

# BAB 1

## PENDAHULUAN

Bab ini menguraikan latar belakang penelitian beserta perumusan masalah, tujuan, dan ruang lingkup penelitian, yang dilanjutkan dengan penjelasan tahapan penelitian dan sistematika penulisan laporan.

### 1.1 Latar Belakang

Evaluasi hasil belajar merupakan komponen penting dalam proses pembelajaran. Melalui evaluasi, dapat diketahui sejauh mana pemahaman peserta didik terhadap materi ajar serta pencapaian tujuan pembelajaran. Ujian atau tes hasil belajar adalah salah satu bentuk evaluasi hasil belajar. Dilihat dari bentuk soalnya, ujian terdiri dari dua macam, yaitu ujian bentuk subjektif dan ujian bentuk objektif. Ujian bentuk subjektif disebut juga dengan ujian esai, dan ujian bentuk objektif dikenal juga dengan tes jawaban pendek (*short answer test*), tes “ya-tidak” (*yes-no test*), dan tes model baru (*new type test*) yang memberi pilihan jawaban kepada peserta ujian, contohnya tes pilihan ganda (*multiple choice test*) (Sudijono, 2005).

Ujian esai memiliki beberapa kelebihan dibandingkan dengan ujian bentuk objektif. Kelebihannya antara lain dapat mencegah timbulnya permainan spekulasi pada peserta ujian dan melalui ujian esai ini dapat diketahui tingkat kedalaman penguasaan peserta ujian terhadap materi yang diujikan. Selain itu, peserta ujian akan terbiasa untuk mengemukakan pendapat dan pikiran dengan gayanya sendiri.

Namun bentuk ujian ini memiliki kekurangan, antara lain kesulitan dalam penilaiannya. Dibutuhkan banyak waktu, pikiran, dan tenaga dalam penilaiannya. Menurut Mason seperti yang disampaikan dalam (Valenti, Neri, & Cucchiarelli, 2003), hampir 30% dari waktu yang dimiliki seorang guru digunakan untuk menilai ujian atau evaluasi siswa. Selain itu, mungkin saja terdapat subjektivitas dari penilai ketika memeriksa dan menilai jawaban esai.

Kekurangan bentuk ujian esai di atas dapat diatasi dengan penggunaan teknologi informasi dan komputasi, yaitu dengan melakukan otomatisasi penilaian jawaban ujian esai. Hal ini mungkin dilakukan karena sekarang proses belajar

mengajar di beberapa tempat sudah ditunjang dengan *e-learning*, yaitu proses belajar mengajar dengan memanfaatkan bantuan teknologi informasi. Melalui jaringan komputer ditambah dengan *learning management system* (LMS), berbagai aktivitas belajar mengajar dapat dilakukan secara *online*, terkomputerisasi, dan melibatkan dokumen digital. Salah satu aktivitas tersebut adalah ujian. Beberapa LMS menyediakan modul ujian berikut sistem penilai otomatisnya. Namun, sampai saat ini tipe ujian yang banyak digunakan adalah ujian bentuk objektif.

Sistem penilai otomatis untuk ujian bentuk subjektif atau esai berbahasa Inggris sudah banyak dikembangkan. Sistem ini terdiri dari sistem yang menilai esai berdasarkan teknik atau gaya penulisan dan berdasarkan konten esai (Valenti, Neri, & Cucchiarelli, 2003). Esai yang dinilai dapat berupa esai karangan (argumentasi, narasi, deskripsi, atau eksposisi) atau esai pendek yang merupakan jawaban ujian bentuk esai. Contoh sistem penilai esai ini antara lain PEG (*Project Essay Grade*), IEA (*Intelligent Essay Assessor*), E-Rater (*Electronic Essay Rater*), dan C-Rater (*Conceptual Rater*).

PEG adalah salah satu sistem yang paling awal dikembangkan. Sistem yang dikembangkan oleh Page ini menilai gaya atau teknik penulisan dari faktor-faktor intrinsik seperti panjang esai, diksi, dan lain-lain. IEA adalah sistem penilai esai berdasarkan metode *Latent Semantic Analysis* (LSA), yang digunakan untuk menilai konten esai. E-Rater dikembangkan oleh Burstein menggunakan metode statistika dan NLP (*Natural Language Processing*) untuk menilai gaya dan teknik penulisan. C-Rater menggunakan NLP untuk menilai jawaban esai pendek seperti yang terdapat pada bagian evaluasi pada buku teks pelajaran atau ujian. Sistem ini menilai konten esai dilihat dari benar tidaknya esai tersebut menjawab pertanyaan (Valenti, Neri, & Cucchiarelli, 2003).

Sementara itu, sistem penilai esai berbahasa Indonesia sudah mulai dikembangkan pada tahun 2005 melalui penelitian Krisnanda (Krisnanda, 2005). Sistem ini menggunakan metode yang sama dengan yang digunakan IEA, yaitu LSA untuk menilai jawaban ujian esai. Jawaban siswa dan kunci jawaban dari dosen direpresentasikan sebagai vektor dalam ruang LSA. Nilai jawaban siswa dihitung dari banyaknya kata kunci yang cocok dengan kunci jawaban, yang

dinyatakan oleh kedekatan antara dua vektor yang mewakilinya. Penelitian ini menunjukkan bahwa semakin banyak kata kunci yang digunakan, maka nilai yang diberikan sistem semakin mendekati nilai yang diberikan manusia. Untuk mengevaluasi sistem, nilai dari sistem dan nilai dari manusia dihitung korelasinya. Pada penelitian ini korelasi nilai sistem dengan penilai manusia mencapai 0.86-0.96.

Penelitian selanjutnya dilakukan oleh Ratna pada tahun 2007 (Ratna, Budiardjo, & Hartanto, 2007). Sistem berbasis web yang diberi nama SIMPLE (SIstiM PeniLaian Esei Otomatis) ini digunakan untuk menilai jawaban ujian esai berdasarkan teknik LSA. Pada publikasi penelitiannya, disebutkan bahwa sistem ini sudah dilengkapi dengan pembobotan kata, urutan kata, dan persamaan kata, namun tidak disebutkan bagaimana implementasi pembobotan dan persamaan katanya.

Uji coba SIMPLE dilakukan untuk ujian esai yang terdiri dari 10 soal dan diikuti dua kelas mahasiswa yang terdiri dari kelas kecil (5 peserta) dan kelas menengah (10 peserta). Hasil uji coba ini menunjukkan SIMPLE menghasilkan persetujuan (*agreement*) nilai dengan penilai manusia sebesar 69.80%-94.64% untuk kelas kecil dan 77.18%-98.42% untuk kelas menengah. Tidak dijelaskan bagaimana persetujuan ini dihitung dan seperti apa bentuk soal yang digunakan. Saran untuk pengembangan selanjutnya antara lain memperhatikan pemilihan kata kunci, urutan kata, dan persamaan arti kata kunci untuk meningkatkan keakuratan sistem.

Penelitian lainnya yang merupakan pengembangan lanjut dari SIMPLE dilakukan masing-masing oleh Octaria, Harisma, dan Hermawandi pada tahun 2008. Ketiga penelitian ini berfokus pada penerapan skema pembobotan kata-kata pada jawaban esai. Octaria menerapkan empat kombinasi pembobotan global dan lokal (Octaria, 2008). Harisma menerapkan tiga tingkatan bobot berbeda yang dapat dipilih dosen untuk setiap kata kunci jawaban, masing-masing bernilai 1, 2, dan 3 (Harisma, 2008). Hermawandi menerapkan skema pembobotan SICBI, yang merupakan akronim dari SQRT-IGFF-COSN-BNRY-IDFB (Hermawandi, 2008).

Penelitian Octaria menunjukkan bahwa hasil terbaik didapatkan dengan menggunakan pembobotan lokal *Binary* dan global *Normal* dengan korelasi nilai

sistem dengan nilai manusia sebesar 0.39, diujikan pada 20 jawaban siswa. Menurut Octaria, faktor yang mempengaruhi rendahnya korelasi antara lain karena sistem tidak bisa mencocokkan kata yang memiliki imbuhan berbeda, padahal kata dasar dan maknanya sama, dan tidak bisa menangkap persamaan makna kata atau sinonim. Misalnya jawaban siswa menggunakan kata “mengirimkan” sedangkan pada kunci jawaban terdapat kata “mengirim”, atau jawaban siswa menggunakan “mengawasi” dan kunci jawaban menggunakan “memonitor” sehingga dianggap tidak cocok, padahal maksudnya sama (Octaria, 2008).

Penelitian Harisma menghasilkan korelasi nilai sistem dengan nilai dari manusia sebesar 0.77 dengan rata-rata selisih nilai 17.36, diujikan pada 10 soal esai untuk 23 mahasiswa. Secara umum nilai yang didapat peserta ujian tidak baik, hal ini mungkin dikarenakan ada kesalahan penulisan kata-kata oleh peserta ujian, yang masih bisa ditolerir manusia dan tidak bisa ditolerir sistem (Harisma, 2008).

Penelitian Hermawandi diujikan pada 10, 15, dan 20 siswa untuk 10 pertanyaan dengan rata-rata selisih nilai dari sistem dan manusia berturut-turut 13.98, 17.84, dan 10.90. Penerapan pembobotan SICBI dapat meningkatkan kualitas sistem, dilihat dari makin kecilnya selisih antara nilai sistem dengan nilai dari manusia. Contohnya untuk 20 siswa, rata-rata selisih nilai dari sistem dan manusia tanpa pembobotan adalah 28.34, sedangkan dengan pembobotan SICBI sebesar 10.90 (Hermawandi, 2008).

Semua penelitian untuk jawaban esai berbahasa Indonesia yang sudah disebutkan di atas menggunakan teknik LSA, yang dapat diartikan sebagai teori induksi pengetahuan dan representasi yang dapat menentukan persamaan makna dari kata atau dokumen dengan menganalisis kumpulan dokumen berukuran besar (korpus). Semua dokumen dan kata pada dokumen kemudian direpresentasikan sebagai titik dalam ruang semantik berdimensi tinggi (misalnya 50-1500) (Landauer, Foltz, & Laham 1998).

Penelitian penilaian esai berbahasa Indonesia yang sudah disebutkan di atas hanya menggunakan kumpulan esai siswa dan kunci jawaban untuk membangun ruang semantik. Padahal, diperlukan data *training* yang cukup besar

dan mewakili konsep yang benar atas jawaban yang akan dinilai untuk membangun ruang semantik sehingga bisa menangkap persamaan makna walaupun kata yang digunakan berbeda. Menggunakan hanya jawaban esai dan kunci dapat menyebabkan ruang semantik yang dibangun bias karena tidak semua esai siswa menyatakan jawaban yang benar. Selain itu, sistem menjadi tidak bisa menangkap persamaan kata seperti yang diungkapkan Octaria (Octaria, 2008).

Keterbatasan data *training* menjadi kendala dalam implementasi LSA untuk penilaian jawaban ujian esai. Selain itu, bila data yang digunakan terlalu umum dalam bidang ilmu yang diujikan, maka bisa jadi jawaban esai untuk soal yang lain dapat diterima untuk sebuah soal yang akan dinilai, karena LSA menilai kedua esai tersebut masih berkaitan.

Dalam kondisi seperti ini, penggunaan *Vector Space Model* (VSM) untuk menilai jawaban esai dapat dipertimbangkan. Apalagi untuk esai yang merupakan jawaban dari pertanyaan yang sudah pasti jawabannya, yaitu sesuai dengan teori yang sudah diajarkan. VSM dapat digunakan untuk mencocokkan kata-kata yang ada pada jawaban siswa dengan kata-kata yang ada pada kunci jawaban tanpa melakukan pengurangan dimensi matriks seperti yang dilakukan LSA.

Bagaimanapun, penelitian mengenai sistem penilai jawaban esai untuk Bahasa Indonesia yang menggunakan LSA dapat menjadi pijakan awal untuk penelitian ini. Sehingga dapat dilakukan perbandingan antara metode LSA dan VSM untuk penilaian jawaban esai berbahasa Indonesia. Fitur tambahan akan ditambahkan dalam penelitian ini, yaitu perluasan kunci jawaban sebagai bentuk *query expansion* pada pencocokan jawaban siswa dengan kunci jawaban, dan penggunaan program pemotong imbuhan kata (*stemmer*). Kedua fitur ini ditambahkan untuk mengatasi kekurangan yang disebutkan dalam penelitian (Octaria, 2008), yaitu agar sistem dapat mengenali persamaan kata dan penggunaan kata dengan imbuhan yang berbeda.

## 1.2. Perumusan Masalah

Masalah yang menjadi fokus utama dalam penelitian ini adalah sebagai berikut:

1. Bagaimana perbandingan efektivitas penilai jawaban ujian esai otomatis

untuk Bahasa Indonesia dengan menggunakan *Vector Space Model* (VSM) dan *Latent Semantic Analysis* (LSA)?

2. Bagaimana pengaruh perluasan kunci jawaban (*query expansion*) dan pemotongan imbuhan kata (*stemming*) terhadap efektifitas sistem?

### 1.3. Tujuan

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengimplementasikan sistem penilai jawaban esai otomatis untuk menilai jawaban esai berbahasa Indonesia menggunakan VSM dan LSA serta membandingkan efektivitas kedua metode tersebut.
2. Mengetahui pengaruh perluasan kunci jawaban dan pemotongan imbuhan kata terhadap efektivitas sistem.

### 1.4. Ruang Lingkup Penelitian

Penelitian ini dibatasi pada pengembangan sistem penilai jawaban ujian esai otomatis untuk esai berbahasa Indonesia menggunakan pendekatan perolehan informasi yaitu metode VSM dan LSA. Sistem LSA yang dikembangkan pada penelitian ini terdiri dari tiga skema sebagai berikut:

- a) Penilaian jawaban esai menggunakan LSA yang dibuat sama dengan sistem yang dikembangkan (Hermawandi, 2008), (Octaria, 2008), dan (Harisma, 2008) dengan cara direkonstruksi berdasarkan keterangan pada publikasi penelitiannya. Ruang semantik pada penilaian ini dibentuk oleh jawaban siswa dan *query*-nya adalah kunci jawaban.
- b) Penilaian jawaban esai menggunakan LSA dengan menambahkan data *training* untuk membentuk ruang semantik, ditambah dengan kunci jawaban. Sedangkan yang menjadi *query* adalah jawaban-jawaban siswa.
- c) Penilaian jawaban esai menggunakan LSA dengan data *training* dan jawaban-jawaban siswa untuk membangun ruang semantik dan yang menjadi *query* adalah kunci jawaban.

Esai yang dinilai dibatasi pada esai yang merupakan jawaban dari pertanyaan pada ujian atau evaluasi, bukan esai karangan yang panjang. Jawaban yang bisa dinilai adalah jawaban yang tidak mementingkan urutan kata dan tidak

mengandung makna negatif seperti yang dinyatakan dengan kata “tidak”, “bukan”, dan sebagainya. Hal ini disebabkan karena VSM dan LSA tidak memperhatikan urutan kata dan tidak bisa membedakan makna negatif. Penilaian jawaban esai dilakukan dengan mencocokkannya dengan kunci jawaban.

Penerapan perluasan kunci jawaban dilakukan dengan menyusun daftar persamaan kata, yang kata-katanya diambil dari kunci jawaban, dan persamaan katanya diambil dari WordNet Bahasa Indonesia.

Pembuatan pemotongan imbuhan tidak termasuk lingkup dari penelitian ini. Penelitian ini menggunakan program pemotong imbuhan untuk dokumen berbahasa Indonesia yang sudah dikembangkan oleh Ichsan dan Adriani (Ichsan & Adriani, 2006).

### **1.5. Tahapan Penelitian**

Penelitian ini dilakukan dengan tahapan sebagai berikut:

#### **1. Studi Pustaka**

Pada tahapan ini, penulis mencari literatur yang berkaitan dengan topik sistem penilai esai otomatis. Literatur yang dipelajari terdiri dari perkembangan sistem penilai esai otomatis di luar negeri dan di Indonesia, *Vector Space Model (VSM)*, *Latent Semantic Analysis (LSA)*, perluasan kunci jawaban atau *query expansion*, dan pemotongan imbuhan kata.

#### **2. Pengumpulan Data Jawaban Esai**

Pada tahapan ini dilakukan pengumpulan jawaban esai yang akan dinilai oleh sistem, yang diambil dari jawaban ujian di lingkungan Fakultas Ilmu Komputer Universitas Indonesia.

#### **3. Analisis dan Perancangan**

Pada tahapan ini dilakukan analisis dan perancangan sistem penilai jawaban esai berbasis VSM dan LSA, berikut dengan perancangan implementasi perluasan kunci jawaban dan penggunaan pemotong imbuhan untuk Bahasa Indonesia.

#### **4. Implementasi dan Uji Coba**

Perancangan yang sudah dibuat sebelumnya diimplementasikan dalam tahapan ini. Berikutnya, dilakukan uji coba pada sistem dengan data

jawaban esai yang sudah dikumpulkan.

#### 5. Analisis Hasil Uji Coba

Pada tahapan ini dilakukan interpretasi terhadap hasil uji coba dan dilakukan analisis untuk mengetahui perbandingan VSM dan LSA serta pengaruh perluasan kunci jawaban dan pemotongan imbuhan kata terhadap efektifitas sistem.

### 1.6. Sistematika Penulisan

Sistematika penulisan laporan ini adalah sebagai berikut:

#### BAB 1 Pendahuluan

Bab ini menjabarkan latar belakang yang mendasari penelitian ini beserta perumusan masalah, tujuan penelitian, dan ruang lingkupnya. Kemudian bagian ini menjelaskan tahapan penelitian dan sistematika penulisan laporan penelitian ini.

#### BAB 2 Landasan Teori

Bab ini menguraikan hasil studi pustaka yang mendasari penelitian ini. Studi pustaka meliputi perkembangan sistem penilai esai otomatis di luar negeri dan di Indonesia, VSM, LSA, perluasan kunci jawaban, dan pemotongan imbuhan kata.

#### Bab 3 Analisis dan Perancangan

Bab ini menguraikan analisis dan perancangan arsitektur sistem penilai jawaban esai berbasis VSM dan LSA berikut dengan perancangan perluasan kunci jawaban dan penggunaan pemotong imbuhan Bahasa Indonesia.

#### Bab 4 Implementasi

Bab ini menjelaskan detail implementasi sistem untuk mengerjakan tugas yang sudah dirancang pada bagian sebelumnya, yang terdiri dari implementasi sistem penilai jawaban esai, penerapan perluasan kunci jawaban, dan penggunaan pemotong imbuhan.

#### Bab 5 Uji Coba dan Analisis

Bab ini menguraikan uji coba yang dilakukan untuk menguji sistem yang telah dibuat beserta analisis hasilnya.



## Bab 6 Penutup

Bab ini terdiri dari kesimpulan dan saran. Pada bagian ini diberikan kesimpulan penelitian, serta diberikan saran-saran untuk penelitian dan pengembangan berikutnya.

