

Pemodelan Penilaian Essay Otomatis Secara Realtime Menggunakan Kombinasi Text Stemming Dan Cosine Similarity

Komang Rinantha

STMIK STIKOM Bali

Jalan Raya Puputan Renon no. 86 Denpasar, 0361-244445

Email: komangrinantha@gmail.com

Abstrak

Ujian merupakan sebuah proses evaluasi dari kumpulan proses pembelajaran. Ujian dalam pendidikan dapat berupa ujian tertulis maupun ujian lisan. Ujian merupakan suatu pekerjaan yang cukup membutuhkan waktu dalam hal pelaksanaan penilaian dari hasil ujian peserta pembelajaran. Penilaian hasil ujian secara manual terkadang memiliki kelemahan yaitu terdapat penilaian yang berbeda terhadap peserta-peserta ajar dengan jawaban yang hampir sama. Hal ini terjadi apabila ujian yang diberikan dalam bentuk pertanyaan dengan sebuah penjelasan atau pertanyaan dalam bentuk essay. Saat ini ujian juga dapat dibantu dengan menggunakan komputer sebagai alat bantu yang lebih dikenal dengan computer asisted test (CAT). Ujian berbasis komputer juga telah banyak dikembangkan untuk pendidikan salah satunya adalah Ujian Nasional Berbasis Komputer (UNBK) yang digunakan untuk siswa SMP dan SMA atau yang sederajat. Selain itu ujian berbasis komputer juga dikembangkan untuk mahasiswa pada tingkat pendidikan tinggi. Soal ujian yang akan dikembangkan adalah soal ujian berbasis web yang bisa memberikan nilai secara otomatis. Soal yang akan dimasukkan kedalam sistem adalah soal dalam bentuk essay. Pemberian nilai otomatis dilakukan dengan menggunakan stemming Porter yang dikombinasikan dengan Cosine similarity. Algoritma Porter digunakan untuk mencari kata dasar dari kata-kata yang digunakan dalam kunci jawaban. Jawaban dari peserta ujian juga akan dicari kata dasarnya dengan menggunakan algoritma Porter. Hasil dari ekstraksi kata dasar pada kunci jawaban dan jawaban siswa akan dibandingkan dengan menggunakan Cosine similarity. Proses pemodelan dilakukan dengan mengembangkan dari model yang telah ada dengan menyesuaikan konsep realtime yang akan direncanakan. Model yang didapatkan adalah proses text stemming dan proses ekstrak kata dari jawaban dilakukan pada komputer client (pengajar dan siswa) dan tidak membebankan ke sever. Server hanya memroses penyimpanan jawaban dan membandingkan jawaban dengan kunci jawaban.

Kata kunci : Pemodelan, Penilaian otomatis, essay, realtime

Latar belakang

Evaluasi pendidikan sangat diperlukan dalam sebuah proses pembelajaran untuk mengetahui tingkat pemahaman peserta ajar terhadap suatu materi. Evaluasi pembelajaran dalam bentuk ujian bukan menjadi satu rujukan utama dalam menentukan pemahaman peserta ajar. Ujian merupakan suatu pekerjaan yang cukup membutuhkan waktu dalam hal pelaksanaan penilaian dari hasil ujian peserta pembelajaran. Sehingga perlu sebuah solusi dalam meminimalkan waktu penilaian hasil ujian.

Penilaian hasil ujian secara manual terkadang memiliki kelemahan yaitu terdapat penilaian yang berbeda terhadap peserta-peserta ajar dengan jawaban yang hampir sama. Hal ini terjadi apabila ujian yang diberikan dalam bentuk pertanyaan dengan sebuah penjelasan atau pertanyaan dalam bentuk essay. Sehingga diperlukan adanya sebuah penilaian yang sama untuk semua mahasiswa dengan tidak mengurangi arti dari ujian/evaluasi itu sendiri.

String similarity merupakan sebuah metode yang digunakan untuk mendapatkan tingkat kesamaan dari sebuah kalimat ataupun paragraf. Dalam pelaksanaannya *string similarity* dapat memberikan penilaian kesamaan sebuah kalimat atau paragraf dengan rentang penilaian dari 0% hingga 100%. Banyak metode yang dapat digunakan dalam *string similarity*, salah satunya adalah *cosine similarity*.

Dalam proses menghitung kesamaan dari kalimat atau paragraf, diperlukan proses menghilangkan kata penghubung dan juga proses menghilangkan imbuhan pada kata-kata yang terdapat dalam kalimat tersebut. Proses untuk menghilangkan kata penghubung disebut dengan proses menghilangkan *stopword*. Proses untuk menghilangkan imbuhan pada kata disebut dengan proses *stemming*. Proses *stemming* yang digunakan adalah *stemming porter* untuk bahasa indonesia. *Stemming Porter* memerlukan waktu yang lebih singkat dalam proses *stemming* apabila dibandingkan dengan Algoritma Nazief & Adriani[1].

Stemming merupakan suatu proses yang terdapat dalam sistem IR (*Information Retrieval*) yang mentransformasi kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (*root word*) dengan menggunakan aturan-aturan tertentu. Sebagai contoh, kata bersama, kebersamaan, menyamai, akan di-*stem* ke *root word*-nya yaitu “sama”. 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, selain sufiks, prefiks, dan konfiks juga dihilangkan[1].

Saat ini ujian juga dapat dibantu dengan menggunakan komputer sebagai alat bantu yang lebih dikenal dengan *Computer Asisted Test (CAT)*. Ujian berbasis komputer juga telah banyak dikembangkan untuk pendidikan salah satunya adalah Ujian Nasional Berbasis Komputer (UNBK) yang digunakan untuk siswa SMP dan SMA atau yang sederajat. Ujian berbasis komputer juga sudah banyak dikembangkan untuk mahasiswa pada tingkat pendidikan tinggi. Sehingga diperlukan adanya sebuah sistem yang dapat memberikan penilaian secara otomatis secara realtime.

Metode perekayasaan

Penelitian sejenis

Fitri [2] menyatakan bahwa system penilaiannya otomatis menggunakan metode *cosine similarity* berjalan dengan baik untuk ujian essay dalam Bahasa Inggris dengan kesesuaian nilai system dengan nilai yang diberikan oleh pengajar adalah rata-rata 89,48%. Dalam penelitian tersebut sebelum menggunakan *cosine similarity*, data jawaban siswa diproses dengan menggunakan pemrosesan teks dan diberikan bobot dengan TF (*Term Frequency*).

Pramukantoro [3] menyatakan bahwa *cosine similarity* digunakan dalam *elearning* karena *cosine similarity* lebih unggul dalam hal komputasi dibandingkan dengan LSA. Proses uji performa system disimpulkan bahwa semakin besar pengguna mengakses sistem, maka semakin besar CPU Usage yang dibutuhkan. Untuk pengujian *page load time* dan *memory usage* tidak ditemukan perbedaan yang cukup signifikan ketika digunakan oleh satu pengguna maupun oleh banyak pengguna.

Nurdiana [4] menyatakan bahwa *cosine similarity* menghasilkan nilai kemiripan yang tertinggi dan menjadi usulan alternatif untuk mencari kemiripan dari teks mining. Keakuratan pengelompokan dokumen teks dan pemilihan kata untuk dijadikan term sangat terpengaruh oleh kelengkapan daftar stoplist dan hasil root kata dari stemming untuk teks Bahasa Indonesia.

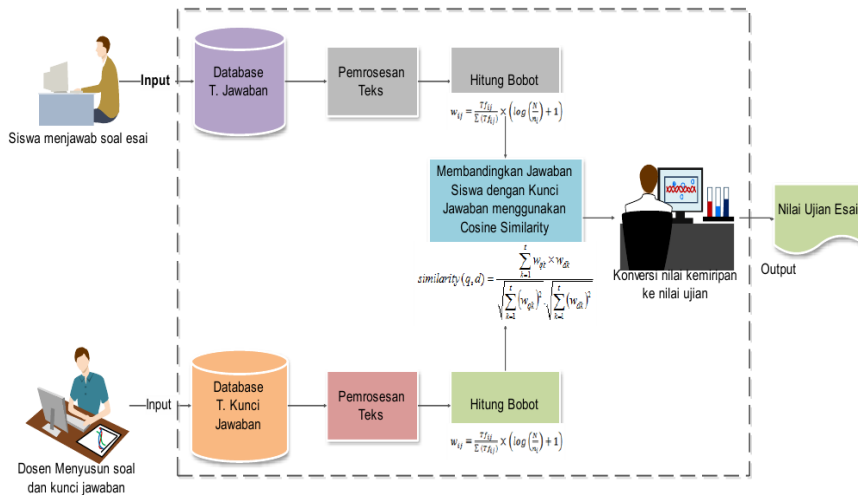
Sulistyo [5] menyatakan bahwa hasil penilaian penilaian menggunakan sistem *text similarity* tidak berbeda signifikan dengan hasil dari pakar. Namun terdapat beberapa tipe soal yang hasilnya berbeda signifikan dikarenakan terdapat karakter unik yang terekam di dalam database serta jawaban pada kunci tidak mengandung *keyword* yang sesuai dengan relevansi jawaban yang benar. Penilaian menggunakan *text similarity* dengan metode TF-IDF menghasilkan *output* yang sesuai dengan spesifikasi dari *user*, namun membutuhkan waktu komputasi yang cukup lama ketika data yang diproses (dokumen) jumlahnya banyak. Ratna [6] menyatakan bahwa sistem *grading* otomatis dilakukan dengan memberikan pembobotan pada kata kunci yang dianggap penting menggunakan metode LSA yang disempurnakan dengan Sistem Penilaian Esei Otomatis (SIMPLE) dengan hasil adalah pada kelas kecil diperoleh nilai kesesuaian dengan *human raters* berkisar 69.80% - 94.64%, sedangkan untuk kelas menengah 77.18% - 98.42%.

Agusta [1] menyatakan bahwa proses *stemming* dokumen teks Bahasa Indonesia menggunakan algoritma Porter membutuhkan waktu lebih singkat dibandingkan dengan dengan algoritma Nazief & Adriani. Proses *stemming* dokumen teks Bahasa Indonesia menggunakan algoritma porter memiliki prosentase keakuratan lebih kecil daripada algoritma Nazief & Adriani.

Dari beberapa penelitian sebelumnya maka didapatkan sebuah hipotesa awal bahwa penilaian essay otomatis dapat dilakukan secara *realtime* dengan menggunakan *stemming* Porter karena waktu yang diperlukan untuk proses *stemming* lebih cepat dibandingkan dengan Nazief & Adriani. *Cosine similarity* dapat digunakan dalam penilaian ujian secara otomatis dengan asumsi bahwa lebih baik daripada LSA (*Latent Semantic Analysis*). Sehingga dapat ditarik sebuah kesimpulan bahwa penilaian dapat dilaksanakan secara *realtime* dengan memanfaatkan *stemming* Porter dan *cosine similarity*.

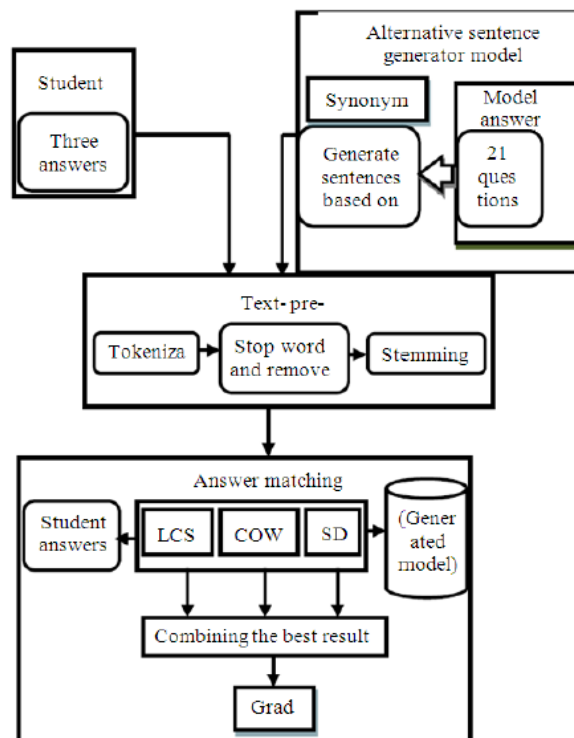
Analisa model yang ada

Studi pendahuluan ini merupakan pembuatan model penilaian essay otomatis secara *realtime* untuk mendapatkan model yang handal untuk dijalankan pada keadaan ujian yang sesungguhnya. Dari penelitian yang pernah dilakukan terdapat beberapa model yang ada :



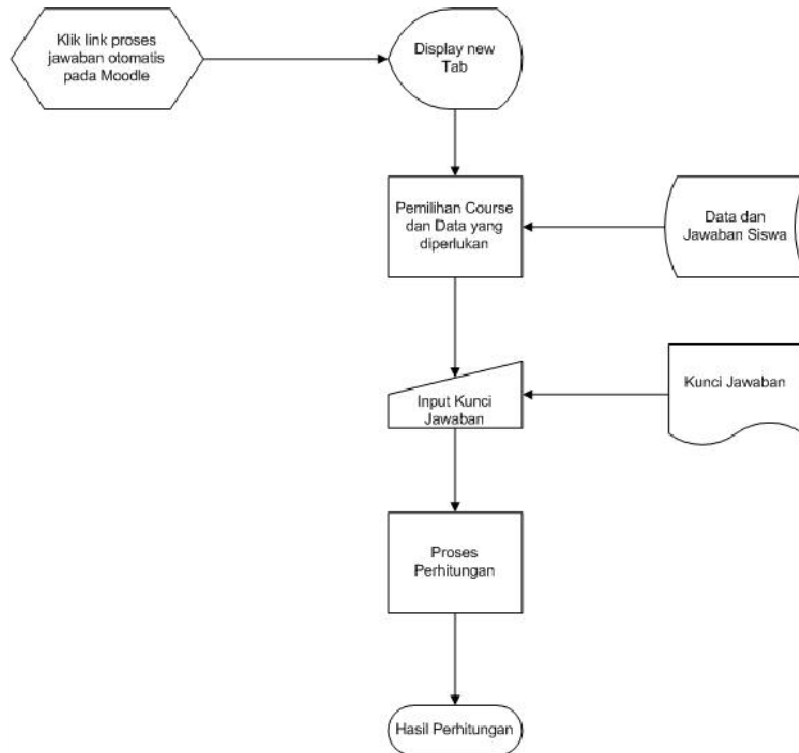
Gambar 1. penilaian ujian essay otomatis[2]

Apabila dilihat dari skema yang ada, maka proses penilaian dilakukan setelah seluruh data jawaban siswa diterima oleh sistem, kemudian dosen melakukan perintah ke sistem untuk melakukan perhitungan. Proses tersebut menunjukkan bahwa nilai tidak dapat dilihat secara *realtime* oleh dosen membutuhkan waktu proses lagi untuk melakukan penilaian.



Gambar 2. automatic essay grading [7]

Dari gambar tersebut terlihat bahwa proses yang dilakukan memang lebih rumit untuk mendapatkan penilaian yang lebih baik terhadap jawaban siswa, namun masih sama dengan yang sebelumnya bahwa proses tersebut tidak dilakukan secara *realtime*.

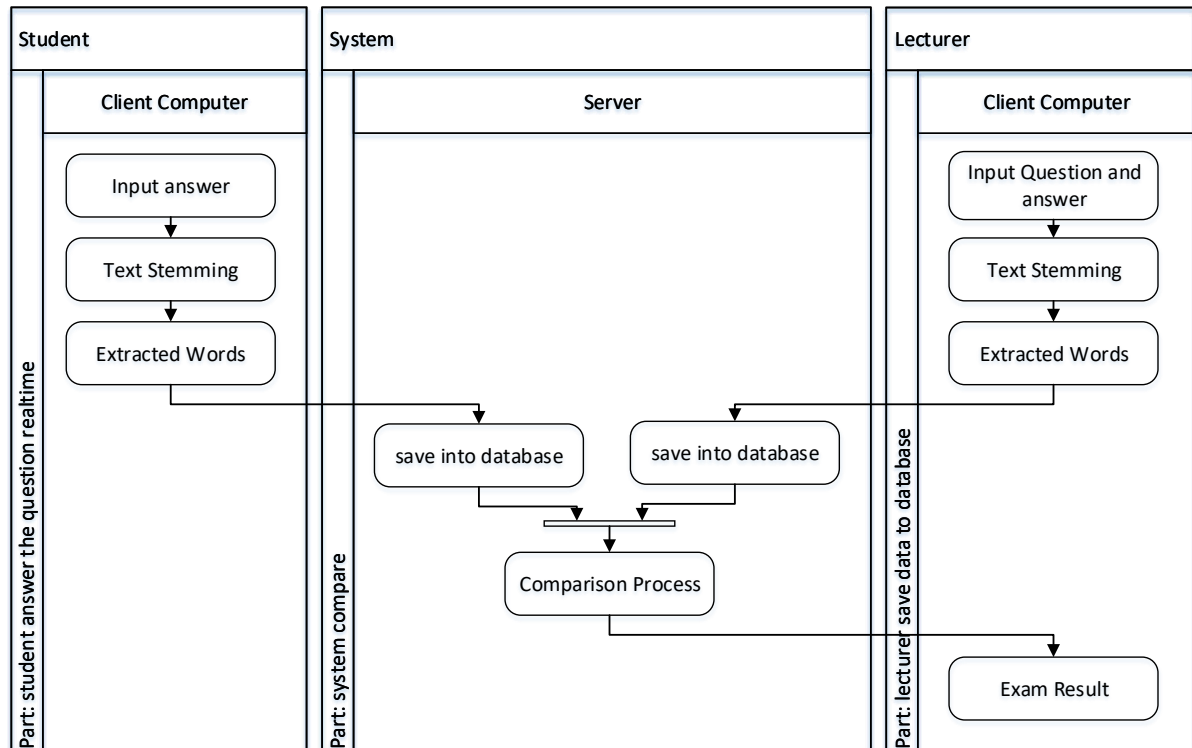


Gambar 3. penilaian ujian bertipe essay [5]

Dari diagram yang ada, terlihat bahwa untuk melakukan penilaian secara otomatis pengguna sistem memilih *course* yang ada dan kemudian diikuti dengan memasukkan kunci jawaban dari soal yang ada. Sehingga bisa ditarik kesimpulan bahwa sistem tidak menilai secara *realtime*.

Hasil dan pembahasan

Dari beberapa model yang ada dapat dibuat sebuah model sistem penilaian essay otomatis secara *realtime* yaitu dengan membagi 2 buah area kerja antara pengajar dan peserta didik agar server aplikasi tidak terbebani secara maksimal. Apabila seluruh proses diberikan pada server maka kinerja dari server akan sangat berat dan mungkin akan membutuhkan sumber daya yang cukup besar. Adapun model yang ditawarkan adalah sebagai berikut :



Gambar 4. model penilaian essay otomatis *realtime*

Pengajar akan memasukkan data soal dan jawaban beserta kuncinya ke dalam sistem pada waktu tertentu. Kemudian data diproses di sisi *client* untuk memperoleh kata-kata kunci dalam jawaban yang kemudian tersimpan di dalam database. Setelah data tersimpan didalam database, system akan digunakan oleh peserta didik untuk melakukan ujian secara *realtime*. Ketika peserta didik menjawab pertanyaan yang ada pada sistem, sistem akan memroses jawaban hingga menemukan kata-kata kunci yang digunakan dalam jawaban tersebut yang kemudian dimasukkan ke dalam *database*. Kata-kata kunci yang didapatkan pada jawaban siswa akan dibandingkan dengan jawaban dari pengajar yang telah diproses sebelumnya dan nilai dari tingkat kesamaan kata-kata akan tersimpan sebagai nilai dari soal ujian yang dikerjakan oleh peserta didik. Hasil ujian akan ditampilkan pada sisi pengajar untuk mengetahui nilai yang didapatkan oleh peserta didik.

Kesimpulan

Dari pelaksanaan studi pendahuluan sistem penilaian otomatis secara *realtime* didapatkan kesimpulan bahwa diharapkan hasil aplikasi *realtime* dapat berjalan dengan baik bila masing-masing komponen sistem dapat menggunakan sumber daya yang ada dengan baik. Apabila server aplikasi menjadi pusat dari proses dan seluruh proses dibebankan ke server, maka sistem *realtime* akan berjalan dengan kurang baik.

Daftar pustaka

- [1] Agusta L. Perbandingan algoritma stemming Porter dengan algoritma Nazief & Adriani untuk stemming dokumen teks bahasa indonesia. Konferensi Nasional Sistem dan Informatika. 2009 Nov:196-201
- [2] Fitri R, Asyikin AN. APLIKASI PENILAIAN UJIAN ESSAY OTOMATIS MENGGUNAKAN METODE COSINE SIMILARITY. POROS TEKNIK. 2015;7(2).
- [3] Pramukantoro ES. Sistem Penilaian Otomatis Jawaban Esai Pada Elearning belajardisini. com. Jurnal Teknologi Informasi dan Ilmu Komputer. 2016 Jun 20;3(2).
- [4] Nurdiana O, Jumadi J, Nursantika D. Perbandingan Metode Cosine Similarity Dengan Metode Jaccard Similarity Pada Aplikasi Pencarian Terjemah Al-Qur'an Dalam Bahasa Indonesia. Jurnal Online Informatika. 2016 Jun 1;1(1).
- [5] Sulisty ME, Saptono R, Asshidiq A. PENILAIAN UJIAN BERTIPE ESSAY MENGGUNAKAN METODE TEXT SIMILARITY. Telematika. 2015 Jul 8;12(2).

-
- [6] Ratna AA, Budiardjo B, Hartanto D. Simple: Sistem Penilai Esei Otomatis untuk Menilai Ujian dalam Bahasa Indonesia. MAKARA. 2009 Aug 28;11(1):pp-5
- [7] Omran AM, Ab Aziz MJ. Automatic essay grading system for short answers in English language. Journal of Computer Science. 2013 Oct 1;9(10):1369.