

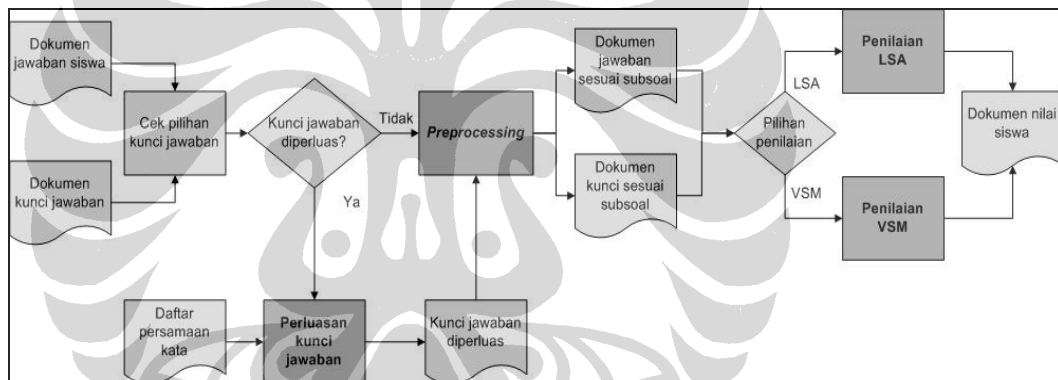
BAB 3

ANALISIS DAN PERANCANGAN

Bab ini menguraikan hasil analisis dan perancangan sistem penilaian jawaban esai otomatis untuk Bahasa Indonesia dengan menggunakan VSM dan LSA. Subbab 3.1 menjelaskan alur sistem secara keseluruhan. Subbab 3.2-3.6 menjelaskan rincian dari tiap bagian pada sistem.

3.1 Rancangan Alur Sistem

Sistem penilai jawaban esai yang dikembangkan terdiri dari satu sistem yang bisa menilai jawaban esai berdasarkan metode VSM dan LSA. Pemilihan VSM atau LSA ditentukan oleh pengguna. Berikut ini adalah rancangan alur sistem penilai jawaban esai otomatis dengan VSM dan LSA.



Gambar 3.1: Alur Sistem Penilai Jawaban Esai Otomatis

Terdapat empat proses utama pada sistem ini, yaitu perluasan kunci jawaban, *preprocessing*, penilaian jawaban esai dengan VSM dan penilaian jawaban esai dengan LSA. Dokumen masukan sistem terdiri dari dokumen jawaban siswa dan dokumen kunci jawaban. Tambahan dokumen yang diperlukan adalah dokumen *training* untuk penilaian LSA, dan dokumen yang berisi daftar persamaan kata untuk proses perluasan kunci jawaban. Dokumen keluaran sistem adalah dokumen nilai siswa.

Perluasan kunci jawaban merupakan pilihan yang dapat dipilih pengguna. Bagian ini memproses dokumen kunci jawaban yang diberikan pengguna dan melakukan perluasan kunci jawaban dengan melihat pada daftar persamaan kata

yang sudah dibuat sebelumnya lalu menambahkan kata-kata persamaannya untuk memperbarui kunci jawaban.

Bagian *preprocessing* menerima masukan berupa kumpulan jawaban siswa dan kumpulan kunci jawaban. Dua dokumen ini dibersihkan dari karakter-karakter dan kata-kata yang tidak penting (*stopwords*) kemudian dibagi menjadi sejumlah dokumen jawaban dan kunci yang berbeda sesuai dengan jumlah subsoal yang terdapat pada jawaban.

Dokumen jawaban dan kunci ini kemudian menjadi masukan dari bagian berikutnya yaitu penilaian VSM atau penilaian LSA. Proses yang dilakukan pada bagian penilaian VSM dan LSA secara umum meliputi pembuatan matriks kata-dokumen dan vektor *query* serta proses penilaian jawaban siswa dengan cara membandingkannya dengan kunci jawaban.

Perincian mengenai setiap bagian sistem beserta dokumen masukan dan keluaran dijelaskan pada lima subbab berikutnya.

3.2 Dokumen Masukan dan Dokumen Keluaran

Dokumen masukan untuk sistem ini terdiri dari dua dokumen teks yaitu dokumen jawaban siswa dan dokumen kunci jawaban untuk satu soal esai. Dalam suatu soal ujian esai, pertanyaan yang diajukan biasanya ada yang terdiri dari beberapa inti pertanyaan. Contohnya adalah sebagai berikut:

“Apa yang dimaksud dengan strategi? Berikan tiga contoh elemen pokok strategi!”

Soal tersebut terdiri dari dua inti pertanyaan, yaitu definsi strategi dan tiga elemen strategi. Dalam penelitian ini, soal seperti contoh di atas dibagi ke dalam beberapa subsoal agar konsep yang disampaikan pada jawaban tidak bercampur, sehingga setiap subsoal mewakili satu konsep yang spesifik. Contoh soal di atas dibagi ke dalam dua subsoal.

Dokumen masukan berupa jawaban siswa diberi tanda pada setiap subsoalnya dengan menggunakan *tag*. Selain untuk menandai subsoal, pada dokumen jawaban siswa digunakan *tag* lain untuk menandai nomor mahasiswa tersebut. *Tag* yang digunakan dalam dokumen jawaban siswa adalah sebagai berikut:

- a) *Tag* `<mhs></mhs>` digunakan untuk menandai nomor mahasiswa dan jawaban mahasiswa tersebut.
- b) *Tag* `<sub>` digunakan untuk menandai awal jawaban untuk setiap subsoal.
- c) *Tag* `</sub>` digunakan untuk menandai akhir jawaban untuk setiap subsoal.

Contoh dokumen jawaban siswa adalah sebagai berikut.

```

<mhs>1</mhs>
<sub>
strategy suatu formulasi yang dirancang sehingga perusahaan
dapat bertahan dan mencapai tujuannya biasanya dilengkapi
dengan perencanaan dan aturan2
</sub>
<sub>
a forecasting => memprediksi bisnis teknologi ekonomi trend dll
b resource allocation => mengalokasikan sumber daya yang
dibutuhkan
c core competency => bisnis utama yang dijalankan
</sub>

<mhs>2</mhs>
<sub>
strategy suatu rencana jangka panjang untuk melaksanakan
bisnis agar mampu bersaing dan mencapai target yang
ditentukan
</sub>

```

Gambar 3.2: Contoh Dokumen Jawaban Siswa

Dokumen masukan berupa kunci jawaban juga diberi *tag* `_{` dan `}` untuk menandai subsoalnya, ditambah dengan *tag* `<subsoal></subsoal>` yang menunjukkan jumlah subsoal. Contoh dokumen kunci jawaban adalah sebagai berikut.

```

<subsoal>2</subsoal>
<sub>
strategi formula menyeluruh bisnis bersaing tujuan rencana
kebijakan diperlukan untuk mencapai tujuan goal
</sub>
<sub>
perkiraan forecasting
alokasi sumber daya resource allocation
kompetensi inti core competence
analisis lingkungan industri industry environment analysis
analisis perusahaan company analysis
rencana bisnis business plan
</sub>

```

Gambar 3.3: Contoh Dokumen Kunci Jawaban

Kedua dokumen tersebut akan menjadi dokumen masukan untuk bagian *preprocessing*. Keluaran dari bagian *preprocessing* adalah dokumen kunci jawaban dan dokumen jawaban sebanyak jumlah subsoal. Isi dokumen ini sudah tidak menggunakan *tag* lagi, dan sudah dibersihkan dari karakter-karakter yang tidak penting. Setiap jawaban siswa pada dokumen jawaban dipisahkan dengan baris kosong. Contoh dokumen jawaban dan kunci jawaban ini adalah sebagai berikut.

Dokumen Jawaban 9.1

strategy suatu formulasi yang dirancang sehingga perusahaan dapat bertahan dan mencapai tujuannya biasanya dilengkapi dengan perencanaan dan aturan

strategy suatu rencana jangka panjang untuk melaksanakan bisnis agar mampu bersaing dan mencapai target yang ditentukan

Dokumen Jawaban 9.2

forecasting memprediksi bisnis teknologi ekonomi trend dll
resource allocation mengalokasikan sumber daya yang dibutuhkan core competency bisnis utama yang dijalankan

forecasting perkiraan ramalan tentang segala faktor yang mempengaruhi bisnis resource allocation mengalokasikan resource

Dokumen Kunci Jawaban 9.1

strategi formula menyeluruh bisnis bersaing tujuan rencana kebijakan diperlukan untuk mencapai tujuan goal

Dokumen Kunci Jawaban 9.2

perkiraan forecasting
alokasi sumber daya resource allocation
kompetensi inti core competence
analisis lingkungan industri industry environment analysis
analisis perusahaan company analysis
rencana bisnis business plan

Gambar 3.4: Contoh Dokumen Keluaran *Preprocessing*

Dokumen jawaban ini kemudian menjadi masukan untuk proses penilaian VSM atau LSA. Hasil akhir dari sistem ini adalah dokumen nilai siswa berupa sebuah dokumen teks, yang setiap barisnya berisi nilai siswa sesuai dengan urutannya pada dokumen jawaban. Contoh dokumen keluaran ini adalah sebagai berikut.

Dokumen Nilai Soal 9

| |
|-----|
| 2.2 |
| 4.8 |
| 1.0 |
| 4.6 |
| 2.1 |
| 1.1 |
| 1.9 |
| 0.5 |
| 3.8 |
| 1.5 |
| 1.8 |
| 1.7 |
| 1.6 |

Gambar 3.5: Contoh Dokumen Nilai Siswa

3.3 Perluasan Kunci Jawaban

Perluasan kunci jawaban dilakukan dengan bantuan sebuah dokumen yang berisi daftar persamaan kata. Daftar ini disusun dengan mengambil seluruh kata yang terdapat pada kunci jawaban. Kata-kata ini kemudian dicari sinonimnya dengan menggunakan WordNet Bahasa Indonesia. Pemilihan sinonim yang relevan ditentukan oleh pengguna atau dosen yang menguasai domain ujian esai. Bentuk dokumen yang berisi daftar persamaan kata ini adalah sebagai berikut.

| |
|---------------------------------------|
| segmentasi, bagian, divisi, pembagian |
| sehingga |
| selama |
| sell-side |
| semua |
| seni |
| serangan |
| sesuai |
| simbol, lambang |
| site |
| situs |
| sniffing |
| spyware |
| strategi, kerangka, rencana, siasat |

Gambar 3.6: Daftar Persamaan Kata

Sinonim dari kata-kata pada kunci jawaban yang didapatkan dari daftar ini kemudian ditambahkan pada dokumen kunci jawaban.

3.4 Preprocessing

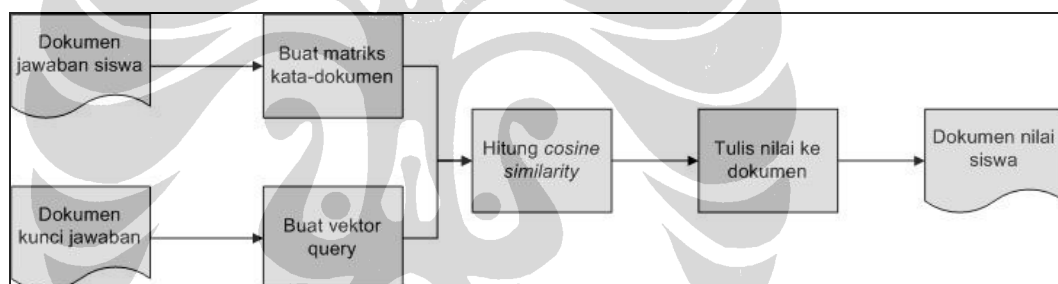
Bagian ini menerima masukan berupa dokumen kunci jawaban dan jawaban siswa. Kedua dokumen ini diproses dengan rincian sebagai berikut:

- Semua huruf dijadikan huruf kecil (*lowercase*).
- Karakter-karakter yang tidak penting seperti [`,",()!"?*:;,] dihilangkan.
- Dibaca semua *tag* yang ada di dalamnya dan *tag* <sub> digunakan untuk memisah jawaban ke dalam setiap dokumen subsoal.

Isi dokumen masukan tersebut kemudian ditulis ke dalam dokumen kunci jawaban dan jawaban siswa sesuai dengan nomor subsoalnya.

3.5 Penilaian VSM

Inti penilaian VSM adalah pencocokkan kata-kata pada kunci jawaban dengan jawaban siswa. Alur penilaian VSM digambarkan dalam diagram berikut ini.



Gambar 3.7: Alur Penilaian VSM

Proses penilaian dilakukan untuk setiap subsoal dengan dokumen masukan yang terdiri dari dokumen jawaban siswa dan dokumen kunci jawaban. Dokumen jawaban siswa dibuat ke dalam bentuk matriks kata-dokumen yang barisnya mewakili kata-kata pada jawaban dan kolomnya mewakili jawaban siswa (satu jawaban dianggap satu dokumen). Dokumen kunci jawaban dibuat ke dalam vektor kunci jawaban yang barisnya mewakili kata-kata dan kolomnya hanya satu (karena kunci jawaban hanya satu dokumen). Setiap elemen matriks dan vektor diberi bobot sesuai kemunculan katanya pada dokumen jawaban dan kunci jawaban.

Nilai siswa dihitung dengan cara membandingkan kedekatan antara vektor kunci jawaban dengan vektor jawaban siswa (yang merupakan kolom-kolom matriks) dengan cara menghitung *cosine similarity*. Jika satu soal terdiri atas lebih

dari satu subsoal, maka nilai *cosine similarity* ini dijumlahkan kemudian dibagi dengan jumlah subsoal dan hasilnya ditulis ke dokumen nilai siswa sesuai dengan nomor urut siswa. Nilai ini dikalikan dengan 10 untuk mendapatkan nilai antara 0-10.

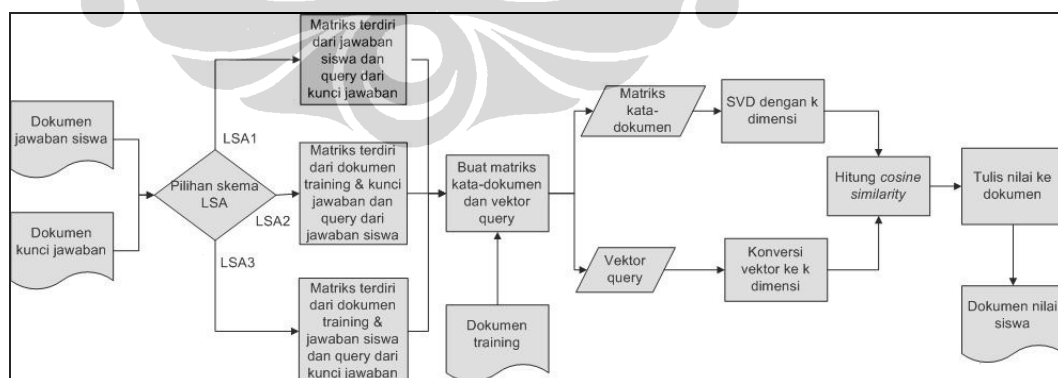
3.6 Penilaian LSA

Penilaian LSA diawali dengan membentuk matriks kata-dokumen yang akan didekomposisi dengan mengurangi dimensinya untuk membentuk suatu ruang semantik LSA. Terdapat tiga skema penilaian LSA sebagaimana yang sudah disebutkan dalam Subbab 1.4. Perbedaan ketiga skema penilaian ini terdapat pada dokumen yang membentuk ruang semantik dan yang menjadi *query*-nya. Rincian perbedaan tersebut terdapat pada Tabel 3.1.

Tabel 3.1: Perbedaan 3 Skema Penilaian LSA

| Skema LSA | Ruang Semantik | Query |
|-----------|--|------------------------|
| LSA1 | Kumpulan jawaban siswa | Kunci jawaban |
| LSA2 | Dokumen <i>training</i> dan kunci jawaban | Kumpulan jawaban siswa |
| LSA3 | Dokumen <i>training</i> dan kumpulan jawaban siswa | Kunci jawaban |

Sama seperti VSM, penilaian LSA dijalankan untuk setiap subsoal. Alur penilaian LSA selengkapnya digambarkan dalam diagram berikut ini.



Gambar 3.8: Alur Penilaian LSA

Terdapat tambahan data yang diperlukan untuk skema penilaian LSA2 dan LSA3, yaitu dokumen *training* untuk membentuk ruang semantik. Pembuatan ruang semantik atau matriks kata-dokumen dan *query* tergantung pada pilihan

pengguna.

Jika pengguna memilih skema penilaian LSA1, maka matriks kata-dokumen dibuat dari kumpulan jawaban siswa dan *query* dibuat dari kunci jawaban. Jika pengguna memilih skema penilaian LSA2, maka matriks kata-dokumen dibuat dari dokumen *training* dan kunci jawaban, dan *query* dibuat dari jawaban-jawaban siswa. Jika pengguna memilih skema penilaian LSA3, maka matriks kata-dokumen dibuat dari dokumen *training* dan jawaban siswa, dan *query* dibuat dari kunci jawaban.

Matriks kata-dokumen yang sudah dibuat kemudian didekomposisi menjadi tiga komponen matriks disertai dengan pengurangan dimensi. Misalnya matriks kata-dokumen adalah matriks A , maka matriks ini didekomposisi (SVD) dengan hanya menggunakan k dimensi sehingga menghasilkan tiga matriks $U_k \Sigma_k V_k^T$.

Setiap vektor pada komponen matriks V_k hasil SVD mewakili vektor jawaban yang akan dibandingkan dengan vektor *query*. Vektor *query* diubah ke dalam k dimensi menggunakan rumus berikut ini.

$$\bar{q} = q^T U_k \Sigma_k^{-1}$$

Nilai siswa didapatkan dengan menghitung *cosine similarity* antara vektor *query* ini dengan vektor jawaban siswa. Penilaian dilakukan untuk setiap subsoal. Jika masih ada jawaban subsoal yang belum dinilai, sistem kembali membuat matriks dari dokumen jawaban kemudian melakukan penghitungan nilai. Nilai untuk tiap subsoal dijumlahkan lalu dibagi dengan jumlah subsoal dan dikalikan dengan 10 agar menghasilkan nilai berskala 10. Nilai untuk setiap siswa secara berurutan kemudian dituliskan ke dalam dokumen nilai siswa seperti yang sudah dijelaskan pada Subbab 3.2. Dokumen inilah yang menjadi hasil akhir sistem.

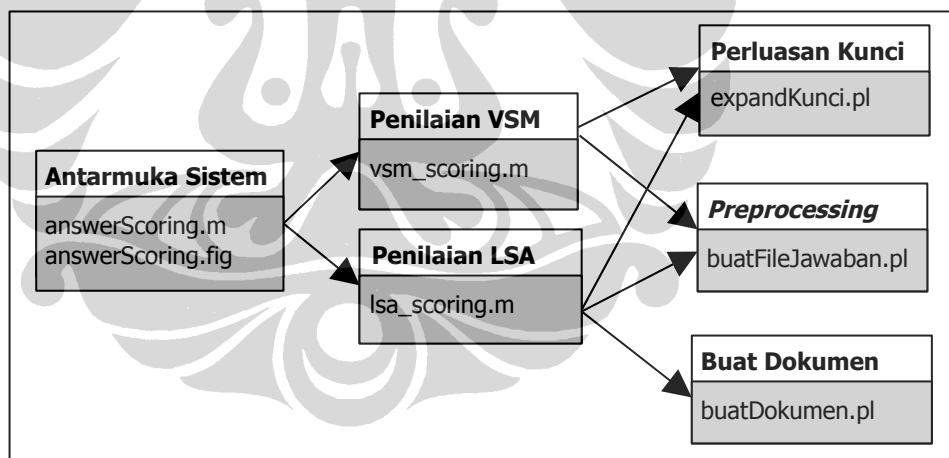
BAB 4 IMPLEMENTASI

Bab ini menguraikan implementasi sistem penilaian jawaban esai yang terdiri dari gambaran umum implementasi, antarmuka sistem, perluasan kunci jawaban, *preprocessing* dokumen, penilaian VSM, dan penilaian LSA.

4.1 Gambaran Umum Implementasi

Sistem ini diimplementasikan dengan bahasa pemrograman Matlab dan Perl. Selain itu juga digunakan Matlab Tool TMG (*Text to Matrix Generator*) (Zeimpekis & Gallopoulos, 2008) untuk membuat matriks kata-dokumen dan vektor *query*. Program lain yang digunakan di luar sistem ini adalah pemotong imbuhan untuk Bahasa Indonesia yang diimplementasikan dengan bahasa pemrograman Java (Ichsan & Adriani, 2006).

Secara keseluruhan, program yang dibuat terdiri dari enam program sebagaimana dijelaskan pada gambar berikut ini.



Gambar 4.1: Pembagian Program pada Sistem

Nama program yang berakhiran “.m” dan “.fig” adalah program Matlab sedangkan nama program yang berakhiran “.pl” adalah program Perl. Rincian dari tiap program akan dijelaskan dalam Subbab berikutnya.

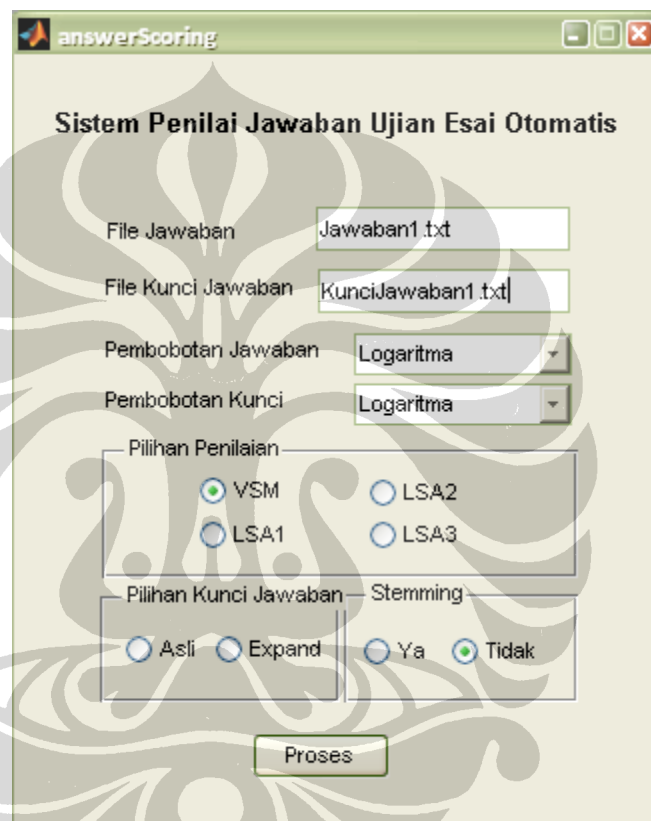
4.2 Antarmuka Sistem

Program ini dibuat untuk menjembatani pengguna dengan sistem. Melalui antarmuka ini, pengguna dapat memilih beberapa pilihan terkait dengan sistem

penilai esai otomatis sebagai berikut.

- a) Pilihan jenis pembobotan jawaban dan kunci jawaban.
- b) Pilihan metode penilaian yang terdiri dari VSM, LSA1, LSA2, dan LSA3.
- c) Pilihan untuk menggunakan perluasan kata kunci.
- d) Pilihan untuk menggunakan pemotong imbuhan untuk Bahasa Inggris (*Porter Stemmer*) yang terintegrasi dengan TMG yang digunakan.

Tampilan antarmuka ini adalah sebagai berikut.



Gambar 4.2: Antarmuka Sistem Penilai Jawaban Esai Otomatis

Selain pilihan-pilihan, masukan yang harus diberikan pengguna adalah nama dokumen kunci jawaban dan dokumen jawaban siswa yang berupa dokumen teks (.txt).

4.3 Program Perluasan Kunci Jawaban

Program ini digunakan untuk perluasan kunci jawaban, yaitu menambahkan sinonim dari kata-kata kunci jawaban untuk membentuk kunci jawaban baru. Sinonim dari kata-kata pada kunci jawaban diambil dari dokumen

daftar persamaan kata yang sudah disiapkan sebelumnya. *Pseudocode* untuk program ini adalah sebagai berikut.

```

input: fileKunci
output: fileKunciExpanded

open(fileKunci)
read(fileKunci)
open(glossary)
open(fileKunciExpanded)

foreach line(fileKunci)
  word=split(line, ' ')
  read(glossary)
  foreach line(glossary)
    array[]=split(line, ',')
    if word equal array[1]
      foreach element of array[]
        write(array[element]) to fileKunciExpanded
      end
    end
  end

return fileKunciExpanded

```

Gambar 4.3: *Pseudocode* Program Perluasan Kunci Jawaban

Program ini membaca dokumen kunci awal lalu membuka dokumen daftar persamaan kata, kemudian membandingkan kata kunci dengan kata pada daftar. Kata pada daftar dan sinonim-sinonimnya dipisahkan dengan tanda koma. Jika terdapat sinonim untuk kata pada jawaban, maka semua sinonim ini ditambahkan ke kunci jawaban.

4.4 Program *Preprocessing*

Bagian *preprocessing* dokumen kunci dan dokumen jawaban siswa dikerjakan dalam program ini. Dokumen kunci jawaban dibaca terlebih dahulu, lalu diambil informasi pada *tag* <subsoal> untuk menentukan dokumen jawaban yang dibuat akan terdiri dari satu dokumen saja atau banyak dokumen sejumlah subsoalnya. *Tag* <sub> digunakan untuk menandai dokumen subsoal mana yang akan digunakan untuk menulis jawaban yang terdapat pada *tag* <sub> tersebut. Dokumen kunci jawaban kemudian disalin ke dokumen kunci sesuai dengan nomor subsoalnya. Setelah itu, dokumen jawaban siswa dibaca lalu disalin juga ke dalam dokumen jawaban sesuai dengan nomor subsoalnya. *Pseudocode* untuk

program ini adalah sebagai berikut.

```

input: fileJawabanAwal, fileKunciAwal
output: fileJawaban, fileKunci, namaFile

open(fileKunciAwal)
read(fileKunciAwal)

foreach line(fileKunciAwal)
  tolowercase(line), remove(['(',')!'?'*:.])
  if line equal <subsoal>
    read jumlahsubsoal
  elsif line equal <sub>
    define which fileKunci to write to
    namaFile[sub] = fileKunci
  else
    write line to fileKunci
  end

open(fileJawabanAwal)

foreach line(fileJawabanAwal)
  tolowercase(line), remove(['(',')!'?'*:.])
  if line equal <sub>
    define which fileJawaban to write to
    namaFile[sub] = fileJawaban
  else
    write line to fileJawaban
  end

return fileJawaban, fileKunci, namaFile

```

Gambar 4.4: *Pseudocode Program Preprocessing*

Semua teks yang ada pada kedua dokumen masukan diproses dengan mengubah hurufnya menjadi huruf kecil semua dan menghilangkan karakter-karakter yang tidak penting. Hasil dari program ini adalah dokumen kunci dan dokumen jawaban sebanyak jumlah subsoal dan *array* yang berisi semua nama dokumen yang dihasilkan.

4.5 Program Penilaian VSM

Program ini membaca masukan dari antarmuka lalu memanggil program Perl untuk menjalankan tugas perluasan kunci jawaban (jika dipilih) dan *preprocessing* seperti yang sudah dijelaskan pada dua subbab sebelumnya, kemudian membuat matriks kata-dokumen dan vektor kunci jawaban, lalu

melakukan proses penilaian sebagaimana yang sudah dijelaskan di Bab 3. *Pseudocode* dari program ini adalah sebagai berikut.

```

input: file_jawaban, file_kunci, pilihan,
         pilihan_dokumen, pilihan_query
output: file_nilai

if pilihan.kunci == expand kunci
  perl(expandKunci.pl, file_kunci)

file_jawaban_kunci = perl(buatFileJawaban.pl,
                          file_jawaban, file_kunci)

for subsoal=1 to jumlahsubsoal
  filejawaban = file_jawaban_kunci[subsoal][1]
  filekunci = file_jawaban_kunci[subsoal][2]
  A = tmg(filejawaban, pilihan_dokumen)
  Q = tmg_query(filekunci, pilihan_query)
  for i = 1 to jumlahsiswa
    d = A(:,i)
    q = Q
    sim=dot(d,q)/norm(d)*norm(q)
    array_nilai[i] += sim
  end
end

for i=1 to length(array_nilai)
  write to file_nilai: array_nilai[i]/jumlahsubsoal*10
end

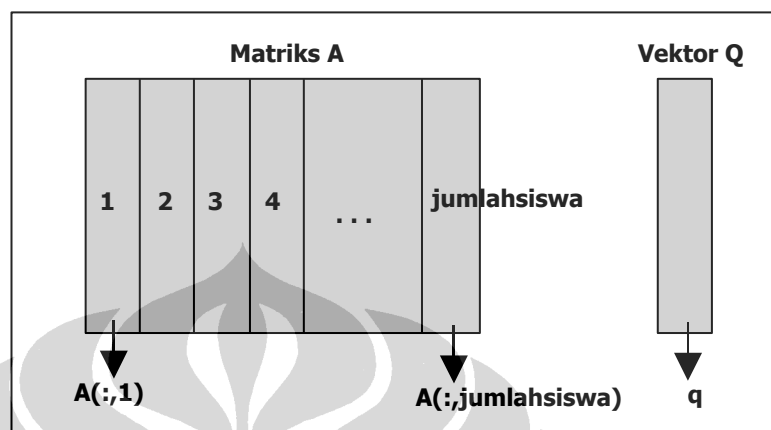
```

Gambar 4.5: Pseudocode Program Penilaian VSM

Setelah membaca masukan dari pengguna, program ini memeriksa pilihan perluasan kunci jawaban. Jika pengguna ingin menggunakan perluasan kunci jawaban, maka program ini menjalankan program Perl `expandKunci.pl`. Jika tidak maka *preprocessing* langsung dilakukan dengan menjalankan program Perl `buatFileJawaban.pl`.

Hasil *preprocessing* adalah dokumen jawaban dan kunci jawaban sebanyak jumlah subsoal. Program *preprocessing* juga mengembalikan daftar nama dokumen jawaban dan kunci jawaban tersebut beserta jumlah siswa. Program ini kemudian membaca dokumen jawaban dan kunci jawaban untuk setiap subsoal dan membuat matriks kata-dokumen (matriks A) dari dokumen jawaban dan vektor *query* (vektor Q) dari kunci jawaban. A dan Q dibuat dengan menggunakan fungsi `tmg` dan `tmg_query` dari TMG. Kedua fungsi ini memerlukan masukan

berupa nama dokumen yang akan dijadikan matriks/vektor dan pilihan-pilihan pembuatan matriks/vektor yang antara lain terdiri dari pilihan jenis pembobotan, penggunaan *Porter Stemmer*, dan penghilangan *stopwords*. Gambaran matriks A dan vektor Q adalah sebagai berikut.



Gambar 4.6: Matriks A dan Vektor Q untuk Penilaian VSM

Kolom-kolom matriks A merepresentasikan jawaban siswa 1 sampai dengan jumlahsiswa (sejumlah siswa yang mengikuti ujian). Penilaian untuk setiap siswa dilakukan dengan iterasi kolom matriks A dari 1 sampai jumlahsiswa. Untuk setiap iterasi, jawaban siswa direpresentasikan sebagai vektor d yang diambil dari satu kolom matriks A. Misalnya untuk siswa pertama, vektor jawabannya diambil dari kolom pertama matriks A atau $A(:,1)$. Nilai di dalam tanda kurung menunjukkan indeks matriks sebagai $A(\text{baris}, \text{kolom})$. Tanda “:” berarti semua baris/semua kolom.

Vektor d dan vektor q kemudian dihitung kedekatannya dengan menggunakan rumus *cosine similarity*. Kedekatan antar dua vektor inilah yang dijadikan sebagai nilai siswa. Nilai ini dijumlahkan untuk semua subsoal kemudian dibagi jumlah subsoal dan dikalikan 10 untuk mendapatkan nilai berskala 10. Nilai untuk semua siswa kemudian ditulis secara berurutan ke dalam dokumen teks.

4.6 Program Penilaian LSA

Program ini mengimplementasikan penilaian LSA untuk semua skema penilaian, yaitu LSA1, LSA2, dan LSA3 seperti yang sudah dijelaskan pada Subbab 3.5. Program ini sama dengan program penilaian VSM dalam hal

menjalankan program perluasan kunci jawaban (jika dipilih), *preprocessing*, dan menulis nilai siswa ke dokumen nilai. Perbedaannya terdapat pada penilaian LSA dan pembuatan dokumen yang menggunakan program buatDokumen.pl. Berikut ini adalah *pseudocode* dari program lsa_scoring.m.

```

input: file_jawaban, file_kunci, pilihan,
          pilihan_dokumen, pilihan_query
output: file_nilai

if pilihan.kunci == expand kunci
  perl(expandKunci.pl, file_kunci)

file_jawaban_kunci = perl(buatFileJawaban.pl,
                          file_jawaban, file_kunci)
for subsoal=1 to jumlahsubsoal
  filejawaban = file_jawaban_kunci[subsoal][1]
  filekunci = file_jawaban_kunci[subsoal][2]
  if pilihan.tipe_lsa == lsa1
    dokumen = perl(buatDokumen.pl, filejawaban,
                  filekunci)
    query = filekunci
  else if pilihan.tipe_lsa == lsa2
    dokumen = perl(buatDokumen.pl, file_training,
                  filekunci, filejawaban)
    query = filejawaban
  else if pilihan.tipe_lsa == lsa3
    dokumen = perl(buatDokumen.pl, file_training,
                  filejawaban, filekunci)
    query = filekunci

  A = tmg(dokumen, pilihan_dokumen)
  Q = tmg_query(query, pilihan_query)
  [U,S,V] = svds(A,rank)
  for i=1 to jumlahsiswa
    if pilihan.tipe_lsa == lsa1
      d = V(i,:)
      q = inv(S)*U'*Q(:,1)
    else if pilihan.tipe_lsa == lsa2
      d=V(length_training+1,:)
      q=inv(S)*U'*Q(:,i)
    else if pilihan.tipe_lsa == lsa3
      d=V(i+length_training,:)
      q=inv(S)*U'*Q(:,1)
    end
    sim=dot(d,q)/norm(d)*norm(q)
    array_nilai[i] += sim
  end

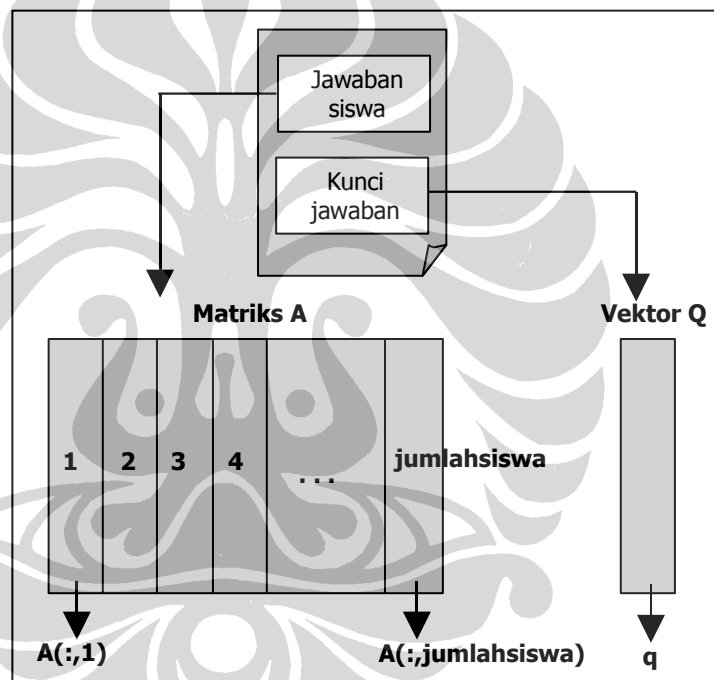
for j=1 to length(array_nilai)
  write to file_nilai: array_nilai[i]/jumlahsubsoal*10
end

```

Gambar 4.7: *Pseudocode* Program Penilaian LSA

Seperti program penilaian VSM, program ini juga membuat matriks kata-dokumen dan vektor *query* menggunakan fungsi *tmg* dan *tmg_query*. Sebelum membuat matriks dan vektor, program ini membuat dokumen terlebih dahulu dengan menjalankan program *buatDokumen.pl*. Program ini hanya bertugas untuk menyatukan dokumen-dokumen masukan ke dalam satu dokumen. Dokumen yang dibuat akan berbeda-beda sesuai dengan skema penilaian yang dibutuhkan. LSA1 hanya memerlukan dokumen jawaban dan kunci jawaban, sedangkan LSA2 dan LSA3 memerlukan tambahan dokumen *training*.

Berikut ini adalah gambaran dokumen beserta matriks dan vektor yang dihasilkan untuk LSA1.



Gambar 4.8: Dokumen, Matriks dan Vektor LSA1

Dokumen dihasilkan hanya dari jawaban siswa dan kunci jawaban. Jawaban siswa kemudian menjadi matriks kata-dokumen (matriks A) dan kunci jawaban menjadi vektor *query* (vektor Q).

Matriks A kemudian didekomposisi dengan dikurangi dimensinya menjadi hanya sebesar k . SVD matriks A menghasilkan tiga komponen matriks U, S, dan V. Vektor Q juga harus dikonversi ke dimensi yang sama dengan yang digunakan matriks A. Untuk menghitung nilai siswa, dilakukan iterasi sebanyak jumlah siswa lalu dihitung *cosine similarity* antara vektor d (jawaban siswa) dengan vektor q

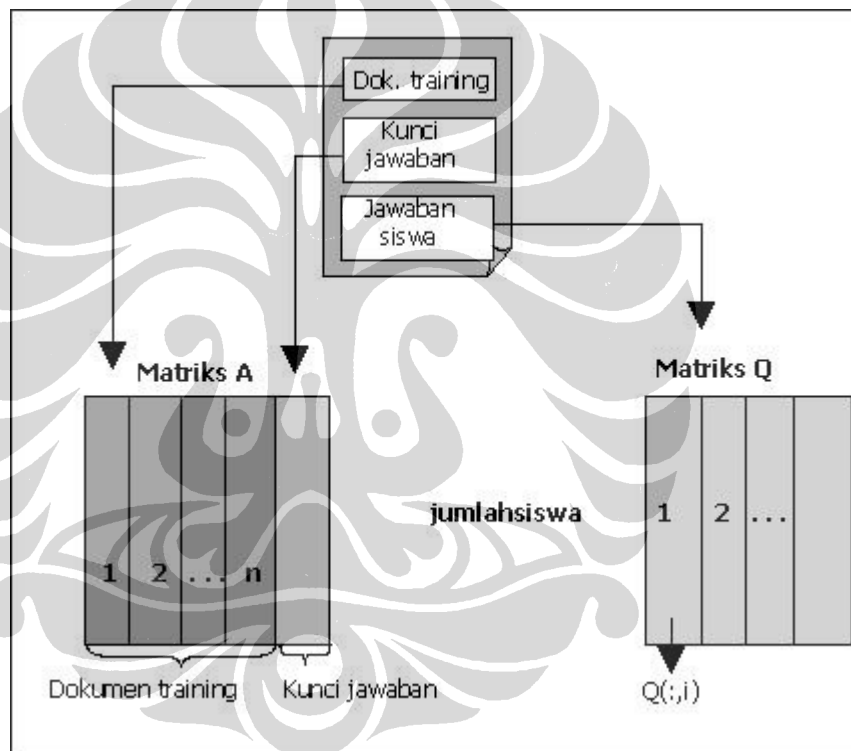
(*query*). Vektor \mathbf{d} dan \mathbf{q} untuk LSA1 adalah sebagai berikut:

$$\mathbf{d} = \mathbf{V}(i,:)$$

$$\mathbf{q} = \text{inv}(\mathbf{S}) * \mathbf{U}' * \mathbf{Q}(:,1)$$

\mathbf{V} adalah komponen matriks kolom hasil SVD matriks \mathbf{A} , sehingga $\mathbf{d} = \mathbf{V}(i,:)$ yang menyatakan baris ke- i matriks \mathbf{V} merupakan vektor jawaban siswa ke- i dalam dimensi k . Vektor \mathbf{q} dikonversi ke dalam dimensi k dengan menggunakan rumus $\mathbf{q} = \text{inv}(\mathbf{S}) * \mathbf{U}' * \mathbf{Q}(:,1)$.

LSA2 menggunakan tambahan dokumen *training* sebagai bahan pembentuk matriks kata-dokumen. Gambaran untuk dokumen dan matriks serta vektor yang dihasilkan untuk LSA terdapat pada Gambar 4.9.



Gambar 4.9: Dokumen, Matriks dan Vektor LSA2

Matriks \mathbf{A} dibentuk dari gabungan dokumen *training* dan kunci jawaban. Pada LSA2, yang menjadi *query* (\mathbf{q}) adalah jawaban-jawaban siswa, sedangkan yang menjadi dokumen (\mathbf{d}) adalah kunci jawaban. Vektor \mathbf{d} dan \mathbf{q} dalam LSA2 adalah sebagai berikut:

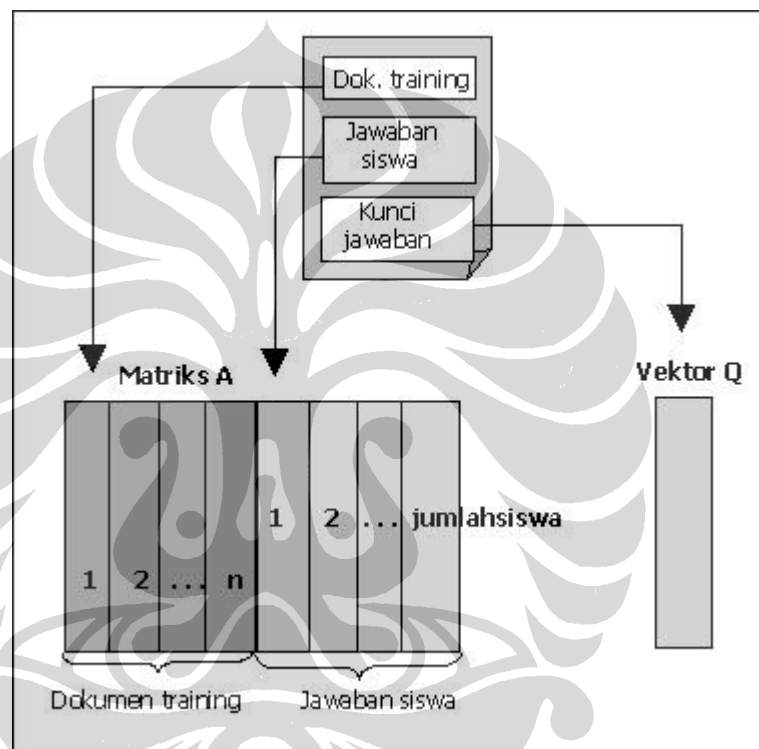
$$\mathbf{d} = \mathbf{V}(\text{length_training}+1,:)$$

$$\mathbf{q} = \text{inv}(\mathbf{S}) * \mathbf{U}' * \mathbf{Q}(:,i)$$

length_training adalah panjang dokumen *training* sehingga vektor \mathbf{d} terdapat

pada kolom ke $\text{length_training}+1$ dari matriks V . Jawaban siswa ke- i direpresentasikan dengan vektor $Q(:,i)$ sehingga q merupakan konversi vektor tersebut ke dimensi k dengan rumus di atas.

LSA3 juga menggunakan tambahan dokumen *training*, tetapi yang membedakannya dengan LSA2 adalah gabungan dokumen *training* beserta jawaban siswa yang dijadikan matriks, sedangkan kunci jawaban menjadi vektor *query*. Gambaran dokumen beserta matriks dan vektor yang dihasilkan untuk LSA3 terdapat pada Gambar 4.10.



Gambar 4.10: Dokumen, Matriks dan Vektor LSA3

Sama seperti LSA1, vektor d pada LSA3 adalah jawaban siswa dan q adalah kunci jawaban. Vektor d dan q ini adalah sebagai berikut:

$$d=V(i+\text{length_training},:)$$

$$q=\text{inv}(S)*U'*Q(:,1)$$

Untuk setiap siswa ke- i , jawabannya terdapat pada kolom $i+\text{length_training}$ dari matriks V . Sedangkan q dihitung dengan rumus yang sama dengan q pada LSA1.

Vektor q dan d kemudian dihitung *cosine similarity*-nya untuk menentukan nilai setiap siswa. Sama seperti penilaian VSM, nilai ini dijumlahkan untuk setiap subsoal kemudian dibagi dengan jumlah subsoal dan dikalikan dengan 10.

BAB 5

UJI COBA DAN ANALISIS

Bab ini menguraikan hasil uji coba sistem penilai jawaban esai otomatis dengan VSM dan LSA. Subbab 5.1 menjelaskan lingkungan uji coba sistem. Subbab 5.2 menjelaskan dokumen yang digunakan dan Subbab 5.3 menjelaskan perincian uji coba. Subbab 5.4 menjelaskan hasil uji coba dan dilanjutkan dengan analisisnya pada Subbab 5.5.

5.1 Lingkungan Uji Coba

Uji coba dilakukan pada komputer dengan *processor* Intel Celeron M 1.73 Ghz, *memory* DDR2 1 GB, dan *harddisk* 80 GB. Perangkat lunak yang dibutuhkan adalah Matlab, Matlab Tool Text to Matrix Generator (TMG), Perl, dan Java dengan Standard Development Kit JDK1.6.0_03. Sistem operasi yang digunakan adalah Windows XP.

5.2 Dokumen Uji Coba

Uji Coba dilakukan dengan 546 jawaban esai yang terdiri dari 13 soal dan dijawab oleh 42 siswa untuk masing-masing soal. Jawaban esai diambil dari jawaban Ujian Akhir Semester *E-Commerce* di Fakultas Ilmu Komputer Universitas Indonesia pada tahun 2008.

Berikut ini adalah pertanyaan yang diajukan dalam ujian esai ini.

Tabel 5.1: Soal Esai untuk Uji Coba Sistem

| Soal | Pertanyaan |
|------|--|
| 1 | Sebutkan lima faktor yang penting dalam kesuksesan B2B <i>Exchanges</i> |
| 2 | Sebutkan lima atribut <i>m-commerce</i> |
| 3 | Bandingkan antara isu legal dan isu etika |
| 4 | Sebutkan dan jelaskan tiga dari lima prinsip perlindungan dasar privasi |
| 5 | Apakah yang dimaksud dengan P2P <i>payments</i> ? Berikan dua contohnya |
| 6 | Sebutkan empat model penerimaan <i>m-commerce</i> |
| 7 | Apa yang dimaksud dengan <i>cookie</i> ? Jelaskan mengapa <i>cookie</i> bisa mengganggu privasi pengunjung website |

| | |
|----|---|
| 8 | Definisikan <i>intellectual property</i> dan tiga jenis yang utama dalam <i>E-Commerce</i> |
| 9 | Definisikan strategi dan sebutkan tiga elemen strategi |
| 10 | Sebutkan tiga fungsi utama pasar |
| 11 | Apa yang dimaksud dengan segmentasi pasar? Bagaimana menggunakan internet untuk melakukan segmentasi pasar? |
| 12 | Jelaskan apa saja jenis <i>cyber crime</i> yang sudah diatur dalam UU ITE |
| 13 | Apakah perbedaan dari G2C, G2B, dan G2G, dan berikan contoh untuk masing-masingnya |

5.3 Perincian Uji Coba

Perincian uji coba yang dilakukan pada sistem adalah sebagai berikut.

- a) Uji coba 1, menggunakan jawaban dan kunci jawaban awal tanpa tambahan apapun.
- b) Uji coba 2, menggunakan jawaban awal dan kunci jawaban yang sudah diperluas (*query expansion*).
- c) Uji coba 3, menggunakan jawaban dan kunci jawaban yang sudah diproses dengan program pemotong imbuhan.
- d) Uji coba 4, menggunakan jawaban awal dan kunci jawaban yang sudah diperluas, ditambah dengan pemotongan imbuhan.

Keempat skema uji coba di atas diterapkan untuk penilaian dengan VSM dan juga ketiga skema penilaian LSA (LSA1, LSA2, dan LSA3).

Variabel lain yang tidak berubah untuk semua skema uji coba di atas adalah sebagai berikut.

- a) Pembobotan jawaban siswa dan kunci jawaban menggunakan pembobotan lokal logaritma. Pembobotan ini dipilih karena bisa memberikan bobot yang proporsional terhadap kata yang banyak muncul sehingga bobotnya tidak terlalu besar. Berbeda dengan TF yang memberikan bobot besar untuk kemunculan kata yang banyak dan biner yang hanya memberi bobot satu berapapun jumlah kemunculan katanya. Pembobotan global tidak digunakan karena pada domain yang spesifik seperti jawaban ujian esai ini, penggunaan pembobotan global yang berguna untuk memberi nilai

beda pada dokumen yang mengandung kata yang jarang muncul kurang cocok untuk menilai jawaban esai. Karena jika banyak siswa menjawab benar, maka beberapa kata sesuai kata-kata pada kunci jawaban akan banyak muncul.

- b) Penghilangan *stopword* untuk Bahasa Indonesia.
- c) *Rank* yang digunakan pada penilaian LSA1 adalah 2 ($k=2$), sesuai dengan penelitian (Hermawandi, 2008), (Octaria, 2008), dan (Harisma, 2008). Sedangkan *rank* yang digunakan untuk LSA2 dan LSA3 adalah 60. *Rank* ini dipilih karena memberikan hasil terbaik pada percobaan awal untuk satu soal. Percobaan awalnya dilakukan untuk $\frac{1}{4}$, $\frac{1}{2}$, dan $\frac{1}{8}$ dimensi matriks dokumen *training* (200) kemudian didapatkan yang tertinggi adalah dengan $k=50$. Kemudian dicoba untuk $k=40$ dan $k=60$. Korelasi nilai sistem dengan nilai manusia dari hasil percobaan awal tersebut adalah sebagai berikut.

Tabel 5.2: Hasil Korelasi dengan Berbagai *Rank* untuk Soal 1

| K=50 | K=100 | K=25 | K=40 | K=60 |
|------|-------|------|------|------|
| 0.61 | 0.29 | 0.34 | 0.59 | 0.63 |

5.4 Hasil Uji Coba

Keluaran dari sistem berupa dokumen teks yang berisi nilai siswa. Nilai ini kemudian dibandingkan dengan nilai yang diberikan oleh penilai manusia dengan cara menghitung korelasi antara keduanya. Korelasi adalah salah cara mengukur efektivitas sistem penilai jawaban esai otomatis, seperti yang digunakan oleh Krisnanda (2008), Octaria (2008), dan Harisma (2008).

5.4.1 Hasil Uji Coba 1

Uji coba 1 merupakan uji coba awal terhadap sistem, dengan menggunakan dokumen awal tanpa pemotongan imbuhan dan perluasan kata kunci jawaban. Hasil uji coba 1 ditunjukkan pada Tabel 5.3.

Tabel 5.3: Hasil Uji Coba 1

| Soal | HR | Rata-rata Nilai | | | | Korelasi | | | |
|------|------|-----------------|------|------|------|----------|---------|---------|---------|
| | | VSM | LSA1 | LSA2 | LSA3 | HR_VSM | HR_LSA1 | HR_LSA2 | HR_LSA3 |
| 1 | 6.57 | 1.59 | 7.55 | 4.26 | 3.12 | 0.72 | 0.45 | 0.66 | 0.55 |
| 2 | 7.12 | 2.36 | 7.28 | 5.19 | 3.02 | 0.81 | 0.65 | 0.83 | 0.69 |
| 3 | 8.6 | 1.75 | 5.2 | 2.09 | 4.8 | 0.41 | 0.37 | 0.38 | 0.44 |
| 4 | 6.69 | 1.94 | 7.24 | 3.82 | 2.47 | 0.82 | 0.69 | 0.63 | 0.72 |

Tabel 5.3: Hasil Uji Coba 1 (sambungan)

| Soal | HR | Rata-rata Nilai | | | | Korelasi | | | |
|------|------|-----------------|------|------|------|----------|---------|---------|---------|
| | | VSM | LSA1 | LSA2 | LSA3 | HR_VSM | HR_LSA1 | HR_LSA2 | HR_LSA3 |
| 5 | 7.31 | 2.25 | 8.18 | 2.93 | 4.64 | 0.46 | 0.52 | 0.32 | 0.48 |
| 6 | 4.74 | 1.55 | 7.05 | 2.97 | 2.87 | 0.74 | 0.77 | 0.73 | 0.74 |
| 7 | 9.62 | 2.09 | 6.85 | 4.59 | 3.04 | 0.32 | 0.16 | 0.3 | 0.3 |
| 8 | 7.95 | 2.1 | 5.72 | 1.85 | 4 | 0.69 | 0.58 | 0.33 | 0.57 |
| 9 | 6.17 | 2.29 | 6.43 | 4.9 | 4.88 | 0.76 | 0.64 | 0.7 | 0.77 |
| 10 | 6.83 | 2.96 | 6.73 | 5.5 | 4.25 | 0.58 | 0.59 | 0.72 | 0.64 |
| 11 | 6.9 | 1.76 | 6.86 | 3.12 | 3.13 | 0.2 | 0.31 | 0.13 | 0.16 |
| 12 | 9.5 | 1.81 | 8.75 | 4.09 | 1.76 | 0.3 | 0 | 0.28 | 0.17 |
| 13 | 9.36 | 3.2 | 6.18 | 5.5 | 4.75 | 0.43 | 0.51 | 0.27 | 0.48 |
| Avg | 7.49 | 2.13 | 6.92 | 3.91 | 3.59 | 0.56 | 0.48 | 0.48 | 0.52 |

Keterangan:

HR: rata-rata nilai dari manusia

VSM: rata-rata nilai dari sistem VSM

LSA1: rata-rata nilai dari sistem LSA1

LSA2: rata-rata nilai dari sistem LSA2

LSA3: rata-rata nilai dari sistem LSA3

HR_VSM: korelasi nilai manusia dengan nilai VSM

HR_LSA1: korelasi nilai manusia dengan nilai LSA1

HR_LSA2: korelasi nilai manusia dengan nilai LSA2

HR_LSA3: korelasi nilai manusia dengan nilai LSA3

Hasil uji coba 1 menunjukkan bahwa rata-rata dari nilai rata-rata VSM (2.13) lebih kecil nilai dari LSA (6.92, 3.91, dan 3.59), tetapi rata-rata korelasi antara nilai dari manusia dengan nilai sistem VSM lebih tinggi dari rata-rata korelasinya dengan nilai sistem LSA. Korelasi nilai manusia dengan nilai VSM rata-rata 0.56, dengan korelasi terrendah 0.2 dan tertinggi 0.82. Korelasi nilai manusia dengan nilai LSA1 rata-rata 0.48, dengan korelasi terrendah 0 dan tertinggi 0.77. Korelasi nilai manusia dengan nilai LSA2 rata-rata 0.48, dengan korelasi terrendah 0.13 dan tertinggi 0.83. Korelasi nilai manusia dengan nilai LSA3 rata-rata 0.52, dengan korelasi terrendah 0.16 dan tertinggi 0.77.

5.4.2 Hasil Uji Coba 2

Uji coba 2 menggunakan perluasan kunci jawaban dengan memasukkan persamaan kata yang relevan pada kunci jawaban. Hasil uji coba 2 ditunjukkan pada Tabel 5.4.

Tabel 5.4: Hasil Uji Coba 2

| Soal | HR | Rata-rata Nilai | | | | Korelasi | | | |
|------|------|-----------------|------|------|------|----------|---------|---------|---------|
| | | VSM | LSA1 | LSA2 | LSA3 | HR_VSM | HR_LSA1 | HR_LSA2 | HR_LSA3 |
| 1 | 6.57 | 1.59 | 7.55 | 4.24 | 3.12 | 0.72 | 0.45 | 0.66 | 0.55 |
| 2 | 7.12 | 2.37 | 7.22 | 5.19 | 3.03 | 0.8 | 0.65 | 0.83 | 0.7 |
| 3 | 8.6 | 1.75 | 5.2 | 2.02 | 4.75 | 0.41 | 0.37 | 0.38 | 0.44 |
| 4 | 6.69 | 2.14 | 7.4 | 4.51 | 2.55 | 0.81 | 0.68 | 0.66 | 0.73 |

Tabel 5.4: Hasil Uji Coba 2 (sambungan)

| Soal | HR | Rata-rata Nilai | | | | Korelasi | | | |
|------|------|-----------------|------|------|------|----------|---------|---------|---------|
| | | VSM | LSA1 | LSA2 | LSA3 | HR_VSM | HR_LSA1 | HR_LSA2 | HR_LSA3 |
| 5 | 7.31 | 2.25 | 8.18 | 2.92 | 4.64 | 0.46 | 0.52 | 0.32 | 0.48 |
| 6 | 4.74 | 1.55 | 7.05 | 2.93 | 2.87 | 0.74 | 0.77 | 0.73 | 0.74 |
| 7 | 9.62 | 2.25 | 6.65 | 4.69 | 2.9 | 0.33 | 0.23 | 0.32 | 0.35 |
| 8 | 7.95 | 2.16 | 5.91 | 2.29 | 4.16 | 0.68 | 0.59 | 0.44 | 0.6 |
| 9 | 6.17 | 2.16 | 6.53 | 5.01 | 4.71 | 0.75 | 0.62 | 0.63 | 0.78 |
| 10 | 6.83 | 2.96 | 6.73 | 5.5 | 4.25 | 0.58 | 0.59 | 0.72 | 0.64 |
| 11 | 6.9 | 1.78 | 6.83 | 3.36 | 3.29 | 0.19 | 0.31 | 0.14 | 0.17 |
| 12 | 9.5 | 1.83 | 8.72 | 3.99 | 1.72 | 0.32 | 0.01 | 0.27 | 0.18 |
| 13 | 9.36 | 3.1 | 6.27 | 5.24 | 4.55 | 0.41 | 0.52 | 0.28 | 0.48 |
| Avg | 7.49 | 2.15 | 6.94 | 3.99 | 3.58 | 0.55 | 0.49 | 0.49 | 0.53 |

Hasil uji coba 2 menunjukkan bahwa rata-rata dari nilai rata-rata VSM (2.15) lebih kecil dari LSA (6.94, 3.99, dan 3.58), tetapi rata-rata korelasi antara nilai dari manusia dengan nilai sistem VSM lebih tinggi dari rata-rata korelasinya dengan nilai sistem LSA. Korelasi nilai manusia dengan nilai VSM rata-rata 0.55, dengan korelasi terendah 0.19 dan tertinggi 0.81. Korelasi nilai manusia dengan nilai LSA1 rata-rata 0.49, dengan korelasi terendah 0.01 dan tertinggi 0.77. Korelasi nilai manusia dengan nilai LSA2 rata-rata 0.49, dengan korelasi terendah 0.14 dan tertinggi 0.83. Korelasi nilai manusia dengan nilai LSA3 rata-rata 0.53, dengan korelasi terendah 0.17 dan tertinggi 0.78.

5.4.3 Hasil Uji Coba 3

Uji coba 3 menggunakan pemotongan imbuhan untuk kunci jawaban, jawaban siswa, dan juga dokumen *training*. Hasil uji coba 3 ditunjukkan pada Tabel 5.5.

Tabel 5.5: Hasil Uji Coba 3

| Soal | HR | Rata-rata Nilai | | | | Korelasi | | | |
|------|------|-----------------|------|------|------|----------|---------|---------|---------|
| | | VSM | LSA1 | LSA2 | LSA3 | HR_VSM | HR_LSA1 | HR_LSA2 | HR_LSA3 |
| 1 | 6.57 | 1.85 | 7.44 | 3.83 | 2.38 | 0.74 | 0.47 | 0.65 | 0.53 |
| 2 | 7.12 | 3.04 | 7.18 | 5.42 | 2.81 | 0.85 | 0.59 | 0.82 | 0.66 |
| 3 | 8.6 | 2.48 | 5.66 | 2.61 | 4.14 | 0.5 | 0.18 | 0.33 | 0.33 |
| 4 | 6.69 | 2.37 | 6.96 | 3.73 | 2.14 | 0.84 | 0.54 | 0.64 | 0.71 |
| 5 | 7.31 | 2.83 | 7.79 | 3.45 | 4.59 | 0.52 | 0.51 | 0.41 | 0.49 |
| 6 | 4.74 | 1.95 | 6.65 | 3.34 | 2.48 | 0.7 | 0.81 | 0.56 | 0.62 |
| 7 | 9.62 | 2.44 | 7.25 | 4.22 | 2.31 | 0.25 | 0.18 | 0.27 | 0.17 |
| 8 | 7.95 | 2.33 | 5.71 | 2.52 | 3.22 | 0.69 | 0.57 | 0.44 | 0.41 |
| 9 | 6.17 | 2.78 | 6.46 | 4.89 | 4.25 | 0.82 | 0.55 | 0.69 | 0.76 |
| 10 | 6.83 | 3.96 | 7.15 | 5.46 | 4.36 | 0.67 | 0.64 | 0.7 | 0.59 |
| 11 | 6.9 | 2.31 | 5.98 | 3.25 | 3.18 | 0.17 | 0.27 | 0.04 | 0.13 |
| 12 | 9.5 | 2 | 7.23 | 4.08 | 1.76 | 0.34 | 0.18 | 0.32 | 0.15 |
| 13 | 9.36 | 3.57 | 6.37 | 5.51 | 4.35 | 0.45 | 0.42 | 0.35 | 0.39 |
| Avg | 7.49 | 2.61 | 6.76 | 4.02 | 3.23 | 0.58 | 0.46 | 0.48 | 0.46 |

Hasil uji coba 3 menunjukkan bahwa rata-rata dari nilai rata-rata VSM (2.61) lebih kecil dari LSA (6.76, 4.02, dan 3.23), tetapi rata-rata korelasi antara nilai dari manusia dengan nilai sistem VSM lebih tinggi dari rata-rata korelasinya dengan nilai sistem LSA. Korelasi nilai manusia dengan nilai VSM rata-rata 0.58, dengan korelasi terendah 0.17 dan tertinggi 0.85. Korelasi nilai manusia dengan nilai LSA1 rata-rata 0.46, dengan korelasi terendah 0.18 dan tertinggi 0.81. Korelasi nilai manusia dengan nilai LSA2 rata-rata 0.48, dengan korelasi terendah 0.04 dan tertinggi 0.82. Korelasi nilai manusia dengan nilai LSA3 rata-rata 0.46, dengan korelasi terendah 0.13 dan tertinggi 0.76.

5.4.4 Hasil Uji Coba 4

Uji coba 4 menggunakan gabungan perluasan kunci jawaban dan pemotong imbuhan untuk jawaban siswa, kunci jawaban, dan dokumen *training*. Hasil uji coba 3 ditunjukkan pada Tabel 5.6.

Tabel 5.6: Hasil Uji Coba 4

| Soal | HR | Rata-rata Nilai | | | | Korelasi | | | |
|------|------|-----------------|------|------|------|----------|---------|---------|---------|
| | | VSM | LSA1 | LSA2 | LSA3 | HR_VSM | HR_LSA1 | HR_LSA2 | HR_LSA3 |
| 1 | 6.57 | 1.87 | 7.44 | 3.8 | 2.39 | 0.75 | 0.47 | 0.65 | 0.54 |
| 2 | 7.12 | 3.03 | 7.1 | 5.35 | 2.81 | 0.85 | 0.59 | 0.82 | 0.66 |
| 3 | 8.6 | 2.48 | 5.5 | 2.55 | 4.07 | 0.5 | 0.36 | 0.43 | 0.37 |
| 4 | 6.69 | 2.51 | 7.17 | 4.21 | 2.2 | 0.82 | 0.6 | 0.67 | 0.74 |
| 5 | 7.31 | 2.79 | 7.81 | 3.44 | 4.4 | 0.53 | 0.51 | 0.45 | 0.52 |
| 6 | 4.74 | 1.95 | 6.65 | 3.34 | 2.48 | 0.7 | 0.81 | 0.56 | 0.62 |
| 7 | 9.62 | 2.49 | 7.12 | 4.42 | 2.13 | 0.27 | 0.18 | 0.29 | 0.22 |
| 8 | 7.95 | 2.45 | 5.8 | 3.11 | 3.22 | 0.69 | 0.57 | 0.54 | 0.47 |
| 9 | 6.17 | 2.94 | 6.88 | 5.01 | 4.23 | 0.86 | 0.64 | 0.71 | 0.8 |
| 10 | 6.83 | 3.96 | 7.15 | 5.46 | 4.36 | 0.67 | 0.64 | 0.7 | 0.59 |
| 11 | 6.9 | 2.24 | 5.99 | 3.26 | 3.24 | 0.17 | 0.27 | 0.06 | 0.15 |
| 12 | 9.5 | 2.03 | 7.24 | 3.94 | 1.74 | 0.35 | 0.18 | 0.3 | 0.16 |
| 13 | 9.36 | 3.61 | 6.71 | 5.12 | 4.01 | 0.41 | 0.4 | 0.31 | 0.35 |
| Avg | 7.49 | 2.64 | 6.81 | 4.08 | 3.17 | 0.58 | 0.48 | 0.5 | 0.48 |

Hasil uji coba 4 menunjukkan bahwa rata-rata dari nilai rata-rata VSM (2.64) lebih kecil dari LSA (6.81, 4.08, dan 3.17), tetapi rata-rata korelasi antara nilai dari manusia dengan nilai sistem VStM lebih tinggi dari rata-rata korelasinya dengan nilai sistem LSA. Korelasi nilai manusia dengan nilai VSM rata-rata 0.58, dengan korelasi terendah 0.17 dan tertinggi 0.86. Korelasi nilai manusia dengan nilai LSA1 rata-rata 0.48, dengan korelasi terendah 0.18 dan tertinggi 0.81.

Korelasi nilai manusia dengan nilai LSA2 rata-rata 0.5, dengan korelasi terendah 0.06 dan tertinggi 0.82. Korelasi nilai manusia dengan nilai LSA3 rata-rata 0.48, dengan korelasi terendah 0.15 dan tertinggi 0.8.

5.4.5 Perbandingan Hasil Uji Coba

Tabel 5.7 menunjukkan hasil keseluruhan uji coba untuk sistem VSM.

Tabel 5.7: Perbandingan Hasil Uji Coba VSM

| Soal | Uji Coba 1 | Uji Coba 2 | Uji Coba 3 | Uji Coba 4 |
|------|------------|------------|------------|------------|
| 1 | 0.72 | 0.72 | 0.74 | 0.75 |
| 2 | 0.81 | 0.8 | 0.85 | 0.85 |
| 3 | 0.41 | 0.41 | 0.5 | 0.5 |
| 4 | 0.82 | 0.81 | 0.84 | 0.82 |
| 5 | 0.46 | 0.46 | 0.52 | 0.53 |
| 6 | 0.74 | 0.74 | 0.7 | 0.7 |
| 7 | 0.32 | 0.33 | 0.25 | 0.27 |
| 8 | 0.69 | 0.68 | 0.69 | 0.69 |
| 9 | 0.76 | 0.75 | 0.82 | 0.86 |
| 10 | 0.58 | 0.58 | 0.67 | 0.67 |
| 11 | 0.2 | 0.19 | 0.17 | 0.17 |
| 12 | 0.3 | 0.32 | 0.34 | 0.35 |
| 13 | 0.43 | 0.41 | 0.45 | 0.41 |
| Avg | 0.56 | 0.55 | 0.58 | 0.58 |

Kolom Uji Coba 1 - Uji Coba 4 berisi korelasi nilai manusia dengan sistem VSM untuk uji coba 1-4. Berdasarkan uji coba 2, perluasan kunci jawaban pada sistem VSM memberikan pengaruh yang berbeda-beda untuk setiap soal, tetapi secara keseluruhan rata-rata korelasi nilai manusia dengan nilai VSM turun dari 0.56 menjadi 0.55. Penggunaan pemotong imbuhan pada uji coba 3 juga memberikan pengaruh yang berbeda-beda untuk setiap soal, namun secara keseluruhan rata-rata korelasi nilai manusia dengan nilai VSM naik dari 0.56 menjadi 0.58. Gabungan penggunaan perluasan kunci jawaban dengan pemotong imbuhan meningkatkan rata-rata korelasi nilai manusia dengan nilai VSM dari 0.56 menjadi 0.58.

Perbandingan hasil uji coba 1 – uji coba 4 untuk penilaian LSA1 ditunjukkan dalam Tabel 5.8 berikut ini.

Tabel 5.8: Perbandingan Hasil Uji Coba LSA1

| Soal | Uji Coba 1 | Uji Coba 2 | Uji Coba 3 | Uji Coba 4 |
|------|------------|------------|------------|------------|
| 1 | 0.45 | 0.45 | 0.47 | 0.47 |
| 2 | 0.65 | 0.65 | 0.59 | 0.59 |
| 3 | 0.37 | 0.37 | 0.18 | 0.36 |

Tabel 5.8: Perbandingan Hasil Uji Coba LSA1 (sambungan)

| Soal | Uji Coba 1 | Uji Coba 2 | Uji Coba 3 | Uji Coba 4 |
|------|------------|------------|------------|------------|
| 4 | 0.69 | 0.68 | 0.54 | 0.6 |
| 5 | 0.52 | 0.52 | 0.51 | 0.51 |
| 6 | 0.77 | 0.77 | 0.81 | 0.81 |
| 7 | 0.16 | 0.23 | 0.18 | 0.18 |
| 8 | 0.58 | 0.59 | 0.57 | 0.57 |
| 9 | 0.64 | 0.62 | 0.55 | 0.64 |
| 10 | 0.59 | 0.59 | 0.64 | 0.64 |
| 11 | 0.31 | 0.31 | 0.27 | 0.27 |
| 12 | 0 | 0.01 | 0.18 | 0.18 |
| 13 | 0.51 | 0.52 | 0.42 | 0.4 |
| Avg | 0.48 | 0.49 | 0.46 | 0.48 |

Kolom Uji Coba 1 - Uji Coba 4 berisi korelasi nilai manusia dengan LSA1 untuk uji coba 1-4. Berdasarkan uji coba 2, perluasan kunci jawaban pada LSA1 memberikan pengaruh yang berbeda-beda untuk setiap soal, tetapi secara keseluruhan rata-rata korelasi nilai manusia dengan nilai LSA1 meningkat dari 0.48 menjadi 0.49. Penggunaan pemotong imbuhan pada uji coba 3 juga memberikan pengaruh yang berbeda-beda untuk setiap soal, namun secara keseluruhan rata-rata korelasi nilai manusia dengan nilai LSA1 turun dari 0.48 menjadi 0.46. Gabungan penggunaan perluasan kunci jawaban dengan pemotong imbuhan pada uji coba 4 memberikan rata-rata korelasi nilai manusia dengan nilai LSA1 sama dengan uji coba 1 yaitu 0.48.

Perbandingan hasil uji coba 1 – uji coba 4 untuk penilaian LSA2 ditunjukkan dalam Tabel 5.9 berikut ini.

Tabel 5.9: Perbandingan Hasil Uji Coba LSA2

| Soal | Uji Coba 1 | Uji Coba 2 | Uji Coba 3 | Uji Coba 4 |
|------|------------|------------|------------|------------|
| 1 | 0.66 | 0.66 | 0.65 | 0.65 |
| 2 | 0.83 | 0.83 | 0.82 | 0.82 |
| 3 | 0.38 | 0.38 | 0.33 | 0.43 |
| 4 | 0.63 | 0.66 | 0.64 | 0.67 |
| 5 | 0.32 | 0.32 | 0.41 | 0.45 |
| 6 | 0.73 | 0.73 | 0.56 | 0.56 |
| 7 | 0.3 | 0.32 | 0.27 | 0.29 |
| 8 | 0.33 | 0.44 | 0.44 | 0.54 |
| 9 | 0.7 | 0.63 | 0.69 | 0.71 |
| 10 | 0.72 | 0.72 | 0.7 | 0.7 |
| 11 | 0.13 | 0.14 | 0.04 | 0.06 |
| 12 | 0.28 | 0.27 | 0.32 | 0.3 |
| 13 | 0.27 | 0.28 | 0.35 | 0.31 |
| Avg | 0.48 | 0.49 | 0.48 | 0.5 |

Kolom Uji Coba 1 - Uji Coba 4 berisi korelasi nilai manusia dengan LSA2 untuk uji coba 1-4. Berdasarkan uji coba 2, perluasan kunci jawaban pada LSA2 memberikan pengaruh yang berbeda-beda untuk setiap soal, tetapi secara keseluruhan rata-rata korelasi nilai manusia dengan nilai LSA2 meningkat dari 0.48 menjadi 0.49. Penggunaan pemotong imbuhan pada uji coba 3 juga memberikan pengaruh yang berbeda-beda untuk setiap soal, namun secara keseluruhan rata-rata korelasi nilai manusia dengan nilai LSA2 tidak berubah dari uji coba 1 yaitu 0.48. Gabungan penggunaan perluasan kunci jawaban dengan pemotong imbuhan pada uji coba 4 memberikan rata-rata korelasi nilai manusia dengan nilai LSA2 naik dari 0.48 menjadi 0.5.

Perbandingan hasil uji coba 1 – uji coba 4 untuk penilaian LSA3 ditunjukkan dalam Tabel 5.10 berikut ini.

Tabel 5.10: Perbandingan Hasil Uji Coba LSA3

| Soal | Uji Coba 1 | Uji Coba 2 | Uji Coba 3 | Uji Coba 4 |
|------|------------|------------|------------|------------|
| 1 | 0.55 | 0.55 | 0.53 | 0.54 |
| 2 | 0.69 | 0.7 | 0.66 | 0.66 |
| 3 | 0.44 | 0.44 | 0.33 | 0.37 |
| 4 | 0.72 | 0.73 | 0.71 | 0.74 |
| 5 | 0.48 | 0.48 | 0.49 | 0.52 |
| 6 | 0.74 | 0.74 | 0.62 | 0.62 |
| 7 | 0.3 | 0.35 | 0.17 | 0.22 |
| 8 | 0.57 | 0.6 | 0.41 | 0.47 |
| 9 | 0.77 | 0.78 | 0.76 | 0.8 |
| 10 | 0.64 | 0.64 | 0.59 | 0.59 |
| 11 | 0.16 | 0.17 | 0.13 | 0.15 |
| 13 | 0.17 | 0.18 | 0.15 | 0.16 |
| 14 | 0.48 | 0.48 | 0.39 | 0.35 |
| Avg | 0.52 | 0.53 | 0.46 | 0.48 |

Kolom Uji Coba 1 - Uji Coba 4 berisi korelasi nilai manusia dengan LSA3 untuk uji coba 1-4. Berdasarkan uji coba 2, perluasan kunci jawaban pada LSA3 secara umum menaikkan korelasi nilai manusia dengan nilai LSA3 dan rata-ratanya meningkat dari 0.52 menjadi 0.53. Penggunaan pemotong imbuhan pada uji coba 3 memberikan pengaruh yang berbeda-beda untuk setiap soal, namun secara keseluruhan rata-rata korelasi nilai manusia dengan nilai LSA3 turun dari 0.52 menjadi 0.46. Gabungan penggunaan perluasan kunci jawaban dengan pemotong imbuhan pada uji coba 4 memberikan rata-rata korelasi nilai manusia

dengan nilai LSA3 turun dari 0.52 menjadi 0.48.

Rata-rata korelasi nilai manusia dengan semua skema penilaian untuk semua skema uji coba ditunjukkan dalam Tabel 5.11.

Tabel 5.11: Perbandingan Semua Uji Coba

| Uji Coba | Korelasi | | | |
|----------|----------|---------|---------|---------|
| | HR_VSM | HR_LSA1 | HR_LSA2 | HR_LSA3 |
| 1 | 0.56 | 0.48 | 0.48 | 0.52 |
| 2 | 0.55 | 0.49 | 0.49 | 0.53 |
| 3 | 0.58 | 0.46 | 0.48 | 0.46 |
| 4 | 0.58 | 0.48 | 0.5 | 0.48 |

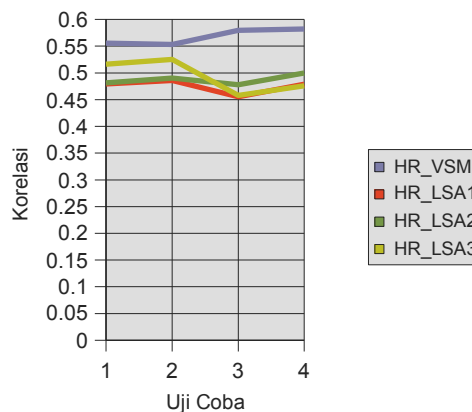
Secara keseluruhan untuk semua skema uji coba, rata-rata korelasi nilai manusia dengan VSM lebih tinggi dari rata-rata korelasi nilai manusia dengan LSA1, LSA2, dan LSA3. Urutan berikutnya, rata-rata korelasi nilai manusia dengan LSA3 lebih tinggi dibandingkan dengan LSA1 dan LSA2 untuk uji coba 1 dan 2. Sedangkan untuk uji coba 3 dan 4, rata-rata korelasi nilai manusia dengan LSA2 lebih tinggi dibandingkan dengan LSA1 dan LSA3.

5.5 Analisis Hasil Uji Coba

Analisis hasil uji coba dilakukan dengan memperhatikan perbandingan antara sistem penilai VSM, LSA1, LSA2, dan LSA3, serta pengaruh penggunaan perluasan kata kunci jawaban, pemotong imbuhan, dan gabungan keduanya.

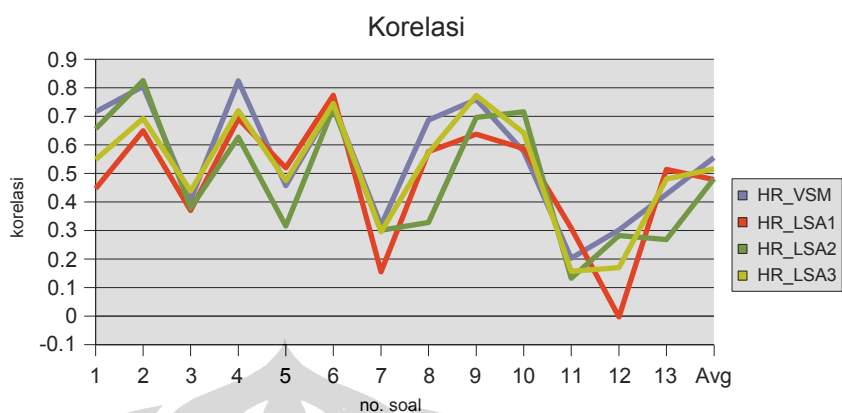
5.5.1 Analisis Perbandingan VSM dan LSA

Perbandingan rata-rata korelasi nilai VSM-manusia dan nilai LSA-manusia dapat dilihat pada Gambar 5.1.



Gambar 5.1: Grafik Perbandingan Rata-rata Korelasi VSM-LSA

Korelasi nilai manusia dengan VSM dan LSA untuk setiap soal digambarkan dalam grafik pada Gambar 5.2 berikut.



Gambar 5.2: Grafik Perbandingan Korelasi VSM-LSA

Secara keseluruhan, rata-rata korelasi antara nilai manusia dengan nilai sistem VSM lebih besar daripada korelasi antara nilai manusia dengan nilai sistem LSA1, LSA2, dan LSA3. Hal ini mungkin terjadi karena kurangnya sumber data (dokumen *training*) untuk membangun ruang semantik pada LSA sehingga hasil LSA tidak optimal. Meskipun dokumen *training* sudah digunakan pada LSA2 dan LSA3, namun rata-rata korelasi yang dihasilkan masih lebih rendah dari VSM. Hal ini mungkin disebabkan karena kurang banyaknya dokumen *training* dan isi dari dokumen tersebut kurang spesifik untuk setiap domain soal yang diujikan.

Antara LSA1, LSA2, dan LSA3, korelasi tertinggi antara nilai manusia dengan sistem diperoleh dari LSA3. Urutan berikutnya setelah LSA3 adalah LSA2, dan yang terendah adalah LSA1. Sebenarnya rata-rata korelasi LSA2 dan LSA1 sama, tetapi korelasi maksimum dan minimum yang dihasilkan LSA2 lebih tinggi dari LSA1. Hal ini terjadi pada uji coba 1 dan 2. Untuk uji coba 3 dan 4, urutannya berubah menjadi LSA2, LSA1, dan LSA3.

LSA3 menghasilkan rata-rata korelasi yang lebih tinggi dari LSA1 dan LSA2 sebelum digunakan pemotong imbuhan. Hal ini mungkin disebabkan karena jawaban siswa yang ikut dimasukkan dalam ruang semantik bisa memperkuat analisis yang dilakukan LSA. Meskipun tidak semua jawaban benar, tetapi ada informasi tambahan yang bisa digunakan dari jawaban-jawaban yang benar. Setelah digunakan pemotong imbuhan, rata-rata korelasi LSA3 lebih rendah dari LSA1 dan LSA2. Hal ini mungkin disebabkan karena kurangnya

konteks sebagai akibat dari pemotong imbuhan. Penjelasan lebih lanjut mengenai hal ini terdapat pada Subbab 5.5.3.

Rata-rata korelasi nilai manusia dengan nilai dari sistem ini memang rendah dan kurang bisa diandalkan untuk menilai jawaban esai. Tetapi, dapat dilakukan pengamatan terhadap bentuk soal atau jawaban yang dapat dinilai dengan baik oleh sistem (korelasi tinggi) dan yang tidak dapat dinilai dengan baik (korelasi rendah).

Korelasi yang tinggi (di atas 0.70) antara nilai manusia dengan VSM didapatkan pada soal nomor 1, 2, 4, 6, dan 9. Sedangkan korelasi yang tinggi antara nilai manusia dengan LSA1 didapatkan pada soal nomor 6, LSA2 pada soal nomor 2, 6, 9, dan LSA3 pada soal nomor 4, 6, 9. Kelima soal (lihat Tabel 5.1) yang menghasilkan korelasi tinggi antara manusia dengan VSM adalah soal yang jawabannya berupa daftar konsep yang sudah pasti, sesuai teori dalam *E-Commerce*, dan tidak banyak cara untuk menyampaikannya. Sedangkan untuk soal lain berupa uraian, meskipun sudah ada kunci jawabannya, tetapi terdapat banyak cara untuk menyampaikannya.

Korelasi yang paling rendah antara nilai manusia dengan semua skema penilaian terjadi pada soal 11 dan 12 (korelasi antara 0 sampai dengan 0.31). Soal nomor 11 adalah soal uraian yang cara menyampaikan jawabannya bisa beragam. Kebanyakan jawaban siswa panjang-panjang sedangkan kunci jawaban singkat. Hal ini mungkin yang menyebabkan rendahnya korelasi nilai manusia dengan nilai sistem, karena saat manusia memberi nilai tinggi, sistem memberi nilai rendah. Soal nomor 12 sebenarnya merupakan soal dengan jawaban berupa daftar konsep yang sudah pasti. Namun korelasi nilai manusia dengan nilai sistem untuk soal ini rendah. Hal ini kemungkinan terjadi karena banyaknya isi kunci jawaban (ada 16 pokok konsep) sehingga jika siswa hanya menyebutkan beberapa saja, sistem memberi nilai yang rendah. Sedangkan penilai manusia memberi nilai 10 meskipun siswa hanya menjawab dengan beberapa pokok saja (misalnya 4).

VSM dapat digunakan untuk bentuk soal yang meminta jawaban pasti seperti istilah atau daftar konsep yang jawabannya tertentu. Sedangkan untuk bentuk soal yang jawabannya pasti tetapi banyak cara menyampaikannya mungkin lebih cocok untuk dinilai dengan LSA, asalkan ada dokumen *training*

yang cukup besar untuk membentuk ruang semantik LSA sehingga LSA dapat digunakan untuk menyimpulkan kesamaan antar kata dan dokumen.

VSM menyerupai cara manusia dalam menilai soal jawaban dengan prinsip *keyword matching* atau pencocokkan kata-kata kunci jawaban dengan jawaban siswa, sedangkan LSA menggunakan analisis yang lebih dalam dari kata-kata pada kunci dan jawaban.

Keyword matching dalam soal-soal tertentu dapat digunakan. Contohnya, untuk menilai soal nomor 1, penilai manusia pertama-tama mencari kata kunci yang merupakan jawaban yang benar seperti “early liquidity”. Hal ini sama dengan yang dilakukan VSM. Tetapi ketika kata kunci ini tidak ada, maka penilai manusia melakukan analisis lebih dalam terhadap makna kata-kata sehingga jawaban siswa seperti “tercapainya jumlah pembeli dan transaksi di awal” dianggap betul meskipun tidak ada kata kunci yang muncul. Ini terjadi karena manusia memiliki pengetahuan awal bahwa jawaban tersebut merupakan pengertian dari “early liquidity”. Analisis lebih dalam ini sama seperti yang dilakukan LSA. Hanya saja LSA memerlukan *training* yang banyak agar dapat menyimpulkan kesamaan antara jawaban tersebut dengan kunci jawaban.

Kekurangan VSM dan LSA adalah sama-sama menggunakan prinsip “*bag-of-words*”, yaitu tidak memperhatikan urutan kata. Urutan kata dan sintaks penulisan memang penting untuk memahami sebuah makna atau wacana, tetapi dalam beberapa kasus, manusia mengekstrak informasi dari wacana tidak hanya bergantung pada urutan kata dan sintaks, melainkan juga dari kombinasi leksikal atau kata-kata yang terdapat di dalamnya. LSA yang digunakan pada IEA terbukti mampu menyamai kemampuan manusia dalam memahami makna tanpa memperhatikan urutan kata (Landauer, et al, 1997).

Dengan adanya kekurangan ini, baik VSM maupun LSA tidak dapat digunakan untuk menilai jawaban yang bergantung pada sintaks atau urutan kata. Contoh pertanyaan dan jawaban di mana urutan kata berpengaruh adalah sebagai berikut.

Pertanyaan: “Sebutkan perbedaan B2B sell-side dan B2B buy-side”.

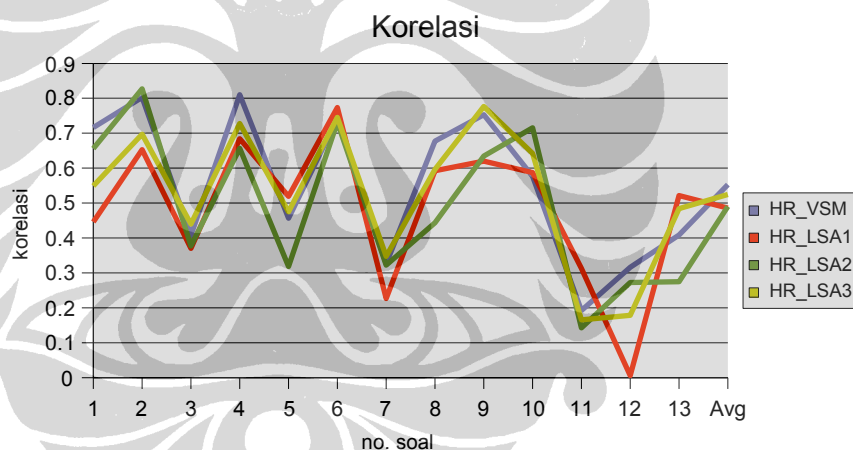
Jawaban: “B2B sell-side terdiri dari banyak pembeli dan satu penjual sedangkan B2B buy side terdiri dari banyak penjual dan

satu pembeli”.

Pada soal seperti ini VSM dan LSA tidak tepat untuk digunakan karena meskipun jawabannya terbalik akan dianggap betul oleh sistem.

5.5.2 Analisis Pengaruh Perluasan Kunci Jawaban

Pengaruh perluasan kunci jawaban pada sistem VSM menurunkan rata-rata korelasi nilainya dengan nilai manusia dibandingkan dengan sebelum menggunakan perluasan kunci, dari 0.56 menjadi 0.55. Sedangkan pada LSA1 terjadi kenaikan rata-rata korelasi nilai dengan nilai manusia dari 0.48 menjadi 0.49. LSA2 juga mengalami kenaikan rata-rata korelasi nilai dengan nilai manusia, dari 0.48 menjadi 0.49. Begitu juga dengan LSA3 yang mengalami kenaikan rata-rata korelasi nilai dengan nilai manusia dari 0.52 menjadi 0.53. Berikut ini adalah grafik yang menggambarkan korelasi nilai VSM, LSA1, LSA2, dan LSA3 setelah perluasan kunci jawaban.



Gambar 5.3: Grafik Perbandingan Korelasi VSM-LSA setelah Perluasan Kunci Jawaban

Perluasan kunci jawaban pada sistem dapat menaikkan kemungkinan kecocokkan kunci jawaban dengan variasi kata yang digunakan siswa pada jawabannya. Pada sistem VSM, perluasan kunci jawaban justru menurunkan korelasinya. Hal ini mungkin terjadi karena banyaknya sinonim yang dimasukkan sebagai perluasan kata kunci, sedangkan yang cocok dengan jawaban siswa hanya satu atau tidak ada yang cocok sama sekali. Panjang vektor kunci jawaban menjadi lebih panjang sehingga bisa menurunkan nilai yang didapat siswa karena

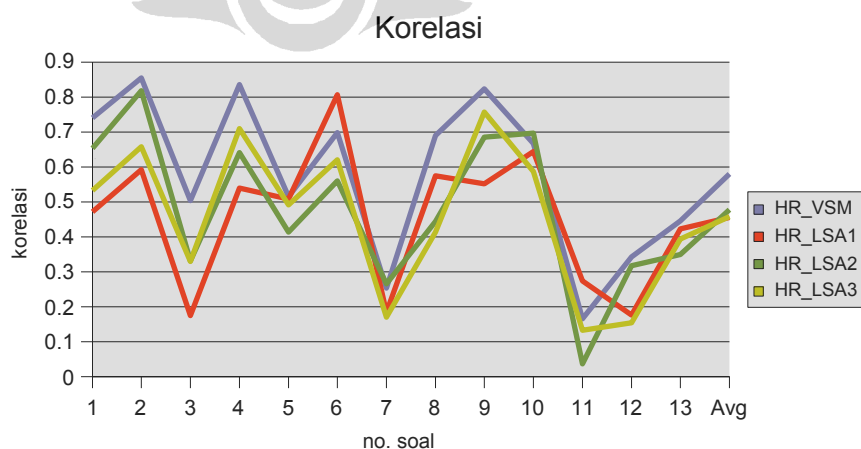
kata yang cocok menjadi lebih sedikit dibandingkan dengan pertambahan panjang vektor kunci jawaban. Akibatnya, perbedaan nilainya dengan nilai dari manusia bisa menjadi lebih besar.

Berbeda dengan VSM, perluasan kunci jawaban pada sistem LSA meningkatkan korelasinya. Hal ini mungkin disebabkan karena penambahan sinonim yang sesuai konteks pada kunci jawaban dapat memperjelas konteks pada kunci jawaban dan jawaban siswa, sehingga analisis kesamaan yang dilakukan LSA lebih baik.

Keterbatasan dari perluasan kunci jawaban ini adalah tidak semua kata pada kunci jawaban bisa ditemukan sinonimnya pada WordNet. Dari 249 kata, hanya 46 kata yang dapat ditemukan sinonimnya dan hanya 29 di antaranya dapat digunakan sebagai sinonim yang sesuai konteks.

5.5.3 Analisis Pengaruh Pemotong Imbuan

Pengaruh pemotong imbuan pada sistem VSM meningkatkan rata-rata korelasi nilainya dengan nilai manusia dibandingkan dengan sebelum menggunakan pemotong imbuan, dari 0.56 menjadi 0.58. Sedangkan pada sistem LSA1 terjadi penurunan rata-rata korelasi nilai dengan nilai manusia dari 0.48 menjadi 0.46. Rata-rata korelasi nilai manusia dengan nilai LSA2 dengan pemotong imbuan tidak berubah (tetap 0.48), sedangkan rata-rata korelasi nilai manusia dengan LSA3 mengalami penurunan seperti pada LSA1, yang tadinya 0.52 menjadi 0.46. Berikut ini adalah grafik yang menggambarkan korelasi nilai VSM, LSA1, LSA2, dan LSA3 setelah dilakukan pemotongan imbuan.

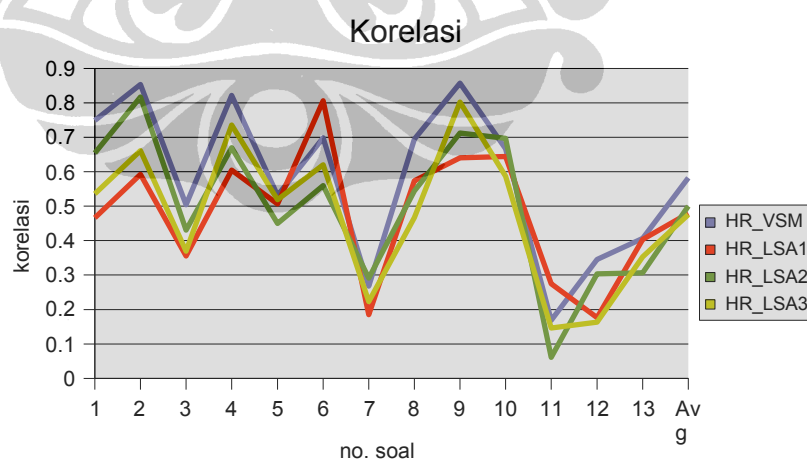


Gambar 5.4: Grafik Perbandingan Korelasi VSM-LSA setelah Pemotongan Imbuan

Pemotongan imbuhan menyebabkan lebih banyak kata yang dapat dicocokkan karena hanya tinggal kata dasarnya saja. Pada VSM, hal ini menguntungkan karena VSM hanya memberi penilaian dari kecocokan kata antara kunci jawaban dengan jawaban siswa. Sedangkan pada LSA yang menganalisis kemunculan kata bukan hanya pada satu jawaban saja melainkan dari seluruh jawaban, hal ini belum tentu menguntungkan. Karena penggunaan pemotong imbuhan bisa saja mengurangi konteks pada jawaban/kunci sehingga mengurangi kualitas analisis. Pengaruh ini sulit untuk diamati secara langsung karena LSA menyimpulkan hubungan kesamaan dengan cara menganalisis kemunculan kata secara global.

5.5.4 Analisis Pengaruh Pemotong Imbuhan dan Perluasan Kunci Jawaban

Penggunaan gabungan pemotong imbuhan dengan perluasan kunci jawaban pada sistem VSM, LSA1, LSA2, dan LSA3 memberikan pengaruh yang bervariasi. Pada sistem VSM, rata-rata korelasi antara nilai manusia dan sistem setelah menggunakan pemotong imbuhan dan perluasan kunci jawaban naik dari 0.56 menjadi 0.58. Rata-rata korelasi nilai manusia dengan LSA2 juga mengalami hal yang sama yaitu dari 0.48 menjadi 0.5. Berikut ini adalah grafik yang menggambarkan korelasi nilai VSM, LSA1, LSA2, dan LSA3 setelah dilakukan perluasan kunci jawaban dan pemotongan imbuhan.



Gambar 5.5: Grafik Perbandingan Korelasi VSM-LSA setelah Perluasan Kunci Jawaban dan Pemotongan Imbuhan

Berbeda dengan VSM dan LSA2, pengaruh perluasan kunci jawaban dan

pemotong imbuhan pada LSA3 menurunkan rata-rata korelasi nilainya dengan nilai manusia menjadi 0.48 dari sebelumnya 0.52. Sedangkan pada LSA1 rata-rata korelasi nilainya dengan nilai manusia tetap 0.48.

Tidak dapat disimpulkan bagaimana pengaruh gabungan perluasan kunci jawaban dan pemotong imbuhan pada sistem LSA karena pengaruhnya berbeda-beda. Tetapi pada VSM pengaruhnya menaikkan rata-rata korelasi. 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 penambahan sinonim untuk 29 kata saja.

