N-gram dokumen yang akan memperhatikan urutan kata dan menjaga kedekatan kata dalam kalimat.

Menurut Russell dan Norvig [12] n-gram didefinisikan sebagai urutan panjang simbol yang tertulis dengan sebutan unigram untuk 1-gram, bigram untuk 2-gram, trigram untuk 3-gram, dan ukuran 4 atau lebih hanya disebut sebagai n-gram. Pemodelan n-gram adalah pemodelan yang memetakan

distribusi probabilitas dari urutan huruf-n. Pemodelan N-gram didefinisikan sebagai urutan rantai Markov. Model N-gram banyak digunakan dalam pengolahan bahasa atau yang biasanya disebut Natural Language Processing (NLP). Dalam penelitian ini digunakan n-gram sebagai pengindeksan. Dari pengertian diatas dapat didefinisikan bahwa:

- N-Gram menunjukkan kombinasi kata berturut-turut
- N gram mempertahankan kedekatan dalam dokumen
- N-gram mewakili frasa ukuran N

N-gram dipilih dikarenakan mempertahankan kedekatan kata, dimana pada term dokumen carbon dioxide memiliki makna yang sama dengan dioxide carbon, sedangkan pada n-gram hal itu memiliki makna yang berbeda. Sehingga dapat dikatakan N-gram menjaga kedekatan kata dalam kalimat. Beberapa contoh disertai penjelasan lebih lengkap dapat dilihat pada Tabel I.

TABEL I. PROSES N-GRAM

INPUT Soal Sejarah:		PREPROCESSING:
usaha suatu negara yang dilakukan oleh suatu negara untuk menguasai negara atau wilayah lain demi kepentingan politik,ekonomi dan budaya. tujuannya adalah untuk mencari kekayaan(gold), mencapai kejayaan (glory), menyebarkan agama nasrani(gospel)		usaha negara negara kuasa negara wilayah penting politik ekonomi budaya tuju cari kaya gold capai jaya glory sebar agama nasran gospel
Unigram	Bigram	Trigram
usaha, negara, kuasa, wilayah, penting dll.	Usaha negara, negara negara, negara kuasa dll	Usaha negara negara, negara negara kuasa dll.

Pada tabel I menunjukkan proses n-gram yang menggunakan trigram (n=3) dimana pada proses GLSA ini proses dilakukan dengan membentuk matriks proses n-gram dimana secara urut baris akan mewakili sebuah kata dari jawaban guru dimulai dari unigram, bigram dan trigram. Pembentukan term dokumen matriks n-gram GLSA dapat dilihat pada Tabel II.

TABEL II. PEMBENTUKAN MATRIKS GLSA

N-gram	Kata yang terbentuk	TF-IDF
Unigram	usaha	0.405465108
Bigram	usaha negara	0.405465108
Trigram	usaha usaha negara	0.405465108
Unigram	negara	0.336472237
Bigram	negara negara	0.405465108
Trigram	negara negara kuasa	0.405465108
...

Pada Tabel II merupakan penjabaran dari contoh n-gram dan kata yang terbentuk serta D1 mewakili nilai frekuensi kemunculan kata dalam n-gram menggunakan algoritma TF-IDF. Dari Tabel II dapat dilihat bahwa proses generate n-gram dilakukan secara berurutan dimulai dari n=1(unigram) hingga n=3(trigram) demikian seterusnya, dengan adanya proses ini tentunya akan menjaga kedekatan kata dalam kalimat namun ketika berada pada kalimat yang memperhatikan urutan kata maka perlu dilakukan kajian ulang pada proses generate n-gramnya.

III. METODE PENELITIAN

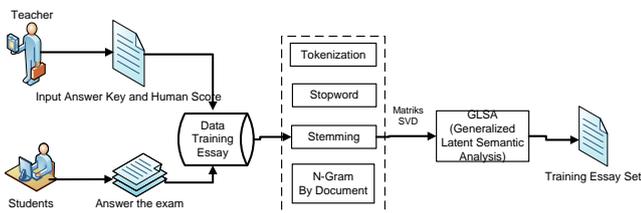
Pada bagian ini akan dijelaskan metode penelitian ini yang meliputi prosedur penelitian, arsitektur sistem, sistem input, preprocessing dan sistem output.

A. Prosedur Penelitian

Penelitian ini pada dasarnya lebih mengedepankan aspek implementasi dan analisis suatu teori, tentang penilaian esai secara otomatis dengan pengembangan algoritma Generalized Latent Semantic Analysis. Seperti pada penelitian-penelitian tentang analisis atau pembuktian suatu teori yang banyak menerapkan metode studi kepustakaan, demikian halnya pada penelitian ini, penulis lebih banyak melakukan studi kepustakaan dengan mengambil referensi dari internet, ebook, paper, jurnal, skripsi, tesis, maupun buku-buku yang berkaitan dengan algoritma pemrograman, penilaian otomatis, stemming khususnya pada bahasa Indonesia, pembentukan kata pada Bahasa Indonesia, dan berbagai literatur ilmiah lainnya.

B. Arsitektur Sistem

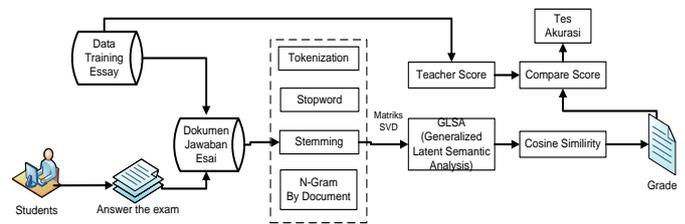
Pada arsitektur sistem ini terdapat dua aktor utama yaitu guru sebagai pembuat soal dan jawaban serta siswa yang menjawab soal ujian, terdapat dua tahapan yaitu data training dan data evaluasi hasil training, semua data ini akan disimpan ke dalam database sistem dan dilakukan proses pengolahan meliputi: preprocessing, pembentukan n-gram hirarki dan proses pengolahan nilai menggunakan cosine similarity dan perhitungan panjang vektor. Secara jelas arsitektur sistem GLSA ini dapat digambarkan pada gambar 3.



Gambar 3. Arsitektur Data Training Esai

Pada gambar 3 menunjukkan proses input training data yang mengambil data training dari input jawaban guru serta input jawaban siswa dalam 1 kelas. Proses selanjutnya adalah semua data training akan dilakukan proses preprocessing terlebih dahulu yang meliputi tokenisasi (memecah kalimat), stopwords, stemming dan pembentukan n-gram dokumen, setelah itu tahap selanjutnya adalah proses pembentukan matriks awal dan matriks svd yang diolah dan akan menghasilkan data training esai.

Pada dasarnya evaluasi jawaban siswa dengan hasil training memiliki proses yang tidak jauh berbeda dengan proses saat melakukan training data, yang membedakan hanya terdapat proses perhitungan kemiripan jawaban dengan data training dan menghasilkan nilai baik yang dinormalisasi ataupun tidak menggunakan normalisasi. Arsitektur proses evaluasi jawaban siswa ditunjukkan pada gambar 4.



Gambar 4. Arsitektur Evaluasi Training GLSA

Gambar 4 menunjukkan bahwa proses input evaluasi dilakukan oleh siswa, selanjutnya dilakukan proses preprocessing yang meliputi proses tokenisasi dimana pada proses ini semua kalimat akan dijadikan token kata kemudian dilakukan proses menghapus stopwords yang membuang semua kata umum yang tidak memiliki makna, proses stemming yang mengubah menjadi kata dasar dan pada akhirnya dilakukan proses n-gram yang akan menghasilkan matriks jawaban siswa dan guru dilanjutkan proses GLSA yang akan membentuk vektor jawaban siswa dan hasil training dan pada tahap akhir dilakukan proses hasil penilaian yang diukur menggunakan cosine similarity dan proyeksi panjang vektor.

C. Sistem Input

Input dari penelitian ini adalah kumpulan dari jawaban essay siswa. Mata pelajaran Sejarah terdapat 5 soal esai dengan rincian: 60 jawaban siswa dan 31 data training yang berjumlah total 455 jawaban. Mata pelajaran Sosiologi terdapat 4 soal esai dengan rincian: 40 jawaban siswa dan 22 data training yang berjumlah total 248 jawaban.

D. Preprocessing

Secara garis besar urutan proses dalam tahap preprocessing ini adalah sebagai berikut: Pertama, text dipisah-pisahkan menjadi token-token atau menjadi potongan-potongan kata dan menghilangkan tanda baca dan simbol pada token tersebut. Sebagai contohnya, simbol tanda baca:

~!@#%&*()_+={ }[]\|:;'"<>./

Kedua, token-token tersebut dirubah hurufnya menjadi huruf kecil semua. Ketiga, menghapus token (kata-kata) yang dianggap tidak penting, sesuai dengan stop list (daftar kata) yang diambil dari rujukan stoplist algoritma Tala[13]. Keempat, proses mengubah token/kata menjadi kata dasar. Kelima adalah proses menghitung frekuensi kemunculan kata melalui n-gram di dalam sekumpulan text.

Hasil output pada tahap preprocessing berbeda-beda yang dapat ditunjukkan pada gambar 5 yang prosesnya melalui export ke bentuk halaman .xls, misal untuk proses case folding dan tokenisasi, output sistem akan menghasilkan kumpulan kata yang terdapat pada setiap jawaban siswa dan jawaban guru. Proses stopwords removal akan menghasilkan kumpulan kata yang dianggap penting. Proses stemming, output sistem akan menghasilkan kumpulan kata dasar. Sedangkan pada proses pembobotan term, output sistem akan menghasilkan nilai frekuensi kemunculan tiap kata berdasarkan n-gram pada semua dokumen yang digunakan untuk input data pada sistem.

Sekumpulan frekuensi tersebut direpresentasikan kedalam sebuah matriks dokumen yang setiap matriksnya mewakili

kumpulan jawaban siswa dan matriks query yang mewakili semua jawaban kunci jawaban guru.

INPUT:
 imperialisme dilihat dari : asal bahasa : bahasa latin yaitu imperase yang artinya memerintah atau imperium artinya untuk memerintah secara umum keinginan, usaha, kegiatan untuk menguasai atau memerintah negara lain demi kepentingan dibidang politik, ekonomi, sosial, dan budaya. sebelum masa revolusi industri, imperialisme kuno bertujuan mencapai "gold, glory, gospel" yaitu paham untuk menguasai kekayaan, kejayaan, dan agama negara lain khususnya gossell dalam bidang penyebaran agama nasrani, kekayaan dan kejayaan lewat perdagangan dan jajahan.

Output:
<Stopword>
 imperialisme bahasa bahasa latin imperase memerintah imperium memerintah usaha kegiatan menguasai memerintah negara kepentingan dibidang politik ekonomi sosial budaya revolusi industri imperialisme kuno bertujuan mencapai gold glory gospel paham menguasai kekayaan kejayaan agama negara gossell bidang penyebaran agama nasrani kekayaan kejayaan perdagangan jajahan

<Stemmer>
 imperialisme bahasa bahasa latin imperase perintah imperium perintah usaha giat kuasa perintah negara penting bidang politik ekonomi sosial budaya revolusi industri imperialisme kuno tuju capai gold glory gospel paham kuasa kaya jaya agama negara gossell bidang sebar agama nasran kaya jaya dagang jajah

Gambar 5. Proses Preprocessing GLSA

Demikian pula, ketika pada input digambar 5 memiliki simbol tanda baca maka pada proses awal semua simbol dan tanda baca itu akan dibuang seperti yang ditunjukkan pada proses stopword. Seperti pada input yang ditunjukkan pada gambar 5 semua kalimat itu berhasil diekstraksi dengan baik melalui proses penghapusan simbol dan proses mendapatkan kata dasar (stemmer).

TABEL III. TEMPLATE INPUT PREPROCESSING

Nama	Kelas	Jawaban	StopWord	Stem	Skor
Amel	XI

Pada akhirnya hasil daripada proses preprocessing akan menkonversi hasil dalam bentuk file dokumen .xls sesuai dengan template excel yang ditunjukkan pada Tabel III.

E. Sistem Output

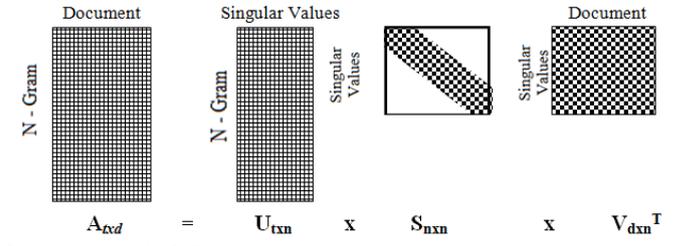
Setelah dilakukan preprocessing, selanjutnya akan dibentuk matriks awal yang memuat kata-kata dimana setiap kata dalam paragraf kalimat direpresentasikan sebagai baris dan kolom matriks. Proses ini dapat ditunjukkan pada gambar 6 dimana terdapat data matriks awal yang mewakili kata yang terbentuk dari n-gram serta kolom dokumen yang mewakili jawaban training guru serta nilai frekuensi kemunculan kata yang dihitung menggunakan tf-idf sebagai perhitungan frekuensi n-gramnya.

Gambar 6. Matriks Awal GLSA

Dengan menggunakan teknik matrik aljabar SVD, matrik didekomposisi menjadi tiga komponen matrik, yaitu dua matriks orthogonal dan satu matriks diagonal singular yang pada tujuannya adalah untuk menjaga informarsi semantik dalam teks dan juga mengurangi teks yang tidak berkaitan dengan jawaban guru.

Proses selanjutnya setelah tahap preprocessing adalah pembentukan matriks Singular Value Decomposition (SVD). Metode SVD merupakan salah satu metode aljabar linear suatu matriks A yang berukuran m x n mempunyai nilai singular yang merupakan akar pangkat dua dari eigen value $A^T A$.

Ilustrasi dari metode SVD dapat ditunjukkan pada gambar 7 yang menjelaskan bagaimana dekomposisi dari nilai singular matriks.



Gambar 7. Matriks SVD

Keterangan :
 A adalah n-gram dari dokumen matriks
 U adalah matriks ortogonal
 S adalah matriks diagonal
 V^T adalah transpose dari matriks ortogonal V

Tahap selanjutnya setelah proses pembentukan matriks SVD adalah dengan pembentukan vektor query. Pembentukan vektor query dibagi menjadi 2 bagian yaitu pembentukan vektor query untuk jawaban guru dan pembentukan vektor query jawaban siswa. Pembentukan vektor query merupakan tahap yang penting sebelum dilakukan proses mencari kemiripan. Hal ini dikarenakan nilai vektor query yang dihasilkan akan dilakukan proses perhitungan baik untuk vektor query jawaban guru maupun vektor query jawaban siswa. Skema pembentukan vektor query dapat ditunjukkan pada gambar 8 dimana W mewakili bobot nilai frekuensi sebanyak n gram yang sudah ditentukan diawal.

Esai jawaban

N-Gram	W_{q1}
	W_{q1}
	W_{q1}
	\vdots
	W_{qn}

Mapel	No	Soal
Sejarah	1	Apakah yang dimaksud dengan imperialisme dan tujuan dari pelaksanaan imperialisme kuno ?
	2	Jelaskan faktor penyebab runtuhnya serikat dagang VOC!
	3	Jelaskan Latar belakang pelaksanaan sistem ekonomi liberal di Indonesia!
	4	Apakah pengaruh perjanjian Tordesilas bagi Portugis dan Spanyol?
	5	Jelaskan latar belakang diberlakukannya politik etis di Indonesia pada masa pemerintahan kolonial Belanda!

Gambar 8. Skema matriks Query

Pada gambar 8 query matriks terbentuk dari 2 bagian penting yaitu esai jawaban sebagai kolom dan kata n-gram sebagai baris. Pembentukan vektor query ini nantinya akan digunakan dalam melihat nilai vektor yang dihasilkan agar dapat dihitung kemiripan antara vektor guru dan vektor jawaban siswa yang terbentuk.

Query Vector guru dihasilkan dari proses input jawaban guru yang dianggap sebagai dokumen dan akan dicocokkan dengan vektor jawaban siswa. Matriks SVD yang sudah direduksi akan digunakan untuk membentuk vektor query dokumen jawaban guru, hal ini dapat ditunjukkan oleh persamaan 3.

$$D_j = D_j^T \times U_k \times S_k^{-1} \quad (3)$$

Keterangan:

D_j : Vektor query dokumen jawaban guru

D_j^T : Transpose vektor query dokumen jawaban guru

U_k : Matriks reduksi Ortogonal

S_k^{-1} : Matriks invers reduksi singular

IV. HASIL DAN PEMBAHASAN

Pada bagian ini akan dijelaskan mengenai analisis data yang digunakan dalam percobaan, analisis preprocessing, analisa nilai GLSA, dan yang terakhir analisis akurasi.

A. Analisis Data

Proses ujicoba menggunakan tiga mata pelajaran (Sosiologi, Sejarah, dan Pendidikan Kewarganegaraan). Mata pelajaran Sejarah terdapat lima soal dengan jumlah jawaban siswa sebanyak 300 jawaban dan 31 data training yang berisi jawaban guru dan jawaban siswa, sehingga totalnya adalah 455 jawaban. Sedangkan mata pelajaran Sosiologi berjumlah total 248 jawaban. Sedangkan mata pelajaran Pendidikan Kewarganegaraan berjumlah total 189 jawaban. Secara detail dijabarkan pada Tabel IV.

TABEL IV. INPUT JAWABAN TRAINING DAN TESTING

No	Mapel	Jumlah Soal	Training	Testing	Total
1	Sejarah	5	155	300	455
2	Sosio	4	88	160	248
3	Pkn	3	66	123	189
Total Data (Σ)			309	583	892

Tabel IV menunjukkan banyaknya dataset Pelajaran (Sejarah, Sosiologi dan Pendidikan Kewarganegaraan) dengan total jawaban sebanyak 892 data training dan data testing. Contoh dari pertanyaan untuk pelajaran Sejarah dapat ditunjukkan pada Tabel V.

TABEL V. INPUT SOAL ESAI

Pada Tabel V menunjukkan data input untuk soal sejarah yang semua soalnya menuntut jawaban uraian panjang dan bukan jawaban singkat. Proses memasukkan data soal dan jawaban siswa sebagai data training dapat ditunjukkan melalui Tabel VI yang mana nilai jawaban siswa diberikan secara koreksi manual oleh guru yang bersangkutan.

TABEL VI. TEMPLATE DATA TRAINING JAWABAN

No	Soal	Bobot	Jawaban1	Nilai
1	Apakah yang dimaksud dengan imperialisme dan?	15	Imperialisme: Usaha suatu negara untuk merebut negara lain demi kepentingan politik, ekonomi, dan budaya....	15
...

Tabel VI menunjukkan template data yang akan dijadikan sebagai data training, dimana terdapat kolom soal, bobot nilai yang diberikan, jawaban soal, serta nilai dari guru atas jawaban tersebut.

Pada gambar 9 menunjukkan interface yang terdiri dari form yang akan digunakan dalam sistem dimana user atau pengguna dapat berinteraksi dengan sistem dengan memasukkan jawaban siswa.



Gambar 9. Form Input Jawaban Siswa

B. Analisis Preprocessing

Pada tahap ini proses preprocessing meliputi dua bagian yaitu menghapus stopword dan melakukan proses stemming, semua proses ini dilakukan seperti yang ditunjukkan pada Tabel VII yang merupakan hasil pengolahan proses preprocessing dan diexport kedalam bentuk file .xls.

TABEL VII. MASUKAN JAWABAN TRAINING DAN TESTING

No	Nama	Kelas	Jawaban1	Stopword	Stemmer	Ni lai
1	Guru	Guru	Imperialis me: Usaha suatu negara untuk...	imperialis me usaha negara merebut negara ...	imperiali sme usaha negara rebut negara ...	15
2	Amel	XI IPS	imperialis me : negara lain demi ...	imperialis me negara lain ...	imperiali sme negara ...	10

Pada proses preprocessing yang dihasilkan pada Tabel VII merujuk pada algoritma Tala [13] yang menggunakan informasi dengan aturan morfologi tanpa menggunakan bantuan kamus kata dasar dan murni berbasis aturan (*purely rule-based stemmer*). Algoritma Tala merupakan pengembangan dari algoritma porter yang diaplikasikan untuk mendapatkan kata dasar text berbahasa Indonesia.

C. Analisis Nilai GLSA

Pada proses analisa melibatkan pembentukan vektor dokumen jawaban siswa dan vektor dokumen data training yang prosesnya ditunjukkan melalui persamaan 3.

Algoritma Pembentukan Query Vector Training

```

01: For i = 1 to p do
02:   a) Hapus Stop-words dari esai E
03:   b) Stem kata-kata dari esai jawaban training
04: EndFor
05: Pilih N-gram untuk index-term dari n-gram oleh
dokumen matriks
06: Bentuk N-Gram dari dokumen matriks, Amnp
07: Decompose matriks Amnp menggunakan SVD, dimana
persamaannya akan menjadi
Amnp = Umnr × Srnr × VrnpT
08: Truncate/potong/reduksi dari U, S dan VT dan
buatlah menjadi persamaan berikut :
Akxk = Umnk × Skxk × Vkxp
09: For j=1 to p Do
10: Bentuklah esai vektor kunci
Dj = DjT × Umnk × Skxk-1
11: EndFor
    
```

Pada tahap selanjutnya sama seperti pembentukan query vector training, matriks SVD yang sudah direduksi akan digunakan untuk membentuk vektor query dokumen jawaban siswa, Pada proses ini setiap jawaban siswa nantinya akan dibentuk query vector yang akan dibandingkan dengan query vector jawaban data training agar dapat ditentukan nilai kemiripannya yang menjadi dasar dalam pemberian nilai secara otomatis oleh sistem. Proses ini dapat ditunjukkan oleh persamaan (4).

$$Q_j = Q_j^T \times U_k \times S_k^{-1} \tag{4}$$

Keterangan :

- Q_j : Vektor query dokumen jawaban siswa
- Q_j^T : Transpose vektor query dokumen jawaban siswa
- U_k : Matriks reduksi Orthogonal
- S_k⁻¹ : Matriks invers reduksi singular

Algoritma Pembentukan Query Vector Siswa

```

01: For i = 1 to p do
02:   a)Hapus Stop-words dari esai E'
03:   b) Stem kata-kata dari esai jawaban siswa
04: EndFor
05: Bentuk dimensi query matriks yang sama dengan
aturan n-gram dari dokumen matriks
06: For j=1 to p Do
07: Bentuklah esai vektor jawaban
Q' j = Q' jT × Umnk × Skxk-1
08: EndFor
    
```

Setelah didapatkan matriks query jawaban guru dan matriks query jawaban siswa, maka tahap terakhir adalah melakukan perhitungan nilai ujian. Proses perhitungan nilai dilakukan dengan menggunakan metode Cosine Similarity yang akan membandingkan kesamaan atau kedekatan antara dokumen acuan (kunci jawaban guru) dengan kandidat dokumen (jawaban siswa).

Tahapan akhir setelah diperoleh nilai Cosine Similarity untuk setiap jawaban esai siswa adalah dilakukan normalisasi nilai yang dihasilkan. Proses normalisasi ditujukan agar nilai yang dihasilkan dapat secara signifikan sesuai dengan nilai human grader/penilai manusia. Proses normalisasi menggunakan metode regresi linear sederhana yang melibatkan bobot soal yang diberikan oleh guru untuk masing-masing soal. Dimana untuk setiap jawaban siswa, sistem akan mendapatkan Cosine Similarity sebanyak jumlah soal. Sebagai ilustrasi dimisalkan terdapat P soal yang diberikan oleh guru kepada siswa.

Input : Cosine similarity = {C₁, C₂,C_p}
 Bobot = {B₁, B₂,B_p}

Sebagai catatan bobot disini diartikan sebagai nilai maksimal untuk soal a, dikarenakan dalam sebuah test esai yang berjumlah n soal, pemberian nilai tidak selalu berjumlah 100 namun terkadang memiliki total bobot yang berbeda.

Persamaan yang digunakan untuk mendapatkan hasil pembobotan dari tes esai yang diberikan dengan n soal dapat ditunjukkan pada persamaan (5)

$$\text{Hasil Pembobotan (H)} = \frac{\sum_{i=1}^p C_i B_i}{\sum_{i=1}^p B_i} \tag{5}$$

Dari persamaan (5) dapat dilihat bahwa hasil pembobotan merupakan nilai dari Cosine Similarity dikalikan dengan bobot soal dan hasil perkalian itu akan dibagi dengan total bobot dari n soal. Proses selanjutnya adalah hasil pembobotan akan dirubah menjadi nilai dengan rentang 0-100 dengan menggunakan Tabel VIII.

TABEL VIII. TABEL NORMALISASI

Hasil Cosine Similarity (x-y)	Nilai Sistem (x'-y')
0,00 - 0,10	1 - 20
0,10 - 0,30	20 - 40
0,30 - 0,60	40 - 70
0,60 - 0,80	70 - 90
0,80 - 1,00	90 - 100

Dari Tabel VIII dapat dilihat bagaimana proses normalisasi nilai melibatkan hasil Cosine Similarity yang memiliki range x hingga y dan nilai sistem yang memiliki range x' hingga y'. Berdasarkan range yang ada maka nilai sistem dapat dirumuskan melalui persamaan (6).

$$\text{Nilai Sistem (S)} = \frac{H-X}{y-x} (y' - x') + x' \quad (6)$$

Keterangan:

H: Hasil pembobotan

x: Nilai rentang bawah Cosine Similarity hasil pembobotan

y: Nilai rentang atas Cosine Similarity hasil pembobotan

x': Nilai rentang bawah dari nilai sistem

y': Nilai rentang atas dari nilai sistem

Penjabaran dari persamaan 6 dapat dijelaskan melalui gambar 10 yang memberikan penjelasan penerapan dari persamaan 6.

Contoh Proses				Skema Korelasi	
Cosine Similarity		Bobot	Ci * Bi		
Soal 1	0,82568	20	16,5136		
Soal 2	0,98523	30	29,5569		
Soal 3	0,64825	20	12,965		
Total		70	59,0355		
Hasil Pembobotan (H)		0,843364286			
Nilai Sistem (S)		92,16821429			
<p>Hasil Pembobotan didapatkan dari rumus: $H = 59,0355 / 70 = 0,843364286$ Nilai Sistem didapatkan dari rumus: $= (0,843364286 - 0,8) / (1 - 0,8) * (100 - 90) + 90$ $= 92,16821429$</p>					

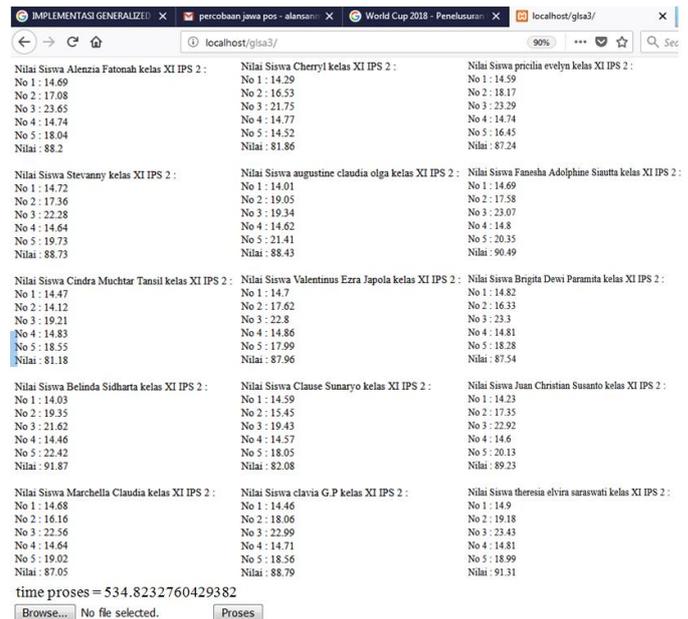
Gambar 10. Penjabaran Proses Nilai GLSA Normalisasi

Algoritma Penilaian Sistem GLSA

- 01: **For** i = 1 to p **do**
- 02: Hitung cosine similarity antara D'i dan Di
- 03: **EndFor**
- 04: Kalkulasi rata-rata cosine similarity untuk semua soal.
- 05: Ubah hasil rata-rata cosine similarity menjadi nilai berdasarkan tabel rentang kemiripan

Pada Algoritma Penilaian GLSA merupakan implementasi dari proses penilaian yang sebelumnya dilakukan proses perhitungan Cosine Similarity menggunakan persamaan (2) kemudian dilakukan proses normalisasi regresi menggunakan persamaan (5) dalam menentukan hasil pembobotan dan persamaan (6) untuk menilai hasil akhir dari sistem untuk setiap soal tes esai yang diberikan guru.

Sistem penilaian yang dilakukan menunjukkan penilaian tidak secara langsung mengambil nilai Cosine Similarity namun dilakukan proses regresi terlebih dahulu dengan menggunakan tabel normalisasi seperti yang ditunjukkan pada tabel 8 dimana hasil yang ditunjukkan dari penggunaan regresi memberi kenaikan nilai.



Gambar 11. Proses Nilai Sistem GLSA

Pada gambar 11 proses analisa nilai GLSA dilakukan dengan data training sebanyak 31 data jawaban training yang terdiri dari data jawaban guru dan siswa. Proses pada analisa disini sangatlah sederhana dikarenakan hanya digunakan untuk proses pengujian sistem dengan menggunakan template seperti pada tabel 6 yang selanjutnya dilakukan proses export terhadap hasil penilaian dari GLSA ke dalam bentuk format .xls.

D. Analisis Akurasi

Pada bagian ini, akan membahas mengenai evaluasi dan analisis tingkat akurasi terhadap penerapan metode GLSA dalam penilaian jawaban tes esai pada soal berbahasa Indonesia. Metode yang digunakan dalam melakukan analisis akurasi menggunakan Standar Error yang merupakan akar dari nilai standar deviasi yang sudah dibagi dengan N. Pengukuran Akurasi sendiri dapat diartikan sebagai hasil pengukuran yang

menunjukkan kedekatan dengan nilai sesungguhnya. Variabel pengukuran yang dipakai dalam penelitian ini adalah hasil penilaian otomatis menggunakan sistem komputer dan hasil penilaian manual yang dilakukan human grader, dalam hal ini guru mata pelajaran. Dalam penelitian ini pengukuran error rate menggunakan persamaan (7) yang menggunakan metode Standar Error yang akan mengukur seberapa tepatnya nilai mean (rata-rata) yang diperoleh dari nilai sistem dan nilai guru.

$$ER = \sqrt{\frac{\sum_{i=1}^n (S_i - G_i)^2}{n}} \quad (7)$$

Keterangan:

ER : Error Rate

S_i : Nilai Sistem untuk siswa ke-i

G_i : Nilai Guru untuk siswa ke-i

M : Nilai Maximal Tes

n : Jumlah Siswa yang mengikuti tes esai

Persamaan (7) menunjukkan bagaimana rumus yang digunakan untuk menghitung error rate, setelah nilai error rate didapatkan maka proses menghitung akurasi cukup mudah yaitu dengan mengurangi nilai error rate dengan 1 yang ditunjukkan oleh persamaan (8).

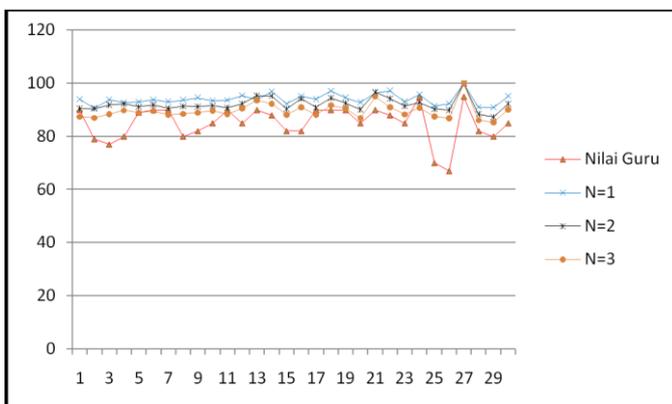
$$\text{Akurasi} = 1 - \text{Error rate} \quad (8)$$

Proses selanjutnya untuk mengetahui perbandingan LSA dan GLSA maka dilakukan pengujian terhadap nilai sistem yang diberikan oleh LSA maupun GLSA, proses pengujian dilakukan dengan parameter selisih nilai sistem dengan nilai guru, selisih lebih kecil menandakan sistem lebih akurat. Proses pengamatan dilakukan terhadap selisih nilai yang dihasilkan oleh LSA dan GLSA terhadap nilai guru, dimana jika semakin kecil perbedaan nilai maka sistem dapat dikatakan akurat terhadap nilai guru, hasil proses pengujian ditunjukkan pada Tabel IX akan ditunjukkan selisih antara LSA dan GLSA serta diberikan keterangan sistem mana yang lebih unggul berdasarkan selisih nilai LSA dan GLSA dengan nilai guru.

TABEL IX. PERBANDINGAN LSA DAN GLSA

No	Nama	Nilai Guru	LSA	GLSA	Selisih LSA	Selisih GLSA	Sistem yang Unggul
1	user1	90	96.80396	94.75003	6.8039590	4.7500270	GLSA
2	user2	79	97.2551	94.97131	18.2551028	15.9713105	GLSA
3	user3	77	96.32001	92.74538	19.3200107	15.7453819	GLSA
4	user4	92	95.04564	92.94996	3.0456366	0.9499642	GLSA
5	user5	89	96.68753	91.57916	7.6875318	2.5791626	GLSA
6	user6	92	96.18479	93.3051	4.1847897	1.3051024	GLSA
7	user7	90	93.64569	92.26479	3.6456877	2.2647873	GLSA
8	user8	89	96.93407	93.42078	7.9340714	4.4207841	GLSA
9	user9	92	94.69692	91.80536	2.6969176	0.1946379	GLSA
10	user10	89	97.53634	95.23078	8.5363374	6.2307789	GLSA
11	user11	92	96.23997	93.4062	4.2399722	1.4061978	GLSA
12	user12	92	97.12323	93.40829	5.1232268	1.4082868	GLSA
13	user13	90	94.54718	93.65493	4.5471767	3.6549303	GLSA
14	user14	97	96.50408	94.36462	0.4959224	2.6353795	LSA
15	user15	92	96.20267	92.85387	4.2026725	0.8538655	GLSA
16	user16	82	96.66393	94.85843	14.6639314	12.8584316	GLSA
17	user17	97	97.41524	93.39809	0.4152352	3.6019103	LSA
18	user18	90	95.28451	92.39167	5.2845078	2.3916700	GLSA
19	user19	92	96.29424	93.51913	4.2942428	1.5191258	GLSA
20	user20	85	97.3344	94.04378	12.3343965	9.0437830	GLSA

Pada gambar 12 merupakan grafik yang menunjukkan perbandingan LSA dan GLSA dengan mengacu pada nilai yang diberikan secara manual oleh guru.



Gambar 12. Grafik Perbandingan LSA(n=1) dan GLSA(n=3)

Proses pengujian akurasi melibatkan penelitian kami sebelumnya [14] yang menggunakan GLSA tanpa

menggunakan data training dan hanya menggunakan data jawaban guru. Proses pengujian akan melibatkan 3 mata pelajaran dengan jumlah data set yang berbeda-beda dan menggunakan 2 skenario yaitu proses pengujian tanpa training data dan proses pengujian melibatkan training data.

TABEL X. HASIL PENGUJIAN GLSA TANPA DATA TRAINING

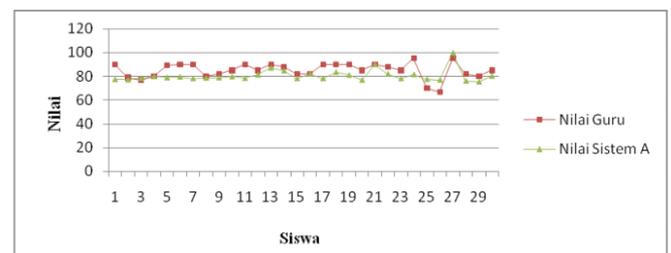
No	Mapel	Jumlah Data set Siswa	Jumlah Data set Jawaban	Total	Akurasi
1	Sejarah	445	10	455	90,76%
2	Sosio	240	8	248	87,79%
3	Pkn	183	6	189	89,10%
Rata-rata					89,21%

Berdasarkan kasus uji pada Tabel X, didapatkan hasil sebagai berikut:

- Pada mata pelajaran Sejarah dengan menggunakan dua jawaban guru (10 data set) dan 89 data jawaban siswa (445 data set) menghasilkan tingkat akurasi mencapai 90,76%.
- Pada mata pelajaran Sosio dengan menggunakan dua jawaban guru (8 data set) dan 60 data jawaban siswa (240 data set) menghasilkan tingkat akurasi mencapai 87,79%.
- Pada mata pelajaran Pkn dengan menggunakan dua jawaban guru (6 data set) dan 61 data jawaban siswa (183 data set) menghasilkan tingkat akurasi mencapai 89,10%.
- Rata-rata dari proses pengujian akurasi tanpa menggunakan training data untuk tiga mata pelajaran adalah 89,21%.

Pada gambar 13 merupakan grafik dari hasil pengujian GLSA tanpa menggunakan data training. Berdasarkan grafik untuk hasil pengujian nilai guru dan nilai sistem tidak terdapat perbedaan nilai yang signifikan.

Proses pengujian selanjutnya adalah dengan menggunakan data training yang disesuaikan dengan data pada Tabel IV dan hasil pengujiannya dapat dilihat pada Tabel XI.



Gambar 13. Grafik Pengujian GLSA Tanpa Data Training

TABEL XI. HASIL PENGUJIAN GLSA DENGAN TRAINING DATA

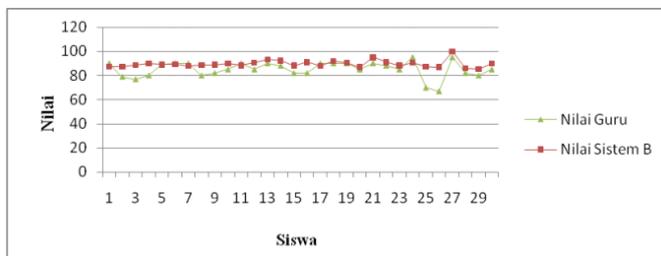
No	Mapel	Jumlah Data set Training	Jumlah Data set Testing	Total Dataset	Akurasi
1	Sejarah	155	300	455	93,33%
2	Sosio	88	160	248	89,58%

3	Pkn	66	123	189	88,27%
Rata-rata					90,39%

Berdasarkan kasus uji pada Tabel XI, didapatkan hasil sebagai berikut:

- a) Pada mata pelajaran Sejarah dengan menggunakan data training sebanyak 31 jawaban guru dan siswa (155 data set) dan 60 data testing jawaban siswa (300 data set) menghasilkan tingkat akurasi mencapai 93,33%.
- b) Pada mata pelajaran Sosio dengan menggunakan data training sebanyak 22 jawaban guru dan siswa (88 data set) dan 40 data testing jawaban siswa (160 data set) menghasilkan tingkat akurasi mencapai 89,58%.
- c) Pada mata pelajaran Pkn dengan menggunakan data training sebanyak 22 jawaban guru dan siswa (66 data set) dan 41 data testing jawaban siswa (123 data set) menghasilkan tingkat akurasi mencapai 88,27%.
- d) Rata-rata dari proses pengujian akurasi dengan menggunakan training data untuk tiga mata pelajaran adalah 90,39%.

Pada gambar 14 merupakan grafik dari hasil pengujian GLSA dengan menggunakan data training. Berdasarkan grafik untuk hasil pengujian nilai guru dan nilai sistem tidak terdapat perbedaan nilai yang signifikan dan korelasi antara nilai sistem dan nilai guru baik.



Gambar 14. Grafik Pengujian GLSA Dengan Data Training

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil pengujian terhadap tingkat akurasi sistem dapat disimpulkan dari dua parameter pengujian yang dilakukan bahwa pengujian dengan menggunakan training data sebanyak 30% dari jumlah dataset akan menghasilkan tingkat akurasi yang lebih baik dibandingkan tidak menggunakan data training. Dari penilaian akurasi terhadap 3 mata pelajaran dengan jumlah data set total mencapai 864 data set jawaban, GLSA dengan menggunakan training data mengalami peningkatan akurasi sebesar 1,18%.

B. Saran

Dari hasil pengujian dan analisa dapat diberikan saran untuk mencoba melakukan penelitian lebih lanjut dan berinovasi dan bervariasi dengan beberapa teknik yang ada pada proses preprocessingnya, penambahan database sinonim kata dan singkatan kata ataupun pada metode perhitungan kemiripannya sehingga tingkat akurasi yang didapatkan akan semakin lebih baik.

REFERENSI

- [1] S. Valenti, F. Neri, and A. Cucchiarelli, "An overview of current research on automated essay grading," *Journal of Information Technology Education*, vol. 2, pp.319-330, 2003.
- [2] T. Kakkonen, N. Myller, E. Sutinen, & J. Timonen, Comparison of Dimension Reduction Methods for Automated Essay Grading. *Educational Technology & Society*, 11(3), pp.275-288, 2008.
- [3] C. U. Lei, K. L. Man, and T. O. Ting, "Using learning analytics to analyze writing skills of students: A case study in a technological common core curriculum course," in *IAENG International Journal of Computer Science* 41(3), 2014.
- [4] S. Dikli, "An Overview of Automated Scoring of Essays," in *The Journal of Technology, Learning, and Assessment* Volume 5 Number 1, Stanford University, 2006.
- [5] A. M. Olney, "Generalizing Latent Semantic Analysis," in *IEEE International Conference on Semantic Computing*. University of Memphis. Memphis, USA, 2009.
- [6] A. A. P. Ratna, H. Artajaya, A. A. Boma, "GLSA Based Online Essay Grading System," in *International Conference on Teaching, Assessment and Learning for Engineering(TALE)*, 26-29 Aug 2013.
- [7] T. K. Landauer, P. W. Foltz and D. Laham, "An Introduction to Latent Semantic Analysis," *Discourse Processes*, vol. 25, no. 2-3, pp. 259-284, 1998.
- [8] M. M. Islam, A. S. M. L. Hoque, "Automatic essay scoring using Generalized Latent Semantic Analysis," 2010 13th International Conference on Computer and Information Technology (ICIT) pp.358-363, 23-25 Dec 2010.
- [9] I. Matveeva, "Generalized Latent Semantic Analysis for Document Representation". Disertasi University of Chicago. Chicago, USA, 2008.
- [10] J. Burstein, C. Leacock, and R. Swartz, "Automated evaluation of essay and short answers," in *Proceedings of the Sixth International Computer Assisted Assessment Conference*, Loughborough University, Loughborough, UK, 2001
- [11] T. K. Landauer, P. W. Foltz, and D. Laham, *An introduction to Latent Semantic Analysis*, 1998
- [12] S. J. Russell, P. Norvig, and J. Canny, *Artificial intelligence: a modern approach*, 2003.
- [13] F. Z. Tala, *A Study of Stemming Effect on Information Retrieval in Bahasa Indonesia*, Insitute for Logic, Language and Computation Universiteit van Amsterdam The Netherlands, 2003.
- [14] Ruslan, Gunawan, S. Tjandra, "Implementasi Generalized Latent Semantic Analysis untuk Penilaian Otomatis Jawaban Esai Siswa pada Tingkat Sekolah Menengah Atas (SMA)", *Prosiding Semantikom 2017* pp.33-40, ISSN 2527-4392, 2017.