BAB 6
PENUTUP

Bab ini menjelaskan kesimpulan yang diperoleh dari penelitian dan saran yang dapat diberikan untuk penelitian mengenai sistem penilaian jawaban esai otomatis.

6.1 Kesimpulan

Sistem penilai jawaban ujian dengan metode VSM dan tiga skema penilaian LSA telah diimplementasikan dalam penelitian ini dan dilakukan uji coba untuk membandingkan empat skema penilaian beserta pengaruh pemotong imbuhan dan perluasan kunci jawaban. Kesimpulan yang diperoleh dari penelitian ini adalah sebagai berikut:

1. Secara keseluruhan, rata-rata korelasi nilai manusia dengan nilai sistem yang dihasilkan VSM lebih tinggi dari tiga skema penilaian LSA. Hal ini mungkin disebabkan karena tidak adanya dokumen *training* yang cukup banyak untuk membangun ruang semantik LSA. Dengan keterbatasan dokumen tersebut, penggunaan VSM mungkin dapat menjadi alternatif untuk menilai jawaban ujian esai.
2. Rata-rata korelasi nilai manusia dengan nilai LSA paling tinggi dihasilkan dengan skema penilaian LSA3, di mana ruang semantik dibangun dari dokumen *training* dan jawaban siswa. Hal ini mungkin disebabkan karena makin banyaknya informasi yang bisa didapatkan dari dokumen tersebut untuk melakukan analisis yang lebih baik. Tetapi hal ini tidak berlaku jika ditambahkan pemotong imbuhan.
3. Pengaruh pemotong imbuhan pada sistem VSM menaikkan rata-rata korelasi nilainya dengan nilai manusia sedangkan pada LSA menurunkan rata-rata korelasi nilainya dengan nilai manusia. Hal ini mungkin disebabkan karena VSM yang bergantung pada pencocokkan kata lebih diuntungkan karena kata yang cocok semakin banyak akibat pemotongan imbuhan. Sedangkan LSA yang menggunakan analisis global dari kemunculan kata pada seluruh jawaban mungkin mengalami perubahan interpretasi disebabkan oleh kurang jelasnya konteks akibat pemotongan

Perbandingan metode..., Heninggar Septiantri, FASILKOM UI, 2009

imbuhan.

4. Pengaruh perluasan kunci jawaban pada sistem VSM sedikit menurunkan rata-rata korelasi nilainya dengan nilai manusia, sedangkan pada LSA menaikkan rata-rata korelasinya dengan nilai manusia. Hal ini mungkin disebabkan karena penambahan sinonim dapat memperjelas konteks suatu kata dalam jawaban sehingga LSA dapat menangkap hubungan antar kata/dokumen dengan lebih baik.
5. Pengaruh gabungan pemotongan imbuhan dan perluasan kunci jawaban tidak dapat disimpulkan karena pengaruhnya berbeda-beda. Rata-rata korelasi nilai manusia dengan nilai VSM dan LSA2 naik, sedangkan pada LSA1 tetap dan LSA3 turun. Mungkin saja pengaruh yang lebih besar diberikan oleh pemotong imbuhan (lebih dominan) karena perubahan pemotongan imbuhan terjadi di seluruh dokumen. Sedangkan pengaruh perluasan kunci jawaban mungkin lebih kecil karena hanya berpengaruh pada penambahkan sinonim untuk 29 kata.

6.2 Saran

Beberapa saran yang dapat diberikan untuk penelitian sistem penilai jawaban esai otomatis untuk Bahasa Indonesia adalah sebagai berikut:

1. Meneliti efektivitas sistem penilai jawaban esai berbasis LSA dengan dokumen *training* yang banyak, sesuai dengan jawaban yang akan dinilai, dan spesifik untuk setiap jawaban esai untuk mendapatkan hasil yang lebih baik.
2. Agar lebih optimal, penelitian sistem penilai jawaban esai dapat dilakukan dengan jawaban esai yang tidak terlalu panjang.
3. Selain untuk jawaban esai, penelitian dapat dilakukan untuk esai karangan dengan menggunakan penilaian LSA.
4. Penelitian lebih lanjut dapat dilakukan dengan pendekatan *Natural Language Processing* untuk penilaian jawaban esai yang lebih akurat.

Perbandingan metode..., Heninggar Septiantri, FASILKOM UI, 2009