

**PENILAI ESAI OTOMATIS UJIAN TENGAH SEMESTER DI SMK ASRAMA
PEMBINA MASYARAKAT JATIMULYO JEMBER MENGGUNAKAN
METODE COSINE SIMILARITY**

Lailiyatul Muharromah¹, Ulya Anisatur R², Mudafiq Riyan Pratama³

¹lailiyatulmuharromah@gmail.com

²ulyaanisatur@unmuhjember.ac.id

³mudafiq.riyan@unmuhjember.ac.id

Teknik Informatika

Universitas Muhammadiyah Jember

Jln.Karimata No.49. Telp(0331)336728.Jember

ABSTRAK

Evaluasi hasil belajar merupakan komponen penting dalam pembelajaran. evaluasi biasanya dilakukan dengan mengadakan ujian yang biasanya menggunakan salah satu jenis soal yang diberikan dalam bentuk esai. Bentuk ujian esai memiliki beberapa kelebihan diantaranya dapat mencegah timbulnya permainan curang. Tetapi ujian esai juga memiliki kekurangan yaitu kesulitan untuk memberikan penilaian, karena masih membutuhkan waktu, tenaga dan pikiran. Untuk bisa membantu pada proses penilain ujian maka dirancang sistem penilai esai jawaban ujian dengan metode cosine similarity sehingga dapat membantu pengajar dalam memberikan penilaian yang lebih tepat dan akurat. Dimana sistem penilai tersebut dapat memilih dokumen yang mirip dan dokumen yang tidak mirip menggunakan ambang batas (Threshold). Setelah dilakukan perhitungan didapatkan hasil threshold terbaik yaitu 0,6 dengan perolehan recall 86% dan precision 100%.

Kata kunci : text mining, ujian esai, cosine similarity

AUTOMATED ESSAY ASSESSMENT MIDTERM IN SMK ASRAMA
PEMBINA MASYARAKAT JATIMULYO JEMBER
USING COSINE SIMILARITY METHOD

Lailiyatul Muharromah¹, Ulya Anisatur R², Mudafiq Riyan Pratama³

¹lailiyatulmuharromah@gmail.com

²ulyaanisatur@unmuhjember.ac.id

³mudafiq.riyan@unmuhjember.ac.id

*Informatics Engineering
University of Muhammadiyah Jember
Jln. Karimata No.49. Telp(0331)336728. Jember*

ABSTRACT

Student assessment is an important component in learning. The evaluation is usually done by holding the test usually uses one of the types of question that was provided in the form of essays. Form of examination essays have some excess of which can prevent the onset of the game cheat. But, essay exam also has drawbacks, i.e difficult to give evaluation, because it still takes time, energy and thoughts. To be able to help in the evaluation process is exam then designed a system of assessors of the essay exam answers with cosine similarity method so that it can help teachers in giving the evaluation that is more precise and accurate. Where the assessor system can choose similar documents and documents that are not similar to using a threshold (Threshold). After calculation of the obtained results the best threshold i.e. 0.6 with acquisition of 86% recall and precision 100%.

Keywords: text mining, essay exam, cosine similarity

1. PENDAHULUAN

Teknologi informasi kini telah mampu membantu manusia dalam segala bidang. Salah satu hasil perkembangan teknologi informasi yang penting adalah dalam bidang pendidikan. Beberapa instansi pendidikan mulai mengembangkan proses dalam evaluasi hasil belajar. Evaluasi hasil belajar merupakan komponen penting dalam proses pembelajaran. Dengan melakukan evaluasi, dapat diketahui sejauh mana pemahaman peserta didik terhadap materi serta pencapaian tujuan pembelajaran. Evaluasi dilakukan dengan mengadakan ujian yang biasanya menggunakan salah satu jenis soal yang diberikan dalam bentuk esai.

Ujian dengan sistem esai merupakan bentuk evaluasi dimana pilihan jawaban tidak disediakan, dan siswa harus menjawab dengan kalimat (Putri Ratna, 2007), sehingga dapat melatih siswa dalam menyampaikan suatu informasi secara subjektif, selain itu ujian esai juga menuntut pemahaman yang lebih baik akan suatu ilmu dan dapat digunakan untuk mengukur tingkat pemahaman siswa akan suatu ilmu secara lebih mendalam (Hamza, 2013). Ujian sistem ini tetap menjadi

pilihan pengajar untuk mengevaluasi tingkat kemampuan pemahaman siswa walaupun kenyataannya tidak mudah untuk memberikan penilaian yang objektif pada jawaban siswa.

Kekurangan penilaian tersebut dapat diatasi dengan penggunaan Teknologi informasi yang dapat dengan mudah melakukan otomatisasi penilaian esai jawaban ujian dengan metode *cosine similarity*. Metode *cosine similarity* merupakan salah satu metode yang dapat digunakan untuk mencocokkan esai jawaban ujian dan kunci jawaban soal. Kelebihan utama dari metode *cosine similarity* adalah tidak terpengaruh terhadap panjang pendeknya suatu dokumen. Karena yang diperhitungkan hanya nilai *term* dari masing-masing dokumen. Untuk memperoleh hasil pencarian dokumen yang maksimal diperlukan sebuah nilai ambang batas (*threshold*) agar sistem dapat memilih mana dokumen yang mirip dan mana yang tidak.

2. METODE PENELITIAN

Cosine Similarity

Kemiripan antar dokumen dihitung menggunakan suatu fungsi ukuran kemiripan yaitu *similarity measure*. Ukuran ini memungkinkan

perangkingan dokumen sesuai dengan kemiripan relevansinya terhadap *query* (Cios, 2007). Salah satu ukuran kemiripan teks yang paling populer adalah *cosine similarity*. Ukuran ini menghitung nilai *cosinus* sudut antara dua vektor. Ukuran ini menghitung nilai *cosinus* sudut antara dua vektor. Jika terdapat dua vektor dokumen d_j dan *query* q , serta t term diekstrak dari dari koleksi dokumen, maka nilai *cosinus* antara d_j dan q didefinisikan sebagai berikut (Cios, 2007).

$$\text{Cosine} \rightarrow \sin(d_j, q) = \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^t (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^t w_{ij}^2 \cdot \sum_{i=1}^t w_{iq}^2}}$$

Langkah-langkah perhitungan *Cosine Similarity* adalah sebagai berikut :

1. Tentukan data set dokumen yang akan dilakukan uji coba.
2. Lakukan tahapan preprocessing dengan melibatkan 5 operasi yaitu tokenizing, filtering, stemming, dan synonym.
3. Lakukan pembobotan lokal dan pembobotan global yaitu *Term Frequency – Inverse Dokument Frequency* (TF-IDF).

$$idf_i = \log\left(\frac{n}{df_i}\right)$$

4. Selanjutnya tahapan menormalkan vektor dokumen sehingga proses

pengukuran kemiripan tidak terpengaruh dari panjang suatu dokumen.

$$\frac{1}{\sqrt{w_1^2 + w_2^2 + \dots + w_m^2}}$$

5. Hitung kemiripan dokumen dengan metode cosine similarity

$$\text{Cosine} \rightarrow \sin(d_j, q) = \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^t (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^t w_{ij}^2 \cdot \sum_{i=1}^t w_{iq}^2}}$$

6. Konversi kemiripan nilai cosine similarity menjadi nilai ujian esai menggunakan rentang nilai versi manusia (human rates).

Tabel 2.1 Tabel perbandingan rentang nilai siswa

perbandingan rentang nilai	
nilai kemiripan	nilai human rates
0,01 - 0,10	10
0,11 - 0,20	20
0,21 - 0,30	30
0,31 - 0,40	40
0,41 - 0,50	50
0,51 - 0,60	60
0,61 - 0,70	70
0,71 - 0,80	80
0,81 - 0,90	90
0,91 - 1	100

7. Tentukan nilai threshold untuk memperoleh hasil pencarian yang maksimal agar sistem dapat memanggil dokumen yang mirip dan dokumen yang tidak mirip.

8. Hitung *recall* (perolehan) dan *precision* (ketepatan) yang berkaitan dengan kemampuan sistem memanggil dokumen yang salah dan dokumen yang benar.

$$\begin{aligned}
 \text{Recall} &= \frac{TP}{TP + FN} \\
 &= \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}
 \end{aligned}$$

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FP} \\
 &= \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}
 \end{aligned}$$

3. HASIL DAN PEMBAHASAN

Setelah program selesai dibuat, selanjutnya tahap mengujian sistem. Pada uji coba penelitian digunakan dua jenis data set yaitu soal ujian beserta kata kunci jawabannya dan esai jawaban ujian. Soal ujian berjumlah 3 pertanyaan, masing-masing disertai 1 kunci jawaban. Sedangkan jawaban ujian berjumlah 10 esai untuk setiap soal. Jadi jumlah total dataset adalah 3 *query* dan 30 dokumen.

Untuk melakukan uji coba pada aplikasi penilai esai otomatis dengan metode cosine similarity maka dilakukan skenario uji coba sebagai berikut:

3.1 Dataset Uji Coba

Percobaan dilakukan dengan meng-inputkan 3 soal dan 3 kunci jawaban yang akan disimpan dalam database. Kemudian dilakukan ujian online oleh 10 siswa sehingga akan didapat 30 jawaban ujian esai.

3.2 Hasil Uji Coba

Dari dataset yang berjumlah 3 soal dan 10 jawaban untuk masing-masing soal maka selanjutnya akan dilakukan tahapan proses cosine similarity untuk mendapatkan hasil nilai cosine similarity sebagai berikut:

Tabel 3.1 Hasil Uji Coba

Soal I	Urutan Kemiripan n	Dokumen	Nilai Cosine Similarity
Soal No.1	1	D2	0,822
	2	D5	0,744
	3	D10	0,744
	4	D6	0,716
	5	D8	0,716
	6	D7	0,674
	7	D4	0,674
	8	D3	0,667
	9	D1	0,646
	10	D9	0,628
Soal No.2	1	D11	0,672
	2	D14	0,613
	3	D15	0,613
	4	D19	0,606
	5	D13	0,606
	6	D20	0,6
	7	D18	0,592
	8	D17	0,592
	9	D12	0,592
	10	D16	0,533
Soal No.3	1	D21	0,899
	2	D25	0,881
	3	D29	0,847
	4	D30	0,847

Soal I	Urutan Kemiripan	Dokumen	Nilai Cosine Similarity
	5	D23	0,847
	6	D22	0,812
	7	D27	0,811
	8	D28	0,753
	9	D26	0,753
	10	D24	0,691

Dari hasil perhitungan diatas didapat nilai cosine similarity kemudian diurutkan dari nilai tertinggi hingga terendah. Untuk memperoleh perankingan nilai.

3.3 Uji coba penentuan Threshold

Untuk menyatakan apakah sebuah esai jawaban bernilai benar atau salah maka diperlukan sebuah nilai ambang batas (*threshold value*). Pada uji coba ini digunakan threshold 0,4 – 0,8 dan kemudian akan ditentukan threshold terbaik dari beberapa threshold tersebut. Untuk mendapatkan threshold terbaik maka diperlukan perhitungan recall dan precision. Perolehan (recall) berhubungan dengan kemampuan sistem untuk memanggil dokumen yang relevan, sedangkan ketepatan (precision) berkaitan dengan kemampuan sistem untuk tidak memanggil dokumen yang tidak relevan.

3.4 Hasil Uji Coba Threshold

Setelah melakukan uji coba penentuan threshold menggunakan threshold 0,4 – 0,8 maka didapat hasil recall dan precision sebagai berikut:

Tabel 3.2 Tabel Hasil Uji Coba Threshold

No	Threshold	Recall	Precision
1	0,4	100%	100%
2	0,5	100%	100%
3	0,6	86%	100%
4	0,7	46%	100%
5	0,8	26%	88%

Semakin besar nilai threshold maka semakin sempit jangkauan penilaian, sebaliknya jika semakin kecil nilai threshold maka semakin luas jangkauan penilaian yang dilakukan.

Menurut Rowley Dalam Hasugin (2003), suatu sistem dikatakan efektif apabila hasil penelusuran mampu menunjukkan ketepatan (precision) yang tinggi sekalipun perolehan recall rendah. Koncisi ideal keefektifan suatu sitem adalah apabila rasio recall dan precision sama besarnya (1:1) menurut Lee Pao (1989).akan tetapi karena rasio dari recall sebenarnya sulit diukur karena jumlah seluruh jumlah dokumen yang relevan dalam database sangat besar, oleh karena itu precision lah yang menjadi salah satu ukuran yang

digunakan untuk menilai keefektifan suatu sistem.

Dari uji coba tersebut maka didapat *threshold* 0,4 dan 0,5 yang dikatakan baik karena perolehan *recall* dan *precision* sebanding (1:1) dan memiliki jangkauan penilaian yang luas, namun hasil yang lebih maksimal yaitu *threshold* 0,6 karena menghasilkan nilai *recall* 86% dan *precision* 100%.

KESIMPULAN

Berdasarkan pengujian yang telah dilakukan, kesimpulan yang diperoleh dari penelitian ini adalah sebagai berikut:

1. Dari kelima pengujian *threshold* yang telah dilakukan, dengan *threshold* 0,4 diperoleh hasil *recall* dan *precision* yang sama yaitu 100%, *threshold* 0,5 diperoleh hasil *recall* dan *precision* yang sama yaitu 100%, *threshold* 0,6 diperoleh hasil *recall* 86% dan *precision* 100%, *threshold* 0,7 diperoleh hasil *recall* 46% dan *precision* 100%, *threshold* 0,8 diperoleh hasil *recall* 26% dan *precision* 88%. Semakin besar nilai *threshold* maka semakin sempit jangkauan penilaian, sebaliknya

jika semakin kecil nilai *threshold* maka semakin luas jangkauan penilaian yang dilakukan oleh sistem. Dari kelima uji coba tersebut, uji coba menggunakan 0,4 dan 0,5 dikatakan baik namun hasil lebih maksimal jika nilai *precision* lebih tinggi dibanding *recall*. Sehingga kelima uji coba tersebut 0,6 dikatakan yang terbaik karena mendapatkan *recall* 86% dan *precision* 100%.

2. Penilai esai otomatis ujian tengah semester di SMK ASRAMA PEMBINA MASYARAKAT Jatimulyo Jember menggunakan metode *cosine similarity* telah berhasil dilakukan. Dimana dari kelima uji coba tersebut didapatkan nilai *threshold* 0,6 yang mendapatkan *recall* 86% dan *recall* 100% dan dapat diartikan bahwa sistem penilai esai otomatis menggunakan metode *cosine similarity* dapat memanggil dokumen yang dianggap benar dan relevan sehingga tidak melakukan pemanggilan terhadap dokumen yang tidak benar dan tidak relevan.

SARAN

Saran yang peneliti ajukan untuk perbaikan dan pengembangan selanjutnya adalah sebagai berikut:

1. Perbaikan untuk pengembangan selanjutnya terhadap proses *synonim*, karena selama melakukan uji coba masih menggunakan tahapan *synonim* yang kurang akurat dan masih menggunakan inputan *synonim* manual.
2. Perbaikan pada sistem uji coba karena jawaban yang tidak terstruktur pada kalimat masih dapat diproses oleh sistem.

DAFTAR PUSTAKA

Amin, Fatkhul. 2011. Implementasi Search Engine (Mesin Pencari) Menggunakan Metode Vektor Space Model. *Dinamika Teknik*, 5(1), 45-58.

Amin, Fatkhul. 2012. Sistem Temu Kembali Informasi Dengan Metode Vector Space Model. *Jurnal Informasi Bisnis*, Online <http://ejournal.undip.ac.id/index.php/jsinbis>

Arianti, Dewa Ayu Rai. Dkk. (2016). Pengukuran Kemiripan Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity. *E-journal Teknik Informatika*, 9(1), 2301-8364.

Febrianto. 2016. *Pengembangan Sistem Ujian Online Berbasis Web*. Skripsi tidak diterbitkan. Yogyakarta: Program Studi Teknik.

Fitri, Rahimi & Asyikin, Arifin Noor. 2015. Aplikasi Penilai Ujian Essay Otomatis Menggunakan Metode Cosine Similarity. *Jurnal Poros Teknik*, 7(2), 54-106.

Gafar, Abdoel. 2008. Penggunaan Internet Sebagai Media Baru dalam Pembelajaran. *Jurnal Ilmiah Universitas Batanghari Jambi*, 8(2), 36-43.

Hasugian, Jonner. 2006. Penelusuran Informasi Ilmiah Secara Online: Perlakuan Terhadap Seorang Pencari Informasi Sebagai Real User. *Jurnal Studi Perpustakaan dan Informasi*, 2(1), 1-13.

Imbar, Radiant Victor. dkk. 2014. Implementasi Cosine Similarity dan Algoritma Smith-Waterman Untuk Mendeteksi Kemiripan

- Teks. *Jurnal Informatika*, 10(1), 31-42.
- Karmayasa, et.al (2009). *Implementasi Vector Space Model dan Beberapa Notasi Metode Term Frequency (TF-IDF) Pada Sistem Temu Kembali Informasi*. Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Udayana.
- Lestari, nisaa putri. 2016. *Uji Recall Dan Precision Sistem Temu Kembali Informasi*. Skripsi diterbitkan. Surabaya: Fakultas Ilmu Sosial Dan Ilmu Politik.
- Nurhuda, Juhri. 2014. *Model Penilai Esai Otomatis Jawaban Ujian Berbahasa Indonesia Menggunakan Cosine Similarity*. Skripsi tidak diterbitkan. Jember: Program Studi Teknik Informatika.
- Perkasa, Dheru Alam. Dkk. (2015). Sistem Ujian Online Essay Dengan Penilaian Menggunakan Metode Latent Sematic Analysis (LSA). *Jurnal Rekayasa Dan Manajemen Sistem Informasi*, 1(1), 1-9.
- Qodratillah, Meity Taqdir. 2008. *Tesaurus Bahasa Indonesia*. Jakarta: Pusat Bahasa Departemen Pendidikan Nasional.
- Rinartha, Komang (2017). *Pemodelan Esai Otomatis Secara Realtime Menggunakan Kombinasi Text Mining Dan Cosine Similarity*. Bali: STMIK STIKOM Bali.
- Roshinta, Trisna Ari & Rahutomo Faisal (2016). *Analisis Aspek-aspek Ujian Esai Daring Berbahasa Indonesia*. Semarang: Seminar Nasional Terapan Riset Inovatif.
- Salim, Mohammad Agus & Anistiyasari Yeni (2017). Pengembangan Aplikasi Penilai Ujian Essay Berbasis Online Menggunakan Algoritma Nazief Dan Andriani Dengan Metode Cosine Similarity. *Jurnal IT-EDU*,02(1),126-135.
- Sudijono, Anas (2005). *Pengantar Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada.
- Sulistyo, Eko Meiyanto. Dkk. (2015). Penilai Ujian Bertipe Essay Menggunakan Metode Text Mining. *Telematika*,12(62), 146-156.
- Wahid, Devid Haryalesmana & SN Azhari (2016). *Peringkasan Sentimen Ekdtraktif di Twitter*

Menggunakan Hybrid TF-IDF
Dan Cosine Similarity. *IJSS*,

10(2), 207-218.