

Sistem Penilaian Otomatis Jawaban Esai Bahasa Indonesia Berdasarkan Kemiripan Kalimat Menggunakan *Syntactic-Semantic Similarity*

Nur Lujeng Kinanti¹, Anita Qoiriah²

¹Jurusan Teknik Informatika/Teknik Informatika, Universitas Negeri Surabaya

²Jurusan Teknik Informatika/Teknik Informatika, Universitas Negeri Surabaya

¹nurkinanti16051204001@mhs.unesa.ac.id

²anitaqoiriah@unesa.ac.id

Abstrak—Ujian esai dapat digunakan untuk mengukur pemahaman dan interpretasi seseorang terhadap materi yang diperoleh. Oleh karena itu, ujian dalam bentuk esai lebih sering digunakan oleh pengajar dari pada ujian dalam bentuk pilihan ganda. Akan tetapi proses penilaian untuk ujian dengan jawaban bentuk esai masih menggunakan proses manual. Kekurangan penilaian secara manual yaitu sering terjadi kesalahan dalam koreksi jawaban karena ketidakkonsistenan penguji dan jumlah soal yang dikoreksi banyak sehingga membutuhkan banyak waktu. Metode *Syntactic-Semantic Similarity (SynSemSim)* merupakan metode pemecahan masalah penghitungan persamaan secara sintaktik berdasarkan kemiripan kata dan semantik berdasarkan arti atau makna sebuah kata dengan mengenali struktur kalimat terlebih dahulu. merupakan metode untuk menyelesaikan permasalahan menghitung kemiripan kalimat secara sintaktik berdasarkan kemiripan kata dan semantik berdasarkan arti atau makna kata dengan mengenali struktur kalimat terlebih dahulu. Pada penelitian ini metode *SynSemSim* diimplementasikan pada sistem penilaian otomatis jawaban esai Bahasa Indonesia dengan membandingkan jawaban dan kunci jawaban. Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa metode *syntactic-semantic similarity* yang diujikan pada 2 materi berbeda menunjukkan hasil uji sistem penilaian otomatis jawaban esai Bahasa Indonesia dibandingkan dengan hasil penilaian manual yaitu rata-rata kemiripan 70,44% .

Kata Kunci— *Automated Essay Scoring, Syntactic-Semantic Similarity, POS Tagging, Levenshtein distance, WordNet.*

I. PENDAHULUAN

Evaluasi adalah proses menentukan tingkat keberhasilan dalam proses pengajaran. evaluasi pembelajaran terdiri dari tes pilihan ganda dan tes uraian atau esai. Dalam penerapannya evaluasi dengan tes uraian atau esai lebih sering dipilih daripada tes pilihan ganda karena tes dalam bentuk esai dapat digunakan untuk mengukur presentase pemahaman dan interpretasi seseorang terhadap materi yang diperoleh. Akan tetapi proses penilaian untuk ujian dengan jawaban bentuk esai masih menggunakan proses manual. Kekurangan penilaian secara manual yaitu sering terjadi kesalahan dalam koreksi jawaban karena ketidakkonsistenan penguji dan jumlah soal yang dikoreksi banyak sehingga membutuhkan banyak waktu. Berdasarkan permasalahan di atas, diperlukan sistem penilaian otomatis jawaban esai untuk mempercepat proses penilaian. Adapun keuntungan penilaian jawaban esai dengan

memanfaatkan sistem yaitu sistem dapat memeriksa jawaban esai lebih detail dibandingkan manusia, sistem lebih konsisten dalam melakukan proses penilaian, dan penilaian sistem bersifat objektif tidak terpengaruh apapun.

Penelitian mengenai penilaian ujian esai sebelumnya telah dilakukan Fifi Novalita Sari, Muhammad Harlanu, dan Djoko Adi Widodo pada tahun 2016 dengan judul “Implementasi Algoritma *Winnowing* Pada *Platform Elearning* Untuk Sistem Penilaian Otomatis Jawaban *Essay*”. Pada penelitian ini terdapat perbedaan yang signifikan antara nilai koreksi manual dan nilai koreksi sistem, dengan selisih rata-rata tertinggi yaitu 9,72. Penelitian ini perlu dikembangkan dengan membuat persamaan untuk katakata yang memiliki makna sama [1].

Selain itu juga terdapat penelitian dari Mohammad Agus Salim pada tahun 2017 yang berjudul “Pengembangan Aplikasi Penilaian Ujian *Essay* Berbasis Online Menggunakan Algoritma Nazief Dan Adriani Dengan Metode *Cosine Similarity*” dan Taufiqurrahman Saleh Afif pada tahun 2018 yang berjudul “Pengembangan Sistem Penilaian Otomatis untuk Pelajaran Kelas X di SMK Negeri 3 Buduran Sidoarjo”. Hasil penelitian menemukan bahwa hasil koreksi implementasi aplikasi penilaian jawaban esai online menggunakan algoritma Nazief dan Adriani dengan metode *cosine similarity* tidak berbeda jauh dibandingkan koreksi manual. Namun, pada penelitian ini masih ada beberapa kata yang memiliki arti yang sama tetapi terdeteksi sebagai jawaban salah [2]. Sedangkan hasil penelitian Taufiqurrahman Saleh Afif menunjukkan bahwa produk sangat sesuai untuk pengujian harian berdasarkan hasil validasi media yaitu 4,64 dengan range rating 4,2 sampai 5,0. Akan tetapi, proses penilaian masih berdasarkan kemiripan kata, belum bisa menemukan nilai dari struktur kalimat [3].

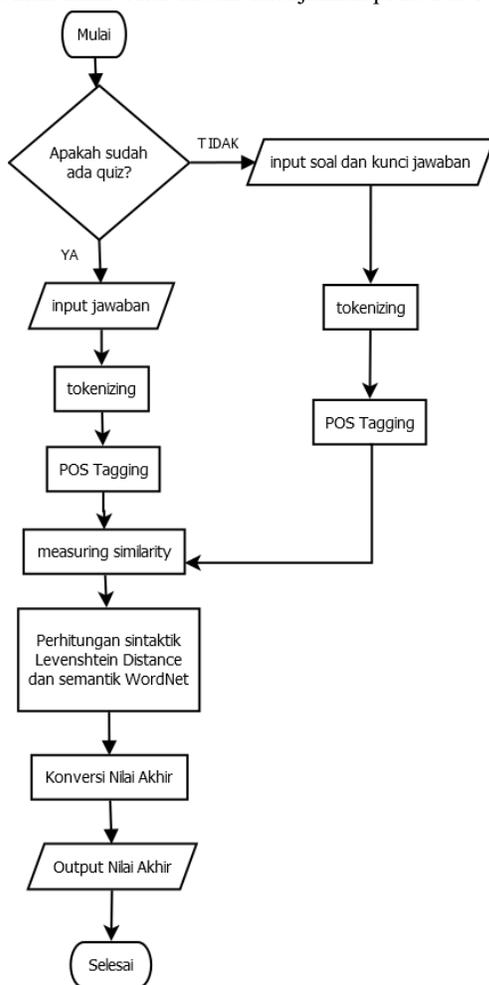
Metode *Syntactic-Semantic Similarity (SynSemSim)* merupakan metode pemecahan masalah penghitungan persamaan secara sintaktik berdasarkan kemiripan kata dan semantik berdasarkan arti atau makna sebuah kata dengan mengenali struktur kalimat terlebih dahulu. Deddy Winarsono, dkk telah melakukan penelitian mengenai sistem penilaian otomatis kemiripan kalimat menggunakan metode *syntactic-semantic similarity* yang berjudul “Sistem Penilaian Otomatis Kemiripan Kalimat Menggunakan *Syntactic-Semantic Similarity* Pada Sistem *E-Learning*”. Pada penelitian ini menggunakan jawaban bahasa Inggris, rata-rata deviasi metode *synsem* pada kasus terburuk adalah 2,61 dan deviasi pada

kasus terbaik adalah 0,19. Dibandingkan dengan hasil penelitian sebelumnya, penggunaan metode *synsemsim* untuk penelitian sistem penilaian otomatis lebih mendekati hasil penilaian manual [4].

Berdasarkan latar belakang permasalahan diatas, peneliti mencoba merancang sistem penilaian otomatis jawaban esai Bahasa Indonesia berdasarkan kemiripan kalimat dengan menggunakan metode *syntactic-semantic similarity*.

II. METODOLOGI PENELITIAN

Hasil dari penelitian ini berbentuk *website* yang dibangun menggunakan *framework django* dengan menerapkan metode *syntactic-semantic similarity*. Perhitungan secara sintaktik menggunakan algoritma *levenshtein distance* dan perhitungan secara semantik menggunakan *wordnet*. Keluaran sistem berupa nilai akhir. Alur sistem ditunjukkan pada Gbr.1



Gbr. 1 Alur sistem menggunakan metode *syntactic-semantic similarity*

sistem mendapatkan jawaban esai yang dimasukan oleh user. Jawaban melalui proses pertama yaitu *tokenizing*, *POS tagging*, *measuring similarity*, pengukuran kemiripan kalimat dengan algoritma *levenshtein distance* dan *wordnet*, dan

konversi nilai akhir. Nilai akhir dari proses tersebut disimpan ke dalam database pada sistem.

A. Automated essay scoring

Automated Essay Scoring (AES) adalah sistem penilaian jawaban esai otomatis yang membandingkan dua jawaban menggunakan metode yang ada. Metode yang digunakan untuk mengimplementasikan *AES* adalah *Corpus-Based Similarity (CBS)* dan *String-Based Similarity (SBS)*. *Corpus-Based Similarity* merupakan metode untuk mengukur kemiripan makna kata berdasarkan kamus yang tersedia. *String-Based Similarity* adalah metode untuk mengukur kemiripan string antara dua teks, terlepas dari arti teks tersebut [5].

Manfaat penggunaan sistem penilaian jawaban esai otomatis yaitu sistem dapat memeriksa jawaban esai lebih detail dibandingkan manusia, sistem lebih konsisten dalam melakukan proses penilaian, dan penilaian sistem bersifat objektif tidak terpengaruh apapun.

B. Syntactic-Semantic Similarity

Syntactic-Semantic Similarity (SynSemSim) merupakan metode untuk menyelesaikan permasalahan penghitungan persamaan secara sintaktik berdasarkan kemiripan kata dan semantik berdasarkan arti atau makna sebuah kata dengan mengenali struktur kalimat dan jenis kata (*part of speech*) terlebih dahulu. Teknik perhitungan secara sintaktik yaitu perhitungan kemiripan antar kata dengan menggunakan algoritma *levenshtein distance*. Sedangkan teknik perhitungan semantik yaitu perhitungan kemiripan arti tiap kata berdasarkan kalimat dengan menggunakan *WordNet*. Metode *synsemsim* terdiri dari empat bagian yaitu :

1. *Tokenizing*, pada tahap ini kalimat jawaban akan dipecah menjadi beberapa kata.
2. *Part of Speech (POS) Tagging*, jawaban yang sudah melalui proses *tokenizing* akan diberi *tag* atau label kata.
3. *Measuring Similarity*, jawaban akan dikelompokkan sesuai dengan hasil *POS tagging* dan parsing menjadi 5 jenis matriks yaitu *noun*, *prepositional*, *adverb*, *adjective*, dan *verb*.
4. *Scoring*, pada tahap ini setiap matriks dilakukan perhitungan secara sintatik dan semantik kemudian dilakukan proses perhitungan *fast heuristic* dan *matching average* untuk menentukan nilai *similarity*.
5. *Konversi Nilai Akhir*, nilai perhitungan *similarity* yang sudah didapatkan pada tahap sebelumnya dikonversikan menjadi nilai jawaban seperti yang ditunjukkan pada tabel 1 [6]. Nilai jawaban setiap soal akan dibagi dengan jumlah soal. Misal jika nilai *similarity* 1 dengan 10 soal, maka nilai akhir soal yang diperoleh yaitu $100/10 = 10$.

TABEL I
KONVERSI NILAI SIMILATIRY KE NILAI JAWABAN

Nilai Similarity	Nilai Jawaban
0,91 – 1	100
0,81 – 0,90	90
0,71 – 0,80	80
0,61 – 0,70	70
0,51 – 0,60	60
0,41 – 0,50	50
0,31 – 0,40	40
0,21 – 0,30	30
0,11 – 0,20	20
0,01 – 0,10	10

C. POS Tagging

Part-of-Speech Tagging atau POS Tagging adalah proses mendeteksi struktur kalimat dengan memberikan label kelas kata atau label tag. Pemberian label kelas kata bertujuan untuk mengelompokan setiap kata berdasarkan role dari masing-masing kata dalam suatu kalimat [7]. Tabel daftar tagset Bahasa Indonesia dapat dilihat pada tabel 2.

TABEL II
DAFTAR TAGSET BAHASA INDONESIA

No.	POS	Nama POS	Contoh
1	OP	Open Parenthesis	{
2	CP	Close Parenthesis	}
3	GM	Slash	/
4	JJ	Adjective	Kaya, Manis
5	RB	Adverb	Sementara, Nanti
6	NN	Common Noun	Sepeda
7	NNP	Proper Noun	Bali, Indonesia
8	NNG	Genitive Noun	Laptopnya
9	VBI	Intransitive Verb	Pergi
10	VBT	Transitive Verb	Memakan
11	IN	Prepositon	Dari, Di, Ke
12	MD	Modal	Dapat
13	CC	Coor-Conjunction	Dan, Atau, Tetapi
14	SC	Subor-Conjunction	Apabila, Ketika
15	DT	Determiner	Para, Ini, Itu
16	UH	Interjection	Wah, Aduh
17	CDO	Ordinal Numerals	Ketiga, Keempat
18	CDC	Collective Numerals	Berdua
19	CDP	Primary Numerals	Lima, Sepuluh
20	CDI	Irregular Numerals	Beberapa
21	PRP	Personal Pronouns	Aku, Mereka
22	WP	WH-Pronouns	Apakah
23	PRN	Number Pronouns	Berbeda-beda
24	PRL	Locative Pronouns	Sini, Situ, Sana
25	NEG	Negation	Bukan, Tidak
26	SYM	Symbols	@#\$%^
27	RP	Particles	Pun, Kah
28	FW	Foreign Words	Foreign, Word

D. Algoritma Levenshtein Distance

Levenshtein distance adalah proses penghitungan matriks yang digunakan untuk menghitung selisih antara dua string. Perhitungan levenshtein distance ditentukan oleh jumlah minimum operasi perubahan yang dihitung dengan menggunakan tabel. Distance adalah jumlah perubahan yang dibutuhkan untuk mengubah satu string ke string lainnya.

Terdapat tiga macam operasi pada levenshtein distance, antara lain (Pamungkas, 2016) [8]:

1. Operasi Penyisipan Karakter (Insertion)

Operasi penyisipan karakter berarti menambahkan karakter ke string. Contoh operasi penyisipan karakter pada Gbr. 2 dengan menggunakan string “naun” menjadi string “namun”, maka memasukkan karakter “m”. Penambahan karakter dapat memasukkan di awal, tengah atau akhir string.

String 1	n	a	m	u	n
String 2	n	a	-	u	n
Insertion			m		

Gbr. 2 Contoh Operasi Penyisipan Karakter
Levenshtein Distance

2. Operasi Penghapusan Karakter (Deletion)

Operasi penghapusan karakter untuk menghapus karakter tertentu dari string. Contoh operasi penghapusan karakter pada Gbr. 3 dengan menggunakan string “tapin” karakter terakhir dihapus sehingga menjadi string “tapi”. Dalam operasi ini, melakukan penghapusan karakter “n”.

String 1	t	a	p	i	-
String 2	t	a	p	i	n
Deletion					n

Gbr. 3 Contoh Operasi Penghapusan Karakter
Levenshtein Distance

3. Operasi Penukaran Karakter (Substitution)

Operasi penukaran karakter adalah operasi untuk menukar beberapa karakter dengan karakter lain. Contoh operasi penukaran karakter pada Gbr. 4 dengan menggunakan string “aklir” menjadi “akhir”. Pada permasalahan ini karakter “l” yang terdapat dibagian tengah string, dilakukan penukaran karakter dengan huruf “h”.

String 1	a	k	h	i	r
String 2	a	k	l	i	r
Substitution			h		

Gbr. 4 Contoh Operasi Penukaran Karakter
Levenshtein Distance

Rumus operasi penghapusan, penyisipan, dan penukaran karakter yang digunakan untuk mendapatkan nilai distance adalah sebagai berikut:

$$1) \text{ Penghapusan} \\ D(s, t) = \min D(s - 1, t) + 1 \quad (1)$$

$$2) \text{ Penyisipan} \\ D(s, t) = \min D(s, t - 1) + 1 \quad (2)$$

- 3) Penukaran
 $D(s, t) = \min D(s - 1, t - 1) + 1, sj \neq ti$ (3)
- 4) Tidak ada perubahan
 $D(s, t) = \min D(s - 1, t - 1), sj = ti$ (4)

Keterangan :

s = string sumber

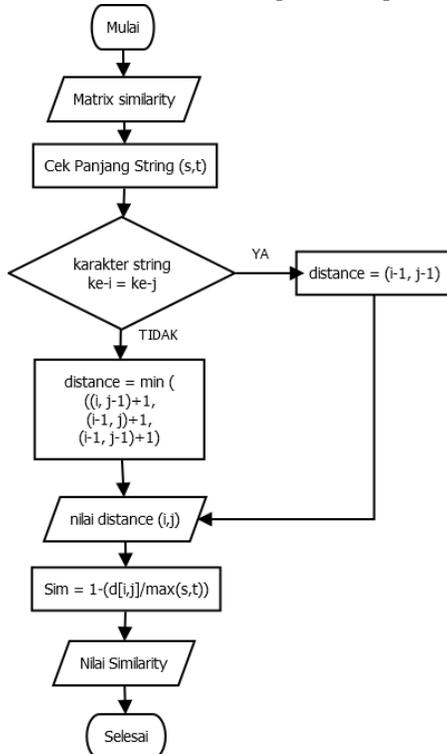
t = string target

D = jarak Levenshtein Distance

s(j) = karakter string sumber ke-j

t(i) = karakter string target ke-i

Alur proses perhitungan sintaktik menggunakan algoritma levenshtein distance dapat dilihat pada Gbr. 5.



Gbr. 5 Alur Algoritma Levenshtein Distance

Gbr. 6 merupakan contoh implementasi perhitungan levenshtein distance untuk mendapatkan nilai distance pada kata “benyam” dan “ephrem” sebagai berikut :

		b	e	n	y	a	m
	0	1	2	3	4	5	6
e	1	1	1	2	3	4	5
p	2	2	2	2	3	4	5
h	3	3	3	3	3	4	5
r	4	4	4	4	4	4	5
e	5	5	4	5	5	5	5
m	6	6	5	5	6	6	5

Gbr. 6 Contoh Implementasi Perhitungan Levenshtein Distance Berdasarkan perhitungan nilai distance pada Gbr. 6, maka nilai distance dari kata “benyam” dan “ephrem” adalah 5.

Perhitungan *similarity levenshtein distance* dilakukan setelah didapatkan nilai jarak dari kedua teks yang akan dibandingkan. Kemudian gunakan rumus berikut untuk menghitung persamaan :

$$Sim = 1 - \frac{d[m,n]}{Max(s,t)} \quad (5)$$

Keterangan :

Sim = similarity

d [m,n] = nilai distance atau jarak yang terletak pada baris ke-m dan kolom ke-n

Max(s,t)= panjang string terbesar antara string awal (s) dan string target (t)

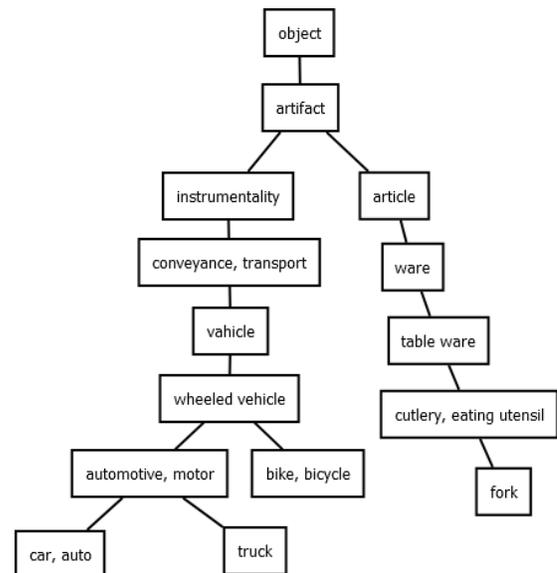
Contoh implementasi rumus perhitungan *similarity levenshtein distance* untuk mencari nilai kemiripan pada kata “benyam” dan “ephrem” sebagai berikut :

$$Sim = 1 - \frac{5}{6} = \frac{1}{6} = 0,1667 \quad (6)$$

E. WordNet

WordNet merupakan database yang terinspirasi teori psikolinguistik mengenai leksikal memori pada manusia [9]. Berbeda dengan kamus yang fokus pada kata, wordnet lebih fokus kepada makna kata. Setiap makna kata dalam wordnet dikelompokan secara terorganisir disebut dengan *synset*. Tiap *synset* berisi daftar sinonim kata dan petunjuk yang mendeskripsikan relasi antar *synset* satu dengan yang lain.

WordNet menyediakan singkatan, definisi umum, dan menyimpan hubungan semantik antar set dengan kumpulan sinonim untuk menghubungkan kamus dan tesaurus. WordNet yang saat ini sudah memiliki *synset* paling banyak adalah Princeton WordNet. Selain itu, Princeton WordNet juga dapat digunakan sebagai acuan untuk membuat wordnet di bahasa lain.



Gbr. 7 Contoh Hirarki Hiponim Pada WordNet

Pada Gbr. 7 panjang *path* antar node dihitung sebagai berikut :

1. *Car* dengan *auto* adalah 1, LCS adalah “*car, auto*”.
2. *Car* dengan *fork* adalah 12, LCS adalah “*artifact*”.

LCS atau *The Least Common Subsumer* dari dua *synsets* adalah node paling dekat dari dua *synset*. Jalan untuk menghitung nilai relasi antar dua kata yang memiliki *senses* sama disebut panjang *path*. Setelah mendapatkan panjang *path* antar dua *synset* dan LCS maka dilakukan perhitungan menggunakan rumus *Wu & Palmer* untuk mengetahui nilai relasi antara dua *synset*.

$$Wu \& Palmer (X, Y) = \frac{2 \times dept(LCS)}{dept(X) + dept(Y)} \quad (7)$$

Keterangan :

X = dokumen

Y = kata kunci

depth = Kedalaman yang diukur dari root

depth (LCS) = Panjang *path* dari konsep umum dari X dan Y ke root

depth X = Panjang *path* X

depth Y = Panjang *path* Y

Contoh implementasi rumus *Wu & Palmer* pada perhitungan nilai relasi antara *car* dengan *bike* :

$$Wu \& Palmer (Car, Bike) = \frac{2 \times 2}{3 + 2} = 0,8 \quad (8)$$

III. HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Sistem penilaian otomatis jawaban esai dapat digunakan sebagai perangkat kuis atau ujian untuk membantu proses penilaian. Berikut adalah tampilan dari sistem penilaian otomatis jawaban esai :

1. Tampilan Halaman *Login*

Gbr. 8 Tampilan Halaman *Login Sistem*

Pada Gbr. 8 menunjukkan tampilan halaman *login* sistem. Halaman ini berisikan *form login* yang terdiri dari *username* dan *password* serta terdapat tombol *button login* untuk masuk ke dalam sistem, *button daftar sebagai mahasiswa* untuk membuat akun baru sebagai mahasiswa, dan *button daftar sebagai dosen* untuk membuat akun baru sebagai dosen.

2. Tampilan Halaman Tambah *Quiz*

Gbr. 9 Tampilan Halaman Tambah *Quiz*

Gbr. 9 merupakan tampilan halaman untuk menambahkan ujian yang hanya bisa dilakukan oleh dosen. Pada halaman ini berisi *form* nama ujian dan jumlah soal yang akan ditambahkan dalam ujian.

3. Tampilan Halaman Tambah Soal dan Kunci Jawaban

Gbr. 10 Tampilan Halaman Tambah Soal dan Kunci Jawaban

Gbr. 10 merupakan tampilan halaman untuk menambahkan soal dan kunci jawaban ke dalam ujian yang telah dibuat. Pada halaman ini terdapat *form* soal dan kunci jawaban.

4. Tampilan Halaman Detail *Quiz*

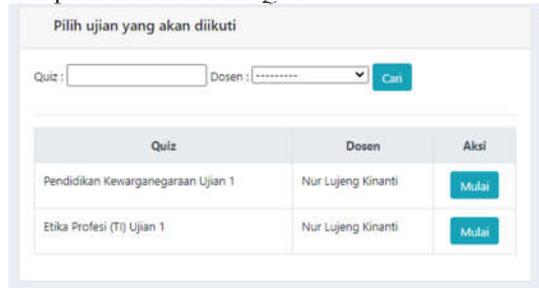
#	Soal	Jawaban	Aksi
1	Jelaskan misi Pendidikan Kewarganegaraan di perguruan tinggi	Untuk membantu mahasiswa memantapkan pribadinya sehingga konsisten mampu mewujudkan nilai-nilai dasar pancasila, rasa kebangsaan dan cinta tanah air dalam menguasai, menerapkan dan mengembangkan ilmu pengetahuan, teknologi dan seni dengan rasa tanggung jawab dan bemosor.	
2	Apa saja landasan ilmiah Pendidikan Kewarganegaraan, sebutkan dan jelaskan!	a. Dasar Pemikiran Pendidikan Kewarganegaraan Setiap warga negara dituntut untuk dapat hidup berguna dan bermakna bagi bangsa dan negaranya, serta mampu mengantisipasi perkembangan dan perubahan masa depannya. Untuk itu itu diperlukan penguasaan ilmu	

Gbr. 11 Tampilan Halaman Detail *Quiz*

Gbr. 11 merupakan tampilan halaman detail *quiz*. Pada halaman ini terdapat tabel yang berisi soal dan

kunci jawaban *quiz* yang telah ditambahkan serta *button* untuk menghapus atau mengubah soal.

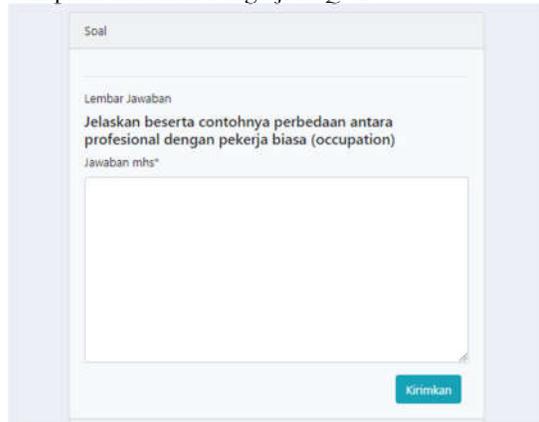
5. Tampilan Halaman Pilih *Quiz*



Gbr. 12 Tampilan Halaman Pilih Quiz

Gbr. 12 merupakan tampilan halaman pilih *quiz* yang menampilkan *quiz* yang tersedia. Pada halaman ini terdapat *form* pencarian berdasarkan nama *quiz* dan nama dosen. Pada *form* terdapat tabel berisikan informasi nama *quiz*, nama dosen, dan tombol *button* mulai untuk memulai *quiz*.

6. Tampilan Halaman Pengerjaan *Quiz*



Gbr. 13 Tampilan Halaman Pengerjaan

Gbr. 13 merupakan tampilan halaman pengerjaan *quiz*. Pada halaman ini terdapat *form* jawaban mahasiswa untuk memasukan jawaban dan *button* kirimkan untuk melanjutkan ke proses perhitungan.

7. Tampilan Halaman Hasil *Quiz*

Soal yang telah dijawab

Nilai Ujian : 82.0			
Soal	Jawaban	Similarity	Nilai
Berikan titik kelemahan dari Kode Etik Profesi	A. Idealisme yang terkandung dalam kode etik profesi adakalanya tidak sejalan dengan fakta yang terjadi di sekitar para profesional. Hal ini cukup menggelfik para profesional untuk berpaling pada kenyataan dan mengabaikan idealisme kode etik profesi. Kode etik profesi tidak lebih dari pajangan tullean belaka. E. Kode etik profesi merupakan himpunan norma moral yang tidak dilengkapi dengan sanksi keras karena keberluannya semata-mata berdasarkan kesadaran profesional, hal inilah yang memberi peluang kepada profesional yang lemah iman untuk berbuat menyimpang dari kode etik profesinya.	0.9	15.0
Jelaskan beserta contohnya perbedaan antara profesional	Pekerja Biasa (Occupation) merupakan suatu kegiatan yang tidak bergantung pada suatu keahlian tertentu. Contoh : Operator, penjaga warnet, tukang ketik di rental, Teknisi Komputer, dll. Profesional adalah orang yang sangat ahli dalam suatu bidang tertentu. Contoh : IT manager, IT consultant, Sistem analyst, consultant ERP (Enterprise Resource Planning).	0.55	10.0

Gbr. 14 Tampilan Halaman Login Sistem

Gbr. 14 merupakan tampilan halaman hasil *quiz*. Pada halaman ini terdapat nilai hasil perhitungan sistem dan tabel yang terdiri dari kolom soal *quiz*, jawaban mahasiswa, dan nilai tiap soal.

B. Pengujian Metode Syntatic-Semantic Similarity

Pengujian sistem dilakukan kepada 30 mahasiswa dengan menerapkan 2 materi yang berbeda yaitu pendidikan kewarganegaraan dan etika profesi informatika. Pengujian dilaksanakan dengan menggunakan 60 data jawaban mahasiswa dan kunci jawaban quiz. Gbr. 15 merupakan salah satu contoh perbandingan hasil penilaian. Pada jawaban yang dimasukan oleh mahasiswa memiliki penggunaan kata yang berbeda sehingga menghasilkan nilai similarity 0,82 dan nilai jawaban 90 dari 10 soal. Hal ini menunjukkan bahwa sistem mampu melakukan penilaian otomatis jawaban esai pada kalimat yang memiliki susunan kata yang berbeda sebagai jawaban benar.

Soal :

“Apa saja landasan ilmiah Pendidikan Kewarganegaraan, sebutkan dan jelaskan!”

Kunci Jawaban :

“Warga negara dituntut hidup berguna dan bermakna bagi negara dan bangsanya, mampu mengantisipasi perkembangan serta perubahan masa depan. Diperlukan bekal ilmu pengetahuan teknologi dan seni yang berlandaskan nilai-nilai keagamaan, nilai-nilai moral, dan nilai-nilai budaya bangsa.”

Jawaban :

“Warga negara berkewajiban menjalani kehidupan yang berguna dan bermakna bagi negara dan bangsanya serta mengantisipasi peristiwa dan perubahan di masa depan. Diperlukan pembekalan pengetahuan teknis dan artistik berdasarkan nilai-nilai agama, moral dan budaya bangsa.”

Nilai Similarity : 0,82

Nilai Jawaban : 90

Gbr. 15 Contoh Hasil Penilaian

Pada tabel III menampilkan 10 sampel dari data perbandingan penilaian jawaban menggunakan sistem dengan penilaian manual.

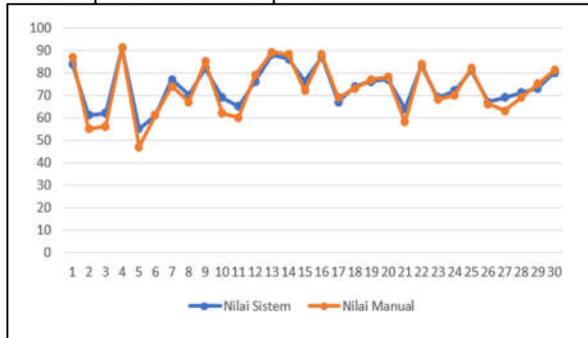
TABEL III
SAMPEL PERBANDINGAN HASIL PENILAIAN SISTEM DAN PENILAIAN MANUAL

Penilaian Sistem	Penilaian Manual
77	75
70	68
82	81
65	63
76	76
72	76
82	85
100	100
68	66
100	100

Penelitian yang dilakukan berfokus pada perbandingan hasil rata-rata nilai menggunakan sistem dan

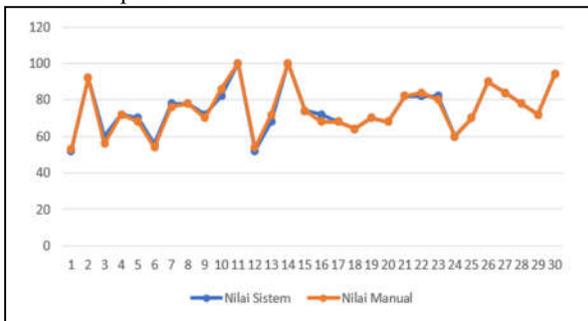
manual. Langkah pertama melakukan uji coba penilaian menggunakan sistem untuk mendapatkan hasil nilai kemudian dilakukan penilaian secara manual oleh pembuat soal. Langkah selanjutnya dilakukan perhitungan untuk mencari rata-rata hasil penilaian sistem dan rata-rata hasil penilaian manual, kemudian akan dihitung selisih rata-rata yang diperoleh.

Gbr. 16 menunjukkan grafik dari hasil perbandingan antara penilaian menggunakan sistem (garis warna biru) dengan penilaian manual (garis warna merah) pada 30 data penilaian *quiz* pendidikan kewarganegaraan. Dapat dilihat dari grafik bahwa garis biru hampir mendekati pola garis merah. Hal ini menunjukkan bahwa hasil penilaian jawaban dengan sistem mendekati penilaian manual dengan selisih terbesar pada skor akhir 8 poin.



Gbr. 16 Grafik Perbandingan Nilai Sistem dengan Nilai Manual Pada Materi Pendidikan Kewarganegaraan

Gbr. 17 menunjukkan grafik dari hasil perbandingan antara penilaian menggunakan sistem (garis warna biru) dengan penilaian manual (garis warna merah) pada 30 data penilaian *quiz* etika profesi. Dapat dilihat dari grafik bahwa garis biru hampir mendekati pola garis merah. Hal ini menunjukkan bahwa hasil penilaian jawaban dengan sistem mendekati penilaian manual dengan selisih terbesar pada skor akhir poin.



Gbr. 17 Grafik Perbandingan Nilai Sistem dengan Nilai Manual Pada Materi Etika Profesi Informatika

Pada tabel IV menampilkan data hasil implementasi sistem penilaian esai otomatis pada materi pendidikan kewarganegaraan menghasilkan selisih rata-rata nilai dari perbandingan penilaian menggunakan sistem dengan penilaian manual sebesar 1,3 dan pada materi etika profesi sebesar 0,16.

TABEL IV
HASIL IMPLEMENTASI SISTEM PENILAIAN ESAI OTOMATIS

No.	Materi	Rata-Rata Nilai Sistem	Rata-Rata Nilai Manual	Selisih Rata-Rata Nilai
1	Pendidikan Kewarganegaraan	73,77	72,47	1,3
2	Etika Profesi Informatika	74,73	74,57	0,16

Berdasarkan pada grafik dan data tabel memperoleh hasil perbandingan antara penilaian jawaban esai menggunakan sistem dengan penilaian manual oleh manusia yaitu rata-rata hasil kesesuaian 70,44%. Hal ini karena kurang tepatnya pemberian label tag pada kata sehingga terdapat jawaban yang salah pada penilaian manual tetap mendapat poin tinggi pada sistem dan dianggap sebagai jawaban benar.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan bahwa metode *syntactic-semantic similarity* yang telah diuji coba pada 2 materi yang berbeda yaitu 10 soal pendidikan kewarganegaraan dan 5 soal etika profesi dengan 60 data jawaban menunjukkan hasil pengujian sistem penilaian otomatis jawaban esai Bahasa Indonesia dibandingkan dengan hasil penilaian manual yaitu rata-rata kemiripan 70,44% . Metode *syntactic-semantic similarity* menghasilkan nilai *similarity* yang tinggi, terutama pada kalimat jawaban yang salah tetap mendapatkan nilai kemiripan rentang angka 0,1 sampai 0,4. Hal ini dikarenakan perbandingan struktur kalimat belum optimal dan kurang lengkapnya *thesaurus* untuk Bahasa Indonesia sehingga mempengaruhi hasil penilaian semantik.

UCAPAN TERIMA KASIH

Puji serta syukur penulis panjatkan kepada Allah SWT yang telah memberi kelancaran dan kemudahan dalam menyelesaikan penelitian ini. Dalam proses penyelesaian penelitian ini tidak lepas dari bimbingan, bantuan, serta dukungan dari berbagai pihak, untuk itu saya ucapkan terima kasih sebesar-besarnya. Semoga semua jasa dan amal kebaikan yang diberikan diberkati dan lindungi Allah SWT dan dapatkan kelimpahan rahmat-Nya.

REFERENSI

- [1] F.N. Sari, M. Harlanu, dan D.A. Widodo, "Implementasi Algoritma *Winnowing* Pada Platform *Elearning* Untuk Sistem Penilaian Otomatis Jawaban *Essay*," *Edu Komputika* 3, vol. 2, 2016.
- [2] M. Agus Salim, "Pengembangan Aplikasi Penilaian Ujian *Essay* Berbasis *Online* Menggunakan Algoritma Nazief Dan Adriani Dengan Metode *Cosine Similarity*," *Jurnal IT-EDU*, vol. 02, hal. 126-135, 2017.
- [3] Taufiqurrahman Saleh Afif, "Pengembangan Sistem Penilaian Otomatis untuk Pelajaran Kelas X di SMK Negeri 3 Buduran Sidoarjo," *Jurnal IT-EDU*, vol. 03, hal. 210 – 217, 2018.
- [4] D. Winarsono, D.O. Siahaan, dan U. Yuhana, "Sistem Penilaian Otomatis Kemiripan Kalimat Menggunakan *Syntactic-Semantic Similarity* Pada Sistem *E-Learning*," *KURSOR*, vol. 5, hal. 75-82, 2009.
- [5] E.S. Pramukantoro, "Sistem Penilaian Otomatis Jawaban Esai Pada *Elearning* Belajardisini.Com," *JTIK*, vol. 3, hal. 248-252, 2016.

- [6] Rahimi Fitri, A.N. Asyikin, "Aplikasi Penilaian Ujian *Essay* Otomatis Menggunakan Metode *Cosine Similarity*," Jurnal Poros Teknik, vol. 07, 2015.
- [7] V. Christanti, Ir. J. Pragantha, dan Victor, "*Part-Of-Speech Tagging* Untuk Bahasa Indonesia Menggunakan *Stanford PosTagging*," A30, 2016, Seminar Nasional Teknologi Informasi.
- [8] B.P. Pratama, S.A. Pamungkas, "Analisis Kinerja Algoritma *Levenshtein Distance* dalam Mendeteksi Kemiripan Dokumen Teks," Jurnal Logika, vol. 02, hal. 131-143, 2016.
- [9] M.I. Pribadi, K.K. Purnamasari, "Pendeteksi Relasi Antar Makna Pada *WordNet* Bahasa Indonesia," ISBN:978-602-60250-1-2), 2017, Seminar Nasional Komputer dan Informatika.