

## **PENILAIAN OTOMATIS TERHADAP JAWABAN ESAI PADA SOAL BERBAHASA INDONESIA MENGGUNAKAN LATENT SEMANTIC ANALYSIS**

**Diyah Yustiana**  
Teknik Informatika  
Sekolah Tinggi Teknik Surabaya  
die.imut@gmail.com

### **ABSTRAK**

Banyak software-software e-learning atau quiz online yang dapat membantu para pengajar untuk melakukan evaluasi pembelajaran (ujian) dengan hasil yang cepat, akan tetapi hampir seluruhnya masih belum memiliki fasilitas penilaian otomatis untuk jenis soal esai. Sehingga dalam melakukan penilaian masih diperlukan peran serta pengajar untuk mengkoreksinya secara manual. Maka saat ini sudah mulai banyak dikembangkan metode-metode agar penilaian jawaban esai ini dapat dilakukan secara otomatis oleh komputer.

Dalam penelitian ini, akan menggunakan Metode LSA dan Euclidean Distance untuk melakukan penilaian jawaban esai berbahasa Indonesia. LSA akan digunakan untuk menangani kesamaan katanya, sedangkan Euclidean Distance untuk mengukur kemiripan antara kunci jawaban dan jawaban siswa. Langkah awal penelitian ini akan mengekstraksi terlebih dahulu masing-masing dokumen tersebut dengan melakukan proses preprocessing, yang meliputi: tokenisasi, stopword removal, stemming, dan term weighting yang pada penelitian ini menggunakan perhitungan TF-IDF suatu kata. Untuk kunci jawaban, akan menghasilkan sebuah matrik query yang merepresentasikan kolom untuk setiap nomor soal dan baris untuk setiap kata. Sedangkan untuk jawaban siswa, akan menghasilkan sekumpulan matrik dokumen yang setiap matriknya merepresentasikan kumpulan setiap nomor jawaban siswa. Dimana kolom matrik dokumen berisikan tiap-tiap siswa yang menjawab soal, sedangkan barisnya berisi term.

Matrik-matrik tersebut akan dijadikan bahan inputan dalam proses LSA. Dimana dalam proses ini matrik-matrik tersebut akan direduksi. Matrik hasil proses LSA inilah yang akan dijadikan vektor dokumen dan vektor query, yang akan diukur kemiripannya dengan menggunakan Euclidean Distance. Hasilnya akan dinormalisasi untuk menghasilkan output skoring 0 sampai dengan 4. Hasil uji coba yang dilakukan menunjukkan sistem ini mampu melakukan proses penilaian secara otomatis terhadap jawaban esai berbahasa Indonesia dengan nilai korelasi sebesar 71%.

Kata kunci: AES, LSA, Euclidean Distance

### **ABSTRACT**

*A lot of e-learning software or online quizzes can help teachers to evaluate faster learning (exams). However, an automatic scoring for essays is still rare to find. In writing, teacher still evaluate manually, therefore, an automatic writing evaluation is needed.*

*This study made use of the LSA and the Euclidean Distance method for the assessment of response essay in Indonesian language. LSA is used to handle the similarity of words, while the Euclidean Distances is used to measure the similarity between answers and an answer key to students. The initial step of this study is to extract each document by preprocessing process, including: tokenization, stopword removal, stemming, and term weighting (in this study using the TF-IDF calculation). For the answer key, a matrix query for each column is produced representing the number of questions and lines for each word. For the student's answers, a set of documents is produced in each matrix representing a collection of student's answers. The column matrix contains documents of each student who answered the question, while the row contains the term.*

*The matrices will be used as input in the LSA process. In this process, matrices will be reduced. Matrix results of the LSA process will be used as a document vector and a query vector measured by using the Euclidean Distance similarity. The result will be normalized to produce scoring output of 0 to 4. Results of experiments show that this system is able to perform automatic assessment process to answer essay in Indonesian language with correlation value of 71%.*

*Keywords: AES, LSA, Euclidean Distance*

## I. PENDAHULUAN

Seiring dengan pesatnya kemajuan dunia teknologi informasi saat ini, maka dunia pendidikanpun sudah turut serta mengalami pergeseran makna. Dimana untuk melakukan sebuah kegiatan pembelajaran saat ini sudah dapat dilakukan tanpa dibatasi oleh ruang dan waktu lagi. Dengan e-learning, kapan dan dimanapun guru dan siswa bisa terus berinteraksi untuk melakukan kegiatan pembelajaran dan juga evaluasinya sekaligus.

Sudah banyak *LMS (Learning Managemen System)* yang bisa digunakan untuk melakukan e-learning saat ini. Salah satu diantaranya adalah Edmodo. Edmodo dianggap sudah mampu melakukan kegiatan pembelajaran jarak jauh sekaligus sebagai media jejaring sosial antara guru dan siswa. Edmodo juga bisa melakukan proses evaluasi dan penilaiannya secara otomatis. Berkaitan dengan evaluasi, Edmodo juga sudah memiliki fasilitas untuk menyimpan kumpulan soal beserta kunci jawabannya, mengacaknya, mengatur waktu pelaksanaan ujian secara otomatis, mengatur lamanya pelaksanaan ujian, dan mengolah hasilnya secara langsung. Jenis soal yang disediakan juga bervariasi, mulai dari: pilihan ganda, benar/salah, menjodohkan, jawaban singkat, sampai dengan jawaban uraian (esai).

Akan tetapi sayangnya, Edmodo masih belum bisa memberikan penilaian secara langsung untuk soal jawaban uraian atau esai. Jadi, untuk jenis soal ini, Edmodo masih perlu melibatkan guru untuk mengkoreksi dan menilai secara langsung. Hal inilah yang biasanya bisa menimbulkan permasalahan baru dalam elearning. Dimana untuk mengkoreksi jawaban esai bukanlah perkara mudah. Beberapa hal yang menyebabkan penilaian jawaban esai tidak akurat, diantaranya adalah: tingkat subyektivitas guru pada saat melakukan pengkoreksian jawaban esai sangat berperan, keakurasian penilaian jawaban esai juga bisa sangat lemah jika jumlah jawaban yang harus dikoreksi banyak, waktu yang diperlukan lebih lama karena harus membaca jawaban siswa satu persatu,

dan guru harus selalu berinteraksi dengan Edmodo atau online selama proses penilaian berlangsung.

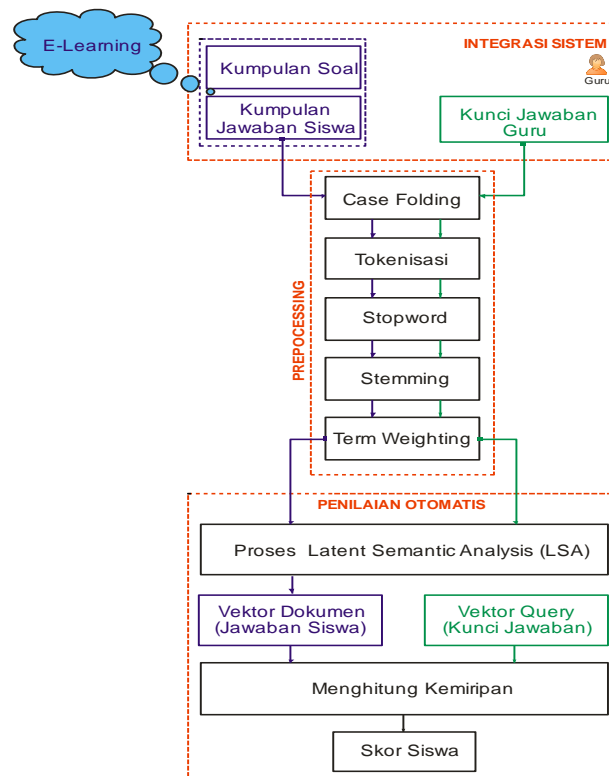
Mengingat latar belakang yang sudah dikemukakan di atas, maka penulis yang kebetulan berprofesi sebagai seorang guru ingin melakukan kajian pustaka dan penelitian lebih jauh tentang sistem penilaian otomatis terhadap dokumen teks berbahasa Indonesia berupa jawaban esai. Sistem yang akan dirancang tidak akan menangani pengambilan jawaban siswa secara langsung dari sistem LMS. Metode yang akan digunakan adalah LSA (Latent Semantic Analysis) dan Euclidean Distance.

Data input yang dipergunakan dalam penelitian ini berupa jawaban esai siswa pada soal berbahasa Indonesia, khususnya pada hasil ujian online beberapa Mata Pelajaran Produktif (Sistem Operasi, Pemrograman Dasar, Pemrograman Web, Jaringan Dasar, dan Simulasi Digital) pada Program Keahlian TKJ Kelas X di elearning SMK Negeri 1 Kediri tahun pelajaran 2013-2014.

Data yang digunakan terdiri dari 25 soal dengan jumlah responden sebanyak 25 siswa yang diambil secara acak dari 64 siswa Kelas X TKJ tahun pelajaran 2013-2014 dengan ketentuan untuk panjang jawaban yang digunakan dalam penelitian ini minimal  $\pm 300$  kata per jawaban siswa. Output yang diharapkan dari sistem ini adalah berupa skor nilai akhir siswa yang berkisar antara angka 0 – 4.

## II. ARSITEKTUR SISTEM

Pada bagian ini akan dijelaskan tahap-tahapan yang akan dilakukan oleh sistem dalam penelitian ini.



**Gambar 1. Arsitektur Sistem**

Jika dilihat dari arsitektur sistem yang dirancang pada gambar 1, maka sistem secara garis besar memiliki 3 tahapan utama, yaitu:

#### 1. Integrasi Sistem

Pada tahapan ini, data jawaban esai siswa di-download dari e-learning yang dipergunakan untuk melakukan ujian online untuk di-ekspor ke dalam database sistem yang akan digunakan sebagai data input. Data langsung diambil dari website [www.edmodo.com](http://www.edmodo.com). Pada tahap ini, guru juga akan memasukkan kunci jawaban dari tiap soal yang dipergunakan dalam inputan data sistem.

#### 2. Preprocessing

Pada tahap ini, akan dilakukan preprocessing terhadap semua jawaban esai siswa yang dipergunakan oleh sistem dan kunci jawabannya secara terpisah. Dalam tahap preprocessing ini, data yang berupa teks tersebut akan dilakukan beberapa proses secara bertahap, yaitu: case folding, tokenisasi, stopword removal, stemming dan pembobotan term dengan menggunakan perhitungan TF-IDF.

Hasil output pada tahap ini adalah berupa matrik-matrik dokumen yang merepresentasikan kumpulan jawaban siswa per nomor soal, dan matrik query yang merepresentasikan kunci jawaban guru. Untuk setiap matriks dokumen mewakili kumpulan satu nomor jawaban esai siswa, dimana untuk masing-masing matrik tersebut berisikan baris yang mewakili kata kunci (term) dan kolom yang mewakili setiap siswa yang mengerjakan soal pada nomor tersebut. Sedangkan untuk matrik query mewakili seluruh kunci jawaban soal, dimana baris mewakili kata kunci dan kolom mewakili jawaban setiap nomor soal.

#### 3. Penilaian Otomatis

Tahap ini adalah tahapan utama dalam sistem yang dirancang, dimana pada tahap ini akan dilakukan proses pengindeksan dengan menggunakan metode LSA pada matriks-matriks dokumen dan matriks query yang dihasilkan dari proses preprocessing sebelumnya. Hasil dari proses LSA tersebut akan dapat dilihat tingkat persamaan makna kata yang terkandung dalam semua jawaban siswa dan kunci jawaban.

Setelah semua matriks dokumen dan matriks query diproses dengan LSA, maka matriks-matriks tersebut akan dirubah ke dalam bentuk vektor dokumen dan vektor query agar dapat diukur tingkat kemiripannya. Vektor dokumen dibentuk berdasar tiap-tiap kolom pada matriks dokumen, yang berarti nilai-nilai vektor tersebut berisi nilai-nilai persiswa pada setiap nomor soal masing-masing. Sedangkan vektor query dibentuk berdasarkan nilai-nilai pada setiap nomor soal.

Pada penelitian ini persamaan yang digunakan untuk mengukur kemiripan 2 vektor adalah Euclidean Distance, yaitu dengan cara menghitung kedekatan antara 2 vector. Setelah itu hasil perhitungan Euclidean Distance tersebut akan dinormalisasi. Semakin kecil skor hasil normalisasi  $\bar{d}(u, v)$ , maka semakin mirip pula kedua vector tersebut. Sebaliknya, semakin besar skor hasil  $\bar{d}(u, v)$ , maka semakin berbeda kedua vektor tersebut. Sifat dari Euclidean ternormalisasi adalah bahwa hasilnya berada pada rentang  $0 \leq \bar{d}(u, v) \leq 2$ . Sehingga untuk memperoleh 5 klasifikasi penilaian yang diinginkan, maka terlebih dahulu harus menghitung skor intervalnya.

Dari hasil perhitungan interval, maka rentang skala klasifikasi skor penilaian yang akan dihasilkan adalah sbb:

Skor 0 = untuk jawaban kategori salah, jika  $\bar{d}(u, v) = 1,61 - 2,00$

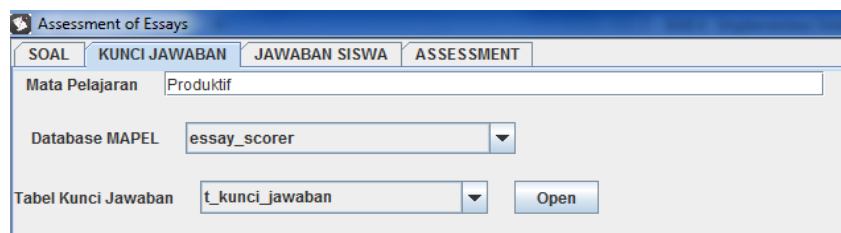
Skor 1 = untuk jawaban kategori kurang, jika  $\bar{d}(u, v) = 1,21 - 1,60$

Skor 2 = untuk jawaban kategori cukup, jika  $\bar{d}(u, v) = 0,81 - 1,20$

Skor 3 = untuk jawaban kategori cukup baik, jika  $\bar{d}(u, v) = 0,41 - 0,80$   
 Skor 4 = untuk jawaban kategori baik, jika  $\bar{d}(u, v) = 0,00 - 0,40$

### III. HASIL DAN PENGUJIAN

Tampilan program aplikasi terdiri dari 4 tab (lihat gambar 2), yaitu: Soal, Kunci Jawaban, Jawaban Siswa, dan Penilaian Otomatis. Pada Tab Kunci Jawaban dan Tab Jawaban Siswa akan dilakukan preprocessing data, perhitungan SVD, dan penyimpanan vektor, yaitu: vektor query yang akan disimpan pada tabel goldenanswersvquery dan vektor dokumen pada tabel studentanswersvquery[no\_soal] pada database sistem "essay\_scorer".



**Gambar 2. Interface Tab Menu Sistem**

Sedangkan pada Tab Assessment (penilaian otomatis) akan dilakukan proses perhitungan Euclidean Distance dan menormalisasikan hasilnya. Sehingga pada tab ini juga dapat menampilkan hasil skor akhir siswa. Pada tab ini juga disediakan fasilitas ekspor hasil tabel perhitungan skor akhir tersebut ke dalam file csv, dengan harapan output sistem tersebut dapat diolah atau dimanfaatkan lebih jauh oleh user.

Untuk mengevaluasi performa sistem yang sudah dibuat akan digunakan perhitungan korelasi, antara hasil penilaian sistem dan hasil penilaian manual. Adapun persamaan yang digunakan untuk menghitung nilai korelasi tersebut adalah sbb:

$$r_{xy} = \frac{N \sum XY - (\sum X) (\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}} \dots \dots \dots (1)$$

Berikut ini adalah contoh perhitungan korelasi terhadap 25 jawaban siswa, khususnya pada soal nomor 1 secara rinci.

**Tabel 1. Korelasi Soal Nomor 1**

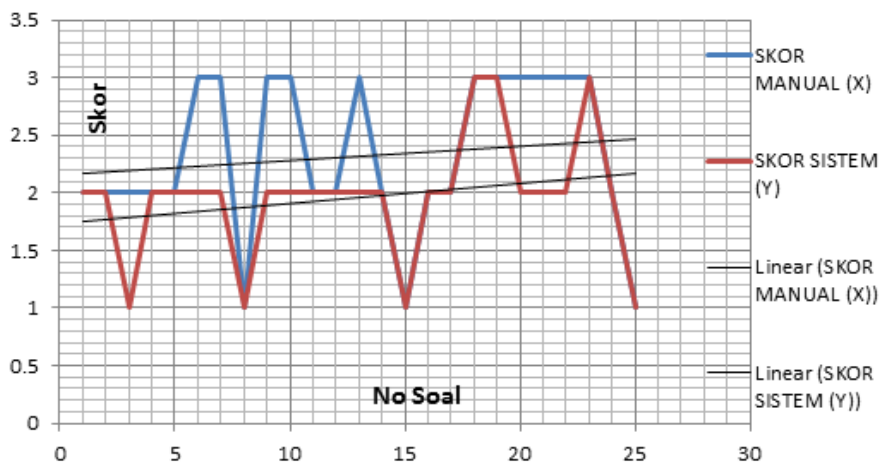
Siswa Ke-	SKOR MANUAL (X)	SKOR SISTEM (Y)	X <sup>2</sup>	Y <sup>2</sup>	XY
1	2	2	4.0000	4.0000	4
2	2	2	4.0000	4.0000	4
3	2	1	4.0000	1.0000	2
4	2	2	4.0000	4.0000	4
5	2	2	4.0000	4.0000	4
6	3	2	9.0000	4.0000	6

7	3	2	9.0000	4.0000	6
8	1	1	1.0000	1.0000	1
9	3	2	9.0000	4.0000	6
10	3	2	9.0000	4.0000	6
11	2	2	4.0000	4.0000	4
12	2	2	4.0000	4.0000	4
13	3	2	9.0000	4.0000	6
14	2	2	4.0000	4.0000	4
15	1	1	1.0000	1.0000	1
16	2	2	4.0000	4.0000	4
17	2	2	4.0000	4.0000	4
18	3	3	9.0000	9.0000	9
19	3	3	9.0000	9.0000	9
20	3	2	9.0000	4.0000	6
21	3	2	9.0000	4.0000	6
22	3	2	9.0000	4.0000	6
23	3	3	9.0000	9.0000	9
24	2	2	4.0000	4.0000	4
25	1	1	1.0000	1.0000	1
$\Sigma$	<b>58</b>	<b>49</b>	<b>146.0000</b>	<b>103</b>	<b>120</b>

Dari data pada tabel 1 dan persamaan di atas, maka dapat dilakukan perhitungan korelasi sebagai berikut:

$$r_{xy} = \frac{25 \cdot 120 - 58 \times 49}{\sqrt{(25 \times 146 - 58^2)(25 \times 103 - 49^2)}} = \frac{158}{\sqrt{(286)(174)}} = 0.7085 \dots \dots \dots (2)$$

Hubungan antar variabel pada tabel 1 dapat diimplementasikan dengan sebuah grafik, lihat gambar 3. Jika titik-titik pertemuan antara 2 variabel membentuk garis lurus menyudut, maka menunjukkan adanya korelasi tinggi positif, artinya ada hubungan yang erat antara kedua variable yang dikorelasikan atau memiliki kesejajaran yang tinggi.



Gambar 3. Grafik Perbandingan Soal Nomor 1

Berdasarkan perhitungan nilai rata-rata korelasi seluruh data input sistem pada tabel 2, maka nilai rata-rata korelasi pada penelitian ini mencapai sebesar 0.71 atau jika dituliskan dalam prosentase mencapai 71%.

**Tabel 2. Rata-Rata Korelasi dari Semua Soal**

Soal No	Nilai Korelasi	Interpretasi
1	0.7083	Cukup
2	0.7119	Cukup
3	0.6356	Cukup
4	0.6765	Cukup
5	0.7459	Cukup
6	0.7964	Cukup
7	0.7083	Cukup
8	0.6356	Cukup
9	0.7964	Cukup
10	0.7559	Cukup
11	0.6218	Cukup
12	0.6356	Cukup
13	0.7083	Cukup
14	0.7559	Cukup
15	0.6356	Cukup
16	0.7119	Cukup
17	0.7161	Cukup
18	0.7119	Cukup
19	0.6356	Cukup
20	0.7161	Cukup
21	0.7119	Cukup
22	0.6765	Cukup
23	0.7083	Cukup
24	0.7964	Cukup
25	0.7119	Cukup
<b>RATA-RATA</b>	<b>0.7054</b>	<b>Cukup</b>

## V. PENUTUP

Berdasarkan sistem yang sudah dibuat beserta hasil uji coba yang telah dilakukan dalam penelitian ini, maka dapat ditarik kesimpulan bahwa adanya beberapa proses yang mempengaruhi secara langsung performa pada LSA, yaitu: stopword removal, pembobotan kata, pengurangan dimensi, metode pengukuran kemiripan, dan banyaknya data (korpus) yang akan diolah. Stopword removal dapat mengurangi banyaknya data yang akan diolah, karena proses ini berfungsi untuk menghilangkan kata-kata yang dianggap tidak penting. Cara pembobotan kata yang digunakan akan dapat mempengaruhi kecepatan sistem, hal ini disebabkan karena pengaruh banyaknya perhitungan yang harus dilakukan dengan banyaknya data yang akan dihitung. Metode pengukuran kemiripan yang dipilih juga akan mempengaruhi hasil akhir sistem, karena berpengaruh langsung dengan penentuan nilai output yang diharapkan.

Dalam penelitian ini, sistem yang dirancang telah berhasil mencapai tingkat korelasi hingga 71% dengan kategori cukup. Maka untuk meningkatkan performa dari

sistem penilaian otomatis ini, disarankan bagi penelitian selanjutnya untuk mencoba: 1) berinovasi dengan beberapa teknik yang ada pada proses preprocessingnya, khususnya pada metode pembobotan kata (bisa dilakukan dengan teknik GfIDF atau Probabilistic IDF), dan metode pengukuran kemiripannya (bisa dilakukan dengan Cosin Similarity atau Language Based Matching); 2) Menggunakan beberapa macam alternatif kunci jawaban, sehingga perhitungan pembobotan kata untuk setiap soal pada kunci jawaban bisa dilakukan dengan menggunakan tema yang sama; 3) Menambahkan sebuah proses integrasi terhadap software-software elearning yang sudah ada, sehingga kinerja sistem ini akan jauh lebih baik dan benar-benar dapat sepenuhnya meringankan tugas para pengajar dalam melakukan proses evaluasi yang cepat dan akurat

## VI. DAFTAR PUSTAKA

- [1] Baker, Kirk. *Singular Value Decomposition Tutorial*. 2005. [http://www.ling.ohio-state.edu/~kbaker/pubs/Singular\\_Value\\_Decomposition\\_Tutorial.pdf](http://www.ling.ohio-state.edu/~kbaker/pubs/Singular_Value_Decomposition_Tutorial.pdf). Diakses tanggal 10 Oktober 2013.
- [2] Fridolin Wild, Chistina, Gerald Stermsek, Gustaf Neumann. *Parameters Driving Effectiveness of Automated Essay Scoring with LSA*. 2008. Department of Information System and New Media. Vienna University in Austria. <https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/2008>.
- [3] Guillermo Jorge-Botana, Jose A. Leon, Ricardo Olmos, dan Inmaculada Escudero. *Latent Semantic Analysis Parameters for Essay Evaluation using Small-Scale Corpora*. 2010. Journal of Quantitative Linguistics 2010. Volume 17. Number 1. Pp 1-29. DOI: 10.1080/09296170903395890.
- [4] Heninggar Septiantri. *Perbandingan Metode Latent Semantic dan Vector Space Model untuk Sistem penilaian Jawaban Esai Otomatis Bahasa Indonesia*. 2009. FASILKOM UI.
- [5] Kukich, K. *Beyond Automated Essay Scoring*. In M. A. Hearst (Ed.). September/October 2000. The debate on automated essay grading. IEEE Intelligent Systems. 27–31.
- [6] Landauer, Foltz, & Laham. *An Introduction to Latent Semantic Analysis*. 1998. Discourse Processes. 25. 259-284.
- [7] Ledy Agusta. *Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Adriani untuk Stemming Dokumen Teks Bahasa Indonesia*. 2009. Konferensi Nasional Sistem dan Informasi. KNS & 109-036. Bali.
- [8] Rama Adhitia, Ayu Purwarianti. *Penilaian Esai Jawaban Bahasa Indonesia Menggunakan Metode SVM-LSA dengan Fitur Generik*. 2009. Journal of Information System Volume 5. Prodi Teknik Informatika ITB.
- [9] Ratna, Budiardjo, Hartanto. *SIMPLE (Sistem Penilaian Esei Otomatis) Untuk Menilai Ujian Dalam Bahasa Indonesia*. 2007. Makara Teknologi. Vol 11 No 1 Hal 5-11.
- [10] Rizqi Bayu Aji P, Abdurrahman Baizal, Yanuar Firdaus. *Automatic Essay Grading System Menggunakan Metode Latent Semantic Analysis*. 2011. ISSN 1907-5022. SNATI Yogyakarta.
- [11] Semire Dikli. *An Overview of Automated Scoring of Essays*. 2006. The Journal of Technology, Learning, and Assessment. Volume 5 Number 1, Stanford University.
- [12] Suharsimi Arikuntoo. *Prosedure Penelitian*. 1998. ISBN 979-518-018-5. Rineka Cipta. Jakarta.
- [13] Tala, Fadillah Z. A. *Study of Stemming Effects on Information Retrieval in Bahasa Indonesia*. Institute for Logic. 2003. Language and Computation Universeit Van Amsterdam.
- [14] Valenti, Neri & Cucchiarelli. *An Overview of Current Research on Automated Essay Grading*. 2003. Journal of Information Technology Education. Volume 2.