

Sistem Penilaian *Reading Text* pada Mata Pelajaran Bahasa Inggris Menggunakan Metode *Jaccard Measure* (Studi Kasus: SMK Al-Qodiri Jember)

¹Amalia Sarifatun Nurma Ida (11 1065 1106), ²Lutfi Ali Muharom, S.Si., M.Si ³Mudafiq Riyan Pratama, S.Kom
Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember
Email : amelpeshek@gmail.com

Abstrak

Penelitian ini menguji tentang penerapan penggunaan metode *jaccard measure* dalam sistem penilaian soal *reading text* pada mata pelajaran Bahasa Inggris di SMK Al-Qodiri Jember. Data set yang digunakan adalah berupa 51 jawaban siswa kelas XI dengan salah satu diantaranya adalah sebagai kunci jawaban guru (*query*). Dengan penggunaan *Text Mining*, perhitungan *TF/IDF*, dan perhitungan *Jaccard Measure* sebagai dasar konstruksi sistem, penelitian ini menghasilkan sebuah sistem yang dapat mewakili penilaian manual dengan tingkat kesalahan sistem sebesar 19.4%. Hal ini disebabkan oleh penggunaan *Text Mining* dan *TF/IDF* dalam mendeteksi *similaritas/ kemiripan* jawaban siswa terhadap jawaban guru (*query*) yang dibandingkan berdasarkan jumlah kata (*term*) yang menghitung bobot per kata akan menghasilkan rentang nilai (*range*) terhadap jawaban secara manual dengan jawaban menggunakan sistem.

Kata Kunci : *reading text, jaccard measure, term frekuensi, sistem penilaian.*

1. PENDAHULUAN

1.1 Latar Belakang

Salah satu cara untuk meningkatkan SDM (Sumber Daya Manusia) adalah melalui pendidikan. Dalam proses pendidikan diperlukan sebuah evaluasi pembelajaran untuk mengetahui sejauh mana tingkat keberhasilan pengajar dalam memberikan materi terhadap peserta didik dalam menerima ilmu dan menerapkannya dengan baik. Ditinjau dari tingkat waktu evaluasi dapat dibedakan menjadi 5 yaitu evaluasi harian (ujian harian), evaluasi mingguan (ujian mingguan yang dilaksanakan setelah pembahasan materi usai), evaluasi tengah semester (ujian tengah semester atau UTS), evaluasi akhir semester (UAS), dan ujian akhir sekolah. Ditinjau dari bentuk soalnya evaluasi dibedakan menjadi dua macam, yaitu evaluasi bentuk subjektif (ujian essay) dan bentuk objektif yaitu memberikan pilihan jawaban kepada peserta ujian, contohnya tes pilihan ganda (*multiple choice test*) (Sukardi, 2007: 159).

Beberapa pengajar (guru) memberikan evaluasi dengan menggunakan metode *essay*. Namun, evaluasi secara *essay* memiliki beberapa kekurangan diantaranya adalah pengajar membutuhkan waktu yang lama dan tenaga yang lebih dalam melakukan proses pengkoreksian dan penilaian.

Berdasarkan permasalahan yang ada, penelitian ini dilakukan guna menguji serta menerapkan *jaccard measure* terhadap sistem penilaian dengan proses pemilahan jawaban berdasarkan tingkat *similaritas/ kemiripan* jawaban siswa dengan kunci jawaban guru (*query*) sehingga dapat meningkatkan efektifitas dan efisiensi kinerja guru dalam

memberikan penilaian terhadap siswa. Karena beberapa alasan inilah yang mendorong penulis untuk merancang Tugas Akhir yang berjudul “Sistem Penilaian *Reading Text* pada Mata Pelajaran Bahasa Inggris Menggunakan Metode *Jaccard Measure* (Studi Kasus: SMK Al-Qodiri Jember)”. Dengan adanya tugas akhir ini, penulis ingin melakukan pengujian kinerja *jaccard measure* dalam penilaian evaluasi belajar dengan objek penelitian berupa jawaban evaluasi pembelajaran semester genap tahun ajaran 2014/2015 mata pelajaran Bahasa Inggris.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, dapat dirumuskan masalah-masalah yang akan dibahas dalam penelitian ini, antara lain:

1. Bagaimana menerapkan pra-proses text mining dan perhitungan *TF/IDF*?
2. Bagaimana mendeteksi jawaban berdasarkan *similaritas / kemiripan* jawaban menggunakan *Jaccard Measure*?
3. Bagaimana melakukan penilaian terhadap jawaban siswa/i menggunakan sistem penilaian berdasarkan tingkat kemiripan/kesamaan terhadap jawaban guru/kunci jawaban (*query*)?
4. Bagaimana sistem penilaian ini dapat mengukur kinerja metode *Jaccard Measure* dalam memberikan penilaian terhadap jawaban siswa menggunakan pengujian tingkat kesalahan sistem serta *white box* texting dan *black box* texting?

1.3 Batasan Masalah

Agar tidak menyimpang jauh dari permasalahan yang telah dirumuskan, maka disusun batasan masalah sebagai berikut:

1. Dataset yang digunakan dalam metode Jaccard Measure merupakan jawaban dari evaluasi pembelajaran secara essay yang dilaksanakan pada semester genap tahun ajaran 2014/2015.
2. Dataset yang digunakan adalah jawaban siswa dari mata pelajaran Bahasa Inggris kelas XI di SMK Al-Qodiri Jember dengan tipe soal reading text.
3. Dataset berjumlah 51, 50 diantaranya merupakan dokumen uji yang didapat dari jawaban siswa dan 1 dokumen diantaranya merupakan jawaban guru/query sebagai pembandingan.
4. Sistem ini tidak melakukan cek arti kata/kalimat dan tidak membaca gambar (picture).
5. Penilaian jawaban berdasarkan atas tingkat similiaritas terhadap query (kunci jawaban/jawaban guru) dan tidak melakukan pembenaran terhadap jawaban serta tidak mendeteksi plagiarisme.

1.4 Tujuan

Tujuan penelitian ini adalah sebagai berikut:

1. Mendiskripsikan proses text mining dan pembobotan TF-IDF dengan perhitungan relevansi similiaritas jawaban menggunakan Jaccard Measure.
2. Memberikan nilai akhir pada jawaban *essay* dengan tipe soal *reading text* berdasarkan hasil perhitungan *Jaccard Measure*.
3. Mempergunakan sistem supaya bisa dioperasikan ditahun ajaran berikutnya.
4. Dapat mengetahui hasil dari kinerja *Jaccard Measure* dalam sistem penilaian jawaban essay dengan tipe soal *reading text* pada mata pelajaran Bahasa Inggris.

1.5 Manfaat

Manfaat yang diperoleh dari hasil penelitian ini adalah sebagai berikut:

1. Dapat mengurangi waktu tenaga pengajar (guru) dalam pengoreksian jawaban *essay* dengan tipe soal *reading text*.
2. Dapat membantu tenaga pengajar dalam mendeteksi/mengetahui kesamaan antar jawaban guru dan peserta didik, sehingga dapat memberikan kemudahan dalam proses penilaian.
3. Peneliti dapat mengetahui kinerja metode *Jaccard Measure* dalam memberikan penilaian pada mata pelajaran Bahasa Inggris.

2. TINJAUAN PUSTAKA

2.1 Sejarah SMK Al-Qodiri Jember

Sekolah Menengah Kejuruan Al-Qodiri didirikan pada tanggal 29 Juli 2011, berada didalam kompleks Pondok Pesantren Al-Qodiri, jalan Manggar 139 A RT : 02 RW : 14 Gebang Patrang Jember Jawa Timur.

2.2 Pengertian Sistem Penilaian

Sistem penilaian adalah suatu rangkaian perintah dan prosedur yang dijalankan secara bersama-sama dan saling berhubungan dengan tujuan mengolah masukan sistem (*input*) dan menghasilkan suatu keluaran (*output*) berdasarkan komponen-komponen sistem tersebut dengan batasan-batasan tertentu untuk mencapai suatu tujuan dan fungsi penilaian.

2.3 Pengertian Mata Pelajaran Bahasa Inggris

Mata pelajaran Bahasa Inggris merupakan mata pelajaran adaptif, yang bertujuan membekali kemampuan peserta didik dalam berkomunikasi Bahasa Inggris dalam konteks material komunikasi yang diperlukan bagi program keahliannya, baik yang bersifat lisan maupun tulis. Di samping itu, mata pelajaran Bahasa Inggris membekali kemampuan berkomunikasi peserta didik dalam kehidupan sehari-hari sesuai dengan tuntutan global. Berdasarkan sistem pembelajaran Kurikulum Satuan Pendidikan (KTSP) yang berlaku di Indonesia, Bahasa Inggris memiliki empat kompetensi dasar antara lain: kompetensi membaca (*reading*), kompetensi menulis (*writing*), kompetensi berbicara (*speaking*), dan kompetensi mendengarkan (*listening*). Secara global, kompetensi membaca dievaluasi dengan soal *reading text*, yaitu berupa teks narasi, paparan proses, berita dan lain sebagainya yang memiliki poin pertanyaan meliputi isi dari bacaan tersebut.

2.4 Pengertian Text Mining

Menurut Feldman dan Sanger (2007: 14-16), *Text Mining* didefinisikan sebagai proses pengetahuan yang intensif dimana penggunaanya melakukan pengumpulan dokumen menggunakan seperangkat alat analisis. *Text mining* dilakukan untuk mengekstrak informasi yang diperlukan dari sumber data melalui identifikasi dan eksplorasi terhadap *database* ataupun data tekstual yang tidak terstruktur. Pada penelitian ini *text mining* memiliki beberapa tahapan yaitu tahap *scanning*, *case folding*, *filtering*, *tokenizing* dan *stemming*.

2.4.1 Case Folding

Tahapan *case folding* adalah suatu proses pembentukan huruf kecil sehingga kalimat atau *text* akan dinormalisasikan menjadi huruf kecil secara keseluruhan.

2.4.2 Tokenizing

Tahap *tokenizing* adalah tahap pemotongan *string input* berdasarkan tiap kata yang menyusunnya. Tokenisasi secara garis besar memecah sekumpulan karakter dalam suatu teks ke dalam satuan kata, bagaimana membedakan karakter-karakter tertentu yang dapat diperlakukan sebagai pemisah kata atau bukan.

2.4.3 Filtering

Tahap *Filtering* adalah tahap mengambil kata-kata penting dari hasil token. Bisa menggunakan algoritma *stop list* (membuang kata yang kurang penting) atau *word list* (menyimpan kata penting) dimana menghilangkan atau membuang kata yang tidak dipakai seperti simbol, tanda baca, kata sambung dan notasi ilmiah seperti angka yang terdapat pada paragraf didokumen tersebut disebut dengan *Stop word removal*.

2.4.4 Stemming

Stemming adalah proses pemetaan dan penguraian berbagai bentuk (variants) dari suatu kata menjadi bentuk kata dasarnya. Proses ini juga disebut sebagai *conflation*. Proses *stemming* secara luas sudah digunakan di dalam pencarian informasi (*information retrieval*) untuk meningkatkan kualitas informasi yang didapatkan (Baskoro, Malik dan Anshari, 2012). Tahapan *stemming* yang akan digunakan dalam penelitian ini adalah algoritma *stemming porter*.

2.5 Algoritma Stemming Porter

Algoritma *Stemming porter* dikembangkan oleh Martin Porter pada tahun 1980 di University of Cambridge. Algoritma *stemming porter* memiliki lima tahapan, dan setiap tahapan memiliki aturan dan ketentuan tersendiri.

Tahap pertama adalah menghilangkan imbuhan jamak, *past participles*, *present participles* dan mengubah akhiran “y” ke “I”, (contoh: “*generalizations*” menjadi “*generalization*”, “*agreed*” menjadi “*agree*”, dan “*happy*” menjadi “*happi*”). Tahap kedua adalah peringkasan imbuhan rangkap menjadi tunggal, (contoh: “*generalization*” menjadi “*generalize*”, “*oscillator*” menjadi “*oscillate*”). Tahap ketiga adalah membuang imbuhan yang belum terbuang pada tahap sebelumnya, (contoh: “*generalize*” menjadi “*general*”). Tahap keempat adalah membuang sisa imbuhan yang masih belum terproses di tahapan sebelumnya, (contoh: “*general*” menjadi “*gener*”, “*oscillate*” menjadi “*oscill*”). Tahap kelima adalah membuang “e” dan huruf konsonan rangkap, (contoh: “*attribute*” menjadi “*attribut*”, “*oscill*” menjadi “*oscil*”).

2.6 TF/IDF (Term Frequency/ Inverse Document Frequency)

Menurut Sulistyono, dkk (2015), *Term Frequency* merupakan frekuensi kemunculan suatu kata (*term*) dalam dokumen. Oleh sebab itu, *tf* memiliki nilai yang bervariasi dari satu dokumen ke dokumen lain tergantung dari tingkat kepentingan sebuah *term* dalam sebuah dokumen. Semakin sering suatu *term* muncul dalam suatu dokumen, *term* tersebut akan memiliki nilai *tf* yang lebih besar daripada *term-term* lain yang jarang muncul. Penggunaan faktor *tf* belum mencukupi dalam menentukan pembobotan. Untuk itu, diperlukan faktor *Inverse Document Frequency (idf)* yang merupakan sebuah metode statistik yang mengkararakteristikan sebuah *term* dalam keseluruhan pengumpulan dokumen. *Idf* merupakan sebuah perhitungan tentang bagaimana *term* didistribusikan secara luas pada pengumpulan dokumen. Semakin sedikit dokumen yang mengandung *term* yang dimaksud, maka nilai *idf* akan semakin *besar*. Jika setiap dokumen mengandung *term*, maka nilai *idf* dari *term* tersebut adalah nol (0).

Dalam penelitian ini, algoritma pembobotan *Term Frequency (tf) – Inverse Document Frequency (idf)* ditetapkan pada tahap *similarity document*. Menurut D.L. Lee, dkk (1997) nilai *tf-idf* diperoleh dengan menggunakan persamaan sebagai berikut:

$$W_{ij} = tf_{ij} \times idf_j = tf_{ij} \times \log\left(\frac{N}{df_j}\right) \dots (1)$$

Keterangan :

- W_{ij} = bobot term ke-j terhadap dokumen (i)
- tf_{ij} = jumlah kemunculan term j dalam dokumen (i)
- N = jumlah dokumen secara keseluruhan
- df_j = jumlah dokumen mengandung term j

Berdasarkan persamaan (1), berapapun besarnya nilai tf_{ij} apabila $N = df_j$ maka untuk perhitungan nilai idf_j akan memperoleh hasil nol (0). Oleh karena itu, dapat ditambahkan nilai 1 pada sisi idf_j , sehingga perhitungan bobot menjadi:

$$W_{ij} = tf_{ij} \times \log\left(\frac{N}{df_j}\right) + 1 \dots (2)$$

Normalisasi *tf-idf* dilakukan dengan formula sebagai berikut:

$$W_{ij} = \frac{tf_{ij} \times idf_j \text{ pembobotan term}}{\sqrt{\sum_{k=1}^n (tf_{ik} \times idf_k \text{ pembobotan term})^2}}$$

Atau dengan persamaan sebagai berikut:

$$W_{ij} = \frac{tf_{ij} \times (\log\left(\frac{N}{df_j}\right) + 1)}{\sqrt{\sum_{k=1}^n (tf_{ik} \times (\log\left(\frac{N}{df_k}\right) + 1))^2}}$$

2.7 Jaccard Measure

Menurut Rolly Intan dan Andrew Defeng (2006), *Jaccard Measure* atau *Jaccard Coeficient* adalah salah satu metoda yang dipakai untuk menghitung *similarity* antara dua objects (items).

Formulasi Jaccard Measure digunakan untuk menghitung similaritas dalam penelitian ini, adalah sebagai berikut:

$$\text{Sim}(X, Y) = \frac{\sum_{i=1}^n (X_{1i} \times Y_{1i})}{\sum_{i=1}^n (X_{1i}) + \sum_{i=1}^n (Y_{1i}) - \sum_{i=1}^n (X_{1i} \times Y_{1i})}$$

dimana X dan Y adalah merupakan dokumen yang akan diuji similaritasnya.

2.8 Penelitian Terkait

Penelitian oleh Lukman Hakim (2014) tentang Sistem Penilaian Otomatis Jawaban *Essay* Bahasa Inggris Menggunakan Metode *Cosine Measure* Pada Sistem E-Learning di SMKN 1 Tanggul telah mendasari penyusun dalam melaksanakan penelitian lanjutan mengenai Sistem Penilaian Otomatis. Hasil yang dicapai sistem dengan metode *cosine similarity* dalam mendeteksi jawaban soal essay dengan kriteria Presisi dan Recall adalah sebesar 50% dan 33,34%. Oleh karena itu penyusun mencoba untuk menerapkan *jaccard measure* atau metode pendeteksi similaritas dalam lingkup soal *reading text* pada mata pelajaran Bahasa Inggris. Jika pada penelitian sebelumnya telah dilakukan dengan pengujian hasil kinerja sistem dibandingkan hasil kinerja secara manual mengalami selisih yang signifikan, diharapkan penelitian kali ini dapat memperoleh hasil yang maksimal dengan melakukan pengujian tingkat kesalahan sistem serta *white box testing* dan *black box testing*. Di dalam penelitian kali ini penyusun memilih soal *reading text* dengan lebih memberikan penekanan terhadap pencarian silimaritas terhadap materi *reading text*.

3. METODOLOGI PENELITIAN

3.1 Arsitektur / Desain Sistem

Adapun model / rancangan / arsitektur / desain sistem yang akan digunakan dalam penelitian ini adalah seperti berikut :



Keterangan :

- Pengumpulan data dari hasil evaluasi kenaikan kelas XI mata pelajaran Bahasa Inggris SMK Al-Qodiri Jember tahun ajaran 2014/2015.
- Pemrosesan *Text Mining*, yang dilakukan dengan 5 tahap. Yaitu *scanning*, *case folding*, *filtering*, *tokenizing* dan *stemming*.
- Hasil dari tahapan *text mining* akan dilakukan perhitungan terhadap *term-term* dengan menggunakan perhitungan *tfidf*.
- Selanjutnya dilakukan perhitungan menggunakan *jaccard measure* guna mengetahui nilai yang diperoleh dari tingkat kemiripan/kesamaan jawaban.

4. HASIL DAN PEMBAHASAN

4.1 Data

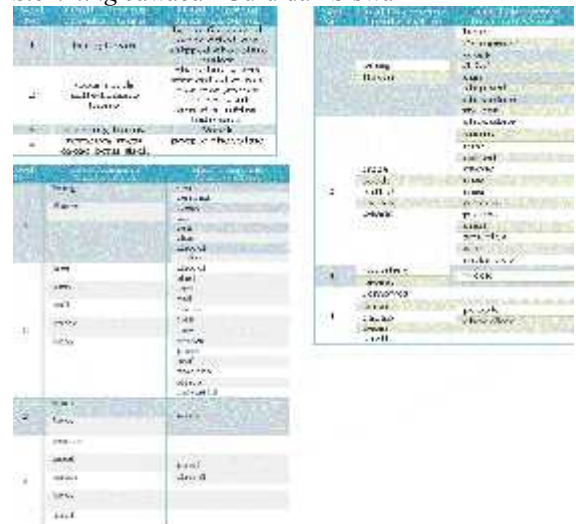
Data yang akan diolah pada proses pengujian didapat dari jawaban *essay* siswa/i SMK Al-Qodiri Jember tahun ajaran 2014/2015 dari masing-masing jurusan yaitu Akuntansi dan Manajemen Dengan jumlah masing-masing dokumen sebagai berikut:

Jurusan	Jawaban Siswa	Jawaban Guru (Kunci Jawaban)
Akuntansi 1	15	1
Akuntansi 2	17	
Multimedia	18	
Total	50	1

4.2 Rekayasa Uji Coba

Pada rekaya uji coba akan dilakukan pengujian terhadap salah satu jawaban siswa. Dalam tahap ini, jawaban siswa dan jawaban guru di input ke dalam proses *text mining* dengan melakukan tahap *filtering*, *tokenizing*, dan *stemming*.

Gambar 4.2 Hasil *Filtering*, *Tokenizing*, dan *Stemming* Jawaban Guru dan Siswa



Selanjutnya dilakukan tahapan pembobotan dengan metode perhitungan *tf-idf* pada masing-masing jawaban siswa dan guru, yang dilanjutkan kepada

tahap pencarian nilai similaritas antara jawaban siswa dengan jawaban guru menggunakan metode *jaccard measure*.

Gambar 4.3 Hasil *Tf-idf*, dan *Jaccard Measure* Jawaban Guru dan Siswa

Setelah diperoleh nilai *jaccard measure* atau nilai similaritas jawaban siswa terhadap jawaban guru, selanjutnya akan dilakukan perhitungan nilai siswa berdasarkan nilai *jaccard measure* yang diperoleh siswa untuk digunakan sebagai nilai akhir. Nilai *jaccard measure* yang diperoleh oleh siswa dari masing-masing soal adalah 0, 0.261, 0, dan 0.

Tabel 4.4 Hasil Perhitungan Nilai Siswa Berdasarkan Nilai *Jaccard Measure*

Soal No	Nilai <i>Jaccard</i>
1	0
2	0.261
3	0
4	0
Total	0.261
Nilai yang diperoleh	0.261 * 100 = 26.1

Dapat diketahui bahwa jawaban dari soal no 1,3, dan 4 tidak memiliki kemiripan dengan jawaban guru, sehingga siswa tersebut memiliki nilai yang rendah. Hal ini dikarenakan jumlah yang diperoleh dari tiap-tiap soal berbeda dan kemiripan jawaban siswa dengan jawaban guru sangatlah berbeda. Dari tabel tersebut telah diperoleh nilai siswa sebesar 26.1 yang akan menjadi nilai akhir yang diperoleh siswa.

4.3 Analisa Pengujian

4.3.1 Pengujian *Precision* dan *Recall*

Pengujian dilakukan dengan menggunakan *range* (jarak), *range* yang digunakan adalah 100-76, 75-51, 50-26 dan 25-0.

Dari dataset ketiga kelas dibagi menjadi dua yaitu data *training* sejumlah 50 dokumen dan data *testing* sejumlah 1 dokumen yaitu dokumen guru(*query*) dengan total 51 dokumen pada semua kelas. Didapatkan hasil dari perhitungan nilai *Precision* dan *Recall* sebagai berikut:

Range	Precision	Recall
100-76	62,50%	76,92%
75-51	83.33%	50%
50-26	43,75%	87,50%
25-0	0%	0%
Rata-rata	35,42%	53,61%

4.3.2 Tingkat Kesalahan Sistem

Pada tahap pengujian peneliti menggunakan sebuah cara untuk mengetahui tingkat kesalahan yang diperoleh dari metode *jaccard measure* dalam memberikan penilaian dengan cara menghitung tingkat kesalahan yang dihasilkan oleh sistem tersebut.

Rumus mencari tingkat kesalahan sistem adalah sebagai berikut (Zakiyya, Muhammad. 2014):

$$\text{Prosentase Kesalahan} = 100 - \left(\frac{NS}{NM} \times 100 \right)$$

Keterangan :

NS = Nilai Sistem, yaitu nilai yang dihasilkan sistem dengan penerapan *Jaccard Measure*.
 NM = Nilai Manual, yaitu nilai yang diberikan oleh guru terhadap siswa.

Penulis menerapkan metode tingkat kesalahan tersebut dalam menguji peneliannya, sehingga diperoleh hasil pada masing-masing kelas dengan tingkat kesalahan sistem sebagai berikut :

Tabel 4.1 Kelas Akuntansi 1

No	Nama Siswa	Nilai Sistem	Nilai Guru	Salah	Tingkat Kesalahan %
1	Khusnul Khotimah	81.2	100	18.8	18.8
2	Lady Susannah	54.4	75	20.6	27.8
3	Alivia Inan Permadani	78.5	100	21.5	21.5
4	Alfiana Yafira	59.7	80	20.3	25.4
5	Firman Rya	65.4	75	9.6	12.8
6	Aseptian Octa	78.5	100	21.5	21.5
7	Anisa Puji Astuti	78.5	100	21.5	21.5
8	Siti Nur Kholida	53.4	75	21.6	28.8
9	Ursulinar Robsoniah	46	50	4	8.0
10	Vivi Andriani	75.5	100	24.5	24.5
11	Febry Septian Zakaria	76.1	0	26.1	-
12	Nyelia Sari	71.2	100	28.8	28.8
13	Prasetya Hendriawan	71.7	85	3.3	3.1
14	Ninuk Sawalidih	55.8	75	19.2	25.7
15	Maulidatul Zahro	58.8	80	8.8	11.0
Total				254.6	274.8
Rata-rata				16.97	18.2

Tabel 4.1 Kelas Akuntansi 2

No	Nama Siswa	Nilai Sistem	Nilai Guru	Selisih	Tingkat Kesalahan %
1	Khoirun Nisa	80.1	75	5.1	6.8
2	Ilana Huzamah	66	73	7.3	20.0
3	Husamul Qaimul	71.7	100	28.8	28.8
4	Almal Yaqin	56.9	75	18.1	24.1
5	Almal Syarif	50.8	100	49.2	49.2
6	Muhammad Heral	58.4	50	8.4	12.8
7	Ayu Purwandari	71.2	100	28.8	28.8
8	Safira Dabbat Amim	23.1	50	26.9	68
9	Muhammad Fahar Ihsa	66	50	16	20.0
10	Nur Aini Lailani Hidayat	33.4	50	16.6	6.8
11	Milham Arfan	23.1	50	26.9	68
12	Firda Firda Nanta	80.1	100	19.9	19.9
13	Fitri Hastawati	71.2	75	3.8	5.1
14	Devi Adagnah	23.6	50	26.4	20.8
15	Amelin Putri	71.2	75	3.8	5.1
16	Isti Kholidah	71.2	75	3.8	5.1
17	Sri Wahyuni	39.9	25	14.9	59.6
Total				202.2	204.3
Rata-rata				11.9	17.9

Tabel 4.1 Kelas Multimedia

No	Nama Siswa	Nilai Sistem	Nilai Guru	Selisih	Tingkat Kesalahan %
1	Alhamdulillah Hidayatillah	57.7	50	7.7	15.4
2	Muhammad Abdulhikmah	67.7	75	7.7	16.9
3	Muhammad Sadiq Maulana	44.5	50	5.5	11.0
4	Robby Muzono	67.7	75	7.7	16.9
5	Rosa Imania Artiska	83.5	100	16.5	16.5
6	Jeda Bangsal A.	39.1	50	10.9	21.8
7	Yunita Aprilia Sari	71.2	100	28.8	28.8
8	Tatang Setyawan	81.5	100	18.5	16.5
9	Siti Nur Azzah	92.3	75	17.3	20.1
10	Almal Saari	76.6	100	23.4	23.4
11	Samsul Arifin	35.6	50	14.4	28.8
12	Fahmi Hidayat	53.1	75	21.6	28.8
13	Muhammad Ammah	37.7	75	37.3	23.1
14	Amriyul Fuzama	76.6	100	23.4	23.4
15	M Rahman Hidayat	53.4	75	21.6	28.8
16	Tika Wahidatulhid	71.2	100	28.8	28.8
17	Anik Marwan	36.4	75	38.6	24.8
18	Sahri Yus	38.8	50	11.2	17.6
Total				306.5	304.1
Rata-rata				17.0	17.0

Pada kelas Akuntansi 1 dengan siswa berjumlah 15 orang, diperoleh nilai prosentase kesalahan paling tinggi yang dihasilkan antara penilaian guru/ penilaian manual dengan penilaian sistem adalah 28.8% dan prosentase kesalahan paling rendah adalah 5.1%. Sehingga didapatkan rata-rata dari kesalahan sistem dibandingkan dengan penilaian manual adalah sebesar 18.3% dengan total selisih nilai rata-rata 16.97%.

Pada kelas Akuntansi 2 dengan siswa berjumlah 17 orang, diperoleh nilai prosentase kesalahan paling tinggi yang dihasilkan antara penilaian guru/ penilaian manual dengan penilaian sistem adalah 59.6% dan prosentase kesalahan paling rendah adalah 5.1%. Sehingga didapatkan rata-rata dari kesalahan sistem dibandingkan dengan penilaian manual adalah sebesar 17.9% dengan total selisih nilai rata-rata sebesar 11.9%.

Pada kelas Multimedia dengan siswa berjumlah 18 orang, diperoleh nilai prosentase kesalahan paling tinggi yang dihasilkan antara penilaian guru/ penilaian manual dengan penilaian sistem adalah 28.8% dan prosentase kesalahan

paling rendah adalah 11.0%. Sehingga didapatkan rata-rata dari kesalahan sistem dibandingkan dengan penilaian manual adalah sebesar 21.9% dengan total selisih nilai rata-rata 17.0%.

Untuk mengetahui rata-rata dari tingkat kesalahan yang diperoleh untuk setiap kelas adalah dengan menjumlahkan nilai kesalahan yang dimiliki sistem pada tiap kelas kemudian dibagikan dengan jumlah kelas yang ada, sehingga diperoleh hasil seperti pada tabel berikut ini:

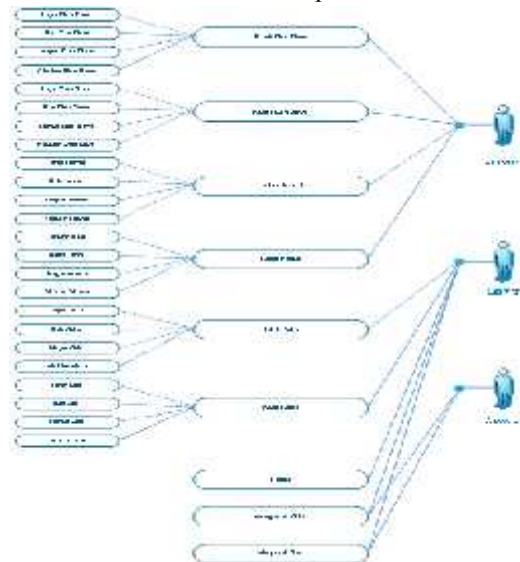
Kelas	Selisih Nilai	Tingkat Kesalahan (%)
Akutansi 1	16.97	18.3
Akutansi 1	11.9	17.9
Manajemen	17.0	21.9
Rata-rata	15.3	19.4

Dari tabel diatas dapat disimpulkan bahwa sistem ini memiliki tingkat kesalahan sebesar 19.4% dengan rata-rata selisih nilai antara nilai sistem dengan nilai yang diberikan oleh guru sejumlah 15.3%.

4.3.3 Black Box dan White Box

1. Pengujian White Box

Pada pengujian sistem pengujian sistem ini menggunakan sistem pengujian *white box*, berikut merupakan *usecase* dari sistem pengujian penerapan *Jaccard Measure* dalam sistem penilaian :

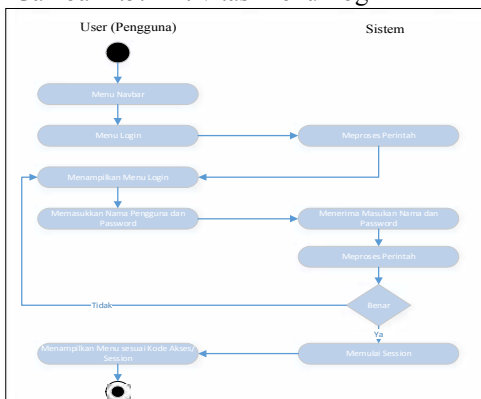


Dari *usecase* diatas dapat dijelaskan bahwa setiap menu yang dapat diakses oleh masing-masing *user* berbeda. Untuk *user* ADMIN dapat mengakses beberapa menu, salah satunya yaitu dapat mengelola data guru (melihat data, menginputkan data, edit/mengubah data, dan menghapus data). Untuk *user* GURU dapat mengakses beberapa menu, salah satunya yaitu mengelola soal (melihat soal,

menginputkan soal, edit/mengubah soal, dan menghapus soal), untuk *user* SISWA dapat menjawab soal yang diberikan, dan melihat hasil nilai ujian.

Setiap menu yang ada terdapat sebuah aktivitas yang dilakukan oleh *user* untuk mengakses menu tersebut yang dapat digambarkan di dalam sebuah diagram aktivitas. Untuk mengakses keseluruhan menu, *user* harus masuk terlebih dahulu (*login*). Setiap kode *user* memiliki akses yang berbeda untuk mengoperasikan sistem. Berikut adalah diagram aktivitas dari menu *login*.

Gambar 4.5. Aktivitas Menu Login



2. Pengujian Black Box

Dalam pengujian alur sistem menggunakan metode *Black Box* untuk menemukan kesalahan dalam sistem, pengujian ini dapat dilihat pada tabel pengujian *Black Box* sebagai berikut :

Tabel 4.5 Pengujian Black Box

No.	State (Bagi)	Aktuator	Aksi	Output	Status
1.	Menu Utama	Tekan pada Tombol	Go to home	Menampilkan home	Heure
2.	Menu Utama	Tekan pada tombol	Go to login	Menampilkan login	Heure
3.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Salah
4.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Heure
5.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Heure
6.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Heure
7.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Heure
8.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Heure
9.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Heure
10.	Menu Login	Apakah	Memeriksa Nama dan Password	Menampilkan pesan	Heure

Berdasarkan dari hasil pengujian dengan menggunakan metode *Black Box* pada data di atas dapat diambil kesimpulan bahwa sistem ini secara fungsional telah memberikan hasil dari setiap proses sesuai dengan yang diharapkan.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari uji coba dan analisa yang telah dijelaskan dalam bab sebelumnya, maka dapat diambil beberapa kesimpulan sebagai berikut :

1. Pada perhitungan TF/IDF yang sudah dilakukan bahwa semakin banyak *term* pada jawaban siswa, maka bobot hubungan antara jawaban siswa terhadap *query* (jawaban guru) akan semakin besar dalam perhitungan *jaccard measure*.
2. Karena sistem ini menerapkan *jaccard measure* dan *text mining*, maka nilai yang didapat oleh sistem memiliki selisih dengan nilai yang diberikan oleh guru yaitu sebesar 162.6.
3. Penilaian yang diberikan merupakan penilaian dari hasil kemiripan guru dengan jawaban siswa dengan menghitung persentase kemiripan jawaban.
4. Pengujian dokumen menggunakan cara menghitung tingkat kesalahan yang dimiliki oleh sistem, sistem memiliki tingkat kesalahan rata-rata sebesar 10.4%.

5.2 Saran

Penulis ingin memberikan beberapa saran yang mungkin dapat membantu dalam pengembangan selanjutnya pada Tugas Akhir ini adalah sebagai berikut :

1. Untuk pengembangan selanjutnya, penulis menyarankan penggunaan metode lain seperti *Rabbin Karp* atau *Levenshtein Distance*.
2. Dalam penelitian ini, penulis hanya berfokus kepada penilaian jawaban, diharapkan pada penelitian selanjutnya terdapat tambahan seperti pendeteksi plagiat, sehingga sistem penilaian dapat memberikan manfaat lebih kepada guru dan siswa.
3. Untuk tahap pengembangan selanjutnya penulis menyarankan peneliti menambahkan deteksi sinonim, antonim, dan pengenalan terhadap Bahasa Natural.

DAFTAR PUSTAKA:

Arrosyida, Rizqa. 2012. *Sistem Penilaian Otomatis Jawaban Essay Menggunakan Deteksi Similarity*. Skripsi Terpublikasi. Surabaya: Universitas Pembangunan Nasional Veteran.

Baskoro, Daniel Oscar; Habib Malik; Muhammad Hafidz Anshari. 2012. *Porter Stemmer*

- Information Retrieval*. Yogyakarta: Universitas Gajah Mada.
- Chuang, Chia-Jui. 2012. *Estimation Of Similarity Indices via Two- Sample Jackknife Procedure*. Taicung: National Chung Hsing University.
- Dikli, Samire. 2006. *Automated Essay Scoring*. Florida: Florida State University.
- Feldman, Ronen; James Sanger . 2007. *Text Mining Hand Book*. New York: Cambridge University Press.
- Gautam, Jyoti; Ela Kumar. 2013. *An Integrated and Improved Approach to Terms Weighting in Text Classification*. Greater Noida: School of Information and Communications Technology Gautam Buddha University.
- Hakim, Lukman. 2014. *Sistem Penilaian Otomatis Jawaban Essay Menggunakan Metode Cosine Measure Pada Sistem E-Learning*, Jember: Universitas Muhammadiyah Jember.
- Harlian, Milkha. 2006. *Text Mining*. Surabaya: Universitas Negeri Surabaya.
- Hasanah, Ainun; Ratih, Ayuninghemi. 2013. *Penerapan Maximum Marginal Relevance Dalam Peringkasan Teks*, Jember: Universitas Muhammadiyah Jember.
- Jogiyanto. 2005. *Analisis dan Desain Sistem Informasi*. Yogyakarta: Andi.
- Kadir, Abdul. 2003. *Pengenalan Sistem Informasi*. Yogyakarta: Andi.
- Kara, Wahiba Ben Abdessalem. 2013. *A New Stemmer To Improve Information*. Tunisia: University Of Tunis.
- Lee, D.L.; H. Chuang; K. Seamons. 1997. *Document Ranking and the Vector Space Model*. IEEE Software, vol. 14, Issue 2, pp. 67 – 75.
- Lemaire, B; Desusu, P. 2001. *A Sistem To Asses The Semantic Content Of Studen Essays*. Jurnal Of Educational Computing Research.
- Ramos, Juan. *Using TF-IDF to Determine Word Relevance in Document Queries*. Rutgers University
- Sudjana, Nana. 1995. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosdakarya.
- Sukardi, 2007. *Evaluasi Pendidikan Prinsip & Operasionalnya*. Jakarta: Alfabeta.
- Sulistiyo, Meiyanto Eko; Ristu Saptno; Adam Asshidiq. 2015. *Penilaian Ujian Bertipe Essay Menggunakan Metode Tex Similarity*. Surakarta: Universitas Sebelas Maret.
- Sutabri, Tata. 2003. *Analisa Sistem Informasi*. Yogyakarta: Andi.
- Waegel, Daniel. 2011. *The Porter Stemmer*. Austin: University of Texas.
- Widodo, Agung; Ratih, Ayuninghemi; Deni, Arifianto. 2012. *Menentukan Topik dari sebuah Tugas Akhir Menggunakan Jaccard Measure*. Jember: Universitas Muhammadiyah Jember.
- Xia, Tian; Yanmei, Chai. 2011. *An Improvement to TF-IDF : Tem Distribution based Term Weight Algorithm*. Journal Of Software; vol.6. No.3.
- Zakiya, Muhammad. 2014. *Sistem Penilaian Otomatis Jawaban Esai Menggunakan Algoritma Levenshtein Distance*. Malang: Universitas Brawijaya Malang.