PENGEMBANGAN KOREKSI ESAI OTOMATIS PADA E-LEARNING DI SMK PLUS AN-NABA SUKABUMI DENGAN MENGGUNAKAN METODE *LATENT SEMANTYC ANALYSIS* (LSA)

Mashun Sofyan[1]

[1] Teknik Informatika – Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
E-mail : Mashunsofyan@gmail.com[1]

ABSTRAK

Pengkoreksian soal esai secara manual yang menghabiskan waktu lebih lama dibandingkan dengan mengoreksi soal esai secara otomatis. Dibutuhkan waktu sekitar 1 menit tiap soal untuk mengkoreksi soal esai secara manual. Jika akan mengkoreksi 30 soal maka dibutuhkan waktu sekitar 30 menit, sedangkan jumlah pelajar di SMK Plus An-Naba Sukabumi sekitar 200 orang.

Berdasarkan data tersebut, perlu dikembangkan sebuah sistem otomatis yang membantu proses penilaian esai. Pengembangan Sistem Penilaian Esai Otomatis pada Tugas Akhir ini menggunakan Metode Latent Semantyc Analysis (LSA) yaitu dengan merepresentasikan kalimat dalam sebuah matriks semantik yang kemudian dilakukan perhitungan matematis dengan mencocokan atau memetakan ada atau tidaknya kata dari kelompok kata pada matriks menggunakan teknik aljabar linear Singular Value Decomposition (SVD).

Berdasarkan hasil data penelitian alpha dan betha dapat diambil sebuah kesimpulan bahwa pengembangan fungsionalitas koreksi esai di SMK Plus An-Naba Sukabumi mampu melaksanakan proses koreksi soal jenis esai dengan lebih cepat.

Kata kunci : LSA, Latent Semantyc Analysis, Singular Value Decomposition (SVD).

1. PENDAHULUAN

SMK `Plus An-Naba Sukabumi merupakan sekolah menengah kejuruan negeri yang beralamat di Jl. Widyakrama No. 112, Sukabumi. SMK Plus An-Naba Sukabumi merupakan salah satu subsistem pendidikan nasional. Tanggung jawab yang terbesar diemban oleh SMK Plus An-Naba Sukabumi adalah membangun bangsa melalui pendidikan teknologi dan kejuruan. Karena itu SMK Plus An-Naba Sukabumi senantiasa berupaya melakukan pembaharuan pendidikan khususnya dalam bidang teknologi dan kejuruan. Dengan mulai diterapkannya kurikulum 2013 yang mengharapkan siswanya untuk berperan aktif dalam proses kegiatan belajar mengajar (KBM) maka penerapan TIK menjadi salah satu kebutuhan bagi setiap murid dan guru.

Berdasarkan hasil wawancara dengan seorang guru di SMK Plus An-Naba Sukabumi yaitu bapak Ruston Pirmasnyah, S.Kom. mengatakan bahwa proses KBM di SMK Plus An-Naba Sukabumi saat ini sudah cukup baik. SMK Plus An-Naba Sukabumi mulai membangun e-learning sejak tahun 2013. Dalam e-learning yang telah dibangun terdapat fitur latihan soal namun terbantas hanya latihan soal dalam bentuk pilihan ganda. Namun e-learning yang telah dibangun belum menyediakan fasilitas koreksi soal esai otomatis yang dibutuhkan oleh para guru di SMK Plus An-Naba Sukabumi. Pengkoreksian soal esai secara manual yang menghabiskan waktu lebih lama dibandingkan dengan mengoreksi soal esai secara otomatis. Dibutuhkan waktu sekitar 1 menit tiap soal untuk mengkoreksi soal esai secara manual. Jika akan mengkoreksi 30 soal maka dibutuhkan waktu sekitar 30 menit, sedangkan jumlah pelajar di SMK Plus An-Naba Sukabumi sekitar 200 orang.

Salah satu metode penilaian esai yang digunakan adalah Latent Semantyc Analysis (LSA) yaitu dengan merepresentasikan kalimat dalam sebuah matriks semantik yang kemudian dilakukan perhitungan matematis dengan mencocokan atau memetakan ada atau tidaknya kata dari kelompok kata pada matriks menggunakan teknik aljabar linear Singular Value Decomposition (SVD).

Dengan ditemukannya permasalahan yang ada, maka dibutuhkan metode yang dapat membantu guru untuk pengoreksian soal esai secara otomatis oleh sistem. Dengan dikembangkan aplikasi LMS ini, diharapkan bisa menjadi sebuah media yang mampu mempercepat proses pengkoreksian soal esai hingga pada akhirnya dapat memudahkan para pengajar di SMK Plus An-Naba Sukabumi. Berdasarkan latar belakang yang telah dipaparkan sebelumnya, maka penelitian tugas akhir ini diberi judul Pengembangan Koreksi Soal Esai Otomatis Pada E-Learning di SMK Plus An-Naba Sukabumi dengan Menggunakan Metode Latent Semantyc Analysis (LSA).

1.1 E-Learning

Perkembangan teknologi informasi yang semakin maju telah mampu membantu manusia

dalam segala bidang. Salah satu bidang yang tak luput pemanfaatan teknologi informasi dalam prosesnya adalah dalam bidang pendidikan. perkembangan yang semakin maju membuat pergeseran paradigm cara pembelajaran dan melahirkan metode-metode baru dalam prose pembelajaran tersebut. proses belajar mengajar saat ini dapat dilakukan dimanapun tanpa bertatap muka secara langsung dalam suatu ruangan. Interaksi keduanya dapat dilakukan dengan menggunakan bantuan media-media elektronik seperti komputer. Proses belajar mengajar dengan menggunakan bantuan media elektronik seperti ini sekarang dikenal dengan sebutan e-learning.

Pada dasarnya konsep e learning dapat dikelompokan berdasarkan basis teknologi yang digunakannya menjadi dua jenis, yaitu computer based training (CBT) yang menggunakan program komputer sebagai media utama yang digunakan oleh pelajar. Dan jenis satunya adalah web based training (WBT) yang merupakan perkembangan lanjutan dari sistem CBT dengan menggunakan tambahan teknologi berbasis internet. Pada jenis WBT penyampaian dan akses materi pelajaran dilakukan melalui internet.

1.2 *Learning Management Sistem (LMS)*

Dalam e-learning ada dua bagian utama yang menopang sistemnya, yaitu Learning Management Sistem dan e-learning content atau materi pelajaran yang akan dipelajari oleh pemakai. *Learning Management Sistem (LMS)* adalah sistem yang membantu dan berfungsi sebagai platform *e-learning content.*

1.3 Penilaian Jawaban Otomatis

Setiap pembelajaran membutuhkan sebuah evaluasi untuk mengetahui sejauh mana daya serap informasi yang dilakukan oleh peserta didik. Secara umum terdapat dua macam bentuk evaluasi yang dapat dilakukan dalam proses belajar mengajar yaitu secara lisan dan tertulis. Ujian secara lisan memiliki kelemahan yaitu penguji dan pelajar yang diuji harus berada pada satu ruang dan satu waktu yang sama dan membutuhkan waktu yang cukup banyak terutama jika peserta ujian tersebut berjumlah banyak. Sedangkan ujian tulisan jauh lebih efisien karna ujian dapat dilakukan dalam waktu bersamaan dengan jumlah peserta ujian yang lebih banyak.

Jenis ujian tulis dapat dibagi menjadi tiga macam, yaitu pilihan ganda, isian singkat, dan esai. Dari ketiga jenis ujian yang dapat dilakukan secara tertulis tersebut ujian yang paling mudah untuk diperiksa secara otomatis adalah pilihan ganda. Pada ujian jenis pilihan ganda pelajar cukup memilih jawaban yang dianggapnya paling benar dari beberapa pilihan jawaban yang ada. namun kelemahan dari model ujian pilihan ganda adalah kurang dapatnya penguji mengetahui sejauh mana pengetahuan pelajar yang diuji dan dimungkinkan adanya peserta jawaban yang mendapatkan nilai tinggi dari factor keberuntungan dengan menebak jawaban yang tersedia. Tingkat akurasi yang dapat dilakukan oleh komputer untuk ujian jenis pilihan ganda dapat mencapai 100% karna bentuk uijan seperti ini sangat mudah diolah komputer cukup dengan membandingkan jawaban pelajar dengan kunci jawaban yang ada dalam database.

Jenis ujian tertulis berikutnya adalah isian singkat. Pada jenis ujian isian singkat pelajar cukup mengisi jawaban atas pertanyaan yang diajukan dalam satu atau beberapa kata. Ujian jenis ini sedikit lebih baik dibandingkan pilihan ganda karna menuntut pelajar untuk mempunyai sebuah jawaban sendiri. Dan jenis ujian tertulis lainnya adalah dengan soal esai. Esai adalah suatu tulisan yang menggambarkan opini penulis tentang subyek tertentu yang coba dinilainya. Atau juga dapat diartikan sebagai karangan proga yang mengupas secara sepintas namun akurat, padat dan berisi mengenai masalah kesusastraan, seni dan budaya dari sudut pandang penulis secara subjektif. Menurut KBBI esai adalah karya tulis atau karangan dalam bentuk prosa yang memaparkan tentang suatu masalah dari sudut pandang penulis secara lugas dan sepintas. Bentuk soal esai merupakan bentuk pengujian yang secara akurat mampu menggambarkan pemahaman pelajar secara mendalam mengenai topik yang diujikan. Selain itu ujian esai juga mampu merangsang kemampuan mengemukakan pendapat dalam bentuk tulisan.

Ujian dalam bentuk esai memiliki kekurangan dari segi sumber daya yang dibutuhkan yaitu manusia dan waktu. Ketika sebuah ujian esai diikuti oleh banyak peserta waktu yang dibutuhkan untuk memeriksa jawaban akan bertambah pula. Untuk mengatasi masalah dalam penilaian ujian esai, sistem penilaian esai secara otomatis menggunakan komputer dapat menjadi sebuah solusi yang cukup baik.

1.4 Text Preprocessing

Text preprocessing adalah langkah awal sebelum dilakukan sebuah pencocokan string. Proses ini melalui beberapa tahap tindakan yang dilakukan pada sebuah string jawaban dan kunci yang diisi oleh guru dan siswa. Tahap-tahap yang dilakukan dalam proses ini adalah :

1. *1. Case folding / toLowerCase*

Case folding adalah tahapan dimana string yang telah diinput oleh user yang awalnya merupakan campuran dari huruf capital dan huruf kecil disamaratakan dengan dijadikan menjadi huruf kecil semua. Hal ini dilakukan karena system komputer membedakan antara “A” dan “a” sehingga antara “saya” dan “Saya” pun akan dinilai berbeda pada system. Tentunya hal ini akan berpengaruh pada

pencocokan string antar kunci dan jawaban dari siswa.

2. Penghilangan tanda baca
   Tahapan ini adalah penghapusan tanda baca sehingga string yang tersisa hanya terdiri dari huruf dan angka saja. Tahapan ini dilakukan karena system yang akan dibangun hanya mencocokan string dan angka belum melibarkan operasi matematika.
3. Penghapusan kata umum (*stopWord)*
   Tahapan *stopWord* adalah penghapusan kata-kata yang dinilai umum sehingga kata-kata tersebut tidak dimasukan kedalam array string yang akan dicocokan.
4. *Tokenizing*
   Tokenizing adalah tahapan pemotongan string dalam teks berdasarkan akta yang menyusunnya dan sisimpan dalam sebuah variable array.
5. *Stemming* (penghapusan awalan dan akhiran)
   Stemming adalah salah satu cara yang digunakan untuk meningkatkan performa pencarian dan pencocokan kata dengan cara mentransformasikan kata kata dalam sebuah dokumen teks ke bentuk dasarnya. Proses stemming dilakukan dengan menghilangkan semua imbuhan baik yang terdiri dari awalan, sisipan maupun akhiran dan kombinasi dari awalan dan akhiran. Ada banyak algoritma yang digunakan untuk melakukan proses stemming. Khusus untuk stemming bahasa indonesia algoritma yang sering digunakan adalah algoritma nazrief dan andriani, arifin dan setiono, algoritma veda, dll.

1.5 Matriks Singular Value Decomposition

SVD merupakan teorema aljabar linear yang disebut mampu untuk memecah blok suatu matriks A menjadi tiga matriks baru yaitu sebuah matriks orthogonal U, matriks diagonal S, dan transpose matriks orthogonal V. Teorema SVD adalah sebagai berikut :

$$Amn = Umm \times Smn \times V^{T}nn \quad (1)$$

Dimana $U^{T}U$= matriks identitas (I), $V^{T}V$ = matriks identitas (I). Kolom matriks U merupakan eigenvektor orthonormal dari AA^{T}. Sedangkan kolom matriks V merupakan eigenvektor orthonormal dari $A^{T}A$. dan S merupakan matriks diagonal akar dari nilai eigen dari matriks U atau V dalam urutan dari yang terbesar.

1.6 Cosine Similarity

Cosine similarity adalah salah satu metode perhitungan similarity yang paling popular untuk diterapkan pada dokumen teks. Kelebihan utama dari metode cosigne similarity adalah tidak terpengaruh pada panjang pendeknya suatu dokumen. Karna yang hanya diperhitungkan nilai term dari masing masing dokumen. Berikut adalah rumus untuk cosine similarity

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

Keterangan :
A = Matriks A
B = Matirks B
||A|| = Panjang Matriks A
||B|| = Panjang Matirks B

1.7 Latent Semantic Analysis (LSA)

Latent Semantic Analysis (LSA) adalah suatu metode perhitungan matematika teknik yang sepenuhnya otomatis dalam penggalian dan penyimpulan hubungan antara kata dan kalimat yang ada dalam paragraf [6]. Algoritma LSA merupakan salah satu algoritma pengembangan dalam bidang ilmu *Information Retrival* yang mampu menghimpun sejumlah besar dokumen dalam basis data dan mengghubungkan relasi antar dokumen dengan mencocokan *query* yang diberikan. Fungsi utama dari LSA ini adalah untuk menghitung kemiripan (*similarity*) dokumen dengan membandingkan representasi vektor tiap dokumen. Penilaian dengan metode LSA lebih kepada kata-kata yang ada dalam tulisan tanpa memperhatikan urutan kata dan tata bahasa dalam tulisan tersebut, sehingga suatu kalimat yang dinilai adalah berdasarkan kata-kata kunci yang ada pada kalimat tersebut.

Untuk menghitung nilai jawaban esai menggunakan metode LSA, caranya adalah dengan cara membuat sebuah dokumen esai yang menjadi acuan (*query*) atau kunci jawaban dan kemudian esai yang akan dinilai dibandingkan dengan esai acuan jawaban, semakin banyak kemiripannya maka semakin besar nilai jawabannya.

Langkah-langkah pemeriksaan esai dengan metode LSA adalah pertama tulisan direpresentasikan ke dalam matriks dimana baris matriks menunjukkan kata-kata kunci (*queries*) pada dokumen tersebut dan setiap kolom menunjukkan kolom dokumen. Setiap sel menunjukkan banyaknya kemunculan kata kunci yang berada pada matriks pada dokumen yang ada di kolom matriks. Kemudian isi dari sel tersebut terlebih dahulu ditransformasikan dimana setiap frekuensi kata dibobotkan dengan sebuah fungsi yang menunjukkan pentingnya sebuah kalimat dalam *paragraph* dan juga derajat yang menunjukkan seberapa pentingnya tipe kata didalam suatu kalimat.

Langkah selanjutnya adalah dilakukan dekomposisi *Singular Value Decomposition* (SVD)

pada matriks kata-dokumen. Pada SVD matriks memuat frekuensi pemunculan kata kunci di dekomposisi menjadi tiga buah matriks yang jika tiga buah matriks tersebut dikalikan maka akan muncul kembali matriks asalnya. Matriks pertamanya mendeskripsikan entitas kolom sebagai nilai vektor *orthogonal* matriks. Matriks keduanya berupa matriks diagonal yang memuat nilai skalar matriks. Secara matematis, faktor yang paling baik adalah menggunakan dimensi terkecil dari matriks awalnya, sehingga rekonstruksi matriks terbaik dihasilkan pada saat nilai faktor lebih kecil dari jumlah faktor yang digunakan.

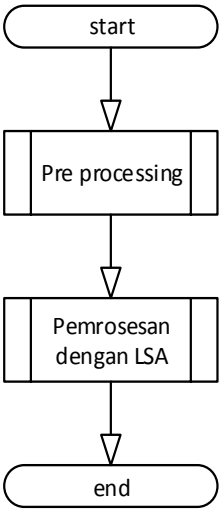
Dimensi dari matriks yang telah disederhanakan dengan menghapus koefisien pada matriks diagonal sebanyak yang diinginkan sampai tersisa koefisien sebanyak dimensi yang terpilih. Tujuan penyederhanaan ini adalah agar terbentuk matriks yang memuat nilai korelasi yang diinginkan ketika tiga buah matriks direkonstruksi. Kemudian penilaian akan dilakukan dengan membandingkan matriks korelasi dari *query* dengan matriks korelasi dari setiap dokumen dengan menggunakan perhitungan *cosinus α*.

2. ISI PENELITIAN

Bagian ini dapat dibagi dalam beberapa sub pokok pembahasan sesuai dengan kebutuhan tulisan. Tidak ada batasan yang baku mengenai jumlah pemerincian sub pokok bahasannya; tetapi setidaknya mengandung: metode, hasil, dan pembahasan.

Pengembangan koreksi esai dimasukan dalam fitur latihan dalam satu halaman bersama soal pilihan ganda. Pengembangan dilakukan karena ditemukan persoalan dalam sistem yang berjalan bahwa pengkoreksian jawaban esai secara manual akan membutuhkan waktu yang lama sebagaimana dijelaskan dalam latar belakang penelitian sehingga membutuhkan fitur tambahan yaitu koreksi soal esai otomatis dalam sistem. Pengembangan ditambahkan dalam bagian latihan soal dengan fitur koreksi soal esai otomatis yang mana dalam sistem sebelumnya sudah teradapat fitur latihan soal namun terabatas hanya soal pilihan ganda saja.

Sistem yang akan dikembangkan dalam penelitian ini adalah sebuah sistem koreksi soal esai otomatis dengan metode *Latent Semantyc Analysis*. Pada prosesnya sebelum sistem melakukan perhitungan dengan metode LSA ada proses proses yang dilakukan terlebih dahulu. Pengembangan koreksi esai ini terdiri dari dua subproses yaitu proses *preprocessing* dan proses LSA itu sendiri. Proses *preprocessing* terdiri dari *case folding*, *tokenizing*, *filtering*, dan *stemming*. Berikut adalah *flowchart* dari masing masing proses.

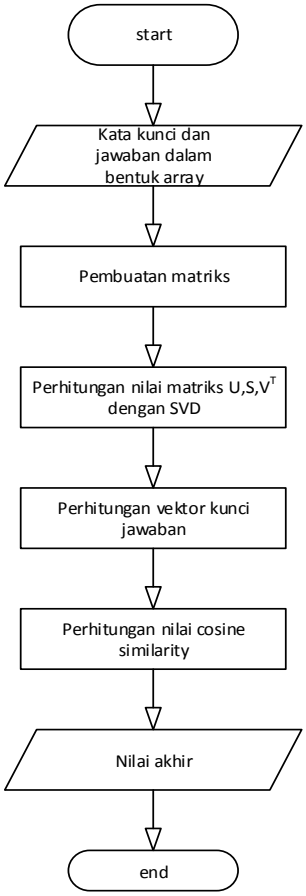


Gambar 1 *Flowchart* Sistem Koreksi Esai Otomatis

Untuk lebih memudahkan dalam pemahaman dalam prosesnya diberikan sebuah contoh pencocokan antara kunci jawaban dan jawaban siswa. Dalam contoh berikut adalah salah satu ulangan yang diadakan yang berjumlah 2 soal esai mata pelajaran bahasa indonesia. Jika nilai sempurna dalam ulangan ini adalah 10 maka setiap soal memiliki bobot 5 dalam penyusunan nilai tersebut.

2.1 Analisis metode LSA

Tahapan yang terdapat dalam proses LSA terdiri dari empat tahap dimana masing-masing tahap ada proses perhitungan matematisnya. Proses ini adalah kelanjutan dari proses *preprocessing* yang telah dilakukan sebelumnya untuk mendapatkan nilai kemiripan antara jawaban dan kunci jawaban.



Gambar 2 *Flowchart* proses LSA

2.2 Perhitungan nilai matriks U, S, VT

Setelah terbentuk sebuah matriks berukuran *m* x *n* maka langkah selanjutnya adalah dengan mencari nilai matriks U, S, V^T dengan metode SVD (*Singular Value Decomposition*). Untuk mendapatkan matriks-matriks berikut melalui beberapa perhitungan yang cukup kompleks mulai dari melakukan transpose matriks, mencari nilai eigen dan mencari vektor eigennya. Namun dalam implementasi program, proses perhitungan dapat menggunakan library dari JAMA yang telah menyediakan fungsi untuk melakukan perhitungan matriks sebagai pengganti MATLAB sehingga didapatkan matriks U, S, V^T.

Dari tabel 3.7 didapatkan sebuah matriks A dengan dimensi 9x4 yang akan dilakukan proses SVD untuk mencari nilai U, S, V^T . langkah pertama untuk mencari matriks tersebut adalah dengan mencari matriks A^T terlebih dahulu.

$$A \begin{vmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix}$$

$$\begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{vmatrix}$$

$$A^T \begin{vmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{vmatrix}$$

Langkah berikutnya adalah dengan mencari nilai $A*A^T$ dan A^T*A sehingga didapatkan matriks berikut:

$$A^T * A \begin{vmatrix} 7 & 4 & 4 & 4 \\ 4 & 5 & 2 & 2 \\ 4 & 2 & 5 & 3 \\ 4 & 2 & 3 & 5 \end{vmatrix}$$

$$A*A^T \begin{vmatrix} 3 & 1 & 2 & 2 & 2 & 1 & 3 & 1 & 2 \\ 1 & 2 & 1 & 0 & 2 & 1 & 2 & 0 & 1 \\ 2 & 1 & 3 & 2 & 2 & 2 & 3 & 1 & 1 \\ 2 & 0 & 2 & 2 & 1 & 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 & 3 & 2 & 3 & 0 & 2 \\ 1 & 1 & 2 & 1 & 2 & 2 & 2 & 0 & 1 \\ 3 & 2 & 3 & 2 & 3 & 2 & 4 & 1 & 2 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 2 & 1 & 1 & 1 & 2 & 1 & 2 & 0 & 2 \end{vmatrix}$$

Setelah didapatkan hasil perkalian antara matriks A dan transpose dari matriks A diatas maka langkah selanjutnya adalah mencari nilai eigen dengan cara mengurangi matriks tersebut dengan *lambda* dikalikan dengan matriks identitas. Yang pertama adalah mencari nilai eigen dari matriks A^T*A.

$$\text{Nilai eigen} = (A^T * A) - \lambda I$$

$$\begin{vmatrix} 7 & 4 & 4 & 4 \\ 4 & 5 & 2 & 2 \\ 4 & 2 & 5 & 3 \\ 4 & 2 & 3 & \end{vmatrix} - \begin{vmatrix} \lambda & 0 & 0 & 0 \\ 0 & \Lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{vmatrix} \begin{vmatrix} 7 - \lambda & 4 & 4 & 4 \\ 4 & 5 - \lambda & 2 & 2 \\ 4 & 2 & 5 - \lambda & 3 \\ 4 & 2 & 3 & 5 - \lambda \end{vmatrix}$$

Dari perhitungan diatas dilanjutkan dengan mencari persamaan polynomial (banyak suku) dari matriks diatas. Caranya adalah dengan mengalikan setiap elemen dikurangi elemen lain dari sisi sebelahnya. Sehingga didapakan persamaan berikut :

$$\lambda^4 - 21\lambda^3 + 300\,\lambda^2 - 858\,\lambda + 1375 = 0$$

Setelah didapatkan sebuah persamaan *polynomial* seperti diatas langkah selanjutnya adalah mencari akar dari persamaan tersebut sehingga didapatkan

$\lambda_1 = 15{,}400$ $\lambda_2 = 1{,}235$

$\lambda_3 = 3.366$ $\lambda_4 = 2{,}000$

Sedangkan untuk matriks $A*A^T$ dengan melakukan proses yang sama didapatkan nilai eigen sebagai berikut.

$\lambda_1 = 15{,}400$ $\lambda_2 = 3{,}366$

$\lambda_3 = 2{,}000$ $\lambda_4 = 1{,}235$

$\lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9 = 0{,}000$

setelah didapatkan nilai eigen dari masing-masing persamaan diatas langkah selanjutnya adalah dengan memasukan nilai λ pada persamaan awal yang kemudian dinormalisasi sehingga didapatkan matriks-matriks berikut. Untuk vector eigen dari matriks A^T*A akan menjadi nilai V yang kemudian akan di transpose menjadi V^T. Dan vektor nilai eigen dari $A*A^T$ akan menjadi nilai matriks U. Sedangkan matriks S didapatkan dari diagonal akar nilai eigen yang sudah diurutkan dari yang terbesar dan bukan nol.

Matriks U

0.386	-0.248	-0.5	-0.025
0.233	0.483	0	-0.533
0.386	-0.248	0.5	-0.025
0.269	-0.489	0	0.251
0.395	0.413	0	0.131
0.279	0.171	0.5	0.408
0.502	-0.006	0	-0.302
0.107	-0.419	0	-0.433
0.279	0.171	-0.5	0.408

Matriks S

3.924	0	0	0
	1.835	0	0
0	0	1.414	0
0	0	0	1.111

Matriks V^T

0.636	0.421	0.457	0.457
0.129	0.769	-0.443	-0.443
0	0	-0.707	-0.707
0.761	-0.482	-0.307	-0.307

Setelah didapatkan ketiga matriks tersebut langkah selanjutnya adalah dengan melakukan penyederhanaan kolom-kolom matriks. Penyederhanaan dilakukan pada matriks U dengan penyederhanaan pada jumlah kolomnya, matriks S pada baris dan kolomnya. Langkah penyederhanaan ini dilakukan sesuai dengan nilai dimensi dari matriks tersebut. sehingga didapatkan matriks U dan S dengan dimensi baru sebagai berikut.

Matriks U

0.386	-0.248
0.233	0.483
0.386	-0.248
0.269	-0.489
0.395	0.413
0.279	0.171
0.502	-0.006
0.107	-0.419
0.279	0.171

Matriks S

3.924	0
0	1.835

2.3 Perhitungan vektor kunci jawaban

Pencarian matriks vektor Q dan tiap-tiap D dilakukan untuk melakukan proses perhitungan *cosine similarity* di tahap berikutnya. Pada langkahnya tahap ini melakukan perkalian matriks dengan himpunan *term* frekuensi masing-masing *query* yang di*transpos*e terhadap matriks U yang kemudian dikalikan kembali dengan invers dari matriks S (S^{-1}). perhitungan berikut adalah contoh dari perhitungan vektor kunci jawaban (Q).

$$\bar{q} = Q^T . U . S^{-1} \qquad (2)$$

$$\bar{q} = \begin{vmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & \end{vmatrix} \begin{vmatrix} 0.386 & -0.248 \\ 0.233 & 0.483 \\ 0.389 & -0.248 \\ 0.269 & -0.489 \\ 0.395 & 0.413 \\ 0.279 & 0.171 \\ 0.502 & -0.006 \\ 0.107 & -0.419 \\ 0.279 & 0.171 \end{vmatrix} \begin{vmatrix} 0.255 & 0.00 \\ 0.000 & 0.545 \end{vmatrix}$$

$$\bar{q} = \bar{q} = | \ 0.617 \ \ -0.135 \ |$$

Jadi hasil perhitungan matriks diatas didapatkan sebuah vektor dari *query* jawaban (Q) adalah $\bar{q}$ = | 0.617 -0.135 |. Dengan melakukan perhitungan yang sama maka masing-masing jawaban akan menghasilkan nilai vektor sebagai berikut :

Nilai vector Q = | 0.617 - 0.135 |
Nilai vector D1 = | 0.407 -0.813 |
Nilai vector D2 = | 0.444 0.470 |
Nilai vector D3 = | 0.445 0.470 |

2.4 Perhitungan nilai cosine similarity

Perhitungan nilai *cosine similarity* dilakukan guna mendapatkan nilai atau tingkat kemiripan antar dua buah vektor. Dalam hal ini vektor yang akan dilakukan perhitungan adalah vektor antara kunci jawaban dan masing masing jawaban siswa. Perhitungan ini dilakukan dengan menggunakan rumus pada persamaan dibawah ini.

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}} \quad \textbf{(3)}$$

$$D1 = \frac{(Q_1 . D_{11}) + (Q_2 . D_{12})}{\sqrt{(Q_1)^2 + (D_{11})^2} \ x \sqrt{(Q_2)^2 + (D_{12})^2}}$$

Dengan :

D = jawaban siswa

Q = *query* jawaban

A = Matriks A

B = Matirks B

||A|| = Panjang Matriks A

||B|| = Panjang Matriks B

Sehingga D1 = $\frac{(0{,}617)(0{,}407)+(-0.135)(-0.813)}{\sqrt{(0{,}617)^2+(0{,}407)^2}\ \sqrt{(-0.135)^2+(-0.813)^2}} = 0.644$

Sehingga didapatkan nilai *cosine similarity* dari D1 sebesar 0.644 . Untuk nilai yang lain juga dilakukan dengan proses yang sama sampai didapatkan masing-masing nilai *cosine similarity* sebagai berikut :

D2 = $\frac{(0{,}617)(0{,}444)+(-0.135)(0{,}470)}{\sqrt{(0{,}617)^2+(0{,}444)^2}\ \sqrt{(-0.135)^2+(0{,}470)^2}} = 0.566$

D3 = $\frac{(0{,}617)(0{,}445)+(-0.135)(0.470)}{\sqrt{(0{,}617)^2+(0{,}445)^2}\ \sqrt{(-0.135)^2+(0.470)^2}} = 0.566$

Setelah didapatkan nilai *similarity* dari masing-masing jawaban maka tahap selanjutnya adalah perkalian dengan bobot dari setiap soal yang diujikan. Dalam soal ini memiliki bobot 5 dari 10 point yang ada dalam ulangan dari dua soal, sehingga ketika jawaban dari siswa memiliki kesamaan sempurna dengan kunci yang diberikan akan memiliki 5 point pada soal tersebut.

Jadi nilai rekomendasi yang didapatkan oleh guru dari fungsionalitas koreksi soal pada contoh soal model ini adalah sebagai berikut :

Nilai rekomendasi = nilai *cosine similarity* * bobot tiap soal

D1 = 0.644 * 5 = 3.2
D2 = 0.566 * 5 = 2.62
D3 = 0.566 * 5 = 2.62

Sehingga menghasilkan Nilai rekomendasi dari 3 siswa yang telah mengerjakan soal esai dengan jawaban berbeda yakin sebagai berikut :
Siswa pertama atau D1 menghasilkan nilai rekomendasi 3.20 dibulatkan dengan 1 angka desimal menjadi 3.2, Siswa kedua atau D2 menghasilkan nilai rekomendasi nilai rekomendasi 2.62 dibulatkan dengan 1 angka desimal menjadi 2.6 dan Siswa ketiga atau D3 menghasilkan nilai rekomendasi Nilai rekomendasi 2.62 dibulatkan dengan 1 angka desimal menjadi 2.6

3 PENUTUP

Berdasarkan dari uraian dan hasil analisis yang telah dilakukan maka dapat diambil kesimpulan bahwa Dengan adanya fungsionalitas koreksi esai otomatis guru mampu melaksanakan proses koreksi soal jenis esai dengan lebih cepat. Sebelumnya dibutuhkan waktu sekitar 1 menit tiap soal untuk mengkoreksi soal esai secara manual.

Jika akan mengkoreksi 30 soal maka dibutuhkan waktu sekitar 30 menit, sedangkan jumlah pelajar di SMK Plus An-Naba Sukabumi sekitar 200 orang. Setelah pengembangan dibutuhkan waktu sekitar 5 detik tiap soal untuk mengoreksi. Jika akan mengkoreksi 30 soal maka membutuhkan waktu sekitar 150 detik atau 2,5 Menit, sedangkan jika akan mengkoreksi 200 soal maka membutuhkan waktu sekitar 1000 detik atau 5 menit

Untuk pengembangan fungsionalitas koreksi esai otomatis, saran untuk penelitian berikutnya agar menggunakan metode selain *Latent Semantyc Analysis* (LSA) dalam melakukan koreksi soal esai otomatis.

DAFTAR PUSTAKA

[1] M. P. B. S. Sahriar Hamza, "Sistem Koreksi Soal Essay Otomatis Dengan Menggunakan Metode Rabin Karp," EECCISS, vol. 7, no. 2, p. 153, 2013.

[2] B. B. D. H. Anak Agung Putri Ratna, "SIMPLE : SISTEM PENILAI OTOMATIS UNTUK MENILAI UJIAN DALAM BAHASA INDONESIA," MAKARA TEKNOLOGI, vol. 11, no. 1, pp. 5-11, 2007.

[3] R. S.Pressman, Software Enginering A Practitioners Approach, New York: McGraw-Hill, 2010.

[4] Munir, "KONTRIBUSI TEKNOLOGI INFORMASI DAN KOMUNIKASI (TIK) DALAM PENDIDIKAN DI ERA GLOBALISASI PENDIDIKAN INDONESIA," Indonesian Comunity on Information and Comunicartion Technology, vol. 2, no. 2, 2009.

[5] D. E. Hartley, Selling e-Learning, ASTD, 2006.

[6] P. W. D. L. Thomas K Launder, An Introduction to Latent Semantic Analysis, 1998.

[7] M. S. Rossa A.S M, Rekayasa Perangkat Lunak terstruktur dan berorientasi objek, Bandung: Informatika, 2013.

[8] A. Nugroho, Analisis dan Perancangan Sistem Informasi dengan Metodologi Berorientasi Objek, Bandung: Informatika, 2005.