

Bab 1. Pendahuluan

1.1 Latar belakang

Penilaian dianggap memainkan peranan penting dalam pendidikan. Kepentingan dalam pengembangan dan dalam penggunaan menilai berbasis komputer telah berkembang pesat dalam beberapa tahun terakhir. Komputer yang menjadi komponen penting dalam proses pembelajaran lebih mudah dalam melakukan penilaian pada ujian pilihan ganda dan isian singkat secara akurat dibandingkan dengan ujian *essay* karena jawaban yang ada harus sama, baik dengan pilihan maupun dengan kata-katanya. Padahal jenis ujian pilihan ganda memiliki kekurangan yaitu dari segi tingkat pemahaman *learner* pada soal ujian. Soal *essay* merupakan bentuk evaluasi dimana pilihan jawaban tidak disediakan dan siswa harus menjawab dengan kalimat mereka sendiri, sehingga jawaban yang dihasilkan dapat sangat bervariasi sesuai dengan pemikiran masing-masing *learner*. Banyak para ahli menganggap bahwa *essay* merupakan metode atau alat yang paling tepat untuk menilai hasil dari kegiatan belajar yang kompleks, menyiratkan kemampuan untuk mengingat, mengintegrasikan ide-ide, keahlian dalam menulis, dan kemampuan untuk melengkapi data hanya dari mengidentifikasi masalah.[2]

Motivasi utama dalam membangun *Automatic Essay Grading System* (Sistem Penilaian *Essay* Otomatis) adalah selain untuk mengurangi waktu *learner* dalam mendapatkan *feedback* untuk ujian mereka, juga menghemat biaya dalam proses penilaian, sebab tidak perlu membiayai orang untuk mengoreksi jawaban *learner*. Dimana asumsi dari kebanyakan sistem adalah nilai yang diberikan oleh manusia merupakan nilai yang sebenarnya. Dengan demikian, tujuan dari sistem adalah untuk meniru proses penilaian manusia dan sistem hanya dapat digunakan jika mampu memberikan penilaian sejelas penilaian manusia.[7]

Salah satu metode yang dapat digunakan untuk menilai *essay* secara otomatis adalah metode yang disebut dengan *Probabilistic Latent Semantic Analysis*. Metode *Probabilistic Latent Semantic Analysis* merupakan metode matematis yang mengekstrak merepresentasikan kalimat atau informasi dan mempunyai ciri khas mementingkan kata-kata konsep (makna) yang terkandung dalam sebuah kalimat tanpa memperhatikan karakteristik linguistiknya dengan menggunakan model statistik. Mengapa menggunakan PLSA? Ada beberapa hal yang mendukung adanya metode ini, yang pertama adalah hasil yang menunjukkan kesamaan makna dengan informasi sebelumnya, yang kedua tingkat akuisisi pengertian dari teks yang mendekati manusia, serta mendefinisikan distribusi probabilitas yang tepat ke beberapa dokumen (multi dokumen) berdasarkan data statistik.[7]

TA ini mengimplementasikan metode PLSA yang menggunakan teknik *Expectation Maximization*. Yang mana algoritma EM adalah prosedur berulang yang efisien untuk menghitung kemungkinan maksimum dalam memperkirakan ada tidaknya kata yang hilang atau tersembunyi. Perhitungan matematis dilakukan dengan memetakan ada atau tidak adanya kata dari kelompok kata pada matriks dokumen. Kemudian melakukan prediksi kata seberapa banyak kata yang menjadi ciri khas atau topik dalam dokumen yang nantinya direpresentasikan dalam bentuk matriks yang akan dimodelkan menggunakan metode PLSA. Dokumen-dokumen jawaban yang telah dimodelkan dicari kemiripannya dengan dokumen-dokumen materi dan dokumen jawaban kunci. Sedemikian sehingga *Automatic Essay Grading System* yang dihasilkan akan memberikan penilaian yang mendekati penilaian secara manual.

1.2 Perumusan Masalah

Pada tugas akhir ini masalah yang akan diselesaikan yaitu:

1. Bagaimana mengimplementasikan metode *Probabilistic Latent Semantic Analysis* pada *Automatic Essay Grading System*.
2. Melakukan perbandingan akurasi antara nilai dari sistem *Automatic Essay Grading* dengan *human raters* (penilaian manual).

1.3 Tujuan

Tujuan tugas akhir ini adalah:

1. Implementasi *Automatic Essay Grading* dengan menggunakan metode *Probabilistic Latent Semantic Analysis*.
2. Menganalisa pengaruh parameter variabel laten metode *Probabilistic Latent Semantic Analysis* yang digunakan pada *Automatic Essay Grading System* dalam menemukan hasil penilaian yang optimal.

1.4 Hipotesa

Pada *Automatic Essay Grading System* untuk parameter variabel laten semakin banyak jumlah variabel laten maka tingkat similarity dengan *human rater* semakin meningkat.

1.5 Batasan masalah

Terdapat beberapa batasan dalam penelitian tugas akhir ini, antara lain:

1. Dokumen text yang digunakan berbahasa Inggris.
2. Tidak menangani *essay arithmetic* (kalimat *essay* murni) atau yang sejenis serta bahasa pemrograman.
3. Human rater yang menilai maksimal 3 orang untuk menjaga objektivitas penilaian.
4. Tidak menangani kasus gramatikal.
5. Penilaian dilakukan secara offline
6. Penilaian yang dilakukan adalah penilaian pada multi dokumen essay.
7. Evaluasi penilaian dilakukan dengan membandingkan nilai keluaran sistem dengan nilai manual dari dosen.

1.6 Metodologi penyelesaian masalah

Metodologi yang dilakukan untuk menyelesaikan permasalahan adalah sebagai berikut:

- a. Studi Literatur :
Pencarian literatur – literatur yang sesuai dengan permasalahan meliputi : mempelajari dan memahami konsep serta landasan teori tentang *Automatic Essay Grading* dan metode *Probabilistic Latent Semantic Analysis*.
- b. Analisis dan Design
Melakukan analisis terhadap *Probabilistic Latent Semantic Analysis* dan parameter-parameter yang perlu dipertimbangkan dalam penyelesaian. Melakukan analisis kebutuhan perangkat lunak, serta perancangan dan desain perangkat lunak.

- c. Implementasi :
Pada tahap ini dilakukan implementasi berdasarkan hasil rancangan,yaitu membangun *Automated Essay Grading* Sistem dengan metode *Probabilistic Latent Semantic Analysis* dan melakukan evaluasi kinerja metode yang digunakan sebagai kajian selanjutnya.
- d. Testing.
Pengujian terhadap *Automatic Essay Grading System* yang telah dibangun, apakah sudah bekerja dengan benar. Skenario uji yang dilakukan adalah dengan memberikan jawaban yang diberikan oleh *learner* kepada beberapa orang penguji, kemudian nilai yang diberikan oleh para penguji tersebut dibandingkan dengan hasil penilaian yang diberikan oleh sistem.
- e. Analisis Hasil.
Melakukan analisis terhadap data dan hasil pengujian dari penggunaan *Automatic Essay Grading System*. Dari hasil tahap ini, ditarik kesimpulan dan diusulkan saran untuk pengembangan lebih lanjut
- f. Pengambilan Kesimpulan dan saran dari hasil analisis hasil uji coba aplikasi dan membandingkan dengan teori – teori pada saat studi literatur.
- g. Penyusunan laporan tugas akhir.