

**MODIFIKASI JARAK LEVENSHTTEIN DALAM
MENENTUKAN KEMIRIPAN JUDUL SKRIPSI**

SKRIPSI

**OLEH
ACHMAD ZAILANIL ULUM
NIM. 16610111**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023**

**MODIFIKASI JARAK LEVENSHTTEIN DALAM
MENENTUKAN KEMIRIPAN JUDUL SKRIPSI**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S. Mat.)**

**Oleh
ACHMAD ZAILANIL ULUM
NIM. 16610111**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023**

**MODIFIKASI JARAK LEVENSHTTEIN DALAM
MENENTUKAN KEMIRIPAN JUDUL SKRIPSI**

SKRIPSI

Oleh
ACHMAD ZAILANIL ULUM
NIM. 16610111

Telah Diperiksa dan Disetujui untuk Diuji

Tanggal 04 Januari 2023

Dosen Pembimbing I

Dosen Pembimbing II



Mohammad Nafie Jauhari, M.Si
NIDT. 19870218 20160801 1 056



Erna Herawati, M.Pd
NIDT. 1976072320 180201 2 222

Mengetahui,

Ketua Program Studi Matematika



Dr. Elly Susanti, M.Sc

NIP. 19741129 200012 2 005

MODIFIKASI JARAK LEVENSHTTEIN DALAM MENENTUKAN KEMIRIPAN JUDUL SKRIPSI

SKRIPSI

Oleh
Achmad Zailanil Ulum
NIM. 16610111

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima sebagai salah satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat.)

Tanggal 05 Mei 2023

Ketua Penguji : Evawati Alisah, M.Pd



Anggota Penguji I : Muhammad Khudzaifah, M.Si



Anggota Penguji II : Mohammad Nafie Jauhari, M.Si



Anggota Penguji III : Erna Herawati, M.Pd



Mengetahui

Ketua Program Studi Matematika



Dr. Erna Susanti, M.Sc

NIP. 19741129 200012 2 005

PERNYATAAN KEASLIAN TULISAN

Saya yang bertandatangan di bawah ini:

Nama : Achmad Zailanil Ulum

NIM : 16610111

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Modifikasi Jarak Levenshtein Dalam Menentukan Kemiripan
Judul Skripsi

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 05 Mei 2023

Yang membuat pernyataan



Achmad Zailanil Ulum

NIM. 16610111

MOTO

“Jangan ikut pembahasa padi, karena waktunya orang berilmulah yang buka suara
dan mengatur”

PERSEMBAHAN

Dengan rasa syukur penulis persembahkan karya ini kepada:

Untuk orang-orang spesial, terkhusus bagi kedua orang tua penulis yaitu bapak Selamat dan ibu Siti Romlah tercinta yang senantiasa memberikan dukungan dan doa yang tiada hentinya demi keberhasilan penulis serta senantiasa memberikan dukungan baik secara moril maupun materil saat proses mengerjakan skripsi ini. Teruntuk istri (Ayu Karomatul Faizah) yang senantiasa sabar dan juga memberikan dukungan baik.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur penulis haturkan kepada Allah SWT yang senantiasa memberikan rahmatnya sehingga penulis dapat menyelesaikan penulisan skripsi ini dengan baik. Tak lupa shalawat serta salam selalu tercurah kepada Baginda Nabi Muhammad SAW, yang telah membimbing umat manusia menuju jalan kebenaran.

Alhamdulillah berkat taufiq serta hidayah-Nya, penulis dapat menyelesaikan skripsi ini dengan judul “**Modifikasi Jarak Levenshtein Dalam Menentukan Kemiripan Judul Skripsi**”. Dalam proses penyelesaian skripsi ini penulis banyak mendapat bimbingan dan bantuan, serta saran dari berbagai pihak. Oleh karena itu penulis menyampaikan ucapan terima kasih kepada:

1. Prof. Dr. M. Zainuddin, MA selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang dan dosen wali penulis.
3. Dr. Elly Susanti, M.Sc, selaku ketua Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Muhammad Nafie Jauhari, M.Si, selaku dosen pembimbing I yang selalu sabar memberikan arahan dan bimbingan kepada penulis.
5. Erna Hermawati, M.Pd, selaku pembimbing II yang telah memberikan arahan dan bimbingan kepada penulis dengan baik.

6. Seluruh dosen dan staf administrasi Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang yang telah memberikan ilmu pengetahuan informasi kepada penulis.
7. Bapak, ibu, dan adik tercinta yang selalu mengingatkan, memberi semangat dan motivasi, serta mendoakan penulis agar diberi kelancaran dalam menyelesaikan penulisan skripsi ini.
8. Teman-teman angkatan 2016 dan teman-teman pondok penulis yang telah memberikan semangat dan motivasi kepada penulis agar segera menyelesaikan penulisan skripsi ini.
9. Semua pihak yang ikut terlibat dalam penyelesaian skripsi ini.

Dengan mengharapkan rida dari Allah SWT, semoga amal baik dari semua pihak telah disebutkan diterima di sisi Allah. Penulis juga berharap semoga skripsi ini bermanfaat kepada orang yang membacanya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 04 Januari 2023
Yang membuat pernyataan,

Achmad Zailanil Ulum
NIM. 16610111

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERNYATAAN KEASLIAN TULISAN	v
HALAMAN MOTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
ABSTRAK	xiv
ABSTRACT	xv
مستخلص البحث	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	5
1.6 Daftar Istilah.....	5
BAB II KAJIAN PUSTAKA	6
2.1 Jarak Levenshtein	6
2.2 <i>Text Processing</i>	10
2.3 <i>Similarity</i>	11
2.4 Skripsi.....	12
2.5 Matriks Jarak Levenshtein Dalam kajian islam	12
BAB III METODE PENELITIAN	15
3.1 Pendekatan Penelitian	15
3.2 Sumber Data	15
3.3 Variabel Penelitian	15
3.4 Tahapan Penelitian	16
BAB IV PEMBAHASAN.....	18
4.1 Tahap Preprocessing	18
4.2 Tahap <i>Processing</i>	19

4.3 Implementasi System	33
BAB V PENUTUP	39
5.1 Kesimpulan.....	39
5.2 Saran.....	39

DAFTAR PUSTAKA

RIWAYAT HIDUP

DAFTAR TABEL

Tabel 2.1 Matrik perhitungan Jarak Levenstein.....	7
Tabel 2.2 Perbandingan huruf ‘T’ dengan ‘M’	8
Tabel 2.3 Perbandingan huruf ‘T’ dengan ‘A’	9
Tabel 2.4 Hasil Jarak Levenshtein	9
Tabel 4.5 Contoh Tahap Modifikasi Jarak Lavenshtein	21
Tabel 4.6 nilai awal Modifikasi Jarak Lavenshtein	21
Tabel 4.7 perbandingan kata ‘MENGHITUNG’ dengan ‘(string kosong)’	21
Tabel 4.8 perbandingan kata ‘JARAK’ dengan ‘(string kosong)’	22
Tabel 4.9 perbandingan kata ‘RATA’ dengan ‘(string kosong)’	23
Tabel 4.10 perbandingan kata ‘RATA’ dengan ‘(string kosong)’	23
Tabel 4.11 perbandingan kata ‘MODIFIKASI’ dengan ‘(string kosong)’	24
Tabel 4.12 perbandingan kata ‘JARAK’ dengan ‘(string kosong)’	24
Tabel 4.13 perbandingan kata ‘LEVENSHTEIN’ dengan ‘(string kosong)’	25
Tabel 4.14 perbandingan kata ‘MODIFIKASI’ dengan ‘MENGHITUNG’	26
Tabel 4.15 perbandingan kata ‘MODIFIKASI’ dengan ‘JARAK’	26
Tabel 4.16 perbandingan kata ‘MODIFIKASI’ dengan ‘RATA’	26
Tabel 4.17 perbandingan kata ‘MODIFIKASI’ dengan ‘RATA’	27
Tabel 4.18 perbandingan kata ‘JARAK’ dengan ‘MENGHITUNG’	28
Tabel 4.19 perbandingan kata ‘JARAK’ dengan ‘JARAK’	28
Tabel 4.20 perbandingan kata ‘JARAK’ dengan ‘RATA’	29
Tabel 4.21 perbandingan kata ‘JARAK’ dengan ‘RATA’	29
Tabel 4.22 perbandingan kata ‘LEVENSHTEIN’ dengan ‘MENGHITUNG’ ...	30
Tabel 4.23 perbandingan kata ‘LEVENSHTEIN’ dengan ‘JARAK’	31
Tabel 4.24 perbandingan kata ‘LEVENSHTEIN’ dengan ‘RATA’	31
Tabel 4.25 perbandingan kata ‘LEVENSHTEIN’ dengan ‘RATA’	32

DAFTAR GAMBAR

Gambar 4.1 Contoh <i>Uppercase</i>	19
Gambar 4.2 Contoh <i>Tokenizing</i>	19
Gambar 4.3 <i>Flowchart processing</i>	33
Gambar 4.4 Daftar Judul Skripsi Pada Database	35
Gambar 4.5 Contoh <i>Input System</i>	35
Gambar 4.6 Daftar Hasil Perhitungan.....	36

ABSTRAK

Ulum, Achmad Zailanil. 2023. **Modifikasi Jarak Levenshtein dalam Menentukan Kemiripan Judul Skripsi**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Mohammad Nafie Jauhari, M.Si. (II) Erna Herawati, M.Pd.

Kata Kunci: Jarak Levenshtein, Modifikasi, Judul Skripsi

Modifikasi Jarak Levenshtein adalah Jarak Levenshtein dengan perhitungan perkata dari dua barisan kalimat, semakin mendekati nol nilai Modifikasi Jarak Levenshtein maka semakin besar kesamaan antara dua kalimat dan sebaliknya. Penelitian ini bertujuan untuk mengetahui nilai kesamaan judul skripsi yang akan diajukan dengan judul skripsi yang sudah ada pada prodi, dengan perhitungan manual pada system progaman bahasa python. Metode yang digunakan dalam penelitian ini adalah penelitian kepustakaan (*library research*). Hasil dari penelitian ini adalah sebuah sistem Modifikasi Jarak Levenshtein yang perhitungannya perkata dari dua barisan kalimat sehingga perhitungannya lebih akurat dari sebelumnya yang menggunakan perhitungan perhuruf dari dua barisan kalimat, misalnya kata 'MATRIKS' dengan 'METRIK' dalam perhitungan jarak Levenshtein sebelumnya adalah 85% yang artinya mirip, sedangkan dalam perhitungan modifikasi jarak Levenshtein persentasenya 0% yang artinya sangat berbeda.

ABSTRACT

Ulum, Achmad Zaialanil. 2023. **The Modification of Levenshtein Distance in Determining Similarity of Thesis Titles**. Thesis. Mathematics Study Program, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisor: (I) Mohammad Nafie Jauhari, M.Si. (II) Erna Herawati, M.Pd.

Keywords: Levenshtein distance, Modified, and thesis title

The Levenshtein Modification distance is the Levenshtein distance with word calculations from two lines of sentences, the closer to zero the value of the modified Levenshtein distance, the greater the similarity between the two sentences and vice versa. This study aims to determine the similarity of the thesis titles that will be submitted with the thesis titles that already exist in the study program, with of python programming system. The method used in this research is library research. The result of this study is a Levenshtein distance modification system which calculates one word from two sentences so that the calculation is more accurate than the previous one which used a letter calculation from two sentences, for example the word 'MATRIKS' and 'METRIK' in the previous Levenshtein distance calculation was 85% which the meaning is similar, while in the calculation of the Levenshtein distance modification the percentage is 0% which means it is different.

مستخلص البحث

العلوم ، احمد جيلان .٢٠٢٣. تعديل المسافة *Levenshtein* في التحديد تشابه عناوين الرسائل
البحث العلمي. قسم الرياضيات، كلية العلوم و التكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية
الحكومية مالانج. المشرف: (١) محمد نافع جوهرى الماجستير ، (٢) ايرنا هيراواتي الماجستير.

الكلمات المفتاحية : مسافة *Levenshtein* ، التعديل ، موضوعان البحث العلمي

مسافة تعديل *Levenshtein* هي مسافة *Levenshtein* بحساب الكلمات الموجودة في السطرين من الجمل ،
وكلمات قربت قيمة مسافة تعديل *Levenshtein* من الصفر، زاد التشابه بين الجملتين والعكس صحيح .
تهدف هذه الدراسة إلى تحديد مدى تشابه عناوين الرسائل التي سيتم تقديمها مع عناوين الرسائل الموجودة بالفعل
في برنامج الدراسة ، مع الحسابات اليدوية ونظام برمجة *Python* . الطريقة المستخدمة في هذا البحث هي البحث
في المكتبات .والنتيجة من هذه الدراسة هي نظام *Levenshtein* لتعديل المسافة الذي يحسب كلمة واحدة من
جملتين بحيث يكون الحساب أكثر دقة من ذي قبل والذي يستخدم حساباً بحرف واحد من الجملتين. على
سبيل المثال ، كانت الكلمات “*MATRIS*” و “*METRIK*” في الحساب السابق لمسافة
Levenshtein هي ٨٥٪، وهو ما يعني متشابهًا ، بينما في الحساب المعدل لمسافة *Levenshtein* تكون
النسبة المئوية ٠٪ ، مما يعني مختلف جدا.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Skripsi merupakan salah satu persyaratan akhir pendidikan akademis mahasiswa yang ditulis dalam bentuk karya tulis ilmiah. Tahapan pertama dalam prosedur skripsi adalah proses pengajuan judul skripsi. Mahasiswa wajib mengajukan judul skripsi dengan tingkat kemiripan yang tidak sama dengan judul skripsi sebelumnya agar tidak terjadi penolakan judul skripsi. Namun, pada tahap pertama dalam prosedur skripsi ini tidak sedikit mahasiswa yang masih ragu akan tingkat kemiripan judul skripsi yang akan diajukan.

Hal tersebut juga berdampak dalam proses validasi judul skripsi sehingga membutuhkan waktu yang lama. Proses pengecekan judul masih dilakukan secara manual dengan memeriksa judul terdahulu hingga sampai sekarang. Sehingga hasil pengecekan tersebut kurang efektif yang mengakibatkan terjadinya kemiripan judul skripsi dikarenakan banyaknya judul skripsi yang sudah ada pada prodi. Oleh sebab itu, diperlukan adanya sistem berbasis komputer dengan menerapkan metode yang tepat sehingga dapat mendeteksi tingkat kemiripan judul skripsi secara otomatis agar dapat mencegah terjadinya penolakan judul skripsi sehingga pelaksanaan pengajuan judul skripsi dapat lebih optimal dan tidak memakan waktu lama.

Sebagai langkah awal untuk mencegah kasus kemiripan judul diperlukan cara mendeteksi kemungkinan terjadinya kemiripan yaitu diperlukan pencegahan dalam meminimalisir tindak plagiarisme dikalangan akademik dimana nantinya

bisa diketahui tingkat kemiripan tugas akhir mahasiswa dengan judul yang sama dengan membandingkan pada skripsi yang sudah ada, dengan memodifikasi Jarak Levenshtein.

Jarak Levenshtein adalah matrik untuk mengukur perbedaan antara dua barisan kata, jarak ini dinamai dengan nama matematikawan Soviet Vladimir Levenshtein, yang menemukan jarak ini pada tahun 1965. Secara informal, Jarak Levenshtein antara dua kata adalah jumlah minimum perubahan karakter yang diperlukan untuk mengubah satu kata menjadi kata lainnya, yang merupakan pengukuran dalam bentuk matriks dan menghitung jumlah perbedaan dua string dengan melalui operasi penambahan (*insert*), penghapusan (*delete*), atau penggantian karakter (*substitute*) pada suatu karakter (Dani, 2006).

Penelitian ini menggunakan Jarak Levenshtein, dikarenakan Jarak Levenshtein merupakan pengukuran dalam bentuk matriks dan menghitung jumlah perbedaan dua string dengan melalui operasi penambahan (*insert*), penghapusan (*delete*), atau penggantian karakter (*substitute*) pada suatu karakter. Dari hasil penelitian membuktikan Jarak Levenshtein bisa menghindari kesamaan string yang berulang sehingga menghasilkan akurasi yang lebih baik (Zhan dkk, 2008).

Dari penelitian lain yang dilakukan oleh (Fadhillah, dkk, 2018) dalam Prosiding Seminar Nasional Ilmu Komputer dan Teknologi Informasi yang berjudul Validasi Pencarian Kata Kunci Menggunakan Algoritma Levenshtein *Distance* Berdasarkan Metode *Approximate String Matching*. Penelitian ini menggunakan dua metode yaitu Jarak Levenshtein dan *Approximate String*

Matching. Hasil penelitian ini yaitu Jarak levenshtein dapat mengoreksi panjang string yang berbeda.

Dari penelitian lain yang dilakukan oleh (Riza AS dan Susi R, 2023) yang berjudul Perbandingan Algoritma *Edit Distance*, *Levenshtein Distance*, *Hamming Distance*, *Jaccard Similarity* Dalam Mendeteksi String *Matching*, Berdasarkan ujicoba yang dilakukan rata-rata waktu pencarian terhadap algoritma *Edit Distance* = 47,55 ms, Jarak Levenshtein = 13,125 ms, *Hamming Distance* = 14,25 ms, dan *Jaccard similarity* = 37,125 ms Ini berarti *Levenshtein Distance* memiliki waktu pencarian tercepat.

Islam menganjurkan umatnya untuk bersungguh-sungguh dalam menuntut ilmu, baik ilmu agama maupun ilmu pengetahuan. Hal ini bisa dilihat dari adanya beberapa ayat Al-Qur'an yang memotivasi untuk menuntut ilmu. Seperti yang dijelaskan dalam surat Al-Mujaadilah ayat 11:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِذَا قِيلَ لَكُمْ تَفَسَّحُوا فِي الْمَجَالِسِ فَافْسَحُوا يَفْسَحِ اللَّهُ لَكُمْ ۗ وَإِذَا قِيلَ انشُرُوا فَاَنْشُرُوا يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ ۗ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ

Artinya:

Hai orang-orang beriman apabila dikatakan kepadamu: "Berlapang-lapanglah dalam majelis", maka lapangkanlah niscaya Allah akan memberi kelapangan untukmu. Dan apabila dikatakan: "Berdirilah kamu", maka berdirilah, niscaya Allah akan meninggikan orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu pengetahuan beberapa derajat. Dan Allah maha mengetahui apa yang kamu kerjakan.

Ayat diatas tidak menyebutkan secara tegas bahwa Allah akan meninggikan derajat orang yang berilmu namun menegaskan bahwa mereka memiliki derajat yang lebih tinggi dari yang sekedar beriman. Kaum beriman dibagi menjadi dua kelompok besar, yang pertama sekedar beriman dan beramal saleh, dan yang kedua beriman dan beramal saleh serta memiliki pengetahuan. Derajat kedua kelompok ini menjadi lebih tinggi, bukan saja karena ilmu yang disandangnya,

tetapi juga amal dan pengajarannya kepada pihak lain baik secara lisan, atau tulisan maupun dengan keteladanan.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, rumusan masalah dari penelitian ini adalah bagaimana nilai Modifikasi Jarak Levenshtein dalam menentukan kemiripan judul skripsi ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan dari penelitian ini adalah mahasiswa dapat membandingkan tingkat kemiripan judul yang akan diajukan dengan judul yang sudah ada pada prodi, sehingga mahasiswa dapat mengantisipasi judul yang akan diajukan tidak terjadi penolakan.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Bahwa metode Jarak Levenshtein dapat digunakan untuk mendeteksi kemiripan judul skripsi yang sering diteliti atau belum.
2. Sebagai antisipasi prodi dalam menerima pengajuan judul skripsi yang sudah ada.
3. Membantu mahasiswa untuk mengantisipasi tidak terjadinya kemiripan judul skripsi yang akan di ajukan sehingga mahasiswa bisah mengubahnya sebelum di ajukan.

1.5 Batasan Masalah

Penelitian ini berfokus pada judul skripsi Mahasiswa Prodi Matematika Tahun 2013-2022 Universitas Islam Negeri Maulana Malik Ibrahim Malang.

1.6 Daftar Istilah

Adapun definisi istilah dalam penelitian ini sebagai berikut:

1. String

String adalah sebuah deret simbol. String dapat dianggap sebagai tipe data dan sering diimplementasikan sebagai struktur data *array* bita (atau kata) yang menyimpan urutan elemen.

2. Algoritma

Algoritma adalah suatu upaya dengan urutan operasi yang disusun secara logis dan sistematis untuk menyelesaikan suatu masalah untuk menghasilkan suatu output tertentu

BAB II

KAJIAN PUSTAKA

2.1 Jarak Levenshtein

Jarak Levenshtein adalah ukuran, dari linguistic komputasi dan ilmu komputer, umumnya digunakan untuk membandingkan dua string berdasarkan urutan karakter. Urutan karakter mendefinisikan struktur string (Edward,dkk 2020). Nilai Jarak Levenshtein dihitung dengan menggunakan matriks. Isian nilai pada matriks tersebut adalah jumlah operasi:

1. Operasi penghapusan karakter ialah proses karakter dari suatu string, contohnya string 'ashar' karakter tengah dihilangkan menjadi string 'asar' pada proses ini dilakukan penghapusan karakter 'h'.
2. Operasi Penyisipan ialah menambahkan karakter ke dalam suatu string, contohnya string 'tdk' menjadi string 'tidak', dilakukan penyisipan karakter 'i' dan 'a' di antara karakter akhir dan awal. Penyisipan tidak hanya bisa dilakukan di tengah string, namun bisa juga disisipkan di awal maupun disisipkan di akhir string.
3. Operasi Penukaran Operasi penukaran karakter merupakan operasi mengganti sebuah karakter dengan karakter lain, contohnya penulis menuliskan string 'kempes' menjadi 'kempis'. Dalam kasus ini karakter 'e' yang terdapat pada string D diubah dengan huruf 'i'.

Rumus operasi yang digunakan untuk mengisi nilai matriks adalah sebagai berikut (Andrei dan Alexandru 2017):

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

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

$$D(s, t) = \min D(i - 1, j - 1) + 1, s[i] \neq t[j] (\text{penukaran}) \quad (3)$$

$$D(s, t) = \min D(i - 1, j - 1), s[i] = t[j](\text{tidak ada perubahan}) \quad (4)$$

s = string sumber

i = baris

t = string target

j = kolom

$s[i]$ = banyak baris kata sumber

D = Jarak levenshtein

$t[j]$ = banyak kolom kata target

Seperti contoh:

Tabel 2. 1 Matrik perhitungan Jarak Levenstein

		String sumber										
			M	A	T	E	M	A	T	I	K	A
String Target		0	1	2	3	4	5	6	7	8	9	10
	T	1										
	E	2										
	K	3										
	N	4										
	O	5										
	L	6										
	O	7										
	G	8										
	I	9										

Pada Tabel 1. diketahui bahwa string sumber “MATEMATIKA” memiliki 10 karakter dan string target “TEKNOLOGI” memiliki 9 karakter. Selanjutnya karakter ke-1 pada masing-masing string dibandingkan dan diketahui bahwa isi karakter ke-1 pada masing-masing string tidak sama, maka nilai matriks yang diberikan sesuai dengan Persamaan (2) yaitu $(1,1) = D(1,1 - 1), +1$. Jadi nilai

Semakin kecil nilai Jarak Levenshtein maka semakin besar kesamaan antara dua string dan sebaliknya.

2.2 *Text Processing*

Text Processing adalah sebuah tahapan yang dilakukan sebelum pengolahan data. Dari Jurnal Informatika Upgris menurut (Jaka Harjanta 2015) yang membahas tentang *preprocessing* pada data *mining*, tahapan *preprocessing* sangat penting dilakukan terhadap aplikasi ataupun hal yang terkait data sebagai data ujinya. Hal ini dikarenakan jika tanpa adanya tahapan *preprocessing* ini, maka hasil akhir yang didapatkan mempunyai nilai yang tidak autentik, terlebih lagi pada data masukan yang memiliki kata-kata yang mempengaruhi.

Pre-processing adalah mengambil teks mentah yang kemudian diproses dan menghasilkan kata-kata yang telah dibersihkan (Anandarajan 2019). Data mentah biasanya datang dengan banyak ketidak sempurnaan seperti inkonsistensi, nilai yang hilang, *noise* dan redundansi. Kinerja algoritma selanjutnya akan terganggu jika disajikan dengan data berkualitas rendah (Ramírez-Gallego 2017). *Preprocessing* menurut (Triawati, 2009) memiliki beberapa tahapan, yaitu:

1. *Case Folding* merupakan tahapan mengubah semua huruf menjadi huruf kecil. Karakter yang diterima yaitu 'a' hingga 'z'.
2. *Tokenizing* merupakan tahap pemotongan kalimat atau string menjadi kata per kata.
3. *Filtering* adalah proses setelah *tokenizing*. Proses ini bertujuan untuk mengambil kata-kata penting dari proses sebelumnya, yaitu *tokenizing*. Terdapat dua algoritma yang bisa digunakan pada proses *filtering* ini, yaitu

wordlist (menyimpan kata-kata penting), dan algoritma *stoplist* (membuang kata yang kurang penting). *Stoplist/stopword* berisi kata-kata yang tidak terlalu penting, seperti *sepasi* dan *tanda baca*.

4. *Stemming* merupakan teknik *preprocessing* yang terdapat pada *Retrieval System* di mana kata-kata atau string diubah menjadi kata dasarnya.

2.3 Similarity

Nilai Similarity adalah nilai yang digunakan untuk mengetahui kemiripan antara string data dan string *input*. Pada proses cek *similarity*, tahapan paling utama adalah melakukan perhitungan mencari persentase nilai string yang sama. Secara umum mencari nilai persentase kesamaan string dengan memanfaatkan Jarak Levenshtein dapat ditulis dengan rumus (Ardhin Primadewi, 2019).

$s[i]$ adalah banyak baris kata dari s

$t[j]$ adalah banyak kolom kata dari j

$$\text{Similarity} = 1 - \frac{\text{Jarak Levenshtein}}{\max(s[i], t[j])} \times 100$$

Dari persamaan diatas, nilai Jarak Levenshtein yang telah dipatakan sebelumnya digunakan untuk menghitung nilai *similarity*.

Comtoh:

s =MATEMATIKA

t =TEKNOLOGI

i = 10

j = 9

Jarak levenshtein = 9

$$\text{Similarity} = 1 - \frac{9}{10} 100 = 10\%$$

Sehingga nilai kesamaan diantara kata “MATEMATIKA” dengan “TEKNOLOGI” adalah 25%, maka dua kata tersebut sangat berbeda.

2.4 Skripsi

Skripsi adalah hal yang tidak asing bagi mahasiswa. Sejak baru menyangkut status mahasiswa saja, mereka sudah dihadapkan dengan berbagai tugas seperti observasi, menganalisis, mengkritisi, dan lainnya yang pada akhirnya adalah pembuatan karya ilmiah sebagai laporan. Hal ini dinyatakan (Wahyuni, 2016) kemampuan pemahaman bacaan, kemampuan berpikir kritis, dan pengaturan diri dapat dieksplorasi pada mahasiswa tingkat awal untuk meningkatkan kemampuan menulis ilmiah.

Pengertian karya ilmiah menurut (Syamsudin, 1994) adalah naskah yang membahas suatu permasalahan tertentu, atas dasar konsepsi keilmuan tertentu, dengan memilih metode penyajian tertentu secara utuh, teratur dan konsisten. Dari pendapat diatas, dapat disimpulkan bahwa, karya ilmiah merupakan sebuah naskah yang membahas segala persoalan persoalan ilmiah berdasarkan fakta, kepustakaan, pengalaman dan observasi yang ditulis menurut metodologi penulisan yang baik dan benar.

2.5 Matriks Jarak Levenshtein Dalam kajian islam

Di dalam Islam sendiri, Jarak Levenshtein yang memiliki banyak manfaat di era yang penuh digital ini sehingga bisa membantu prodi dan untuk mahasiswa

untuk lulus tepat waktu. Hal-hal seperti membatu urusan orang lain dalam kebaikan telah diatur dalam Al-Qur'an, hadits, maupun pendapat para ulama.

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا تَحْلُوا شَعَائِرَ اللَّهِ وَلَا الشُّهُرَ الْحَرَامَ وَلَا الْهَدْيَ وَلَا الْقَلَائِدَ وَلَا آمِينَ الْبَيْتِ الْحَرَامِ يَنْتَعُونَ
فَضْلًا مِنْ رَبِّهِمْ وَرِضْوَانًا ۚ وَإِذَا حَلَلْتُمْ فَاصْطَادُوا ۚ وَلَا يَجْرِمَنَّكُمْ شَنَاَنُ قَوْمٍ أَنْ صَدُّوكُمْ عَنِ الْمَسْجِدِ
الْحَرَامِ أَنْ تَعْتَدُوا ۚ وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ ۚ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ ۚ وَاتَّقُوا اللَّهَ ۚ إِنَّ اللَّهَ
شَدِيدُ الْعِقَابِ

Artinya:

Wahai orang-orang yang beriman! Janganlah kamu melanggar syiar-syiar kesucian Allah, dan jangan (melanggar kehormatan) bulan-bulan haram, jangan (mengganggu) hadyu (hewan-hewan kurban) dan qala'id (hewan-hewan kurban yang diberi tanda), dan jangan (pula) mengganggu orang-orang yang mengunjungi Baitulharam; mereka mencari karunia dan keridaan Tuhannya. Tetapi apabila kamu telah menyelesaikan ihram, maka bolehlah kamu berburu. Jangan sampai kebencian(mu) kepada suatu kaum karena mereka menghalang-halangi dari Masjidilharam, mendorongmu berbuat melampaui batas (kepada mereka). Dan tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan permusuhan. Bertakwalah kepada Allah, sungguh, Allah sangat berat siksaan-Nya.. (QS. Al-Ma'idah : 2)

Berdasarkan ayat quran di atas, perintah bagi umat islam untuk selalu tolong menolong dalam kebajikan dan taqwa. Dikarnakan Jarak levenshtein yang kami teliti nantinya berguna untuk membatu prodi dan mahasiswa untuk mempercepat validasi judul skripsi dengan harapan bisa lulus tepat waktu. Hal ini seperti perkataan Rasulullah di dalam sebuah hadis yang berbunyi

عَنْ أَبِي هُرَيْرَةَ رَضِيَ اللَّهُ عَنْهُ عَنِ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ مَنْ نَفَسَ عَنْ مُؤْمِنٍ كُرْبَةً مِنْ كُرْبِ
الدُّنْيَا ، نَفَسَ اللَّهُ عَنْهُ كُرْبَةً مِنْ كُرْبِ يَوْمِ الْقِيَامَةِ ، وَمَنْ يَسَّرَ عَلَى مُعْسِرٍ ، يَسَّرَ اللَّهُ عَلَيْهِ فِي الدُّنْيَا
وَالْآخِرَةِ ، وَمَنْ سَتَرَ مُسْلِمًا ، سَتَرَهُ اللَّهُ فِي الدُّنْيَا وَالْآخِرَةِ ، وَاللَّهُ فِي عَوْنِ الْعَبْدِ مَا كَانَ الْعَبْدُ فِي عَوْنِ
أَخِيهِ ، وَمَنْ سَلَكَ طَرِيقًا يَلْتَمِسُ فِيهِ عِلْمًا ، سَهَّلَ اللَّهُ لَهُ بِهِ طَرِيقًا إِلَى الْجَنَّةِ ، وَمَا اجْتَمَعَ قَوْمٌ فِي بَيْتِ
مِنْ بُيُوتِ اللَّهِ يَتْلُونَ كِتَابَ اللَّهِ ، وَيَتَدَارَسُونَهُ بَيْنَهُمْ ، إِلَّا نَزَلَتْ عَلَيْهِمُ السَّكِينَةُ ، وَعَشِيَتْهُمْ الرَّحْمَةُ ،
وَحَقَّتْهُمُ الْمَلَائِكَةُ ، وَذَكَرَهُمُ اللَّهُ فِيمَنْ عِنْدَهُ ، وَمَنْ بَطَأَ بِهِ عَمَلُهُ ، لَمْ يُسْرِعْ بِهِ نَسَبُهُ

Artinya:

“Dari Abu Hurairah radhiyallahu ‘anhu, dari Nabi shallallahu ‘alaihi wa sallam bersabda, “Barangsiapa yang menghilangkan kesusahan dari kesusahan-kesusahan dunia orang mukmin, maka Allah akan menghilangkan kesusahan dari kesusahan-kesusahan hari kiamat. Barangsiapa yang memberi kemudahan orang yang kesulitan (utang), maka Allah akan memberi kemudahan baginya di dunia dan akhirat. Siapa yang menutup aib seorang muslim, maka Allah akan menutup aibnya di dunia dan di akhirat. Siapa saja yang menolong saudaranya, maka Allah akan menolongnya sebagaimana ia menolong saudaraya. Barangsiapa yang menempuh perjalanan dalam rangka menuntut ilmu, maka Allah akan mudahkan baginya jalan menuju surga. Tidaklah berkumpul sekelompok orang di salah satu rumah Allah (masjid) untuk membaca Kitabullah dan saling mempelajarinya di antara mereka, melainkan akan turun kepada mereka ketenangan, rahmat meliputinya, para malaikat mengelilinginya, dan Allah menyanjung namanya kepada Malaikat yang ada di sisi-Nya. Barangsiapa yang lambat amalnya, maka tidak akan bisa dikejar oleh nasabnya (garis keturunannya yang mulia).” (HR. Muslim dengan lafal ini) [HR. Muslim, no. 2699].

Oleh karena itu, penulis ingin membuat suatu aplikasi yang bertujuan untuk mengecek tingkat bobot kemiripan antara judul skripsi satu dan lainnya. Program ini dibuat dalam skala yang kecil dan ditujukan untuk bahasa pemrograman Python, dan pada aplikasi ini dibuat berdasarkan jarak levenshtein.

BAB III

METODE PENELITIAN

3.1 Pendekatan Penelitian

Pada penelitian ini, pendekatan yang digunakan adalah menggunakan pendekatan eksperimen deskriptif kuantitatif. Pendekatan eksperimen yaitu dengan mengendalikan serta memanipulasi variabel bebas, pada hal ini adalah string sumber dan string target, kemudian mengamati variabel terikatnya yang dalam hal ini akan disusun dalam *preprocessing* dan selanjutnya akan dibawa dalam tahapan proses yang terdiri dari *system* modifikasi Jarak Levenshtein dan *similarity*. Sedangkan pendekatan deskriptif kuantitatif dengan menganalisis data serta menyusun data yang sudah ada sesuai dengan kebutuhan.

3.2 Sumber Data

Data yang digunakan dalam penelitian ini adalah judul skripsi yang sudah terdaftar pada prodi matematika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang di dapat pada tanggal 05 Mei 2023. Data judul skripsi berjumlah 113 dengan atribut nama dan judul.

3.3 Variabel Penelitian

Variabel penelitian yang digunakan untuk penelitian ini adalah string sumber merupakan judul skripsi yang ada pada file excel sebagai data base dan string target adalah judul skripsi yang akan kita ajukan.

3.4 Tahapan Penelitian

Adapun tahapan penelitian ini yang terdiri dari lima bagian utama, studi kuantitatif bertujuan untuk memahami cara kerja Jarak Levenshtein dan juga tahapan *preprocessing* apa saja yang ingin diterapkan pada pembuatan program ini. Setelah studi pustaka, berlanjut pada pembuatan aplikasi, yaitu dengan memodifikasi Jarak Levenshtein. Kemudian mengumpulkan data judul skripsi yang sudah terdaftar pada prodi matematika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang di dapat pada tanggal 05 Mei 2023. Setelah terkumpul dan diuji coba, hal yang perlu dilakukan adalah mengevaluasi sistem, apakah sudah berjalan sesuai rencana awal atau ada yang perlu diperbaiki lagi.

Pada penelitian pembuatan aplikasi untuk menghitung kesamaan atau disebut bobot *similarity* dengan memodifikasi Jarak Levenshtein ini memiliki dua tahapan, yaitu pertama tahapan *preprocessing* dan yang terakhir yaitu tahapan intinya, yaitu *processing*. Tahapan *preprocessing* ini bertujuan untuk menyeleksi hal-hal yang dapat mengurangi akurasi pengukuran, seperti token dan lain sebagainya. Berikut lebih jelasnya:

1. Mengumpulkan sumber data yang yang berisi judul skripsi yang sesuai kriteria, sebagai *database*.
2. Tahapan *preprocessing*, maka sistem akan mengambil data berupa kata-kata yang berisi judul skripsi kemudian akan di-filter oleh sistem seperti yang diinginkan.
3. Setelahnya, tiap judul skripsi yang ada pada sumber dihitung jarak antar kata-nya dengan judul skripsi yang akan diajukan, hasilnya disebut nilai Modifikasi Jarak Levenshtein.

4. Setelah mengetahui nilai Modifikasi Jarak Levenshtein, nilai tersebut sebagai acuan untuk tahapan *similarity*, sehingga di bagian akhir nantinya.
5. Menghasilkan output berupa nilai Modifikasi Jarak Levenshtein dan persentase *similarity*.

BAB IV

PEMBAHASAN

Pada bab IV ini berisi pembahasan modifikasi sistem Jarak Levenshtein. Jarak Levenshtein atau sering disebut dengan perbandingan dua string atau suatu sistem untuk menghitung dua kata, dan sebelum menuju ke pembahasan Modifikasi Jarak Levenshtein ini, kita bahas preprocessing sebagai berikut.

4.1 Tahap Preprocessing

Pada sub-bab ini peneliti akan membahas lebih detail bagaimana tahap *preprocessing* yang akan digunakan untuk menentukan sumber data judul skripsi dan judul skripsi yang akan di ajukan. Pada kasus ini, peneliti hanya menggunakan dua tahapan *preprocessing*, dikarenakan untuk menyesuaikan objek penelitian yang akan dilakukan. Pada penjabaran di bab 2 memiliki 4 tahapan, namun pada kasus ini yang digunakan hanyamenggunakan dua tahapan *preprocessing* yaitu *Case Folding* dan *Tokenizing*.

Data tersebut akan diproses memenuhi syarat pada *preprocessing* atau tidak, sehingga dapat dikatakan jika pada tahapan *preprocessing* tidak memenuhi syarat, maka tahapan *processing* tidak akan berjalan sesuai keinginan, di mana tujuannya adalah untuk mempermudah perhitungan, agar karakter kata menjadi kapital. Namun untuk sumber data, sebelum ketahapan ini, ada tahapan *read file*, dimana *input database* harus beresentensi *file excel* (xlsx).

1. Case Folding

Case Folding adalah tahapan *preprocessing* yang digunakan untuk mengubah huruf-huruf pada objek yang dalam kasus ini adalah kolom excel yang

diinginkan dan inputan judul yang akan di ajukan menjadi huruf kapital (*upercase*). Huruf yang bisa diubah hanyalah a-z.

Nama Zailanilael Dendy
Ulum



NAMA ZAILANILAEEL DENDY
ULUM

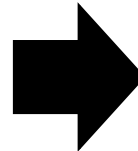
2. *T*

Gambar 4.1 Contoh *Filter upercase*

okenizing

Tokenizing merupakan tahap pemotongan kalimat menjadi kata per kata.

Uin Maulana Malik Ibrahim malang
ada di jalan Gajayana



UIN
MAULANA
MALIK
IBRAHIM
MALANG
ADA
DI
JALAN
GAJAYANA

Gambar 4.2 Contoh *Tokenizing*

4.2 Tahap *Processing*

Pada tahap *processing* menggunakan penggabungan dua metode yaitu algoritma Modifikasi Jarak Levenshtein yang digabungkan dengan *Persentase Similarity*.

1. Modifikasi Jarak Levenshtein

Dalam Modifikasi Jarak Levenshtein, akan melewati tahapan penggantian, penghapusan, penyisipan dan tidak ada perubahan jika sama. Untuk mengimplementasikan rumus Modifikasi Jarak Levenshtein, pada kasus ini perhitungan menggunakan *python* sehingga dapat dihitung panjang *s* dan *t*, misal

m dan n , kemudian membuat matriks berukuran m dan n , anggap m adalah ‘banyak baris kata sumber’ dan n adalah ‘banyak kolom kata target’.

$s[i]$ = banyak baris kata sumber

$t[j]$ = banyak kolom kata target

Priksa $s[i]$ untuk $1 < i < m$ dan $t[j]$ untuk $1 < j < n$:

A. Jika $s[i] = t[j]$

Maka entrinya adalah ambil nilai dari diagonal atas kirinya atau $D[i - 1, j - 1]$ sehingga nilai dari $D[i, j] = \min D [i - 1, j - 1] = 0$

B. Jika $s[i] \neq t[j]$

Maka entrinya adalah cari nilai minimum dari:

- Nilai yang terletak tepat di atasnya, ditambah satu, yaitu $D[i, j - 1] + 1$.
- Nilai yang terletak tepat dikirinya, ditambah satu, yaitu $D[i - 1, j] + 1$.
- terletak pada tepat di diagonal atas sebelah kirinya, ditambah satu, yaitu $D[i - 1, j - 1] + 1$.

Sehingga nilai minimum dari $D[i, j] = (\min \text{ dari } (D[i, j - 1], D[i - 1, j], D[i - 1, j - 1])) + 1$

Hasil entri matriks pada baris ke i dan kolom ke j , yaitu $D[i, j]$, dan akan di ulang sampai nilai ditemukan berada di akhir hitungan entri pojok kanan bawah.

Dan lebih jelasnya untuk perhitungan manualnya pada contoh:

Misalkan membandingkan kalimat berikut

String Sumbe : Menghitung Jarak Rata-Rata

String Target : Modifikasi Jarak Lenshtein

Kemudian akan dibuat sebuah matriks

Tabel 4.5 contoh tahap matriks Modifikasi Jarak Lavenshtein

		String sumber			
		MENGHITUNG	JARAK	RATA	RATA
String target					
	MODIFIKASI				
	JARAK				
	LEVENSHTein				

Kotak $[0,0]$ Nilainya *default* 0 untuk kotak paling pojok atas kiri

Tabel 4.6 nilai awal Modifikasi Jarak Lavenshtein

		String sumber			
		MENGHITUNG	JARAK	RATA	RATA
String target		0			
	MODIFIKASI				
	JARAK				
	LEVENSHTein				

Kotak $[0,1]$ karena $s[i] \neq t[j]$ yaitu ‘MENGHITUNG’ \neq ‘(string kosong)’ maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = tidak ada

Kotak sebelah atas ($D[i - 1, j]$) = tidak ada

Kotak sebelah kiri ($D[i, j - 1]$) = 0

Nilai Minimum = 0

Jadi nilai kotak $D[0,1] = \text{Nilai Minimum} + 1 = 0 + 1 = 1$

Tabel 4.7 perbandingan kata ‘MENGHITUNG’ dengan ‘(string kosong)’

		String sumber			
		MENGHITUNG	JARAK	RATA	RATA
s	t	0	1		

	MODIFIKASI					
	JARAK					
	LEVENSHTTEIN					

Kotak $[0,2]$ karena $s[i] \neq t[j]$ yaitu 'JARAK' \neq '(string kosong)' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = tidak ada

Kotak sebelah atas ($D[i - 1, j]$) = tidak ada

Kotak sebelah kiri ($D[i, j - 1]$) = 1

Nilai Minimum = 1

Jadi nilai kotak $D[0,2] = \text{Nilai Minimum} + 1 = 1 + 1 = 2$

Tabel 4.8 perbandingan kata 'JARAK' dengan '(string kosong)'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2		
	MODIFIKASI					
	JARAK					
	LEVENSHTTEIN					

Kotak $D[0,3]$ Karena $s[i] \neq t[j]$ yaitu 'RATA' \neq '(string kosong)' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = tidak ada

Kotak sebelah atas ($D[i - 1, j]$) = tidak ada

Kotak sebelah kiri ($D[i, j - 1]$) = 2

Nilai Minimum = 2

Jadi nilai kotak B atau $D[0,3] = \text{Nilai Minimum} + 1 = 2 + 1 = 3$

Tabel 4.9 perbandingan kata 'RATA' dengan '(string kosong)'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	
	MODIFIKASI					
	JARAK					
	LEVENSHTTEIN					

Kotak [0,4] Karena $s[1] \neq t[j]$ yaitu 'RATA' \neq '(string kosong)' maka diambil nilai minimum dari kotak diagonal atas kiri, kotak kiri, dan kotak kanan dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = tidak ada

Kotak sebelah atas ($D[i - 1, j]$) = tidak ada

Kotak sebelah kiri ($D[i, j - 1]$) = 3

Nilai Minimum = 3

Jadi nilai kotak $D[0,4] = \text{Nilai Minimum} + 1 = 3 + 1 = 4$

Tabel 4.10 perbandingan kata 'RATA' dengan '(string kosong)'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI					
	JARAK					
	LEVENSHTTEIN					

Kotak [1,0] Karena $s[i] \neq t[j]$ yaitu 'MODIFIKASI' \neq '(string kosong)' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = tidak ada

Kotak sebelah atas ($D[i - 1, j]$) = 0

Kotak sebelah kiri ($D[i, j - 1]$) = tidak ada

Nilai Minimum = 0

Jadi nilai kotak $D[1,0] = \text{Nilai Minimum} + 1 = 0 + 1 = 1$

Tabel 4.11 perbandingan kata ‘MODIFIKASI’ dengan ‘(string kosong)’

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1				
	JARAK					
	LEVENSHTTEIN					

Kotak [2,0] Karena $s[i] \neq t[j]$ yaitu ‘JARAK’ \neq ‘(string kosong)’ maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = tidak ada

Kotak sebelah atas ($D[i - 1, j]$) = 1

Kotak sebelah kiri ($D[i, j - 1]$) = tidak ada

Nilai Minimum = 1

Jadi nilai kotak $D[2,0] = \text{Nilai Minimum} + 1 = 1 + 1 = 2$

Tabel 4.12 perbandingan kata ‘JARAK’ dengan ‘(string koson)’

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1				
	JARAK	2				
	LEVENSHTTEIN					

Kotak [3,0] Karena $s[i] \neq t[j]$ yaitu 'LEVENSHTEIN' \neq '(string kosong)' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = tidak ada

Kotak sebelah atas ($D[i - 1, j]$) = 2

Kotak sebelah kiri ($D[i, j - 1]$) = tidak ada

Nilai Minimum = 2

Jadi nilai kotak $D[3,0] = \text{Nilai Minimum} + 1 = 2 + 1 = 3$

Tabel 4.13 perbandingan kata 'LEVENSHTEIN' dengan '(string kosong)'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1				
	JARAK	2				
	LEVENSHTEIN	3				

Kotak [1,1] Karena $s[i] \neq t[j]$ yaitu 'MODIFIKASI' \neq 'MENGHITUNG' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = 0

Kotak sebelah atas ($D[i - 1, j]$) = 1

Kotak sebelah kiri ($D[i, j - 1]$) = 1

Nilai Minimum = 0

Jadi nilai kotak $D[1,1] = \text{Nilai Minimum} + 1 = 0 + 1 = 1$

Tabel 4.14 perbandingan kata ‘MODIFIKASI’ dengan ‘MENGHITUNG’

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1			
	JARAK	2				
	LEVENSHTTEIN	3				

Kotak [1,2] Karena $s[i] \neq t[j]$ yaitu ‘MODIFIKASI’ \neq ‘JARAK’ maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = 1

Kotak sebelah atas ($D[i - 1, j]$) = 2

Kotak sebelah kiri ($D[i, j - 1]$) = 1

Nilai Minimum = 1

Jadi nilai kotak B atau $D[1,2] = \text{Nilai Minimum} + 1 = 1 + 1 = 2$

Tabel 4.15 perbandingan kata ‘MODIFIKASI’ dengan ‘JARAK’

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2		
	JARAK	2				
	LEVENSHTTEIN	3				

Kotak [1,3] Karena $s[i] \neq t[j]$ yaitu ‘MODIFIKASI’ \neq ‘RATA’ maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = 2

Kotak sebelah atas ($D[i - 1, j] = 3$)

Kotak sebelah kiri ($D[i, j - 1] = 2$)

Nilai Minimum = 2

Jadi nilai kotak $D[1,3] = \text{Nilai Minimum} + 1 = 2 + 1 = 3$

Tabel 4.16 perbandingan kata ‘MODIFIKASI’ dengan ‘RATA’

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	
	JARAK	2				
	LEVENSHTTEIN	3				

Kotak [1,4] Karena $s[i] \neq t[j]$ yaitu ‘MODIFIKASI’ \neq ‘RATA’ maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1] = 3$)

Kotak sebelah atas ($D[i - 1, j] = 4$)

Kotak sebelah kiri ($D[i, j - 1] = 3$)

Nilai Minimum = 3

Jadi nilai kotak $D[1,4] = \text{Nilai Minimum} + 1 = 3 + 1 = 4$

Tabel 4.17 perbandingan kata ‘MODIFIKASI’ dengan ‘RATA’

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2				
	LEVENSHTTEIN	3				

Kotak [2,1] Karena $s[i] \neq t[j]$ yaitu 'JARAK' \neq 'MENGHITUNG' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = 1

Kotak sebelah atas ($D[i - 1, j]$) = 1

Kotak sebelah kiri ($D[i, j - 1]$) = 2

Nilai Minimum = 1

Jadi nilai kotak $D[2,1] = \text{Nilai Minimum} + 1 = 1 + 1 = 2$

Tabel 4.18 perbandingan kata 'JARAK' dengan 'MENGHITUNG'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2			
	LEVENSHTein	3				

Kotak [2,2] Karena $s[i] = t[j]$ yaitu 'JARAK' = 'JARAK' maka nilai dari $D[2,2]$ Diambil dari kotak diagonal kiri atas atau $D[i - 1, j - 1] = 1$

Jadi nilai kotak $D[2,2] = 1$

Tabel 4.19 perbandingan kata 'JARAK' dengan 'JARAK'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2	1		
	LEVENSHTein	3				

Kotak [2,3] Karena $s[i] \neq t[j]$ yaitu 'JARAK' \neq 'RATA' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = 2

Kotak sebelah atas ($D[i - 1, j]$) = 3

Kotak sebelah kiri ($D[i, j - 1]$) = 1

Nilai Minimum = 1

Jadi nilai kotak $D[2,3] = \text{Nilai Minimum} + 1 = 1 + 1 = 2$

Tabel 4.20 perbandingan kata 'JARAK' dengan 'RATA'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2	1	2	
	LEVENSHTAIN	3				

Kotak [2,4] Karena $s[i] \neq t[j]$ yaitu 'MODIFIKASI' \neq 'RATA' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

Kotak diagonal kiri ($D[i - 1, j - 1]$) = 3

Kotak sebelah atas ($D[i - 1, j]$) = 4

Kotak sebelah kiri ($D[i, j - 1]$) = 2

Nilai Minimum = 2

Jadi nilai kotak $D[2,4] = \text{Nilai Minimum} + 1 = 2 + 1 = 3$

Tabel 4.21 perbandingan kata 'JARAK' dengan 'RATA'

		String sumber

			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2	1	2	3
	LEVENSHTEIN	3				

Kotak [3,1] Karena $s[i] \neq t[j]$ yaitu 'LEVENSHTEIN' \neq 'MENGHITUNG' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

$$\text{Kotak diagonal kiri } (D[i - 1, j - 1]) = 2$$

$$\text{Kotak sebelah atas } (D[i - 1, j]) = 2$$

$$\text{Kotak sebelah kiri } (D[i, j - 1]) = 3$$

$$\text{Nilai Minimum} = 2$$

$$\text{Jadi nilai kotak B atau } D[3,1] = \text{Nilai Minimum} + 1 = 2 + 1 = 3$$

Tabel 4.22 perbandingan kata 'LEVENSHTEIN' dengan 'MENGHITUNG'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2	1	2	3
	LEVENSHTEIN	3	3			

Kotak [3,2] Karena $s[i] \neq t[j]$ yaitu 'LEVENSHTEIN' \neq 'JARAK' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

$$\text{Kotak diagonal kiri } (D[i - 1, j - 1]) = 2$$

$$\text{Kotak sebelah atas } (D[i - 1, j]) = 1$$

Kotak sebelah kiri ($D[i, j - 1]$) = 3

Nilai Minimum = 1

Jadi nilai kotak B atau $D[3,2]$ = Nilai Minimum +1 = 1 + 1 = 2

Tabel 4.23 perbandingan kata 'LEVENSHTEIN' dengan 'JARAK'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2	1	2	3
	LEVENSHTEIN	3	3	2		

Kotak [3,3] Karena $s[i] \neq t[j]$ yaitu 'LEVENSHTEIN' \neq 'RATA' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini

Kotak diagonal kiri ($D[i - 1, j - 1]$) = 1

Kotak sebelah atas ($D[i - 1, j]$) = 2

Kotak sebelah kiri ($D[i, j - 1]$) = 2

Nilai Minimum = 1

Jadi nilai kotak B atau $D[3,3]$ = Nilai Minimum +1 = 1 + 1 = 2

Tabel 4.24 perbandingan kata 'LEVENSHTEIN' dengan 'RATA'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2	1	2	3
	LEVENSHTEIN	3	3	2	2	

Kotak [3,4] Karena $s[i] \neq t[j]$ yaitu 'LEVENSHTEIN' \neq 'RATA' maka diambil nilai minimum dari kotak diagonal kiri atas, kotak sebelah kirinya, dan kotak atasnya, dalam hal ini yaitu:

$$\text{Kotak diagonal kiri } (D[i - 1, j - 1]) = 2$$

$$\text{Kotak sebelah atas } (D[i - 1, j]) = 3$$

$$\text{Kotak sebelah kiri } (D[i, j - 1]) = 2$$

$$\text{Nilai Minimum} = 2$$

$$\text{Jadi nilai kotak } D[3,4] = \text{Nilai Minimum} + 1 = 2 + 1 = 3$$

Tabel 4.25 perbandingan kata 'LEVENSHTEIN' dengan 'RATA'

		String sumber				
			MENGHITUNG	JARAK	RATA	RATA
String target		0	1	2	3	4
	MODIFIKASI	1	1	2	3	4
	JARAK	2	2	1	2	3
	LEVENSHTEIN	3	3	2	2	3

2. Similarity

Pada similarity ini berguna untuk membantu mengetahui seberapa besar persentase kesamaan judul yang akan di ajukan, di situ pengguna tanpa menghitung manual berapa max kata dalam judul sehingga dalam perbandingan mendekati nol atau sebaliknya.

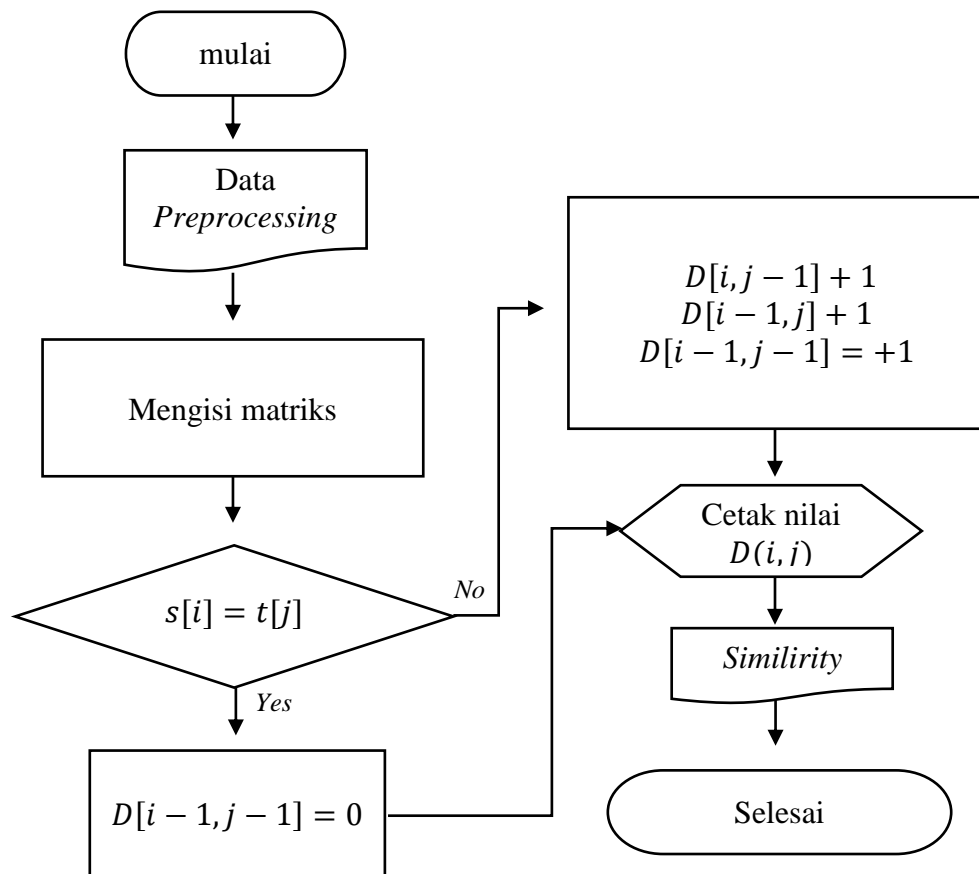
Dari hasil contoh diatas nilai modifikasi jarak levenshtein adalah 3 dengan menggunakan similarity penulis bisa tahu apakah nilai 3 itu mendekati nol atau tidak, berikut perhitungan manualnya

$$\begin{aligned} \text{Similarity} &= 1 - \frac{3}{4} \times 100 \\ &= 0,25 \times 100 \end{aligned}$$

= 25%

4.3 Implementasi System

Rancangan sistem telah dibuat diimplementasikan menggunakan bahasa pemrograman python.



Gambar 4.3 flowchart processing

Untuk pembacaan sumber data kami menggunakan *openpyxl* untuk mengambil judul skripsi dari setiap sel yang berada dalam kolom judul dan bersamaan dengan tahap *preprocessing* yang berupa *uppercase* dan *tokenizing*.

Setelah tahap *preprocessing* selesai, selanjutnya ketahapan Jarak Levenshtein Kemudian peneliti telah memodifikasi yang sebelumnya perhitungan perhuruf di ubah menjadi perkata sedemikian hingga akan menemukan nilai Modifikasi Jarak Levenshtein dan dapat dihitung persentase *similarity* dengan mengambil nilai

Modifikasi Jarak Levenshtein dan mengambil *max* string baris atau kolom yang sudah tersimpan di *database* atau data inputan judul skripsi yang akan dibandingkan, untuk dihitung persentasenya dari tiap judul. Perhitungan akan berulang bersama dengan Modifikasi Jarak Levenshtein, Sehingga pembaca bisa mengetahui tingkat kemiripan dengan nilai persentase yang didapat dari Modifikasi Jarak Levenshtein, dan di simpan pada kolom excel yang diinginkan.

Pengujian System

Langkah pertama uji coba yaitu memasukkan 100 data disimpan dalam *database* yang didapat dari *e-thesis* jurusan matematika Universitas Islam Negeri Maulana Malik Ibrahim Malang. Kemudian menginputkan judul yang akan di uji. Selanjutnya melakukan proses pengecekan dengan menggunakan metode modifikasi jarak levenshtein setelah diperoleh nilainya dari setiap judul yang tersimpan didalam *database* selanjutnya menghitung nilai plagiarisme setiap judul dan menghasilkan nilai plagiarisme dari semua data yang tersimpan didalam *database*.

No	NAMA	JUDUL
1	Muhammad ArisAbdili	Indeks eksentrisitas Zagreb pertama dan kedua graf koprima dari grup matriks upper unitriangular atas ring bilangan bulat modulo prima
2	Aisyah Rafi'Aaddani	Penerapan graf berarah dan berbobot untuk mengetahui influencer yang paling berpengaruh dalam penyebaran informasi pada media sosial Twitter
3	Indri FatikhuAflikh	Implementasi algoritma ElGamal dan fungsi hash SHA-256 pada data electronic voting (e-voting)
4	NurAgustin	Hierarchical agglomerative clustering pada klasifikasi produk susu formula bayi di bawah umur 1 tahun di Hypermart Kota Malang
5	Sri WidyaAgustin	Peramalan harga saham PT Kimia Farma Tbk di masa pandemi COVID-19 dengan metode jaringan syaraf tiruan
6	Mohammad UlilAlbab	Penerapan metode transformasi double laplace pada penyelesaian persamaan Klein Gordon
7	Fanny MillenniaAman	Implementasi regresi zero inflated negative binomial pada kasus penyakit campak Jawa Timur Tahun 2020
8	Mohamad FebryAndri	Penerapan metode noiseless steganography dan elliptic curves cryptography digital signature algorithm pada pengamanan pesan teks
9	Dio AlifArfiansyah	Degree Distance dan Gutman Index pada Graf Total dari Ring Komutatif Z ₂ ^p
10	Mega NurAzizah	Sifat-sifat inklusi pada ruang Grand Morrey
11	Siti RohmahAzizah	Ketaksamaan operator integral fraksional yang diperumum pada ruang Morrey tak homogen yang diperumum
12	SyahliaAzradillah	Modifikasi affine cipher menggunakan algoritma blum-blum shub dalam mengamankan pesan
13	Nazla AyuniBanat	Jumlah jarak eksentrik pada digraf Cayley dari grup dihedral-2n
14	Muhammad Al Himni	Implementasi hybrid cryptography berbasis Galois field dalam mengamankan pesan
15	Fasya IlyahCahyadina	Analisis dinamik model populasi perokok konvensional dan elektrik
16	Niken DwiCahyanti	Modifikasi Vigenere Cipher menggunakan grup simetri untuk mengamankan pesan teks
17	Aditya DwiCahyono	Penerapan Algoritma A-Star untuk mencari rute terpendek destinasi wisata budaya Kota Yogyakarta
18	Lisa SherlyCholliqa	Estimasi parameter suku bunga hull-white menggunakan metode jackknife dan ordinary least square
19	Tika Ma'rifatulChusni	Implementasi model ARMA dengan metode bootstrap: Studi kasus harga saham penutupan PT Kimia Farma Tbk
20	MiftahulDewanto	Ortogonalitas-G di ruang norm-n
21	Ilfli NurDiana	Elemen Engel kiri dari grup dihedral
22	Jihan Nanda PutriFadi	Perbandingan cadangan premi bulanan pada Asuransi Jiwa DwiJguna menggunakan Metode New Jersey Dan Canadian
23	Iffah NurFadlilah	Pengelompokan data tweet kecelakaan menggunakan pendekatan text mining dan algoritma Birch
24	Anisah NurilFahmi	Sistem penjadwalan ujian skripsi menggunakan algoritma Greedy
25	AchmadFauzi	Keefektifan algoritma kruskal dan prim dalam menyelesaikan optimasi jaringan listrik penyalang Sunan Ampel Kota Pasuruan
26	Moh. MuhsisFirdaus	Analisis perilaku model SEIR penyebaran covid-19 dengan faktor vaksinasi
27	NurilFirdausi	Aplikasi teori permainan dan sistem inferensi fuzzy mamdani pada strategi pemasaran marketplace makanan
28	IkaFitrianiSih	Analisis dinamik aktivasi sistem imun terhadap infeksi HIV
29	Putri LisaroHadi	Enkripsi dan Dekripsi Pesan Menggunakan Kurva Eliotik pada Affine Cipher dengan Metode Koblitz

Gambar 4.4 Daftar judul skripsi pada database

Judul skripsi yang telah tersimpan di database tersebut yaitu berjumlah 114 data yang akan dibandingkan oleh sistem dengan judul skripsi yang akan di uji yakni “MODIFIKASI JARAK LEVENSHTTEIN DALAM MENENTUKAN KEMIRIPAN JUDUL SKRIPSI” yang mana judul tersebut telah terdeteksi kesamaan dengan persentase yang berbeda beda, output dari sistem telah disimpan pada database, sesuai dengan perintah yang diinginkan

```
Python 3.7.0 (v3.7.0:1bf9cc5093, Jun 27 2018, 04:59:51) [MSC v.1914 64 bit (AMD64)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
RESTART: E:\nafi\New folder\print revisi\skripsi\levenshtein-distance\main.py
Apa nama dari file Excel (.xlsx) yang akan digunakan?
judul 5 mei.xlsx
Kolom mana yang menyimpan semua judul tesis?
c
Apa judul tesis kamu?
MODIFIKASI JARAK LEVENSHTTEIN DALAM MENENTUKAN KEMIRIPAN JUDUL SKRIPSI
Kolom mana yang akan menyimpan nilai Levenshtein Distance?
d
Kolom mana yang akan menyimpan nilai similarity?
e
Membaca judul 5 mei.xlsx...
```

Gambar 4.5 contoh input system

Sehingga untuk hasil dari modifikasi levenshtein berada pada kolom “D” dan untuk persentase kesamaan berada di kolom “E”

	A	B	C	D	E
1	No	NAMA	JUDUL	Levenshtein Distance	Similarity
2	1	Muhammad ArisAbdil	Indeks eksentrisitas Zagreb pertama dan kedua graf koprim	19	0
3	2	Aisyah Rafi'Addani	Penerapan graf berarah dan berbobot untuk mengetahui ir	17	5,55555556
4	3	Indri FatikhuAflikh	Implementasi algoritma ElGamal dan fungsi hash SHA-256 p	14	0
5	4	NurAgustin	Hierarchical agglomerative clustering pada klasifikasi prod	18	0
6	5	Sri WidyaAgustin	Peramalan harga saham PT Kimia Farma Tbk di masa pande	17	0
7	6	Mohammad UliAlbab	Penerapan metode transformasi double laplace pada peny	10	0
8	7	Fanny MillenniaAman	Implementasi regresi zero inflated negative binomial pada	14	0
9	8	Mohamad FebryAndré	Penerapan metode noiseless steganography dan elliptic cu	15	0
10	9	Dio AlifArfiansyah	Degree Distance dan Gutman Index pada Graf Total dari Rir	12	0
11	10	Mega NurAzizah	Sifat-sifat inklusi pada ruang Grand Morrey	8	0
12	11	Siti RohmahAzizah	Ketaksamaan operator integral fraksional yang diperumum	13	0
13	12	SyahlaAznadillah	Modifikasi affine cipher menggunakan algoritma blum-blum	10	9,09090909
14	13	Nazla AyuniBanat	Jumlah jarak eksentrik pada digraf Cayley dari grup dihedra	9	10
15	14	Muhammad Al Himni	Implementasi hybrid cryptography berbasis galois field dal	9	0
16	15	Fasya IlyaCahyadina	Analisis dinamik model populasi perokok konvensional da	8	0
17	16	Niken DwiCahyanti	Modifikasi Vigenere Cipher menggunakan grup simetri unt	9	10
18	17	Aditya DwiCahyono	Penerapan Algoritma A-Star untuk mencari rute terpendek	13	0
19	18	Lisa SherlyCholiqa	Estimasi parameter suku bunga hull-white menggunakan n	13	0
20	19	Tika Ma'rifatulChusniy	Implementasi model ARMA dengan metode bootstrap: Stu	15	0
21	20	MiftahulDewanto	Ortogonalitas-G di ruang norm-n	8	0
22	21	Ilfi NurDiana	Elemen Engel kiri dari grup dihedral	8	0
23	22	Jihan Nanda PutriFadi	Perbandingan cadangan premi bulanan pada Asuransi Jiwa	14	0
24	23	Iftah NurFadlilah	Pengelompokan data tweet kecelakaan menggunakan pen	11	0
25	24	Anisah NurilFahmi	Sistem penjadwalan ujian skripsi menggunakan algoritma C	8	0
26	25	AchmadFauzi	Keefektifan algoritma kruskal dan prim dalam menyelesaikan	14	6,66666667
27	26	Moh. MuhlisFirdaus	Analisis perilaku model SEIR penyebaran covid-19 dengan	10	0
28	27	NurilFirdausi	Aplikasi teori permainan dan sistem inferensi fuzzy mamdi	13	0
29	28	IkaFitrianingsih	Analisis dinamik aktivasi sistem imun terhadap infeksi HIV	8	0
30	29	Putri LisaroHadi	Enkripsi dan Dekripsi Pesan Menggunakan Kurva Eliotik pa	13	0

Gambar 4.6 Daftar Hasil Perhitungan

Integrasi System Terhadap Islam

Penelitian ini modifikasi jarak levenshtein dalam menentukan kemiripan judul skripsi, sehingga dapat diketahui tingkat kemiripan judul skripsi yang sama, yang perlu mendapatkan perhatian bagi mahasiswa tingkat akhir untuk mengantisipasi penolakan judul skripsi yang akan diajukan. Hal ini penting dilakukan untuk memperbaiki serta mempersiapkan segala kebutuhan yang diperlukan untuk pelaksanaan validasi judul skripsi pada prodi dimasa yang akan datang, agar tidak mengalami kerugian serta kendala yang tidak diinginkan selama pelaksanaannya.

Proses memperbaiki serta mempersiapkan kebutuhan ini selaras dengan kandungan isi Surat Al-Anfal Ayat 60:

وَأَعِدُّوا لَهُمْ مَا اسْتَطَعْتُمْ مِنْ قُوَّةٍ وَمِنْ رِبَاطِ الْخَيْلِ تُرْهَبُونَ بِهِ وَعَدُّوا اللَّهَ وَعَدُوَّكُمْ وَأَخْرَيْنَ مِنْ دُونِهِمْ لَا تَعْلَمُونَهُمُ اللَّهُ يَعْلَمُهُمْ ۗ وَمَا تُنْفِقُوا مِنْ شَيْءٍ فِي سَبِيلِ اللَّهِ يُوَفَّ إِلَيْكُمْ وَأَنْتُمْ لَا تُظْلَمُونَ

Artinya:

Persiapkanlah untuk (menghadapi) mereka apa yang kamu mampu, berupa kekuatan (yang kamu miliki) dan pasukan berkuda. Dengannya (persiapan itu) kamu membuat gentar musuh Allah, musuh kamu dan orang-orang selain mereka yang kamu tidak mengetahuinya, (tetapi) Allah mengetahuinya. Apa pun yang kamu infakkan di jalan Allah niscaya akan dibalas secara penuh kepadamu, sedangkan kamu tidak akan dizalimi.

قَالَ الْإِمَامُ أَحْمَدُ: حَدَّثَنَا هَارُونُ بْنُ مَعْرُوفٍ، حَدَّثَنَا ابْنُ وَهْبٍ، أَخْبَرَنِي عَمْرُو بْنُ الْحَارِثِ، عَنْ أَبِي عَلِيٍّ ثَمَامَةَ بْنِ شُعَيْبٍ، أَنَّهُ سَمِعَ عُقْبَةَ بْنَ عَامِرٍ يَقُولُ: سَمِعْتُ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ يَقُولُ وَهُوَ عَلَى الْمِنْبَرِ: {وَأَعِدُّوا لَهُمْ مَا اسْتَطَعْتُمْ مِنْ قُوَّةٍ} أَلَا إِنَّ الْقُوَّةَ الرَّمْيُ، أَلَا إِنَّ الْقُوَّةَ الرَّمْيُ

Artinya:

Imam Ahmad mengatakan, telah menceritakan kepada kami Harun ibnu Ma'ruf telah menceritakan kepada kami Ibnu Wahb, telah menceritakan kepada kami Amr ibnul Haris, dari Abu Ali Sumamah ibnu Syafi (saudara lelaki Uqbah ibnu Amir). Ia pernah mendengar Uqbah ibnu Amir mengatakan bahwa ia pernah mendengar Rasulullah Saw. bersabda di atas mimbarinya: "Dan siapkanlah untuk menghadapi mereka kekuatan apa saja yang kalian sanggupi." Ingatlah, sesungguhnya kekuatan itu terletak pada pasukan pemanah. Ingatlah, sesungguhnya kekuatan itu terletak pada pasukan pemanah.

{وَأَعِدُّوا لَهُمْ مَا اسْتَطَعْتُمْ}

Artinya:

Dan siapkanlah untuk menghadapi mereka apa saja yang kalian sanggupi

Berdasarkan ayat tersebut dapat dipahami bahwasannya mempersiapkan segala sesuatunya merupakan hal penting yang perlu diutamakan dalam kehidupan sehari-hari. Terlebih lagi dalam pelaksanaan skripsi dibutuhkan persiapan yang matang agar tidak ada kesulitan dimasa yang akan datang. Adapun tahapan pertama dalam prosedur skripsi adalah proses pengajuan judul skripsi, mahasiswa wajib mengajukan judul skripsi dengan tingkat kemiripan judul yang tidak sama

dengan judul skripsi sebelumnya agar tidak terjadi penolakan judul skripsi, oleh karena itu diperlukan persiapan judul skripsi yang tidak sama dengan judul skripsi yang sudah ada pada prodi, maka dari itu dibutuhkan sistem yang dapat mengidentifikasi kesamaan judul skripsi yang suda ada pada parodi.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan pembahasan, implementasi modifikasi jarak Levenshtein dalam menentukan kemiripan judul skripsi lebih akurat dari jarak Levenshtein sebelumnya yang membandingkan perhuruf menjadi perkata, misalnya kata 'MATRIKS' dengan 'METRIK' dalam perhitungan jarak Levenshtein sebelumnya adalah 85% yang artinya mirip, sedangkan dalam perhitungan modifikasi jarak Levenshtein persentasenya 0% yang artinya sangat berbeda.

5.2 Saran

Pada penelitian ini, Untuk pengembangan Modifikasi Jarak Levenshtein dalam menentukan kesamaan judul skripsi ini di kemudian hari diperlukan beberapa tambahan untuk mendapatkan hasil yang lebih baik, diantaranya:

1. Diperlukan uji coba dengan sumber data yang lebih banyak, dikarenakan semakin lengkap sumber data, maka akan mempengaruhi tingkat akurasi.
2. Dapat dikembangkan dengan menggunakan metode *similarity* lain untuk dijadikan perbandingan.
3. Bisa menggunakan *database* menggunakan *web scraping*.

DAFTAR PUSTAKA

- Agus Budi., Viny Christianti., Jimmy. (2018). *Deteksi Kemiripan Source Code Dengan Metode Fingerprint Based Distance dan Levenshtein Distance*. Journal of Computer Science and Information Systems. ISSN: 2549-2829.
- Al-Qur'an, L. P. M. (2021). *Qur'an Kemenag* in Microsoft Word versi 2.0.
- Andrei Ene and Alexandru Ene. (2017) *An application of Levenshtein algorithm in vocabulary learning*, Proc. 9th Int. Conf. Electron. Comput. Artif. Intell. ECAI 2017, vol. pp. 1–4,
- Ardhin Primadewi, Agus Setiawan, Muhamad Hendra Febiawan. (2019) *Sistem Pendeteksi Dini Plagiarisme Menggunakan Algoritma Levenshtein Distance*. Jurnal Komtika (Komputasi dan Informatika). Universitas Muhammadiyah Magelang
- Candra Triawati. *Metode Pembobotan Statistical Concept Based untuk Klustering dan Kategorisasi Dokumen Berbahasa Indonesia*. http://digilib.ittelkom.ac.id/index.php?option=com_content&view=article&id=590:textmining&catid=20:informatika&Itemid=1 [21 Januari 2023]
- Chung Edward ,A. Bhaskar, and Krishna N.S. Behara. (2020) *A novel approach for the structural comparison of origindestination matrices: Levenshtein distance*, *Transp. Res. Part C Emerg. Technol.*, vol. 111, no. May 2019, pp. 513–530,
- Dani, T.G. Limandra & L.R.E Adiseputra. 2006. *Deteksi Kemiripan Kode Program dengan Metode Preprocessing dan Perhitungan Leveinshtein Distance*. Prosiding Seminar Ilmiah Sumber ISSN : 1411-6286.
- Departemen Agama RI. *Al-Qur'an dan Terjemahannya*. (Jakarta: CV Penerbit Diponegoro, 2014), hal. 184
- Fadhillah, N., Aziz, H., & Lantara, D. (2018). *Validasi Pencarian Kata Kunci Menggunakan Algoritma Levenshtein Distance Berdasarkan Metode Approximate String Matching*. Prosiding Seminar Nasional Ilmu Komputer dan Teknologi Informasi , Vol. 3, No. 2, hal 129-133.
- Jaka Harjanta, Aris Tri (2015). *Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining*. Jurnal informatika Upgris, Vol.1, No. 1. I.P-ISSN 2460-4801

- Levenshtein, Vladimir I. (February 1966). *Binary Codes Capable Of Correcting Deletions, Insertions, And Reversals*. Soviet Physics Doklady 10 (8): 707-710
- M. Anandarajan, C. Hill, and T. Nolan. (2019) *Text Preprocessing BT - Practical Text Analytics: Maximizing the Value of Text Data*. Springer International Publishing. pp45-59
- Makmun Agus, (2018) *Studi Kinerja Algoritma Similaritas Untuk Identifikasi Dan Pemetaan Pernyataan Swot Universitas Muhammadiyah Surakarta*. Sekripsi Universitas Muhammadiyah Surakarta
- Riza AS dan Susi R. (2023) *Perbandingan Algoritma Edit Distance, Levenshtein Distance, Hamming Distance, Jaccard Similarity Dalam Mendeteksi String Matching*. jurnal sistem informasi, universitas suryadarma
- S. Ramírez-Gallego, B. Krawczyk, S. García, M. Woźniak, and F. Herrera, (2017) *A Survey On Data Preprocessing For Data Stream Mining: Current status and future directions*, Neurocomputing, vol. 239, pp. 39–57
- Syaikh Muhammad bin Al-‘Utsaimin. *Arba'in* hal 38
- Syamsudin, M. (1994). *Dasar Dasar dan Metode Penulisan Ilmiah*. Surakarta: Sebelas Maret University Press.
- Wahyuni, E. S. (2016). *Pengaruh Kemampuan Berpikir Kritis, Pemahaman Bacaan, Dan Pengaturan Diri Terhadap Kemampuan Menulis Ilmiah*. Ranah, 4 (1), 120-132
- Widiatry, W., Sari, N., Pranatawijaya, V., dan Adidyana Anugrah Putra, P. (2019). *Penerapan Algoritma Levenshtein Distance Untuk Pencarian Pada Sistem Informasi Perpustakaan Fakultas Kedokteran Universitas Palangka Raya*. Jurnal SAINTEKOM, 9(1), 66-82.
- Zhan, Su, Byung-Ryul Ahn, Ki-Yol Eom, Min-Koo Kang, Jin-Pyung Kim, dan Moon-Kyun Kim. (2008). *Plagiarism Detection Using the Levenshtein Distance and Smith-Waterman Algorithm*. The 3rd International Conference on Innovative Computing, Information and Control (ICICIC),

RIWAYAT HIDUP



Penulis bernama lengkap Achmad Zailanil Ulum. Terlahir 25 tahun silam di Kabupaten Malang pada bulan pertama. Penulis memulai pendidikan formal secara berturut-turut di SDN 01 Riinginsari Sumbermanjing Wetan, MTs An-Nur Bululawang, dan MA An-Nur Bululawang. Kemudian pada tahun 2016, penulis ditakdirkan menjadi mahasiswa Jurusan Matematika di UIN Maulana Malik Ibrahim Malang. Sedangkan pendidikan nonformal ditempuh di Pondok Pesantren An-Nur Bululawang saat sekolah menengah pertama dan di Ma'had Sunan Ampel Al 'Aly saat tahun pertama menjadi mahasiswa. Penulis tertarik pada dunia alam dan pernah menjadi salah satu pelopor penggerak pramuka saat SMA. Kemudian aktif di organisasi Resimen Mahasiswa pada tahun-tahun pertama kuliah. Memiliki pengalaman kemiliteran, pengetahuan alam, bertanggung jawab, memahami personalia atau SDM, senang mempelajari kehidupan, dan penuh dengan rasa ingin tahu perihal bagaimana manusia berpikir dan merasa serta hal-hal yang berkaitan dengan dua hal tersebut.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayuna No.50 Dinoyo Malang Telp. / Fax. (0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Achmad Zalanil Ulum
NIM : 16610111
Fakultas / Jurusan : Sains dan Teknologi / Matematika
Judul Skripsi : Modifikasi Jarak Levenstein Dalam Menentukan Kemiripan Judul Skripsi
Pembimbing I : Mohammad Nafie Jauhari, M.Si.
Pembimbing II : Erna Herawati, M.Pd.

No	Tanggal	Hal	Tanda Tangan
1.	04 Januari 2023	KONSULTASI BAB 2	1.
2.	18 Januari 2023	REVISI BAB 2 dan 3	2.
3.	25 Januari 2023	KONSULTASI KAJIAN AGAMA	3.
4.	27 Januari 2023	ACC BAB 1, 2 dan 3	4.
5.	01 Februari 2023	ACC KAJIAN AGAMA	5.
6.	27 Februari 2023	KONSULTASI REVISI SEMPRO	6.
7.	06 Maret 2023	KONSULTASI BAB 4 dan 5	7.
8.	15 Maret 2023	REVISI BAB 4 dan 5	8.
9.	17 Maret 2023	ACC BAB 4 dan 5	9.
10.	17 April 2023	KONSUL REVISI SEMHAS	10.
11.	10 Mei 2023	REVISI KESIMPULA dan ABSTRAK	11.
12.	17 Mei 2023	ACC REVISI SEMHAS	12.
13.	19 Mei 2023	REVISI KAJIAN AGAMA	13.

Mengetahui,
Ketua Program Studi Matematika



Dr. Elly Susanti, M.Sc
NIP. 197411292000122005