

## ANALISA HASIL SISTEM PENILAIAN ESAI OTOMATIS MENGGUNAKAN METODE DICE SIMILARITY DAN QUERY EXPANSION

- Lingkungan Uji Coba

Pada sub bab berikut akan dijelaskan tentang spesifikasi perangkat lunak dan perangkat keras yang dipergunakan pada saat uji coba sistem penilaian esai otomatis menggunakan metode Dice Similarity dan Query Expansion.

Sistem Operasi : Windows 10

CPU : Core i3 generasi 7

RAM : 4 GB

- Hasil Skenario

Penelitian ini dilakukan di SMAN 1 Baureno pada siswa kelas XI MIPA-3 yang berjumlah 36 siswa dengan 20 soal dan kunci jawaban mata pelajaran Biologi dengan perbandingan nilai menggunakan Query Expansion dan Dice Similarity, dengan nilai tanpa Query Expansion dan menggunakan Dice Similarity, dengan nilai dari Guru.

Pada grafik nilai siswa terlihat bahwa nilai guru lebih besar daripada nilai sistem penilaian jawaban esai otomatis menggunakan Dice Similarity dan Query Expansion dengan nilai sistem menggunakan Dice Similarity tanpa Query Expansion, jadi nilai yang mendekati nilai guru yaitu nilai sistem yang menggunakan Dice Similarity dan Query Expansion.

Uji coba dilakukan kembali menggunakan data jawaban yang dikerjakan oleh siswa SMPN 1 Megaluh kelas IX sejumlah 50 siswa, dengan masing-masing siswa mengerjakan 10 soal. Soal dan kunci jawaban menggunakan mata pelajaran Biologi yang berasal dari buku Panduan Pendidik Ilmu Pengetahuan Alam untuk SMP/MTS terbitan JP BOOKS. Berikut merupakan perbandingan nilai menggunakan Query Expansion dan Dice Similarity, dengan nilai tanpa Query Expansion dan menggunakan Dice Similarity, dengan nilai dari Guru.

Berdasarkan Gambar 4.4 terlihat bahwa nilai guru mengalami fluktuatif data, yang mana nilai guru tidak selalu lebih tinggi dengan nilai sistem. Nilai sistem menggunakan metode Dice Similarity dan Query Expansion selalu lebih tinggi daripada nilai sistem menggunakan metode Dice Similarity tanpa Query Expansion. Sehingga dapat disimpulkan, rata-rata nilai sistem menggunakan metode Dice Similarity dan Query Expansion lebih mendekati nilai guru dibandingkan dengan nilai sistem menggunakan metode Dice Similarity tanpa Query Expansion.

Berdasarkan kedua pengujian yang telah dilakukan, dapat diketahui bahwa pada sistem

penilaian jawaban esai otomatis menggunakan Dice Similarity dan Query Expansion sudah dapat mengatasi masalah pada kata-kata yang berbeda namun memiliki arti atau makna yang sama. Kekurangan dari sistem penilaian esai otomatis tersebut adalah membutuhkan waktu komputasi yang lama daripada sistem penilaian esai otomatis menggunakan metode Dice Similarity tanpa Query Expansion. Waktu yang dibutuhkan untuk penggunaan metode Dice Similarity tanpa Query Expansion dari hasil uji coba yang dilakukan menggunakan 20 soal yang dijawab oleh 36 siswa menghabiskan waktu 214 detik, sedangkan penggunaan metode Dice Similarity dan Query Expansion membutuhkan waktu yang lebih lama yaitu 404 detik.

Setelah itu dilakukan skenario pengujian pada sistem, maka dihitung nilai Error Rate dari pengujian sistem menggunakan Dice Similarity dan Query Expansion, dengan pengujian sistem menggunakan Dice Similarity tanpa Query Expansion. Akurasi yang dihasilkan sistem penilaian esai otomatis menggunakan Dice Similarity dan Query Expansion 82,87% dengan nilai error rate sebesar 0,1713. Sedangkan akurasi yang dihasilkan sistem penilaian esai otomatis menggunakan metode Dice Similarity tanpa Query Expansion sebesar 80,68% dengan nilai error rate sebesar 0,1932.

Pengujian kedua dilakukan dengan data lain menggunakan Dice Similarity dan Query Expansion menghasilkan nilai error rate sebesar 0,1198 dan akurasi sebesar 88,02%. Sedangkan dengan menggunakan metode Dice Similarity tanpa Query Expansion menghasilkan nilai error rate sebesar 0,1529 dan akurasi sebesar 84,71%.

Berdasarkan hasil pengujian yang telah dilakukan, dapat diketahui bahwa penambahan Query Expansion pada metode Dice Similarity dapat meningkatkan nilai akurasi antara penilaian oleh sistem dengan penilaian yang dilakukan oleh guru. Hal ini dapat ditunjukkan dari persentase nilai akurasi dari hasil pengujian pertama yaitu sebesar 82,87% untuk metode Dice Similarity dan Query Expansion dan 80,68% untuk metode Dice Similarity tanpa Query Expansion. Hasil akurasi dari pengujian kedua memiliki nilai akurasi sebesar 88,02% untuk metode Dice Similarity dan Query Expansion dan 84,71% untuk metode Dice Similarity tanpa Query Expansion. Nilai akurasi menunjukkan bahwa sistem mampu melakukan penilaian dengan output nilai yang hampir mirip dengan penilaian oleh guru.

- Analisa Hasil

Berdasarkan kedua uji coba yang telah dilakukan pada sistem penilaian esai otomatis menggunakan metode Dice Similarity dan Query Expansion terdapat beberapa kekurangan pada proses pengimplementasian, seperti tidak bisa mendeteksi struktur kata dalam suatu kalimat, serta sistem juga tidak bisa mendeteksi kata-kata typo, atau kata-kata yang disingkat sehingga hal itu dapat mempengaruhi tingkat akurasi sistem.

Sistem ini pun tidak bisa mendeteksi kemiripan token dengan kata istilah, seperti kata organisme merupakan istilah biologi yang berarti makhluk hidup, namun dalam sistem hanya

memproses tiap token sehingga makhluk hidup tidak bisa didefinisikan sebagai arti dari organisme.

Sistem penilaian esai otomatis menggunakan metode Dice Similarity dan Query Expansion hanya dapat melakukan penilaian kemiripan berdasarkan kata bukan berdasarkan kalimat, sehingga jika ada jawaban siswa dan kunci jawaban memiliki maksud yang sama namun dengan kata-kata yang berbeda dan tidak ada kata yang sama maka sistem akan menampilkan nilai kemiripan kecil, seperti contoh pada kunci jawaban dari soal nomor 8 yaitu "Supinasi adalah gerakan otot untuk memutar lengan atau kaki, sehingga telapak tangan mengarah ke atas. Sementara itu, pronasi adalah gerakan otot lengan bawah, sehingga telapak tangan mengarah ke bawah atau belakang." dan jawaban siswa A. Rizal Sahputra pada soal nomor 8 yaitu "Supinasi = menengadahkan, Pronasi = menelungkup.", nilai kemiripan yang dihasilkan sangat kecil, karena pada kasus tersebut dalam jawaban siswa tidak ada kata yang sama dengan kata dalam kunci jawaban, walaupun dalam penilaian subyektif itu dianggap benar karena memiliki maksud dan arti yang sama.

Sistem penilaian esai otomatis menggunakan metode Dice Similarity dan Query Expansion menilai kemiripan berdasarkan kata, sehingga jika banyak kata yang sama akan menghasilkan nilai kemiripan yang lebih tinggi, dan sebaliknya.

- Kesimpulan

Berdasarkan hasil pengujian dan analisis yang telah dilakukan terhadap sistem penilaian esai otomatis menggunakan metode Dice Similarity dan Query Expansion, dapat ditarik kesimpulan sebagai berikut :

Tingkat akurasi dari sistem penilaian esai otomatis menggunakan metode Dice Similarity dan Query Expansion cukup tinggi. Hal ini dibuktikan dari besarnya nilai akurasi yang dihasilkan sistem dengan penilaian oleh guru. Hasil pengujian sistem dengan data 20 soal yang dijawab oleh 36 siswa menunjukkan bahwa akurasi terbaik dalam penilaian jawaban esai otomatis menggunakan metode Dice Similarity dan Query Expansion sebesar 82,87%.

Penambahan Query Expansion kedalam sistem ini mampu meningkatkan akurasi sistem. Hal ini ditunjukkan dari hasil uji coba yang dilakukan dengan menambahkan metode Query Expansion mendapatkan nilai akurasi sebesar 82,87% dengan waktu komputasi 404 detik. Sedangkan hasil uji coba sistem tanpa penambahan metode Query Expansion mendapatkan nilai akurasi sebesar 80,68% dengan waktu komputasi 214 detik.

- Saran

Penilaian jawaban esai otomatis menggunakan Dice Similarity dan Query Expansion pada penelitian ini masih memiliki akurasi yang belum begitu baik dan memiliki kekurangan waktu komputasi yang dibutuhkan lama. Sistem dapat menggunakan metode Query Expansion selain sinonim kata yang mungkin dapat memiliki akurasi yang lebih baik dan waktu komputasi yang

lebih cepat.

Sistem penilaian esai otomatis menggunakan studi kasus Biologi selanjutnya juga dapat menerapkan dengan tambahan kamus istilah yang mungkin dapat meningkatkan nilai akurasi. Selain itu sistem juga dapat menggunakan metode text mining lainnya selain Dice Similarity. Sistem juga dapat menerapkan pada soal, mata pelajaran ataupun studi kasus lain dengan data jawaban yang bersifat definitif untuk menghasilkan akurasi yang lebih baik.