

PENGGUNAAN CONSINE SIMILARITY DALAM PENILAIAN JAWABAN UJIAN ESAI OTOMATIS

Mochamad Iqbal Mardiana¹, Slamet Risnanto²
^{1,2} Teknik Informatika, Universitas Sangga Buana

¹ korespondensi: iqbalmar.id@gmail.com

ABSTRACT

Learning evaluation using the description or essay method is one way to find out to what extent the students' ability to master the learning material. Evaluation of learning using essays is a test that requires students to write their answers and is assessed manually by the teacher and requires the teacher to assess the exam manually. It will take quite a lot of time if the test is carried out simultaneously by students, not to mention the student's writing which is difficult to read, the assessment is less consistent and this can affect the quality of the assessment of the exam itself. To overcome these problems, an automated essay test answer scoring system was created as an option for conducting essay grading. The assessment is carried out through text processing methods and then weighted term frequency - inverse document frequency (TF-IDF) and consine similarity algorithm. Based on the results of the study that the automatic essay test answer scoring system using the consine similarity method went well with the average value given by the teacher of 88.89%.

Keywords: Essay Exam, TF- IDF, Consine Similarity, Text Preprocessing

ABSTRAK

Evaluasi pembelajaran dengan metode uraian atau esai merupakan salah satu cara untuk mengetahui sampai mana kemampuan siswa menguasai materi pembelajaran. Evaluasi pembelajaran menggunakan esai merupakan ujian yang mengharuskan siswa untuk menulis jawaban mereka dan dinilai secara manual oleh guru dan mengharuskan guru menilai ujian tersebut dengan cara manual. Akan memakan waktu yang cukup banyak jika ujian tersebut dilaksanakan secara serentak oleh siswa, belum lagi tulisan siswa yang sulit untuk dibaca, penilaian yang kurang konsisten dan hal tersebut dapat mempengaruhi kualitas penilaian dari ujian itu sendiri. Untuk mengatasi masalah-masalah tersebut, sistem penilaian jawaban ujian esai otomatis dibuat sebagai pilihan untuk melakukan penilaian esai. Penilaian yang dilakukan melalui metode pemrosesan teks lalu pembobotan term frequency - inverse document frequency (TF-IDF) dan algoritma consine similarity. Berdasar hasil penelitian bahwa sistem penilaian jawaban ujian esai otomatis dengan menggunakan metode consine similarity berjalan dengan baik dengan nilai rata-rata yang diberikan oleh pengajar sebesar 88,89%.

Kata Kunci: Ujian Esai, TF- IDF, Consine Similarity, Text Preprocessing

PENDAHULUAN

Ujian dengan sistem esai merupakan hal yang sering dilakukan di lingkungan pendidikan bahkan di lingkungan kerja sebagai bahan evaluasi pembelajaran ataupun evaluasi kerja karyawan, Ujian esai juga menuntut pemahaman yang lebih baik akan suatu ilmu dan dapat digunakan untuk mengukur tingkat pemahaman manusia akan suatu ilmu. Saat kondisi pandemi saat ini banyak kegiatan belajar mengajar dilakukan secara online

mulai dari pertemuan sampai ujian, proses ujian yang biasa dilaksanakan adalah jenis ujian pilihan ganda ataupun ujian singkat dengan alasan kemudahan dalam proses penilaian. Dalam menjalankan proses pembelajaran secara online perangkat komputer menjadi hal yang penting karena menjadi perangkat pendukung untuk melakukan penilaian jenis ujian esai. Hal ini merupakan landasan lahirnya berbagai macam bentuk penilaian yang didesain untuk

membantu tenaga pengajar dalam melakukan penilaian terhadap jawaban esai secara komputerisasi (1). Metode yang digunakan merupakan metode yang membandingkan kesamaan jawaban yaitu metode cosine similarity. Cosine similarity merupakan metode pengklasifikasian dengan cara membandingkan kesamaan atau kesamaan antara tulisan dengan kata kunci pertama, kemudian cara membandingkan kesamaan antara judul dokumen dengan kata kunci kedua, begitu seterusnya (2). Ujian esai adalah tes yang disusun dalam bentuk pertanyaan terstruktur dan mahasiswa menyusun, mengorganisasikan sendiri jawaban tiap pertanyaan itu dengan bahasa sendiri. Soal esai merupakan bentuk evaluasi dimana pilihan jawaban tidak disediakan, dan mahasiswa harus menjawab dengan kalimat, sehingga jawaban dapat sangat bervariasi sesuai dengan pemikiran masing-masing peserta ujian (3).

METODE PENELITIAN

Cosine similarity merupakan metode pembandingan untuk kesamaan atau kemiripan kata pada jawaban esai terhadap kunci jawaban yang memiliki waktu yang cepat untuk mengukur kesamaan teks (4).

Selain *cosine similarity* Pemrosesan teks (*text preprocessing*) merupakan proses mempersiapkan teks menjadi data untuk menganalisis pada langkah selanjutnya, yaitu tahap machine learning. Dokumen pada umumnya mempunyai struktur yang sembarangan atau tidak terstruktur. Oleh

karena itu, diperlukan suatu proses yang dapat mengubah bentuk data yang sebelumnya tidak terstruktur ke dalam bentuk data yang terstruktur. Proses perubahan ini dikenal dengan istilah *text preprocessing* (5).

Cosine similarity merupakan ukuran kesamaan antara dua buah vektor dalam sebuah ruang dimensi yang didapat dari nilai cosinus sudut dari perkalian dua buah vektor yang dibandingkan karena cosinus dari 00 adalah 1 dan kurang dari 1 untuk nilai sudut yang lain, maka nilai similarity dari dua buah vektor dikatakan mirip ketika nilai dari cosine similarity adalah 1 (6).

Berikut merupakan rumus cosine similarity:

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

Sebelum cosine similarity tersebut dihitung, ada beberapa tahapan yang harus dilakukan yaitu *text preprocessing* dan juga perhitungan *tf* dan *idf* jika hendak menghitung *cosine similarity* secara manual (7).

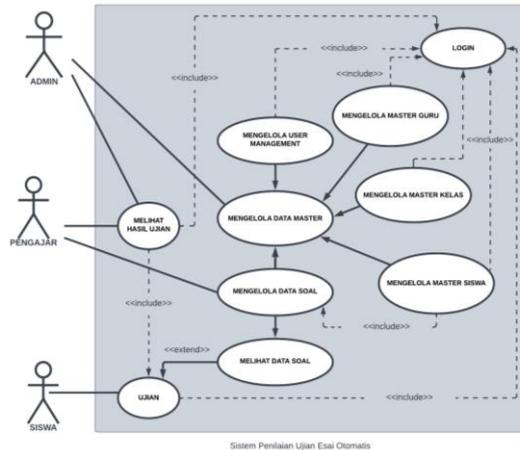
HASIL DAN PEMBAHASAN

Sistem penilaian jawaban ujian esai otomatis ini dibangun menggunakan MySQL dengan Bahasa pemrograman PHP dan juga Python yang dibuat menggunakan Visual Studio Code dan menggunakan metode *cosine similarity*. Bagian ini menjelaskan tentang gambaran sistem yang akan dibuat berupa *Use Case Diagram*, *Class Diagram*, *Sequence Diagram*, dan *User Interface*

Use Case Diagram

Use case diagram yang dibuat menunjukkan hubungan antara actor, use case dan

komponen lainnya (8). Berikut merupakan use case yang dapat dilihat pada Gambar 1.



Gambar 1: Use Case Diagram Sistem Penilaian Ujian Esai Otomatis

Untuk deskripsi interaksi antara pengguna dan sistem penilaian ujian esai otomatis dapat dilihat pada Tabel 1.

Tabel 1: Usecase Diagram Sistem Penilaian Jawaban Ujian Esai Otomatis

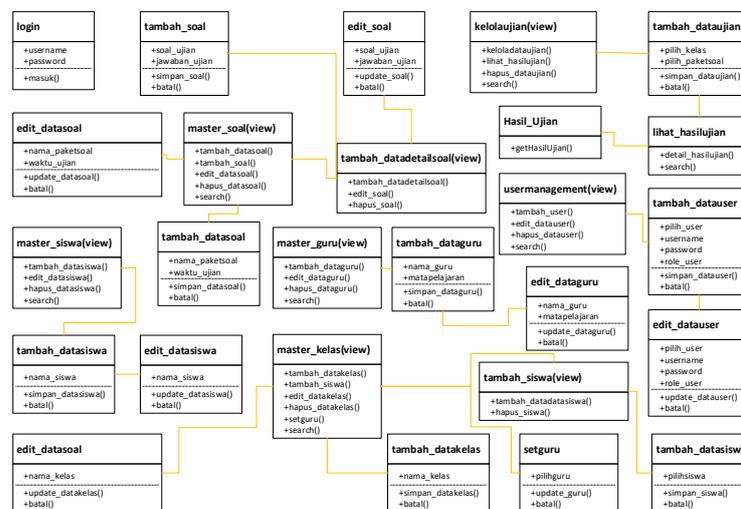
No	Use Case	Deskripsi
1	Login	Validasi untuk mengakses halaman penilaian ujian esai otomatis
2	Mengelola Data Master	Merupakan fitur menu yang memiliki 4 sub yang terdiri dari master kelas, master mata pelajaran, master guru, master siswa
3	User Management	User Management berisi tentang data hak akses yang aktif. User Management ini memiliki opsi berupa edit, delete, search, reload dan opsi tampilan dari diri otomatis
4	Master Siswa	Master siswa merupakan sub menu dari data master yang memiliki isi form data siswa diantaranya NIS, nama siswa, email dan kelas. Operasi yang digunakan berupa edit, delete, search, reload dan tambah data
5	Master Kelas	Bagian ini merupakan sub menu dari data master yang berisi seluruh data yang berisi data seluruh kelas, memiliki pilihan aksi untuk menambahkan data, edit, delete, search dan reload

No	Use Case	Deskripsi
6	Data Soal	Data soal berisikan data soal dan kunci jawaban yang telah diinput oleh guru. Operasi yang digunakan berupa buat soal, <i>edit</i> , <i>delete</i> , <i>search</i> dan <i>reload</i>
7	Hasil Ujian	<i>Use case</i> ini berisi tentang data setelah siswa melaksanakan ujian. Operasi yang dapat digunakan berupa <i>reload</i> , <i>search</i> , lihat hasil, dan cetak
8	Ujian	Ujian ini merupakan <i>use case</i> yang dapat diakses hanya oleh siswa untuk melaksanakan ujian <i>online</i> . Didalam bagian ujian ini terdapat data nama ujian, mata pelajaran, guru, jumlah soal, dan waktu ujian. Operasi ini dapat dilakukan setelah siswa memasukan token ujiab dan dapat melihat hasil ujian sesuai akses guru jika ujian telah selesai

Class Diagram

Berikut merupakan penjabaran dari class diagram sistem penilaian ujian esai otomatis yang diawali dengan login sampai mengolah

detail data secara keseluruhan mulai dari soal (9), management user, guru, siswa, kelas, dan pengolahan detail lainnya yang bisa dilihat pada Gambar 2.

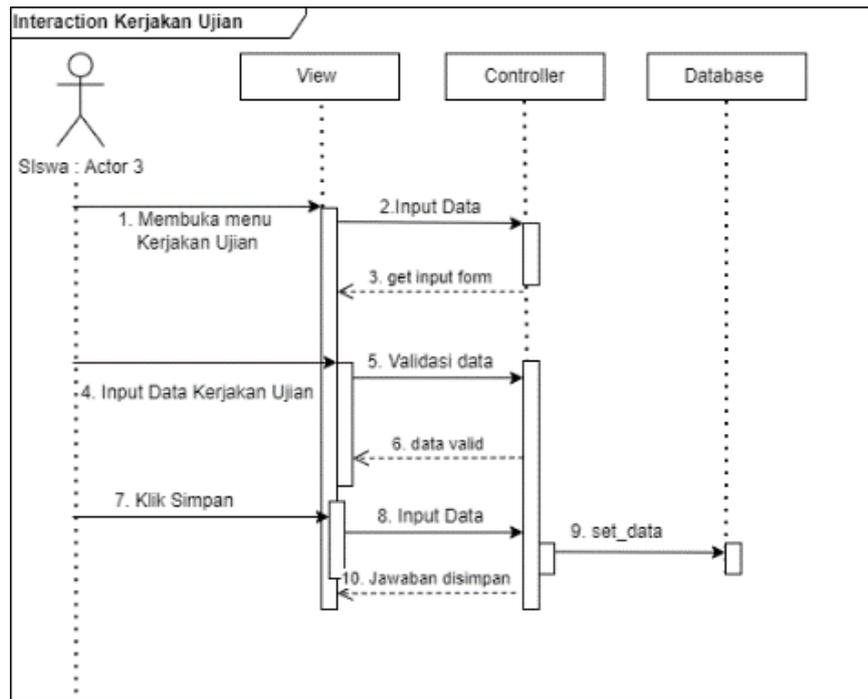


Gambar 1.: Class Diagram Sistem Penilaian Jawaban Ujian Esai Otomatis

Sequence Diagram

Sequence diagram dapat digunakan untuk kebutuhan dalam membuat visualisasi interaksi elemen satu dengan yang lainnya.

Berikut merupakan *sequence diagram* pengerjaan ujian siswa yang dapat dilihat pada Gambar 3.

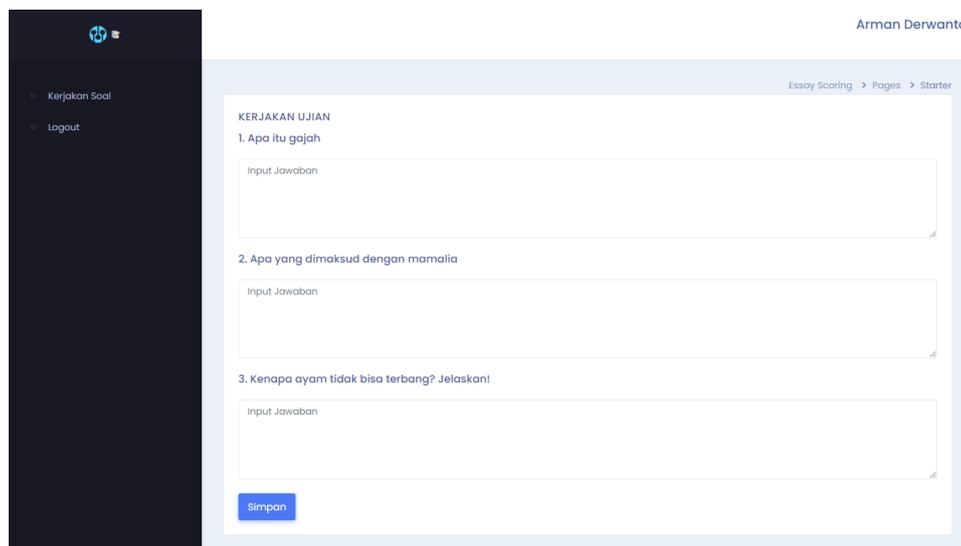


Gambar 2: Sequence Diagram Pengerjaan Ujian

User Interface

Untuk membuat sistem penilaian jawaban ujian esai otomatis dan sebagai dasar bagian dari desain sistem, Dibuatlah user interface yang bertujuan agar interaksi pengguna dengan sistem yang dibuat dapat digunakan

seefektif dan sesederhana mungkin(10). Berikut merupakan gambar rancangan user interface dari sistem penilaian jawaban ujian esai otomatis yang dapat dilihat pada Gambar 4.



Gambar 3: Rancangan User Interface Sistem Penilaian Jawaban Ujian Esai Otomatis

KESIMPULAN

Kesimpulan yang dapat diambil dari sistem penilain jawaban ujian esai otomatis. Latar

belakang pembuatan sistem penilaian jawaban ujian esai otomatis ini didasari dengan banyaknya sistem penilaian yang banyak

dipakai untuk pilihan ganda saja, dan hasil dari pengujian sistem penilaian jawaban ujian esai otomatis ini mudah-mudahan mampu berjalan dengan baik. Penerapan metode *cosine similarity* pada sistem penilaian jawaban ujian esai otomatis ini berpengaruh terhadap teknis penyelenggaraan ujian mulai dari pembuatan soal, pelaksanaan ujian, dan juga penilaian ujiannya. Perancangan sistem penilaian jawaban ujian esai otomatis dapat membantu pengajar ataupun siswa dalam melakukan pelaksanaan ujian agar berjalan secara efektif dan efisien.

DAFTAR PUSTAKA

1. Hamza Sahriar; Sarosa M; Budi Santoso Purnomo. Sistem Koreksi Soal Essay Otomatis Dengan Menggunakan Metode Rabin Karp. Jurnal EECCIS. 2013 Dec;7:153–8.
2. Wahyuni RT, Prastiyanto D, Suprpto DE. Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi. Semarang; 2017.
3. Perkasa DA, Saputra E, Fronita M, Informasi JS, Sains F, Teknologi D, et al. Sistem Ujian Online Essay Dengan Penilaian Menggunakan Metode Latent Sematic Analysis (LSA). Jurnal Rekayasa dan Manajemen Sistem Informasi. 2015;1(1):1–9.
4. Pramukantoro ES, Fauzi MA. Comparative analysis of string similarity and corpus-based similarity for automatic essay scoring system on e-learning gamification. In: 2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS). 2016. p. 149–55.
5. Feldman R, Sanger J. The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge University Press; 2007. 410 p.
6. Ariantini DAR, Lumenta ASM, Jacob A. Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity. E-Journal Teknik Informatika, No 1. 2016;9.
7. Hickman L, Thapa S, Tay L, Cao M, Srinivasan P. Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. Organizational Research Methods. 2022;25(1).
8. Kurniawan TA. Pemodelan Use Case (UML): Evaluasi Terhadap beberapa Kesalahan dalam Praktik. Jurnal Teknologi Informasi dan Ilmu Komputer. 2018 Mar 2;5(1):77.
9. Berardi D, Calvanese D, de Giacomo G. Reasoning on UML class diagrams. Artificial Intelligence. 2005;168(1–2).
10. Miraz MH, Ali M, Excell PS. Adaptive user interfaces and universal usability through plasticity of user interface design. Vol. 40, Computer Science Review. 2021.