


# Deteksi Kemiripan Judul Skripsi Menggunakan Algoritma Smith Waterman

Uswatun Nurul Hasanah<sup>a,1,\*</sup>, Ramdan Satra<sup>a,2</sup>, Fitriyani Umar<sup>a,3</sup>

<sup>a</sup>Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Muslim Indonesia, Jl.Urip Sumohardjo KM.05, Makassar dan 90231, Indonesia

<sup>1</sup>uzwhatunh@gmail.com; <sup>2</sup>ramdan@umi.ac.id; <sup>3</sup>fitriyani.umar@umi.ac.id

INFORMASI ARTIKEL	ABSTRAK
Diterima : 15 – 02 – 2020 Direvisi : 20 – 02 – 2020 Diterbitkan : 28 – 08 – 2020	Tujuan penelitian ini yaitu membangun sebuah sistem berbasis android untuk mendeteksi presentase kemiripan judul dan membantu mahasiswa dalam mengetahui presentase kemiripan judul skripsi yang akan diajukan. Penelitian ini menggunakan Algoritma Smith Waterman. Data yang digunakan adalah judul skripsi mahasiswa Teknik Informatika FIKOM UMI tahun 2018-2019, dengan jumlah data sampel sebanyak 4 judul. Hasil dari penelitian ini adalah mendeteksi kecocokan judul skripsi yang diajukan dengan judul skripsi yang pernah ada sebelumnya dengan presentasi kemiripan 100% diperoleh pasangan D1 dan D1 dengan skor 68, pasangan D1 dan D2 memiliki presentase kemiripan 34% dengan skor -10, serta pasangan D1 dan D3 memiliki skor -16 dengan presentasi kemiripan 31%. Sehingga dapat disimpulkan bahwa dari ketiga sampel yang paling mirip adalah pasangan sampel D1 dan D1.
<b>Kata Kunci:</b> Penelitian Penyeleksian Pendanaan Proposal LP2S UMI	
	This is an open access article under the <a href="#">CC-BY-SA</a> license
	

## I. Pendahuluan

Dalam suatu instansi akademik, Penelitian atau biasa dikenal dengan istilah tugas akhir bagi mahasiswa merupakan sks wajib yang tersaji pada KRS setiap program studi atau jurusan di seluruh perguruan tinggi di dunia [1]. Tugas akhir juga dikenal dengan istilah Skripsi, tahapan dalam pembuatan skripsi diawali dengan ide atau judul, Judul skripsi merupakan hal yang tidak asing didengar di kalangan mahasiswa, karna proses untuk mendapatkan judul atau ide sangat sulit terlebih dengan kondisi alumni yang sudah tidak sedikit lagi. Sehingga pada penelitian ini tentu saja tidak lepas dari judul skripsi. Seperti halnya yang terjadi di Fakultas Ilmu Komputer Universitas Muslim Indonesia. Judul-judul skripsi yang di usulkan oleh mahasiswa harus dikumpulkan di ruangan sekretaris prodi kemudian diproses apakah judul yang dikumpulkan memenuhi syarat atau tidak.

Salah satu syarat diterimanya sebuah judul skripsi adalah tingkat kesamaan judul skripsi yang diajukan dengan judul skripsi yang telah ada sebelumnya. Namun saat ini, yang menjadi salah satu kendala bagi mahasiswa yaitu belum adanya sebuah bank atau tempat yang dapat menampung judul skripsi secara digital. Bank judul skripsi bertujuan agar mahasiswa dapat mengetahui judul yang akan diajukan sudah ada atau belum, juga mengetahui apakah judul skripsi yang diajukan memiliki kemiripan dengan yang lain. Untuk mengetahui tingkat kemiripan sebuah judul skripsi tidak bisa dilakukan dengan hanya membandingkan kata-kata yang terdapat pada judul yang diajukan dengan judul yang ada sebelumnya, tetapi perlu metode khusus yang dapat menangani tingkat kemiripan tersebut.

Seiring perkembangan teknologi, mulai bermunculan metode-metode yang dapat diterapkan pada aplikasi untuk mengukur tingkat akurasi kemiripan judul skripsi mahasiswa, salah satunya adalah algoritma Smith Waterman adalah perluasan algoritma pencocokan string pada teks atau barisan sebagai salah satu implementasi program dinamis. Algoritma ini akan membandingkan keseluruhan panjang 2 sekuen (sederetan pernyataan-pernyataan yang urutan dan pelaksanaan eksekusinya runtut) yang terbagi menjadi sub bagian untuk mendapatkan kesamaan tertinggi antara kedua sekuens[2], [3].

Penelitian oleh [2] Metode yang diterapkan dalam sistem ini adalah algoritma smith Waterman. Proses pendeteksian menggunakan algoritma smith waterman menggunakan 3 tahap yaitu tahap penentuan skor, tahap penentuan jalur traceback dan tahap penentuan skor akhir. Hasil dari penelitian ini ditemukan bahwa Algoritma Smith Waterman dapat digunakan dalam mengetahui persentase kemiripan tugas kuliah mahasiswa. Dimana

setiap karakter akan dipasang kemudian dibandingkan. Nilai yang dihasilkan dalam sistem pendeteksi kemiripan tugas kuliah mahasiswa merupakan skor kemiripan tugas, semakin banyak karakter yang sama, maka semakin tinggi nilai skor yang dihasilkan (skor akhir bernilai positif). Sebaliknya, semakin banyak karakter yang tidak sama (mismatch), maka semakin rendah nilai skor yang dihasilkan (Skor akhir bernilai negatif).

Penelitian oleh [4] menerapkan algoritma Smith Waterman dalam sistem pendeteksi kesamaan dokumen. Algoritma ini sudah dikenal luas dalam bidang bioteknologi untuk pendeteksian kesamaan DNA. Algoritma ini akan membandingkan dua dokumen dalam sebuah matriks. Dari matriks ini, nantinya akan dilakukan traceback untuk mencari letak kesamaan dan apakah terdapat penghapusan dan penyisipan dari dokumen aslinya. Hasil yang akan didapat oleh pengguna adalah seberapa besar persentase dari kesamaan dokumen dan menunjukkan kepada pengguna letak kesamaannya. Algoritma ini berhasil diterapkan untuk sistem pendeteksian kesamaan dokumen, hanya saja tidak cocok digunakan untuk perbandingan dokumen dengan data yang lebih besar dan diperlukan waktu yang agak lama untuk mendapatkan hasil pendeteksian dari kedua dokumen yang dibandingkan.

Penelitian oleh [5] menjelaskan bahwa algoritma Smith Waterman dapat digunakan untuk mengatasi plagiarisme yang saat ini banyak terjadi. Dengan menerapkan konsep algoritma Smith Waterman, suatu kalimat dalam dokumen teks akan diibaratkan sebagai sebuah sekuens gen yang dibentuk oleh rangkaian kata dan akan dibandingkan dengan kalimat di dokumen teks lain. Sebelumnya, dokumen teks akan melalui tahapan penghilangan noise. Setelah itu diproses menggunakan algoritma Smith Waterman sehingga diperoleh suatu presentase kesamaan dokumen yang menandakan seberapa besar kemiripan dokumen tersebut dengan dokumen lain.

Penelitian oleh [3] menggunakan algoritma Smith-Waterman untuk mengidentifikasi kesamaan yang paling signifikan (local alignment) dari dua buah rangkaian sequence string proses (rangkaian/susunan) nucleotide atau protein sequences sehingga kemiripan antara dua dokumen tersebut dapat terlihat. Algoritma Nazief & Andriani merupakan algoritma stemming pada text-preprocessing sebagai algoritma pendukung dalam proses penentuan nilai kemiripan dokumen teks. Perbandingan dua sequence yang diuji dengan menggunakan preprocessing mendapatkan nilai perhitungan similarity dengan tingkat kemiripan lebih besar dibandingkan dengan tanpa preprocessing. Hal tersebut disebabkan karena penghilangan noise pada dokumen dapat meningkatkan nilai similarity tingkat kemiripan pada dokumen. Selain itu, pengujian menggunakan preprocessing membutuhkan waktu yang lebih lama dibandingkan tanpa preprocessing, karena terdapat proses tokenization, stopword removal dan algoritma stemming Nazief & Andriani yang dilakukan pada suatu dokumen. Algoritma Nazief & Andriani membantu dalam proses preprocessing (stemming) yang mendukung algoritma Smith-waterman dalam proses deteksi kemiripan dokumen teks. Dari rata-rata pengujian pada dokumen asli dan dokumen uji menunjukkan tingkat kemiripan yang dihasilkan lebih dari 50%, yang berarti hasil tersebut dapat dinyatakan terjadi tindak plagiarisme.

Penelitian oleh [6], Untuk menangani teks berukuran besar yang banyak, diperlukan suatu aplikasi yang dapat mengotomatisasi proses pendeteksian kemiripan teks. Sedangkan untuk mengatasi struktur teks yang tidak konsisten dan kompleks, diperlukan konsep text mining. Berfokuskan pada teks berbahasa Indonesia dan berkonsepkan text mining, penelitian ini bertujuan untuk mengembangkan sebuah aplikasi yang mengimplementasikan cosine similarity yang berguna untuk mengukur kesamaan teks berdasarkan kemunculan kata-kata dalam teks tersebut dan algoritma Smith-Waterman yang berfungsi untuk menghitung kemiripan teks berdasarkan urutan kata. Sebagai tambahan, penelitian ini menggunakan algoritma Nazief-Adriani untuk mengubah kata-kata berbahasa Indonesia menjadi kata-kata dasarnya (stemming) yang nantinya digunakan dalam tahap preprocessing. Ruang lingkup penelitian ini adalah sebagai berikut: kata-kata yang dapat di-stem hanyalah kata-kata yang berbahasa Indonesia dan aplikasi dapat memproses teks yang berada di dalam berkas berekstensi .doc,.docx, serta .txt.

Beberapa penelitian terkait algoritma Smith Waterman yang pernah dilakukan hanya diperlukan text judul skripsi, membandingkan judul skripsi, mengidentifikasi keseluruhan karakter citra angka maupun huruf. Dalam penelitian yang dilakukan oleh [7] pada penelitian tersebut diperlukan dokumen text untuk mengetahui skor kemiripannya. Menurut [8] pada penelitian tersebut mengecek kesalahan huruf pada kata dalam penyetikan. Sedangkan menurut [9] penelitian tersebut menggunakan kamera untuk mengidentifikasi kecocokan karakter citra berupa angka atau nomor.

## II. Metode

### A. Teori Algoritma Smith Waterman

Algoritma Smith Waterman metode yang dapat mengidentifikasi local similarities (penjajaran sekuens) yaitu proses penyusunan dua local sequences (rangkaian/ susunan atau rentetan). Berdasarkan fungsi proses

penyejajaran sekuens, algoritma ini dapat diterapkan dalam pemrograman komputer untuk pendeteksian kesamaan atau pengukuran tingkat kemiripan sebuah teks dengan teks lain[7].

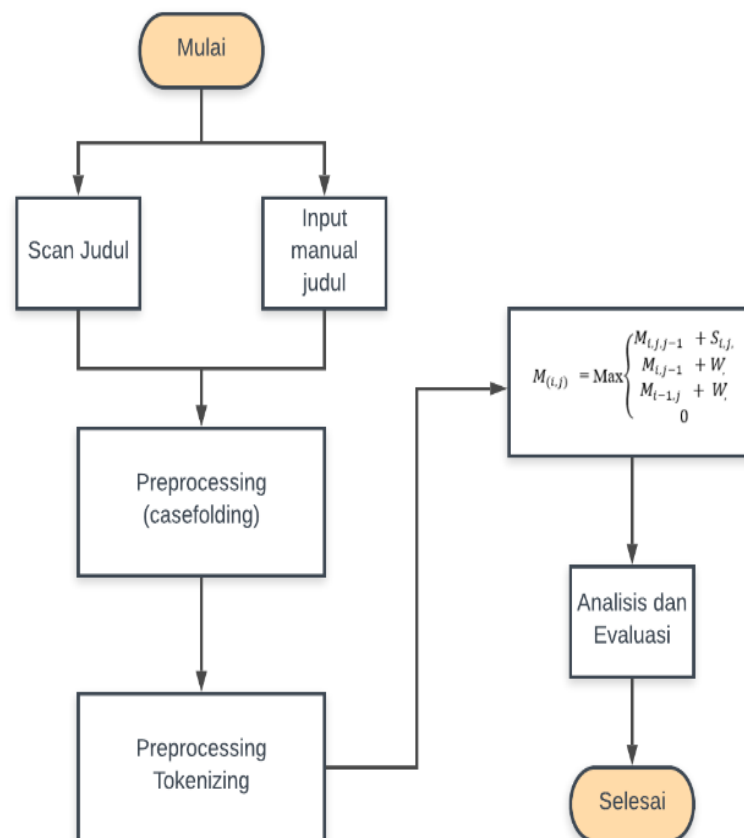
Algoritma Smith Waterman menyejajarkan dua urutan berdasarkan kecocokan/ketidakcocokan (juga dikenal sebagai substitusi), penyisipan, dan penghapusan. Baik penyisipan dan penghapusan adalah operasi yang memperkenalkan celah, yang diwakili oleh tanda hubung. Algoritma Smith-Waterman yang digunakan dalam proses mengidentifikasi local similarities (penyejajaran sekuens) yaitu proses penyusunan dua local sequence. Berikut ini adalah cara kerja dari algoritma Smith-Waterman dalam menemukan kesamaan dari dua sequence. Inisialisasi matriks dengan kolom pertama sebagai sequence A dan baris pertama sebagai sequence B[10].

#### B. Alur Penelitian

Perancangan sistem dibuat sebagai tahapan untuk mempersiapkan proses implementasi sistem pendeteksi kemiripan judul skripsi dan untuk menggambarkan secara jelas proses-proses yang diinginkan di dalam sistem.

Dalam perancangan sistem ini, akan membahas mengenai karakteristik user yang dapat mengakses sistem dan proses-proses apa saja yang dapat dilakukan di dalam sistem. Tiap proses interaksi user dengan sistem akan digambarkan dengan menggunakan flowchart. Use Case Diagram, Activity Diagram, Sequence Diagram dan Database.

Selain membahas mengenai user dan proses yang dapat dilakukan di dalam sistem. Pada perancangan sistem ini juga membahas sketsa rancangan tampilan aplikasi yang akan dibangun (interface sistem). Alur (Flowchart) dari sistem pendeteksi kemiripan judul skripsi adalah sebagai berikut:



Gambar 1. Flowchart Sistem

Berdasarkan flowchart pada gambar 1 proses pendeteksian kemiripan, dapat dilihat bahwa sebelum melakukan pendeteksian kemiripan judul skripsi, mahasiswa harus terlebih dahulu menginputkan judul skripsi baik melalui scan menggunakan kamera ataupun secara mengetikkan manual.

**Tabel 1. Data Sampel Judul Skripsi Mahasiswa**

No.	Tahun	No Buku	Nama Mahasiswa	Judul Skripsi
1	2019	1188	Suryadi Saputra	Rancangan E-learning Pada SMPN 11 Makassar
2	2019	1195	Dharma Wansyah	Perancangan E-Repository Jurnal Ilmiah Berbasis Web
3	2018	1186	Alim Pratama Dahlan	Rancangan Alat Pengukur PH Air Tambak Berbasis <i>Mobile</i>

Setelah data uji telah dikumpulkan, tahap selanjutnya melakukan *casefolding*, yakni mengubah semua huruf *capital* menjadi huruf kecil. Berikut data hasil *casefolding* ditunjukkan pada tabel 2.

**Tabel 2. Data Hasil Casefolding**

Judul Skripsi
Rancangan E-learning pada SMPN 11 Makassar
Perancangan E-repository jurnal ilmiah berbasis web
Rancangan alat pengukur pH air tambak berbasis mobile

Selanjutnya, dilakukan proses tokenizing, yakni menghapus semua karakter spesial atau karakter yang tidak dibutuhkan. berikut data hasil tokenizing ditunjukkan pada tabel 3.

**Tabel 3. Data Hasil Tokenizing**

Judul Skripsi
Rancangan E-learning pada SMPN 11 Makassar
Perancangan E-repository jurnal ilmiah berbasis web
Rancangan alat pengukur pH air tambak berbasis mobile

Setelah proses tokenizing kemudian proses deteksi kemiripan judul skripsi, pada proses deteksi kemiripan judul skripsi mahasiswa digunakan 4 data judul skripsi yang dimana 1 judul skripsi yang akan diajukan mahasiswa dan 3 judul skripsi yang sudah ada. Berikut kode masing-masing judul skripsi ditunjukkan pada tabel 4.

**Tabel 4. Kode Masing-Masing Judul Skripsi**

Kode	Judul Skripsi yang diajukan Mahasiswa
D1	Rancangan E-learning pada SMPN 11 Makassar
Kode	Judul Skripsi yang sudah ada
D2	Rancangan E-learning pada SMPN 11 Makassar
D3	Perancangan E-Repository Jurnal Ilmiah Berbasis Web
D4	Rancangan Alat Pengukur pH air tambak berbasis Mobile

Untuk mendapatkan skor dari masing-masing data yang dideteksi, rumus yang digunakan adalah sebagai berikut :

$$H_{ij} = \max \left\{ \begin{array}{l} H[i-1, j-1] + S(i, j), \\ H[i, j-1] + W, \\ H[i-1, j] + W \end{array} \right\} \quad (1)$$

$$H_{ij} = H[i-1, j-1] + S(i, j) \quad (1)$$

Dimana:

- H : Skor dari Matriks
- i : Baris
- j : Kolom

S : Skor Perbandingan Karakter

W : Gap Perbandingan Karakter

Dengan Skor:

S (match) = 2 yaitu jika kedua karakter yang dibandingkan sama

W = -1 yaitu jika karakter dipasangkan dengan gap

S(mismatch) = -1 yaitu jika kedua karakter yang dibandingkan tidak sama.

Jika kedua karakter yang dibandingkan match (sama), maka menggunakan persamaan (ii), jika kedua karakter yang dibandingkan tidak sama (mismatch) maka menggunakan persamaan ke (i).

**III. Hasil dan Pembahasan**

**A. Proses pencocokan Judul Skripsi**

Data yang digunakan dalam penelitian merupakan data judul skripsi mahasiswa FIKOM UMI yang diajukan pada tahun 2018-2019 dengan jumlah sebanyak 130. Untuk data sampel judul skripsi yang pernah diajukan mahasiswa ditunjukkan pada tabel 3.1. data tersebut digunakan untuk mengukur tingkat kemiripan judul skripsi yang akan diajukan mahasiswa dengan judul skripsi yang sudah ada sebelumnya. Berikut merupakan tabel hasil skor perhitungan menggunakan algoritma smith waterman:

a. Tabel Skor Perhitungan Sampel 1

Kode Judul D1 : Rancangan E-learning Pada SMPN 11 Makassar

Kode Judul D2 : Rancangan E-learning Pada SMPN 11 Makassar

	r	a	n	c	a	n	g	a	n	e	l	e	a	r	n	i	n	g	p	a	d	a	s	m	p	n	m	a	k	a	s	s	a	r	
r	0	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	2	
a	0	1	4	3	2	1	-1	-2	1	-1	-2	-2	1	1	-1	-2	-2	-2	-2	1	-1	1	-1	-2	-2	-2	1	-1	1	-1	-2	1	1		
n	0	-1	3	6	5	4	3	2	1	3	2	1	-1	-1	-1	3	2	-1	-2	-3	-1	-2	-1	-2	-3	-1	-2	-1	-2	-1	-2	-3	-1	-1	
c	0	-1	2	5	8	7	6	5	4	3	2	1	-1	-2	-2	2	1	-1	-2	-3	-2	-3	-2	-3	-4	-4	-2	-3	-2	-3	-4	-2	-2		
a	0	-1	1	4	7	10	9	8	7	6	5	4	3	1	-1	1	-1	-2	-3	-4	-1	-2	-1	-2	-3	-4	-3	-4	-1	-2	-1	-2	-3	-2	-3
n	0	-1	-1	3	6	9	12	11	10	9	8	7	6	5	4	1	-1	1	-1	-2	-2	-3	-2	-3	-4	-5	-2	-3	-2	-3	-2	-3	-4	-3	-4
g	0	-1	-2	2	5	8	11	14	13	12	11	10	9	8	7	6	5	4	3	2	1	-1	-2	-3	-4	-5	-3	-4	-3	-4	-5	-4	-5	-4	-5
a	0	-1	1	1	4	7	10	13	16	15	14	13	12	11	10	9	8	7	6	5	4	3	1	-1	-2	-3	-4	-5	-2	-3	-2	-3	-4	-3	-4
n	0	-1	-1	3	3	6	9	12	15	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	-1	-2	-3	-4	-3	-4	-5	-4	-5
e	0	-1	-2	2	2	5	8	11	14	17	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	-1	-2	-3	-4	-5
l	0	-1	-2	1	1	4	7	10	13	16	19	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	-1	-2
e	0	-1	-2	-1	-1	3	6	9	12	15	18	21	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2
a	0	-1	1	-1	-2	1	5	8	11	14	17	20	23	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5
r	0	2	1	-1	-2	-1	4	7	10	13	16	19	22	25	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8
n	0	1	-1	3	2	1	1	6	9	12	15	18	21	24	27	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11
i	0	-1	-2	2	1	-1	-1	5	8	11	14	17	20	23	26	29	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
n	0	-1	-2	-1	-2	-2	1	4	7	10	13	16	19	22	25	28	31	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17
g	0	-1	-2	-2	-3	-3	-1	3	6	9	12	15	18	21	24	27	30	33	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20
p	0	-1	-2	-3	-4	-4	-2	2	5	8	11	14	17	20	23	26	29	32	35	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23
a	0	-1	1	-1	-2	-2	-3	1	4	7	10	13	16	19	22	25	28	31	34	37	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26
d	0	-1	-1	-2	-3	-3	-4	-1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	41	40	39	38	37	36	35	34	33	32	31	30	29
a	0	-1	1	-1	-2	-1	-2	-2	1	5	8	11	14	17	20	23	26	29	32	35	38	41	44	43	42	41	40	39	38	37	36	35	34	33	32
s	0	-1	-1	-2	-3	-2	-3	-3	-1	4	7	10	13	16	19	22	25	28	31	34	37	40	43	46	45	44	43	42	41	40	39	38	37	36	35
m	0	-1	-2	-3	-4	-3	-4	-4	-2	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	47	46	45	44	43	42	41	40	39	38
p	0	-1	-2	-3	-4	-4	-5	-5	-3	2	5	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	49	48	47	46	45	44	43	42	41
n	0	-1	-2	-1	-2	-3	-2	-3	-4	-1	4	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	51	50	49	48	47	46	45	44
m	0	-1	-2	-2	-3	-4	-3	-4	-5	-2	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	53	52	51	50	49	48	47
a	0	-1	1	-1	-2	-1	-2	-3	-2	-3	2	5	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	53	56	55	54	53	52	51	50
k	0	-1	-1	-2	-3	-2	-3	-4	-3	-4	1	4	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55	58	57	56	55	54	53
a	0	-1	1	-1	-2	-1	-2	-3	-2	-3	-1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60	59	58	57	56
s	0	-1	-1	-2	-3	-2	-3	-4	-3	-4	-2	2	5	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	53	56	59	62	61	60	59
s	0	-1	-2	-3	-4	-3	-4	-5	-4	-5	-3	1	4	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55	58	61	64	63	62
a	0	-1	1	-1	-2	-2	-3	-4	-3	-4	-4	-1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60	63	66	65
r	0	2	1	-1	-2	-3	-4	-5	-4	-5	-2	2	5	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	53	56	59	62	65	68	

Gambar 2. Tabel Skor Perhitungan Sampel 1

b. Tabel Skor Perhitungan Sampel 2









Tabel 8. Skor Dan Presentase Akhir Perhitungan algoritma Smith Waterman

No	Judul Skripsi Mahasiswa		Skor	Presentase(%)
	Diajukan	Sudah ada		
1	D1	D1	68	100 %
2	D1	D2	-10	34 %
3	D1	D3	-16	

Nilai presentase tertinggi adalah pasangan D1 dan D1 dengan skor 68 presentasi kemiripan 100%, dari skor tersebut membuktikan bahwa pada untuk pasangan D1 dan D1 karakter yang bernilai sama (match) lebih banyak dibandingkan pasangan karakter yang bernilai tidak sama (mismatch). Sedangkan untuk pasangan D1 dan D2 dengan skor -10 presentase kemiripan 34% membuktikan bahwa pasangan karakter yang tidak sama (mismatch), dan pasangan D1 dan D3 dengan skor -16 presentasi kemiripan 31% membuktikan bahwa beberapa karakter yang tidak sama (mismatch) sehingga dapat disimpulkan bahwa dari ketiga sampel yang paling mirip adalah sampel 1. Nilai skor yang didapatkan dari perhitungan menggunakan algoritma smith waterman tidak dimasukkan dalam perhitungan untuk mendapatkan persentase kemiripan agar nilai presentase kemiripan yang dihasilkan dapat lebih akurat.

#### IV. Kesimpulan dan saran

Berdasarkan penelitian yang dilakukan dalam membuat sistem pendeteksi kemiripan judul skripsi mahasiswa menggunakan algoritma smith waterman maka penulis dapat menyimpulkan bahwa algoritma Smith Waterman dapat digunakan dalam mengetahui skor skor dan presentase kemiripan judul skripsi mahasiswa. Dimana setiap karakter akan dipasangkan kemudian dibandingkan untuk diketahui persentase kemiripan judul yang akan diajukan dan Nilai yang dihasilkan dalam sistem pendeteksi kemiripan judul skripsi mahasiswa merupakan skor kemiripan judul skripsi, semakin banyak karakter yang sama (match), maka semakin tinggi nilai skor yang dihasilkan (Skor akhir bernilai positif). Sebaliknya, semakin banyak karakter yang tidak sama (mismatch), maka semakin rendah nilai skor yang dihasilkan (Skor akhir bernilai negatif).

Dari kesimpulan diatas, maka saran-saran yang kiranya dapat dipertimbangkan untuk pengembangan aplikasi berikutnya. Adapun saran yang bisa disampaikan yaitu Sistem dapat melibatkan skor akhir yang didapatkan melalui tahapan algoritma smith waterman agar nilai presentase kemiripan yang dihasilkan dapat lebih akurat dan penggunaan model dari sistem pendeteksi kemiripan judul skripsi mahasiswa dengan metode yang lain, untuk dapat membandingkan hasil kinerja sistem sehingga dapat ditemukan model yang paling baik dari sistem pendeteksi kemiripan judul skripsi mahasiswa.

#### Daftar Pustaka

- [1] Mauliddin, "Analisis Kesalahan Penggunaan Uji Statistik pada Skripsi Mahasiswa," *J. Mat. dan pembelajaran*, vol. 5, no. 2, pp. 141–158, 2017.
- [2] L. Amsir, S. Anraeni, P. Lestari, and L. Belluano, "Penerapan Algoritma Smith Waterman Untuk Mengukur Kemiripan Tugas Kuliah Mahasiswa Implementation of Smith Waterman Algorithm In Measuring of Student Task Study," *Pros. Semin. Nas. Komun. dan Inform.*, no. 2009, pp. 63–68, 2019.
- [3] B. Sari and Y. Sibaroni, "Deteksi Kemiripan Dokumen Bahasa Indonesia Menggunakan Algoritma Smith- Waterman dan Algoritma Nazief & Andriani," *Ind. J. Comput.*, vol. 4, no. 3, pp. 87–98, 2019, doi: 10.21108/indojc.2019.4.3.365.
- [4] F. B. Djafar, A. Lahinta, and L. Hadjaratie, "Penerapan Algoritma Smith-Waterman dalam Sistem Pendeteksi Kesamaan Dokumen," Universitas Negeri Gorontalo, 2003.
- [5] H. Darmawan, "Analisis dan Implementasi Algoritma Smith-Waterman pada Proses Identifikasi Kesamaan Dokumen," Institut Teknologi Telkom Bandung, 2009.
- [6] R. V. Imbar, Adelia, M. Ayub, and A. Rehata, "Implementasi Cosine Similarity dan Algoritma Smith-Waterman untuk Mendeteksi Kemiripan Teks," *J. Inform.*, vol. 10, no. 1, pp. 31–42, 2014.
- [7] F. Thalib and R. Kusumawati, "Pembuatan Program Aplikasi untuk Pendeteksian Kemiripan Dokumen Teks dengan Algoritma Smith-Waterman," 2014.
- [8] Y. Rochmawati and R. Kusumaningrum, "Studi Perbandingan Algoritma Pencarian String dalam Metode Approximate String Matching untuk Identifikasi Kesalahan Pengetikan Teks," *J. Buana Inform.*, vol. 7, no. 2, pp. 125–134, 2016, doi: 10.24002/jbi.v7i2.491.
- [9] Y. R. Prayogi and S. N. Budiman, "Color Grading Systems to Classify Ripeness of Apple Mango Fruit," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 3, no. 2, pp. 57–61, 2018, doi:

- 
- 10.25139/inform.v3i2.1010.
- [10] I. C. Bu'ulolo, N. Simamora, S. Tampubolon, and A. Pinem, "Sequence Alignment Menggunakan Algoritma Smith Waterman," *J. Integr. Ed. Khusus*, vol. II, no. 2, pp. 2–7, 2010.