

E-learning Automated Essay Scoring System Menggunakan Metode Searching Text Similarity Matching Text

Rahmat Ahmad^{*1}, Wardi¹, Dewiani¹

¹Departemen Teknik Elektro, Fakultas Teknik, Universitas Hasanuddin
Jl. Poros Malino Km.6, Bontomarannu, Kabupaten Gowa, Sulawesi Selatan, 92171

*Email: rahmat.ahmad111185@gmail.com

DOI: 10.25042/jpe.052018.07

Abstrak

E-learning adalah teknologi informasi yang telah banyak diterapkan pada bidang pendidikan berupa *website* karena dapat diakses tanpa batas waktu dan tempat. Perkembangan *E-learning* ini merupakan hal yang logis karena mengikuti zaman yang mengedepankan teknologi informasi dan komunikasi. Pembelajaran konvensional terkhusus evaluasi pembelajaran bentuk esai, dimana guru memeriksa jawaban esai siswa secara manual yang membutuhkan waktu dan tenaga lebih. *Automated Essay Scoring (AES)* merupakan salah satu fitur *E-learning* yang dapat memudahkan guru dalam mencocokkan jawaban esai siswa dengan kunci jawaban. *Text similarity* adalah metode *searching* pencocokan kata dengan *matching text* berdasarkan kondisi (*term*) yang telah ditentukan. Bahasa yang digunakan pada AES ini adalah bahasa Indonesia yang mempunyai morfologi kata yang berbeda dengan bahasa lainnya. Hasil yang didapatkan pada penelitian ini yaitu penggunaan term frequency (*tf*) lebih rendah dibandingkan dengan menggunakan inverse dokumen frequency (*idf*) dengan akurasi yang tinggi.

Abstract

E-learning Automated Essay Scoring System using Searching Text Similarity Matching Text Method. E-learning is an information technology that has been widely applied to the field of education in the form of websites because it can be accessed indefinitely and place. The development of E-learning is logical because it follows the age of putting forward information and communication technology. Conventional learning is especially an evaluation of essay form learning, where teachers check students' essay answers that require more time and effort. Automated Essay Scoring (AES) is one of the features of E-learning that can facilitate the teacher in matching the answer of student essay with answer key. Text similarity is a method of matching text that is recall word based on specified condition (term). The language used in AES is Indonesian which has a different word morphology with other languages. The results obtained in this study is the Error rate using the term frequency (tf) is lower with high accuracy.

Kata Kunci: *Automated essay scoring, e-learning, keyword, searching method, text similarity, term frequency*

1. Pendahuluan

Dunia pendidikan sangatlah penting dalam kehidupan kita. Sejalan dengan itu, proses pembelajaran haruslah mengikuti perkembangan teknologi. Pendidikan yang dikolaborasikan dengan teknologi dapat memberikan suatu efek yang realistis sehingga output yang diharapkan dapat sesuai yang diinginkan yaitu hasil yang maksimal. Mulai dari model pembelajaran konvensional sampai ke model pembelajaran komputerisasi. Kehadiran media komputersisasi akan sangat membantu dalam proses belajar mengajar [1]. Guru diharapkan mampu

mentransfer ilmunya ke siswa dengan maksimal pula.

Salah satu media pembelajaran yang sangat populer adalah *E-learning*, dimana dalam mengakses pelajaran maupun evaluasi pembelajaran dapat dilakukan tanpa ada batas waktu dan tempat. Guru dan siswa dapat dengan mudah berinteraksi melalui sehingga proses transfer ilmu berjalan dengan baik [2]. Interaksi yang dimaksud yaitu guru memberikan evaluasi pembelajaran dalam bentuk soal esai dan siswa menjawab berdasarkan *skill* yang mereka miliki. Esai yang dimaksud adalah esai dalam bahasa Indonesia yang mempunyai banyak variasi



imbunan. Fitur *Automated Essay Scoring (AES)* yang ditambahkan pada *E-learning* ini berfungsi sebagai pemberi nilai akhir siswa sehingga guru dapat mengefisienkan waktu dan tenaga terhadap pemeriksaan jawaban esai siswa [3].

Penelitian sebelumnya yang berhubungan dengan *Automated Essay Scoring (AES)* yaitu penelitian Li Bin, dkk. Tujuan penelitiannya adalah penilaian teks dengan menggunakan metode KNN dengan nilai akurasi sebesar 76% [4]. Penelitian selanjutnya oleh Li Y dan Yan Y. Penelitian ini bertujuan untuk melihat nilai akurasi dalam *AES*. Hasil akurasi yang didapatkan sebesar 86% [5]. Penelitian berikutnya oleh Shehab, dkk., Tujuan dari penelitian ini adalah membandingkan hasil penilaian guru dengan *AES*. Hasil yang diperoleh pada penelitian ini yaitu Penilaian otomatis dengan nilai akurasi sebesar 90% mendekati nilai dari guru [6]. Selanjutnya penelitian tentang perbandingan algoritma stemming porter dan nazief & adriani yang dilakukan oleh Agusta,. Tujuan penelitian ini adalah membandingkan kedua algoritma stemming porter dan nazief & adriani yang mana jika dilihat dari segi kecepatan maka Algoritma Stemming Porter membutuhkan waktu yang lebih sedikit tapi kurang terhadap presentase akurasi sedangkan Algoritma Nazief & Adriani membutuhkan waktu lebih lama tapi dalam hal presentase akurasi lebih tinggi[7]. Khusus Stemming bahasa Indonesia berbeda dengan stemming bahasa Inggris. Stemming bahasa Indonesia lebih kompleks, karena untuk mendapat *root word* harus menghilangkan variasi imbuhan yang terdapat pada kata tersebut.

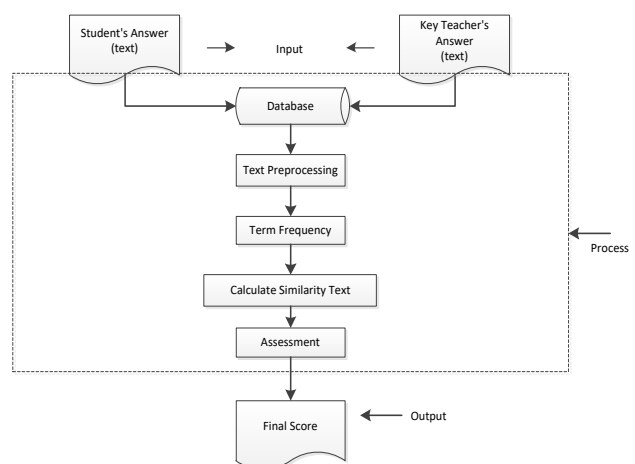
Selain metode yang disebutkan diatas terdapat metode lain yang dapat digunakan untuk melakukan penilaian jawaban esai otomatis yaitu dengan metode *searching text similarity* [8]. Metode digunakan untuk menghitung kesamaan atau kemiripan dari dua buah dokumen. Untuk menyamakan frekuensi setiap kata yang terdapat didalam kalimat yang ada digunakan persamaan Tf atau *Term Frequency*, *Term Frequency* merupakan suatu faktor yang menentukan bobot kata yang didasarkan pada jumlah frekuensi kata dalam sebuah dokumen. Nilai jumlah kemunculan suatu kata (tf) diperhitungkan dalam pemberian

bobot terhadap suatu kata. Semakin besar jumlah kemunculan suatu kata (tf tinggi) dalam dokumen, semakin besar pula bobotnya dalam dokumen atau akan memberikan nilai kesesuaian yang semakin besar [9]. Metode ini bertujuan mencocokkan kedua kalimat pada suatu dokumen yang nantinya akan dibandingkan, kemudian hasil perbandingan akan digunakan sebagai nilai *similarity*.

Berdasarkan uraian di atas maka diperlukan suatu *E-learning AES* yang bertujuan untuk mengefisienkan waktu dan tenaga guru dalam penilaian jawaban esai siswa menggunakan metode *searching text similarity matching text*.

2. Metodologi Penelitian

Penelitian ini merupakan penelitian kuantitatif deduktif yang bersifat eksperimen melalui pembuatan aplikasi *E-learning* yang merujuk pada referensi tentang sistem penilaian otomatis menggunakan metode *searching text similarity matching text*. Adapun Gambar 1 desain arsitektur sistem pada penelitian ini dapat dilihat sebagai berikut:



Gambar 1. Design Architecture System AES

2.1. Text Pre-processing

Text pre-processing adalah suatu proses pengubahan bentuk data yang belum terstruktur menjadi data yang terstruktur sesuai dengan kebutuhan, untuk proses *mining* yang lebih lanjut (*sentiment analysis*, peringkasan, *clustering* dokumen, *etc.*) [10]. Bagaimana cara kerja *text pre-processing* dapat dijelaskan sebagai berikut:



a. Case Folding

Tidak semua dokumen teks konsisten dalam penggunaan huruf kapital. Oleh karena itu, peran *Case Folding* dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar (biasanya huruf kecil atau *lowercase*). Sebagai contoh, user yang ingin mendapatkan informasi “KOMPUTER” dan mengetik “KOMPUTER”, “KomPUter”, atau “komputer”, tetap diberikan hasil *retrieval* yang sama yakni “komputer”. *Case folding* adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf ‘a’ sampai dengan ‘z’ yang diterima. Karakter selain huruf dihilangkan dan dianggap *delimiter*.

b. Tokenizing

Tahap *tokenizing* / *parsing* adalah tahap pemotongan *string input* berdasarkan tiap kata yang menyusunnya. Selain itu, spasi digunakan untuk memisahkan antar kata tersebut. Sebagai contoh : "saya belajar *Information Retrieval*" maka akan dihasilkan : "saya", "belajar", "information", "*retrieval*".

c. Filtering

Tahap *Filtering* adalah tahap mengambil kata-kata penting dari hasil token. Bisa menggunakan *algoritma stoplist* (membuang kata kurang penting) atau *wordlist* (menyimpan kata penting). *Stoplist/stopword* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*. Contoh *stopwords* adalah “yang”, “dan”, “di”, “dari” dan seterusnya.

d. Stemming

Stemming merupakan suatu proses yang terdapat dalam sistem *Information Retrieval (IR)* yang mentransformasi kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (*root word*) dengan menggunakan aturan-aturan tertentu. Proses *stemming* pada teks berbahasa Indonesia berbeda dengan *stemming* pada teks berbahasa Inggris. Pada teks berbahasa Inggris, proses yang diperlukan hanya proses menghilangkan sufiks. Sedangkan pada teks berbahasa Indonesia semua kata imbuhan baik itu sufiks dan prefiks juga dihilangkan. Sebagai contoh kata

“bersama”, “kebersamaan”, “menyamai”, akan distem ke *root word*-nya yaitu “sama”. Namun, seperti halnya *stopping*, kinerja *stemming* juga bervariasi dan sering tergantung pada domain bahasa yang digunakan. Pada penelitian ini, *stemming* akan dilakukan dengan algoritma Nazief-Adriani.

2.2. Term Weight

Term Weight (bobot Kata) merupakan faktor yang menentukan bobot *term* pada suatu dokumen berdasarkan jumlah kemunculannya dalam dokumen tersebut. Pada dokumen yang besar, skema yang paling sukses dan secara luas digunakan untuk pemberian bobot term adalah skema pembobotan atau *Term Weighting TF-IDF*. Kelemahan *scoring* dengan *Jaccard coefficient* adalah tidak disertakannya frekuensi suatu *term* dalam suatu dokumen, maka diperlukan skoring dengan kombinasi *Term Weighting TF-IDF*.

Hal yang perlu diperhatikan dalam pencarian informasi dari koleksi dokumen yang heterogen adalah pembobotan *term*. *Term* dapat berupa kata, frase atau unit hasil indexing lainnya dalam suatu dokumen yang dapat digunakan untuk mengetahui konteks dari dokumen tersebut, maka untuk setiap kata tersebut diberikan indikator, yaitu *term weight*.

- *Term Frequency (tf)*

Term Frequency (tf) adalah frekuensi dari kemunculan sebuah term dalam dokumen yang bersangkutan. Semakin besar jumlah kemunculan suatu term (*TF* tinggi) dalam dokumen, semakin besar pula bobotnya atau akan memberikan nilai kesesuaian yang semakin besar.

- *Inverse Document Frequency*

Inverse Document Frequency (IDF) merupakan sebuah perhitungan dari bagaimana term didistribusikan secara luas pada koleksi dokumen yang bersangkutan. *IDF* menunjukkan hubungan ketersediaan sebuah term dalam seluruh dokumen. Semakin sedikit jumlah dokumen yang mengandung term yang dimaksud, maka nilai *IDF* semakin besar. Perhitungan *IDF* mengacu pada persamaan 1 sebagai berikut:

$$idf_j = \log \left(\frac{D}{df_j} \right) \quad (1)$$



Dimana:

D : jumlah semua dokumen dalam koleksi
 df_j : jumlah dokumen yang mengandung term t_j

• **TF-IDF**

Term Frequency-Inverse Document Frequency (TF-IDF) merupakan metode penggabungan dua konsep untuk perhitungan bobot, yaitu *Term Frequency (TF)* merupakan frekuensi kemunculan *term (t)* pada dokumen (d). dan *Document Frequency (DF)* adalah banyaknya kalimat dimana suatu kata (t). Adapun persamaan 2 merupakan perhitungan umum *TF-IDF*:

$$W_{ij} = tf \times idf \tag{2}$$

Dimana :

W_{ij} : bobot kata/term ke-j dan dokumen ke-i

2.3. Cosine Similarity

Cosine similarity adalah ukuran kesamaan yang lebih umum digunakan dalam *information retrieval* dan merupakan ukuran sudut antara vektor dokumen D_a (titik (ax, bx)) dan D_b (titik (ay, by)). Tiap vektor tersebut merepresentasikan setiap kata dalam setiap dokumen (teks) yang dibandingkan dan membentuk sebuah segitiga, sehingga dapat diterapkan hukum kosinus untuk menyatakan bahwa:

$$\cos C = \frac{a^x + b^2 - c^2}{2ab} \tag{3}$$

Dimana:

$$a^2 = a_x^2 + a_y^2, b^2 = b_x^2 + b_y^2 \tag{4}$$

Dan

$$c^2 = (b_x - a_x)^2 + (a_y - b_y)^2 \tag{5}$$

Menggantikan nilai-nilai persamaan di atas untuk a, b dan c , sehingga didapatkan:

$$\cos C = \frac{a_x b_x + a_y b_y}{\sqrt{a_x^2 + a_y^2} \times \sqrt{b_x^2 + b_y^2}} \tag{6}$$

Ketika dua dokumen identik, sudutnya adalah nol derajat (0°) dan kesamaannya adalah satu (1); dan ketika dua dokumen tidak identik sama sekali, sudutnya adalah 90 derajat (90°) dan kesamaannya adalah (0) [10].

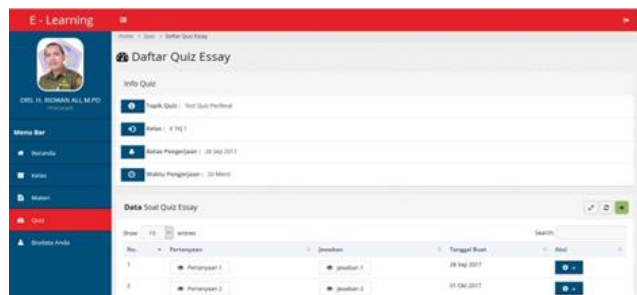
3. Hasil dan Analisis

3.1. Result

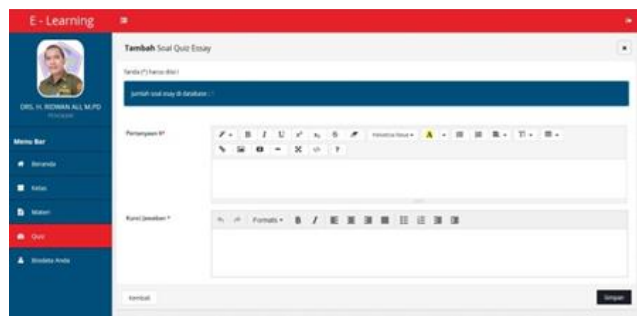
Berdasarkan hasil rancangan sistem yang telah dilakukan maka dihasilkan media pembelajaran elektronik (*E-learning*) dengan fitur penilaian otomatis untuk jawaban esai. Bagian-bagian utama pada hasil penelitian ini sebagai berikut:

3.1.1 Page Teacher

Pada halaman ini pengguna guru bertindak sebagai pembuat soal esai, berupa pertanyaan dan kunci jawaban yang nantinya akan melihat tingkat kemiripan dengan jawaban siswa. Gambar 2 dan Gambar 3 memperlihatkan hasil daftar dan pembuatan soal.



Gambar 2. Page guru



Gambar 3. Page pembuatan soal kuis

3.1.2 Page Student

Pada halaman ini pengguna siswa bertindak sebagai pemberi jawaban untuk soal esai yang telah dibuat oleh pihak guru, hasil jawaban siswa akan dinilai secara otomatis menggunakan metode cosine similarity dari kunci jawaban guru. Gambar 4 dan 5 memperlihatkan hasil *page* siswa.





Gambar 4. Result Page siswa



Gambar 5. Kuis siswa

3.2. Analisis

Pada bagian ini telah melalui pengujian-pengujian dan analisis yang dilakukan terhadap aplikasi yang telah dikembangkan. Berikut ini adalah bahan pengujian dengan menggunakan empat teks jawaban siswa:

Soal : Apa yang dimaksud dengan Pertanyaan : *periferal* ?

Kunci Jawaban : Komponen tambahan yang berfungsi untuk mendukung kerja komputer sehingga fungsi kerja komputer menjadi maksimal

Jawaban Siswa 1 : Komponen tambahan yang berfungsi untuk mendukung kerja komputer sehingga fungsi kerja komputer menjadi maksimal

Jawaban Siswa 2 : Semua hal yang berkaitan dengan komputer

Jawaban Siswa 3 : Suatu unit yang mempunyai kaitan dengan komputer berdasarkan fungsi tertentu dari komputer

Jawaban Siswa 4 : bagian-bagian komputer yang terdiri dari periferil utama dan

periferil pendukung

Jawaban Siswa 5 : belajar bagian-bagian komputer

Kelima teks di atas telah dilakukan pengujian antara satu dengan yang lainnya menggunakan aplikasi ini dan hasil perbandingannya ditunjukkan pada Tabel 1 pengujian menggunakan *term frequency (tf)*, *inverse document frequency (idf)* dan *tf-idf*. Karena pada metode *text mining* dan penelitian-penelitian sebelumnya perihal pencarian kemiripan kata selalu menggunakan *tf-idf* untuk menghitung bobot kata (*term weight*) tiap dokumen. Selain itu juga penelitian ini bertujuan untuk mengetahui tingkat akurasi kemiripan menggunakan metode *text similarity* antara kunci jawaban dari Guru dan jawaban dari Siswa.

Tabel 1. Perbandingan AES dengan Human menggunakan *Term Weight*

	Automated Essay Scoring (%)			Human Rates (%)
	tf	idf	tf-idf	
Jawaban Siswa 1	100	89	89	100
Jawaban Siswa 2	39	0	0	40
Jawaban Siswa 3	59	0	0	70
Jawaban Siswa 4	57	32	32	65
Jawaban Siswa 5	39	0	0	40

Berdasarkan hasil pengujian di atas dapat disimpulkan bahwa *error rate* antara penggunaan *term frequency (tf)* lebih rendah dibandingkan dengan menggunakan *inverse dokumen frequency (idf)* dan *term frequency-inverse document (tf-idf)*. Hal itu terjadi karena pada perhitungan *idf* dan *tf-idf* menggunakan faktor pembagi *n* dari seluruh dokumen dalam perhitungannya

4. Kesimpulan

Berdasarkan hasil penelitian yang telah diuraikan maka dapat diambil kesimpulan bahwa *E-learning AES System* menggunakan metode



searching text Similarity Matching text dapat dipakai sebagai media dalam mengevaluasi hasil jawaban teks siswa terhadap kunci jawaban guru karena mempunyai akurasi yang tinggi dalam pencocokan kesamaan teks. *Error rate* penggunaan *term frequency (tf)* lebih rendah dibandingkan dengan menggunakan *inverse dokumen frequency (idf)* dan *term frequency-inverse document (tf-idf)*. Pada penelitian kedepannya diharapkan sistem penilaian memperhatikan kata yang memiliki kesamaan arti (sinonim).

Referensi

- [1] Bhavna K Pradeep C and Manmohan S 2013 Comparative Analysis Study of E-Learning and Traditional Learning in Technical Institution (India) International Conference on Communication Systems and Network Technologies.
- [2] Khatri B Chouskey P and Singh M 2013 Comparative Analysis Study of E-Learning and Traditional Learning in Technical Institution (India) International Conference on Communication Systems and Network Technologies Comparative P 770-773.
- [3] Astutik S Cahyani A D and Sophan M K 2012 Sistem Penilaian Esai Otomatis Pada E-Learning dengan Algoritma Winnowing vol 12 no 2 (Indonesia) Jurnal Informatika P 47-52.
- [4] Bin L Jun L Min Y J and Ming Z Q 2008 Automated Essay Scoring Using the KNN Algorithm (China) International Conference on Computer Science and Software Engineering Automated P 735-738.
- [5] Li Y and Yan Y 2012 An effective automated essay scoring system using support vector regression Yali (China) Fifth International Conference on Intelligent Computation Technology and Automation P 65-68.
- [6] Shehab A Elhoseny M and Hassanien A E 2016 A Hybrid Scheme for Automated Essay Grading Based on LVQ and NLP Techniques (Egypt: IEEE Journal) P 65-75.
- [7] Agusta L 2009 Perbandingan Algoritma Stemming Porter Dengan Algoritma Nazief & Adriani Untuk Stemming Dokumen Teks Bahasa Indonesia (Bali) Konferensi Nasional Sistem dan Informatika KNS&I P 196-201.
- [8] Ewees A A Eisa M Refaat M M 2014 Comparison of cosine similarity and k-NN for automated essays scoring vol 3 Issue 12 (Egypt) International Journal of Advanced Research in Computer and Communication Engineering P 8669-8673.
- [9] Yates B and Neto R 1999 Modern Information Retrieval (USA: Addison-Wesley Longman Publishing Co. Inc).
Konchady M 2006 Text Mining Application Programming (Boston: Charles River Media).

