

PENERAPAN ALGORITMA LEVENSHTTEIN DISTANCE, N-GRAM DAN SYNONYM REPLACEMENT PADA SISTEM PENILAIAN ESAI OTOMATIS DI UNIVERSITAS BUDI LUHUR

Muhammad Akbar Bagaskoro¹, Mufti^{2*}

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹1911502324@student.budiluhur.ac.id, ^{2*}mufti@budiluhur.ac.id
(* : corresponding author)

Abstrak-Ujian merupakan salah satu cara bagi para pengajar untuk menilai kemampuan peserta didiknya. Universitas Budi Luhur telah menerapkan ujian secara online menggunakan sistem e-learning. Sistem e-learning ini sangat memudahkan para pengajar dalam memberikan nilai ujian secara otomatis, tetapi hanya untuk ujian berjenis pilihan ganda dan belum memfasilitasi ujian berjenis isian atau esai. Berdasarkan latar belakang tersebut, penulis mencoba menerapkan algoritma Levenshtein Distance, N-Gram, dan Synonym Replacement pada pembuatan sistem penilaian esai otomatis dengan harapan sistem ini dapat memudahkan para pengajar dalam memberikan nilai, yang sebelumnya harus dilakukan secara manual dan memakan banyak waktu. Dalam penerapannya, penulis mengombinasikan algoritma Levenshtein Distance dengan N-Gram dan melakukan pengujian sebanyak empat kali, yaitu dengan N-Gram 1 (Unigram), N-Gram 2 (Bigram), N-Gram 3 (Trigram), dan N-Gram 4 (Quadrigram). Dataset yang digunakan dalam penelitian ini terdiri dari delapan soal dan 284 record jawaban yang diambil dari hasil ujian 71 mahasiswa pada matakuliah Bahasa Indonesia. Hasil dari pengujian menunjukkan bahwa sistem penilaian esai otomatis dapat menghasilkan akurasi tertinggi sebesar 81,7% pada penggunaan N-Gram 1.

Kata Kunci: Levenshtein Distance, N-Gram, Synonym Replacement, Penilaian, Esai

APPLICATION OF LEVENSHTTEIN DISTANCE ALGORITHM, N-GRAM AND SYNONYM REPLACEMENT IN THE AUTOMATIC ESSAY ASSESSMENT SYSTEM AT BUDI LUHUR UNIVERSITY

Abstract- Exams are one of the ways for educators to assess their students' abilities. Budi Luhur University has implemented online exams using an e-learning system. This e-learning system greatly facilitates educators in automatically grading multiple-choice exams, but it does not yet support fill-in-the-blank or essay-type exams. Based on this background, the author attempted to apply the Levenshtein Distance algorithm, N-Gram, and Synonym Replacement in creating an automated essay scoring system, with the hope that this system would make it easier for educators to grade essays, which previously had to be done manually and was time-consuming. In its implementation, the author combined the Levenshtein Distance algorithm with N-Gram and conducted testing four times, using N-Gram 1 (Unigram), N-Gram 2 (Bigram), N-Gram 3 (Trigram), and N-Gram 4 (Quadrigram). The dataset used in this research consisted of eight questions and 284 answer records taken from exams of 71 students in the Indonesian Language course. The results of the testing showed that the automated essay scoring system achieved the highest accuracy of 81.7% with the use of N-Gram 1.

Keywords: Levenshtein Distance, N-Gram, Synonym Replacement, Assessment, Essay

1. PENDAHULUAN

Ujian adalah metode untuk mengevaluasi pemahaman siswa [1]. Ujian dapat dilakukan secara offline menggunakan kertas atau online melalui e-learning [2]. Dalam e-learning, umumnya hanya soal pilihan ganda dan isian singkat yang dinilai otomatis, sedangkan soal esai masih dinilai secara manual, yang memakan waktu dan sering subjektif [3].

Universitas Budi Luhur menggunakan sistem e-learning namun hanya otomatis untuk soal pilihan ganda. Penilaian esai yang dilakukan manual memerlukan waktu lama dan tidak efisien [4].

Metode yang dapat digunakan untuk penilaian esai otomatis adalah algoritma Levenshtein Distance, yang mengukur jarak antar string berdasarkan perubahan karakter yang diperlukan [5]. Penelitian ini juga menggunakan Synonym Replacement dan N-Gram untuk meningkatkan akurasi [6].

Penelitian ini bertujuan untuk mengembangkan sistem penilaian esai otomatis berbasis website dengan algoritma Levenshtein Distance untuk mempermudah proses penilaian di Universitas Budi Luhur, serta

meningkatkan akurasi dan efisiensi dibandingkan penilaian manual. Penelitian ini berbeda dari studi sebelumnya yang fokus pada penggunaan Levenshtein Distance untuk kesalahan ejaan dan pencarian teks [7].

2. METODE PENELITIAN

2.1 Soal Esai

Soal esai adalah soal yang jawabannya diberikan dalam bentuk uraian atau cerita. Jenis pertanyaan yang biasanya digunakan untuk mengawali soal ini adalah jelaskan, bandingkan, uraikan, terangkan, dan sebagainya [8].

2.2 Ujian Esai

Salah satu metode pengukuran yang digunakan adalah ujian esai, yang memungkinkan siswa untuk menyampaikan dan menyatukan jawaban mereka dalam bentuk kata-kata mereka sendiri. Sebagaimana dinyatakan oleh Djiwandono (2008), jenis ujian esai mengacu pada hasilnya, dan harus berupa esai yang ditulis dalam berbagai gaya, termasuk deskriptif dan argumentatif, sesuai dengan topik bahasan [9].

2.3 Penilaian

Penilaian (grading) adalah proses penyematan atribut atau dimensi atau kuantitas (berupa angka/huruf) terhadap hasil asesmen dengan cara membandingkannya terhadap suatu instrumen standar tertentu. Hasil dari penilaian berupa atribut/dimensi/kuantitas tersebut digunakan sebagai bahan evaluasi [10].

2.4 Text Preprocessing

Text preprocessing merupakan langkah yang mengelola teks tidak beraturan atau tidak terstruktur agar lebih terstruktur sehingga teks tersebut dapat diolah kembali untuk memperoleh informasi dan pola teks dalam suatu dokumen tertentu [11]. Adapun tahapan *text preprocessing*, sebagai berikut:

- a. *HTML Tag Removal* adalah proses yang bertujuan untuk menghilangkan semua tag *HTML*.
- b. *Space Removal* adalah proses yang bertujuan untuk menghapus spasi yang lebih dari satu menjadi satu spasi.
- c. *Punctuation Removal* adalah proses penghilangan tanda baca dalam sebuah teks.
- d. *Case Folding* adalah proses yang mengubah semua huruf besar menjadi huruf kecil.
- e. *Stemming* adalah proses mereduksi suatu kata menjadi bentuk dasar atau akar kata dengan memotong bagian akhir atau awal kata.
- f. *Stopword Removal* merupakan proses menghilangkan kata-kata tidak deskriptif. Contoh stopword adalah “di”, “ke”, “dari”, “dan” dan sebagainya. Stopword Removal merupakan proses menghilangkan kata-kata yang tidak relevan pada teks.
- g. *Tokenizing* merupakan proses memecah text menjadi potongan-potongan kecil yang disebut token

2.5 Synonym Replacement

Synonym Replacement adalah teknik yang mengganti sejumlah kata tertentu dalam teks dengan sinonimnya. Daftar sinonim yang tersedia diambil dari database tesaurus Tematis Bahasa Indonesia oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia [12].

2.6 N-Gram

N-gram adalah model probabilistik yang dirancang untuk memprediksi huruf atau kata berikutnya dalam sekumpulan text. Dalam prosesnya, n-gram melakukan pemotongan huruf atau kata sepanjang nilai n. Berdasarkan jenis potongannya, n-gram dibagi menjadi n-gram perhuruf dan n-gram perkata. Pada penelitian ini penulis menggunakan jenis n-gram perkata [13].

2.7 Algoritma Levenshtein Distance

Levenshtein Distance adalah sebuah matriks yang mengukur jarak (perubahan) antara dua string. Nilai jarak diambil dari jumlah minimum perubahan yang diperlukan oleh suatu string untuk menjadi string lainnya. Perubahan yang dimaksud terdiri dari proses penyisipan (insertion), penghapusan (deletion), atau penukaran (substitution) [14].

2.8 Rancangan Pengujian

Pengujian yang dilakukan dalam penelitian ini dimulai dengan melakukan simulasi ujian dari 71 mahasiswa. Setelah semua ujian mahasiswa disimulasikan, nilai yang dihasilkan oleh sistem akan dibandingkan dengan nilai yang diberikan secara manual oleh dosen. Dari hasil perbandingan ini, akan dihitung berapa jumlah nilai yang

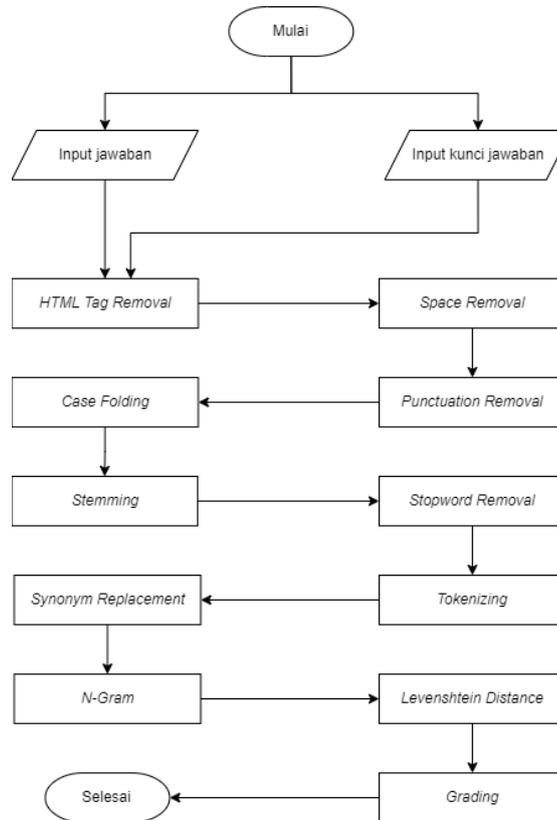
sesuai antara sistem dan penilaian manual dosen, dan kemudian dibuat presentasi akurasi menggunakan rumus 3 [15]

$$\text{Akurasi} = \frac{\text{Jumlah Identifikasi Benar}}{\text{Jumlah Data}} \times 100\% \quad (3)$$

3. HASIL DAN PEMBAHASAN

3.1 Penerapan Metode

Dalam mengimplementasikan algoritma Levenshtein Distance, N-Gram, dan Synonym Replacement, penulis merancang proses atau langkah-langkah yang diterapkan dalam sistem penilaian esai otomatis. Detail dari rancangan proses tersebut dapat dilihat pada Gambar 1.



Gambar 1. Penerapan Metode

3.1.1 Input Data

Tahapan pertama adalah memasukan data kunci jawaban dan jawaban mahasiswa menggunakan form input yang tersedia pada sistem. Kedua inputan data dapat dilihat pada Tabel 1.

Tabel 1. Inputan Data Kunci Jawaban dan Jawaban Mahasiswa

Kunci Jawaban	Jawaban Mahasiswa
<p>Pembelajaran ini diharapkan mereka memiliki perhatian, kepekaan dan menghargai Bahasa dan sastra yang memperkaya kerohanian dan mentalitas</p>	<p>Diharapkan pembelajaran ini dapat meningkatkan perhatian, kepekaan, dan penghargaan mereka terhadap Bahasa dan sastra, yang akan memperkaya kejiwaan dan mentalitas mereka</p>

3.1.2 Preprocessing

Tahapan *Preprocessing* yaitu terdiri dari proses *HTML Tag Removal*, *Space Removal*, *Punctuation Removal*, *Stemming*, *Stopword Removal*, dan *Tokenizing*. Proses dari tahapan *Preprocessing* dapat dilihat pada tabel 2

Tabel 2. Proses dari tahapan *preprocessing*

Ket	<i>HTML Tag Removal</i>	<i>Space Removal</i>	<i>Punctuation Removal</i>	<i>Case Folding</i>	<i>Stemming</i>	<i>Stopword Removal</i>	<i>Tokenizing</i>
Kunci Jawaban	Pembelajaran ini diharapkan mereka memiliki perhatian, kepekaan dan menghargai. ...dst	Pembelajaran ini diharapkan mereka memiliki perhatian, kepekaan dan menghargai. ...dst	Pembelajaran ini diharapkan mereka memiliki perhatian, kepekaan dan menghargai. ...dst	pembelajaran ini diharapkan mereka memiliki perhatian, kepekaan dan menghargai. ...dst	ajar ini harap mereka milik perhatian, peka dan harga bahasa ...dst	ajar harap milik perhatian, peka harga ...dst	["ajar", "harap", "milik", "perhatian", "peka", "harga"]
Jawaban	Diharapkan pembelajaran ini dapat meningkatkan perhatian, kepekaan, dan penghargaan. ...dst	Diharapkan pembelajaran ini dapat meningkatkan perhatian, kepekaan, dan penghargaan. ...dst	Diharapkan pembelajaran ini dapat meningkatkan perhatian, kepekaan, dan penghargaan. ...dst	diharapkan pembelajaran ini dapat meningkatkan perhatian, kepekaan, dan penghargaan mereka. ...dst	harap ajar ini dapat tingkat perhatian, kepekaan, dan harga, dan ...dst	harap ajar tingkat perhatian, kepekaan, harga hadap ...dst	["harap", "ajar", "tingkat", "perhatian", "kepekaan", "harga", "hadap"]

3.1.3 Synonym Replacement

Tahapan ketiga setelah proses *preprocessing* adalah proses *Synonym Replacement*. Pada tahap ini, setiap kata dalam jawaban mahasiswa yang bersinonim dengan kata-kata di kunci jawaban akan diganti dengan kata yang ada dalam kunci jawaban. Bisa kita lihat pada Tabel 3 kolom Jawaban Mahasiswa, terdapat kata yang diberi warna kuning artinya memiliki kesamaan arti dengan salah satu kata pada kolom Kunci Jawaban yang diberi warna biru. Lalu kata tersebut diubah menjadi kata yang ada pada kolom Kunci Jawaban.

Tabel 3. *Synonym Replacement*

Token Kunci Jawaban	Token Jawaban Mahasiswa	Hasil Akhir
[..., "sastra", "kaya", "rohani", "mentalitas"]	[..., "harga", "hadap", "bahasa", "sastra", "kaya", "jiwa", "mentalitas"]	[..., "harga", "hadap", "bahasa", "sastra", "kaya", "rohani", "mentalitas"]

3.1.4 N-Gram

Tahapan keempat setelah proses *Synonym Replacement* adalah pembuatan *N-Gram* untuk kunci jawaban dan juga jawaban mahasiswa. Pada tahap ini, setiap token akan ditambahkan dengan token didepannya sejumlah *N* yang digunakan. Pada sub bab ini penulis mencontohkan dengan menggunakan *N* berjumlah dua atau *Bigram*. Hasil proses *N-Gram* bisa dilihat pada Tabel 4

Tabel 4. Hasil *N-Gram* Jawaban Mahasiswa

Jawaban Mahasiswa	Hasil Akhir
["harap", "ajar", "tingkat", "perhatian", "kepekaan", "harga", "hadap", "bahasa", "sastra", "kaya", "rohani", "mentalitas"]	["harapajar", "ajarharga", "hargaperhatian", "perhatian,kepekaan", "kepekaan,harga", "hargahadap", "hadapbahasa", "bahasasastra", "sastra,kaya", "kayarohani", "rohanimentalitas"]

3.1.5 Levenshtein Distance

Tahapan kelima setelah proses pembuatan *N-Gram* adalah penerapan Levenshtein Distance. Penerapan algoritma Levenshtein Distance yang terdiri dari beberapa tahap yaitu:

1. Menghitung Jarak Antar Kata

Pada tahap ini, setiap kata dalam jawaban mahasiswa dibandingkan dengan kata dalam kunci jawaban menggunakan matriks dua dimensi. Nilai perubahan untuk setiap perbandingan dihitung dengan matriks Levenshtein Distance. Misalnya, perubahan dari kata "ajar" ke "harap" adalah 3, dibagi dengan panjang kata terpanjang ("harap"), menghasilkan nilai 0,75.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		harap	ajar	harga	perhati	peka	harga	hadap	bahasa	sastra	kaya	rohani	mentalitas
2	ajar	0,75	0	0,75	0,83	1	0,75	0,75	0,8	0,8	1	0,8	0,89
3	harap	0	0,75	0,5	0,67	0,75	0,5	0,25	0,6	0,6	0,5	0,8	0,78
4	milik	1	1	1	1	1	1	1	1	1	1	1	0,78
5	perhati	0,67	0,83	0,67	0	0,67	0,67	0,83	0,67	1	0,83	0,5	0,67
6	peka	0,75	1	0,75	0,67	0	0,75	0,75	0,8	0,8	0,67	0,8	0,78
7	harga	0,5	0,75	0	0,67	0,75	0	0,75	0,6	0,6	0,5	1	0,78
8	bahasa	0,6	0,8	0,6	0,67	0,8	0,6	0,6	0	0,6	0,6	0,6	0,78
9	sastra	0,6	0,8	0,6	1	0,8	0,6	0,8	0,6	0	0,6	1	0,78
10	kaya	0,5	1	0,5	0,83	0,67	0,5	0,5	0,6	0,6	0	0,8	0,78
11	rohani	0,8	0,8	1	0,5	0,8	1	0,8	0,6	1	0,8	0	0,78
12	mentalitas	0,78	0,89	0,78	0,67	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0

Gambar 1. Matriks perhitungan jarak semua kata jawaban dan kunci jawaban

21		h	a	r	a	p
22	a	0	1	2	3	4
23	j	1	1	2	3	4
24	a	2	1	2	2	3
25	r	3	2	1	2	3

Gambar 2. Matriks perhitungan jarak antar kata

2. Menentukan Nilai Terkecil

Pada tahap ini, disetiap baris matriks dicari nilai terkecilnya seperti yang ditunjukkan Gambar 3. Setelah tiap baris ditentukan nilai terkecilnya, nilai-nilai tersebut dijumlahkan lalu dibagi dengan jumlah baris pada matriks, sehingga diperoleh hasil $0,78 \div 11 = 0,07$.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		harap	ajar	harga	perhati	peka	harga	hadap	bahasa	sastra	kaya	rohani	mentalitas
2	ajar	0,75	0	0,75	0,83	1	0,75	0,75	0,8	0,8	1	0,8	0,89
3	harap	0	0,75	0,5	0,67	0,75	0,5	0,25	0,6	0,6	0,5	0,8	0,78
4	milik	1	1	1	1	1	1	1	1	1	1	1	0,78
5	perhati	0,67	0,83	0,67	0	0,67	0,67	0,83	0,67	1	0,83	0,5	0,67
6	peka	0,75	1	0,75	0,67	0	0,75	0,75	0,8	0,8	0,67	0,8	0,78
7	harga	0,5	0,75	0	0,67	0,75	0	0,75	0,6	0,6	0,5	1	0,78
8	bahasa	0,6	0,8	0,6	0,67	0,8	0,6	0,6	0	0,6	0,6	0,6	0,78
9	sastra	0,6	0,8	0,6	1	0,8	0,6	0,8	0,6	0	0,6	1	0,78
10	kaya	0,5	1	0,5	0,83	0,67	0,5	0,5	0,6	0,6	0	0,8	0,78
11	rohani	0,8	0,8	1	0,5	0,8	1	0,8	0,6	1	0,8	0	0,78
12	mentalitas	0,78	0,89	0,78	0,67	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0

Gambar 3. Menentukan nilai terkecil pada baris matriks

3. Menghitung Persentase Kemiripan

Pada tahap ini, persentase kemiripan dihitung menggunakan rumus berikut:
 $= 100\% - (0,07 * 100\%)$
 $= 93\%$

3.2 Pengujian

Dalam pengujian algoritma Levenshtein Distance, N-Gram dan Synonym replacement yang diterapkan pada sistem penilaian esai otomatis ini, penulis menggunakan dataset yang terdiri dari 8 soal beserta kunci jawabannya dan 284 record jawaban mahasiswa. Untuk mengetahui penggunaan jumlah n-gram yang tepat, penulis melakukannya dengan 4 jenis pengujian yaitu:

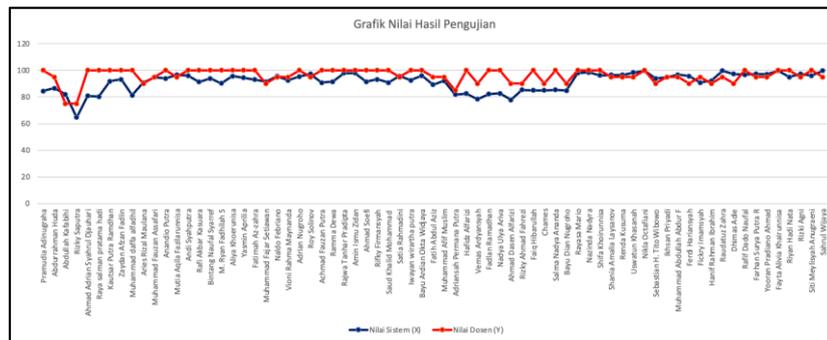
3.2.1 Pengujian menggunakan n-gram berjumlah 1 (Unigram)

Hasil pengujian menggunakan n-gram berjumlah 1 dapat dilihat pada tabel 5.

Tabel 6. Hasil Pengujian Dengan N-Gram 1

No.	Nomor Induk Mahasiswa	Nilai Sistem (X)	Nilai Dosen (Y)	Grade Nilai Sistem	Grade Nilai Dosen	Keterangan
1.	Nim 1	84,5	100	A-	A	Tidak Sama
2.	Nim 2	86,5	95	A	A	Sama
....35.	Nim 71
Jumlah		6501,5	6825			
Jumlah Data Valid (Sama)					58	
Jumlah Data Valid (Tidak Sama)					13	

Berdasarkan tabel 5, Total dari nilai yang diberikan sistem berjumlah 6501,5 sedangkan dosen 6825. Adapun Grafik hasil pengujian menggunakan n-gram 1 dapat dilihat pada gambar 4. Untuk nilai yang dihasilkan sistem diberi warna biru sedangkan nilai dosen diberi warna merah.



Gambar 4. Grafik Hasil Pengujian Dengan N-Gram 1

Tahap selanjutnya setelah mendapatkan total data yang valid adalah melakukan perhitungan akurasi menggunakan rumus berikut:

$$\text{Akurasi} = \frac{\text{Jumlah Identifikasi Benar}}{\text{Jumlah Data}} \times 100\%$$

$$\text{Akurasi} = \frac{58}{71} \times 100\%$$

$$\text{Akurasi} = 81,7\%$$

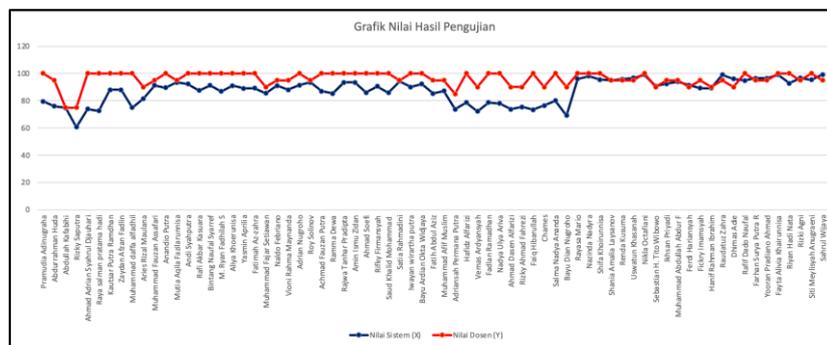
3.2.2 Pengujian menggunakan n-gram berjumlah 2 (Bigram)

Hasil pengujian menggunakan n-gram berjumlah 2 dapat dilihat pada tabel 6.

Tabel 6. Hasil Pengujian Dengan *N-Gram 2*

No.	Nomor Induk Mahasiswa	Nilai Sistem (X)	Nilai Dosen (Y)	Grade Nilai Sistem	Grade Nilai Dosen	Keterangan
1.	Nim 1	79,5	100	B+	A	Tidak Sama
2.	Nim 2	76	95	B+	A	Tidak Sama
....35.	Nim 71
Jumlah		6210,25	6825			
Jumlah Data Valid (Sama)					52	
Jumlah Data Valid (Tidak Sama)					19	

Berdasarkan tabel 6, Total dari nilai yang diberikan sistem berjumlah 6210,25 sedangkan dosen 6825. Adapun Grafik hasil pengujian menggunakan n-gram 2 dapat dilihat pada gambar 5. Untuk nilai yang dihasilkan sistem diberi warna biru sedangkan nilai dosen diberi warna merah.


Gambar 5. Grafik Hasil Pengujian Dengan *N-Gram 2*

Tahap selanjutnya setelah mendapatkan total data yang valid adalah melakukan perhitungan akurasi menggunakan rumus berikut:

$$\text{Akurasi} = \frac{52}{71} \times 100\% = 73,2\%$$

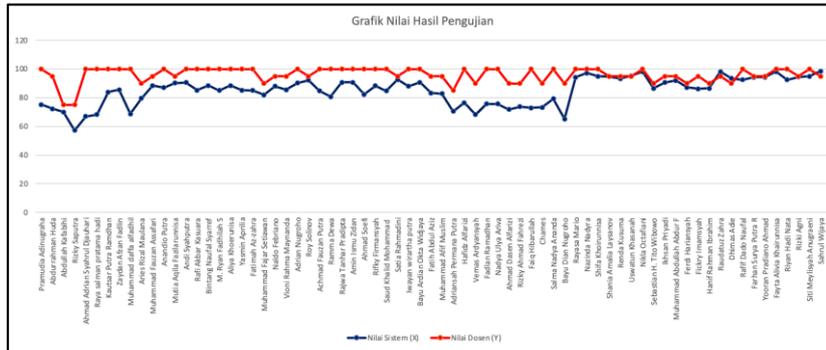
3.2.3 Pengujian menggunakan *n-gram* berjumlah 3 (*Tigram*)

Hasil pengujian menggunakan n-gram berjumlah 3 dapat dilihat pada tabel 8.

Tabel 7. Hasil Pengujian Dengan *N-Gram 3*

No.	Nomor Induk Mahasiswa	Nilai Sistem (X)	Nilai Dosen (Y)	Grade Nilai Sistem	Grade Nilai Dosen	Keterangan
1.	Nim 1	75,25	100	B+	A	Tidak Sama
2.	Nim 2	72,25	95	B	A	Tidak Sama
....35.	Nim 71
Jumlah		6030,75	6825			
Jumlah Data Valid (Sama)					44	
Jumlah Data Valid (Tidak Sama)					27	

Berdasarkan tabel 7, Total dari nilai yang diberikan sistem berjumlah 6030,75 sedangkan dosen 6825. Adapun Grafik hasil pengujian menggunakan n-gram 3 dapat dilihat pada gambar 6. Untuk nilai yang dihasilkan sistem diberi warna biru sedangkan nilai dosen diberi warna merah.



Gambar 6. Grafik Hasil Pengujian Dengan N-Gram 3

Tahap selanjutnya setelah mendapatkan total data yang valid adalah melakukan perhitungan akurasi menggunakan rumus berikut:

$$\text{Akurasi} = \frac{44}{71} \times 100\% = 61,9\%$$

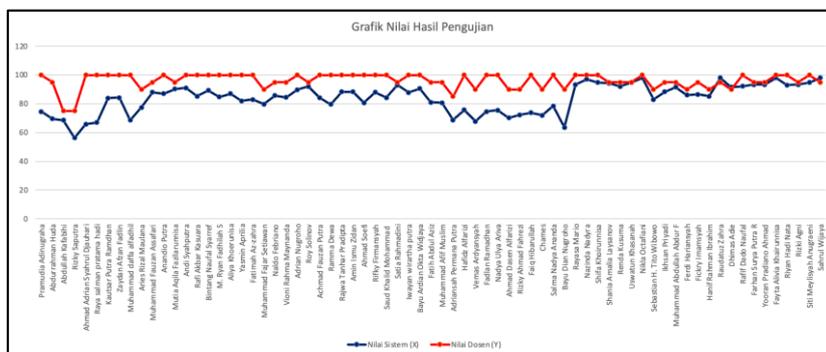
3.2.4 Pengujian menggunakan n-gram berjumlah 4 (Quadrigram)

Hasil pengujian menggunakan n-gram berjumlah 4 dapat dilihat pada table 8.

Tabel 8. Hasil Pengujian Dengan N-Gram 4

No.	Nomor Induk Mahasiswa	Nilai Sistem (X)	Nilai Dosen (Y)	Grade Nilai Sistem	Grade Nilai Dosen	Keterangan
1.	Nim 1	74,5	100	B	A	Tidak Sama
2.	Nim 2	69,5	95	B-	A	Tidak Sama
....35.	Nim 71
	Jumlah	5963,5	6825			
	Jumlah Data Valid (Sama)				38	
	Jumlah Data Valid (Tidak Sama)				33	

Berdasarkan tabel 8, Total dari nilai yang diberikan sistem berjumlah 5963,5 sedangkan dosen 6825. Adapun Grafik hasil pengujian menggunakan n-gram 4 dapat dilihat pada gambar 7. Untuk nilai yang dihasilkan sistem diberi warna biru sedangkan nilai dosen diberi warna merah.



Gambar 7. Grafik Hasil Pengujian Dengan N-Gram 4

Tahap selanjutnya setelah mendapatkan total data yang valid adalah melakukan perhitungan akurasi menggunakan rumus berikut:

$$\text{Akurasi} = \frac{38}{71} \times 100\% = 53,5\%$$

3.3 Analisa Hasil Pengujian

Hasil Dari 4 jenis pengujian yang dilakukan, akurasi tertinggi didapatkan dari pengujian jenis pertama yaitu penggunaan n-gram berjumlah 1 (Unigram) dengan nilai persentasi 81,7%. Beberapa poin yang dapat disimpulkan dari pengujian ini antara lain:

- a. Jumlah *N-Gram*
Variasi jumlah n-gram mempengaruhi akurasi algoritma Levenshtein Distance. Pengujian menunjukkan bahwa n-gram 1 menghasilkan akurasi tertinggi (81,6%), diikuti n-gram 2 (73,2%), n-gram 3 (61,9%), dan n-gram 4 (53,5%). Dengan kata lain, semakin kecil jumlah n-gram, semakin tinggi akurasinya..
- b. Variasi Jawaban
Berdasarkan dataset, ada soal yang memungkinkan improvisasi, seperti membuat kalimat efektif dan menentukan kalimat utama. Semakin tinggi jumlah n-gram, algoritma Levenshtein Distance semakin peka terhadap variasi jawaban, yang terlihat dari penurunan akurasi pada pengujian n-gram 2, 3, dan 4.

4. KESIMPULAN

Berdasarkan hasil penerapan algoritma Levenshtein Distance, N-Gram dan Synonym Replacement pada pembuatan sistem penilaian esai otomatis, penulis menyimpulkan sebagai berikut:

- a. Algoritma Levenshtein Distance, N-Gram dan Synonym Replacement berhasil diterapkan dalam pembuatan sistem penilaian esai otomatis di Universitas Budi Luhur.
- b. Sistem ini dapat menghasilkan akurasi tertinggi yaitu 81,6% dengan penggunaan n-gram berjumlah 1.
- c. Penggunaan soal-soal esai yang memungkinkan mahasiswa memberikan jawaban secara variatif, sangat mempengaruhi penggunaan n-gram yang tinggi.
- d. Sistem ini tidak cocok digunakan oleh soal yang memiliki jawaban bersifat opini.
- e. Sistem ini dapat memudahkan mahasiswa dan dosen untuk melaksanakan ujian esai.

Penulis menyadari adanya kekurangan dalam penelitian ini. Agar sistem ini lebih sempurna dan berfungsi lebih baik lagi, ada beberapa poin saran yang dapat penulis sampaikan diantaranya:

- a. Pengujian dengan dataset soal-soal esai yang diambil dari matakuliah selain Bahasa Indonesia.
- b. Melakukan penelitian lanjutan untuk menentukan jenis soal yang cocok untuk setiap jumlah n-gram.
- c. Menerapkan metode multi kunci jawaban agar jawaban bervariasi bisa mendapat penilaian sepadan
- d. Menerapkan metode pengambilan data yang tepat agar soal-soal yang digunakan memiliki cara penilaian manual yang terstruktur. Dengan demikian, nilai-nilai dosen menjadi lebih konsisten sehingga perhitungan akurasi sistem dapat dilakukan dengan lebih adil.

DAFTAR PUSTAKA

- [1] I. Mufiid, S. Lestanti, and N. Kholila, "Aplikasi Penilaian Jawaban Esai Otomatis Menggunakan Metode Synonym Recognition Dan Cosine Similarity Berbasis Web," *Jurnal Mnemonic*, vol. 4, no. 2, pp. 31–37, 2021, doi: 10.36040/mnemonic.v4i2.4067.
- [2] A. A. Suhandana, N. Marcheta, and M. Rosalina, "Optimalisasi Sistem Monitoring pada Ujian Online Berbasis Website," *Jurnal Pendidikan Tambusai*, vol. 7, no. 2, pp. 4468–4478, 2023, [Online]. Available: <https://jptam.org/index.php/jptam/article/view/6419>
- [3] F. E. Kurniawati and W. M. Pradnya, "Implementasi Algoritma Winnowing Pada Sistem Penilaian Otomatis Jawaban Esai Pada Ujian Online Berbasis Web," *Jurnal Teknik Komputer AMIK BSI*, vol. 8, no. 2, pp. 169–175, 2020, doi: 10.31294/jtk.v4i2.
- [4] H. Tsaqif, "Sistem Koreksi Jawaban Uraian Singkat Otomatis Menggunakan Metode Winnowing Algorithm dan Synonym Recognition," *Universitas Islam Negeri Maulana Malik Ibrahim*, 2021.
- [5] M. Dzaki Razaan Faza, "Penerapan Algoritma Pencocokan String pada Penilaian Jawaban Pendek," 2021.
- [6] W. M. P. Dhuhiha and F. Zone, "Perbandingan Kinerja Algoritma Multinomial dan Bernoulli Naïve Bayes dalam Mengklasifikasikan Komentar Cyberbullying," *Komputika: Jurnal Sistem Komputer*, vol. 12, no. 2, pp. 109–117, 2023.

- [7] M. R. Batisya and K. Jevanda BS, “Implementasi Algoritma Levenshtein Distance Untuk Misspelled Word Pada Pencarian Lagu Melayu,” *Jurnal Informatika*, vol. 10, no. 1, pp. 72–78, 2023, doi: 10.31294/inf.v10i1.15208.
- [8] L. R. Wachidah, Y. Laila, A. Irmawati, and S. Amin, “Implementasi Penggunaan Tes Essay dalam Evaluasi Pembelajaran Daring pada Siswa Kelas VII SMP Negeri 1 Tlanakan,” *GHANCARAN: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, p. 3, 2021.
- [9] educchannel.id, “Pengertian Instrumen Tes Esai | EduChannel Indonesia,” EduChannel Indonesia. Accessed: Jun. 24, 2024. [Online]. Available: <https://www.educhannel.id/artikel/Asesmen-dan-Evaluasi-Pembelajaran/pengertian-instrumen-tes-esai.html>
- [10] dpa.uui.ac.id, “Pengantar Asesmen, Penilaian, dan Evaluasi Pembelajaran - Direktorat Pengembangan Akademik Universitas Islam Indonesia | DPA UUI,” Direktorat Pengembangan Akademik Universitas Islam Indonesia. Accessed: Jun. 25, 2024. [Online]. Available: <https://dpa.uui.ac.id/pengantar-asesmen-penilaian-evaluasi/#>
- [11] H. Ashari, D. Arifianto, H. Azizah, and A. Faruq, “Perbandingan Kinerja Algoritma Multinomial Naive Bayes (MNB, Multivariate Bernoulli dan Rocchio Algoritim Dalam Klasifikasi Konten Berita Hoax Berbahasa Indonesia Pada Media Sosial,” *Http://Repository.Unmuhjember.Ac.Id*, pp. 1–12, 2020.
- [12] I. Athiyyah Rahma and L. Hulliyyatus Suadaa, “Penerapan Text Augmentation Untuk Mengatasi Data Yang Tidak Seimbang Pada Klasifikasi Teks Berbahasa Indonesia Studi Kasus: Deteksi Judul Clickbait Dan Komentar Hate Speech Pada Berita Online,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 10, 2023, doi: 10.25126/jtiik.2023107325.
- [13] Fitriyani and T. Arifin, “Penerapan Word N-Gram Untuk Sentiment Analysis Review Menggunakan Metode Support Vector Machine (Studi Kasus: Aplikasi Sambara),” *SISTEMASI : Jurnal Sistem Informasi*, vol. 9, pp. 610–621, 2020.
- [14] M. F. Azhri, D. Swanjaya, and R. K. Niswatin, “Penerapan Algoritma Levenshtein Distance pada Aplikasi Asisten Guru Bahasa Inggris,” *Seminar Nasional Inovasi Teknologi*, pp. 155–160, 2019.
- [15] M. N. Cholis, E. Yudaningtyas, and M. Aswin, “Pengaruh Penggunaan Synonym Recognition dan Spelling Correction pada Hasil Aplikasi Penilaian Esai dengan Metode Longest Common Subsequence dan Cosine Similarity,” *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)*, vol. 3, no. 2, pp. 138–142, 2019, doi: 10.30743/infotekjar.v3i2.1061.