

**PENERAPAN ALGORITMA *LEVENSHTEIN DISTANCE*
UNTUK *MISSPELLED WORD* PADA PENCARIAN LAGU
MELAYU BERBASIS *WEBSITE***

RINGKASAN SKRIPSI



**MARCEL RINO BATISYA
1813008**

**PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS KATOLIK MUSI CHARITAS
PALEMBANG
2023**

PENERAPAN ALGORITMA LEVENSHTTEIN DISTANCE UNTUK MISPELLED WORD PADA PENCARIAN LAGU MELAYU BERBASIS WEBSITE

Marcel Rino Batisya

Program Studi Informatika Universitas Katolik Musi Charitas

Jl. Bangau No.60 Palembang

E-mail: marcelrinobatisya10@gmail.com

ABSTRACT

Today, people cannot live without music, both indoors and outdoors. Music is a very useful part of human life. It can be used to relax, relieve stress or change the mood. Music is easy to find on many websites today. However, people who like music, especially Malay songs, sometimes forget or misunderstand the title of the song they often hear and sing on various websites. This made it difficult for them to find the title of the Malay song they wanted. This method was not successful because the website could not explain how the spelling of the word in question corresponds to the meaning of the song title. Therefore, the authors rely on the Levenshtein Distance Algorithm to create web-based applications. This algorithm will correct spelling mistakes in song titles, either due to typing speed or typos from users. The Levenshtein Distance between two strings is determined by the least number of transform operations needed to convert one string into another; this matrix, called distance, is used to measure this difference. The operations are insert, delete, and swap. Tests show that correcting song titles that have misspelled words using the Levenshtein Distance algorithm to add and remove characters is effective.

Keywords: Music, Malay songs, Website, Levenshtein Distance Algorithm

I PENDAHULUAN

Musik merupakan suatu karya seni suara dalam bentuk lagu atau komposisi musik. Musik sendiri mengkomunikasikan pikiran dan perasaan penciptanya melalui beberapa aspek unsur musik. Unsur musik tersebut meliputi ritme, melodi, harmoni, bentuk, struktur lagu, dan ekspresif secara keseluruhan (Jamalus, 1988). Menurut Banoe (2003:288) musik merupakan sebuah karya seni yang mengkaji dan mendeskripsikan banyak suara ke dalam pola yang dapat dipahami dan dimengerti oleh

manusia. Pada saat ini musik juga merupakan salah satu karya seni yang sangat pesat perkembangannya. Hal ini diakui oleh pengamat musik Wendi Putranto dalam kutipan Pratiwi (2022) bahwa musik yang diminati oleh orang-orang pun menjadi semakin beragam sehingga membuat Cross Genre atau musik lintas genre mulai menduduki posisi utama di kancah musik secara umum.

Salah satu Cross Genre yang ada di Indonesia yaitu musik ber genre pop melayu yang merupakan genre terbesar perkembangannya di Indonesia. Menurut Rahayu

Kertawigun (2019) bahwa genre musik Dangdut dalam artian melayu di Indonesia memiliki peluang booming di era digital saat ini. Hal ini disebabkan karena musik ber genre Melayu yang masuk dalam kategori musik Melayu standar yang bisa disandingkan dengan musik ber genre EDM (Electronic Dance Music). Sehingga melahirkan musik melayu modern yang justru terlahir fresh dan mudah masuk ke usia remaja sampai dengan lanjut usia (Randy, 2019).

Seperti penyanyi Arif Putra, meskipun Ia masih remaja tetapi sudah banyak melantunkan musik Melayu modern. Orang-orang mengenalnya dengan nama Arif. Berdasarkan tanggal lagu dan Insight terpopuler di Youtube, Arif yang bernama panjang Arif Putra masuk ke dalam peringkat 2 terpopuler selama 34 minggu. Arif Putra ini merupakan penyanyi yang berasal dari Padang, Sumatera Barat. Sejalan dengan tempat asalnya, Padang merupakan salah satu tempat berkembangnya orkes melayu yang disebut sebagai lagu-lagu Melayu Deli (Widiarko, 2023).

Kepopuleran inilah yang membuat penulis mengangkat nama Arif Putra sebagai subjek penelitian. Sehingga user baik usia remaja maupun lanjut usia yang menggemari Arif Putra dapat melihat seluruh lagu karya Arif Putra. Akan tetapi user terkadang salah menafsirkan judul lagu yang sering didengar dan dinyanyikan di berbagai website yang disebabkan karena ada terjadinya human error atau typo. Tom Stafford, seorang psikolog di University of Sheffield di Inggris mengemukakan bahwa typo merupakan contoh tipikal dari bentuk kesalahan yang dilakukan

oleh manusia. Otak manusia tidak selalu mampu menangkap hal-hal setepat yang seharusnya (Pratama, 2022).

Terjadinya typo tersebut membuat website lain terkadang tidak dapat mencari judul lagu dengan maksimal, sehingga membuat hambatan bagi user sebagai penikmat musik dalam melakukan pencarian judul lagu. Berdasarkan permasalahan tersebut, penulis melakukan penelitian tentang lagu-lagu ber genre Melayu berbasis website yang dipadukan dengan sistem pencarian yang dapat memudahkan user untuk mencari lagu melayu yang diinginkan dengan menggunakan algoritma yang belum diterapkan pada website lain yakni algoritma levenshtein distance. Algoritma ini berguna untuk membenarkan penulisan ejaan kata, sehingga memudahkan user dalam pencarian lagu Melayu. Adapun algoritma ini dibuat oleh Vladimir Levenshtein pada tahun 1965 untuk menghitung jumlah minimum operasi pengubahan, penghapusan, penyisipan karakter yang digunakan untuk menemukan pemisahan antar dua string (Darnita and Muntahanah, 2019).

II TINJAUAN PUSTAKA

2.1 Studi Pustaka

Adapun Penelitian sebelumnya menurut (Darnita and Muntahanah, 2019) bahwa algoritma *Levenshtein Distance* dapat membantu mengatasi permasalahan pada kesalahan ejaan kata dengan mekanisme penambahan, penyisipan, dan penghapusan. Kemudian menurut (Yulianto et al., 2018) bahwa algoritma *Levenshtein Distance* memiliki tingkat akurasi

pemeriksaan ejaan kata sebesar 86%. Hal ini yang melatar belakangi penulis menggunakan algoritma *Levenshtein Distance* untuk mengoreksi kesalahan ejaan kata pada website yang memiliki *output* berupa judul penelitian "Penerapan Algoritma *Levenshtein Distance* Untuk *Misspelled Word* Pada Pencarian Lagu Melayu Berbasis *Website*".

2.2 Studi Literatur

2.2.1 Information Retrieval

Information retrieval (IR) adalah proses menemukan dokumen dan bahan terstruktur lainnya yang memenuhi kebutuhan informasi koleksi komputer yang besar. (Manning et al., 2009). *Information retrieval* (IR) adalah tentang bagaimana dokumen, halaman Web, katalog online, struktur dan catatan semi-terstruktur, dan objek multimedia ditampilkan, disimpan, diatur, dan tersedia. Cara informasi disajikan dan disatukan harus memudahkan pengguna untuk menemukan informasi yang mereka inginkan. (Baeza Yates and Ribeiro-Neto, 1999). Pada artikel ini, kita akan melihat IR dari dua perspektif: ilmu komputer dan humaniora. Pengindeksan dan pengambilan (IR) adalah cabang ilmu komputer yang berkaitan dengan peningkatan hasil pencarian dengan membuat indeks yang lebih efektif, memproses kueri pengguna dengan cepat, dan menggunakan algoritma peringkat yang lebih baik. Memahami perilaku pengguna, kebutuhan manusia, dan menentukan pemahaman yang memengaruhi struktur dan operasi sistem pengumpulan adalah bagian dari sisi manusia IR. (Budiharto,

2016). Tipe data *Unstructured* dan *semi-structured* juga dapat dicari lebih mudah dengan penggunaan IR.

2.2.2 Algoritma Levenshtein Distance

Menurut Pratama (2016:133) *Levenshtein Distance* adalah matriks *string* yang digunakan untuk mengukur perbedaan antara dua *string*. Nilai *distance* antara dua *string* ini adalah jumlah operasi perubahan yang diperlukan untuk mengubah satu *string* menjadi *string* lainnya dengan jumlah pekerjaan paling sedikit. Ini adalah penyisipan (*insertion*), penghapusan (*deletion*), dan penukaran (*substitution*) (penggantian). *Levenshtein Distance* adalah algoritma yang dapat digunakan untuk menemukan kesamaan antara dua *string* yang dapat dijiplak.

2.2.3 Lagu Melayu

Menurut Ritonga (2008:21) Kesenian musik melayu adalah cara orang-orang untuk secara teratur membagikan perasaan atau pikiran mereka melalui suara. Dalam beberapa hal, suara adalah bagian terpenting dari musik. Suara musik yang bagus berasal dari cara ritme, melodi, dan harmoni semuanya bekerja bersama. Ritme adalah bagaimana suara disatukan dalam waktu, panjang, pendek, dan kecepatan, dan inilah yang membuat setiap karya musik unik.

2.2.4 Website

Menurut Andriyan et al (2020:2) Sebuah *website* atau situs dapat dianggap sebagai sekumpulan halaman yang menampilkan informasi data teks, data gambar diam atau data gambar bergerak, data

animasi, suara, video, dan kombinasi dari semua ini, baik statis maupun bergerak, untuk membuat serangkaian bangunan yang semuanya terhubung ke halaman web.

2.2.5 Uji Statistik

Pengujian Statistik menggunakan Statistik bersifat *Deskriptif*. Statistika *deskriptif* adalah teknik yang digunakan untuk meringkas/menafsirkan data dan menampilkannya dalam bentuk yang dapat dimengerti oleh setiap orang. Hal ini melibatkan proses kuantifikasi dari penemuan suatu fenomena. Berbagai statistik sederhana, seperti rata-rata, dihitung dan ditampilkan dalam bentuk tabel dan grafik. (Wahyudi and Djamaris, 2018)

III METODOLOGI PENELITIAN

3.1.1 Jenis Penelitian

Penelitian ini merupakan penelitian terapan. Tujuan dari penelitian ini adalah untuk menerapkan Algoritma *Levenshtein Distance* Untuk *Misspelled Word* Pada Pencarian Lagu Melayu Berbasis *Website*.

3.1.2 Tempat Dan Waktu Penelitian

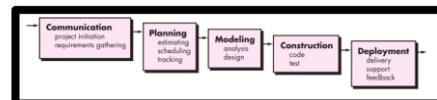
- Tempat : Jalan Melati Maskarebet Palembang
- Waktu : September 2022 - Januari 2023

3.1.3 Metode Pengembangan Sistem

Waterfall adalah metodologi penelitian yang digunakan di sini. Model air terjun adalah teknik yang telah teruji waktu untuk membuat perangkat lunak secara metodis dan

berurutan. Sebenarnya, judul yang tepat untuk model ini adalah "*Linear Sequential Model*". "*Classic life cycle*" atau "*waterfall*" adalah nama umum untuk jenis ini. Meskipun umumnya dianggap kuno, model generik dalam rekayasa perangkat lunak yang disajikan Winston Royce pada tahun 1970-an sejauh ini merupakan model paling populer di lapangan *Software Engineering* (SE). Metodologi yang digunakan di sini adalah metadis dan berurutan. Di *waterfall*, setiap tahap berturut-turut harus menunggu tahap sebelumnya selesai sebelum melanjutkan. (Pressman, 2015:42)

Pressman mengidentifikasi tahapan berikut dalam Model Waterfall:



Gambar 3.1 Waterfall Pressman
Sumber : (Pressman, 2015:42)

- Communication (Project Initiation & Requirements Gathering)*

Komunikasi dengan klien sangat penting sebelum memulai pekerjaan teknis untuk memahami dan mencapai tujuan yang ingin dicapai. Dialog ini membantu menentukan fitur dan fungsi *software* dan sangat penting dalam inisiasi proyek, yang mencakup hal-hal seperti analisis masalah, pengumpulan data, dan definisi fitur. Jurnal, publikasi, dan *world wide web* semuanya dapat ditambang untuk informasi tambahan.

- Planning (Estimating, Scheduling, Tracking)*

Langkah berikutnya adalah perencanaan, yang meliputi pendeskripsian perkiraan untuk tugas-tugas teknis, bahaya, sumber daya, produk kerja, penjadwalan kerja, dan *tracking* proses kerja sistem.

c. *Modeling (Analysis & Design)*

Tahap ini merupakan fase perencanaan dan pemodelan arsitektur sistem, dengan fokus pada perancangan struktur data, arsitektur *software*, *interface* pengguna, dan kerangka kerja algoritmik. Tujuannya adalah untuk memberikan gambaran mental yang lebih lengkap tentang apa yang harus dicapai.

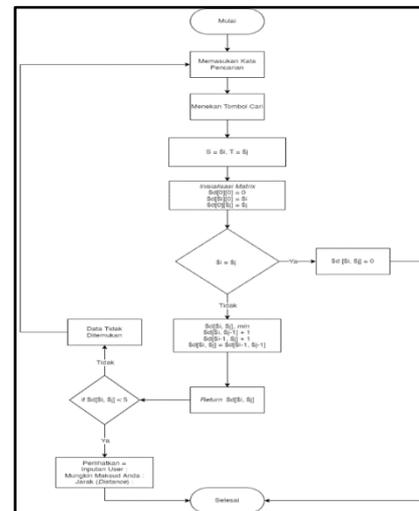
d. *Construction (Code & Test)*

Pada titik ini dalam proses pembangunan, bentuk desain diubah menjadi semacam kode atau bentuk/bahasa yang dapat dibaca mesin. Pengujian dilakukan pada sistem dan kode setelah ditulis. Tujuannya adalah untuk mengidentifikasi potensi masalah sehingga dapat diperbaiki di jalan.

e. *Deployment (Delivery, Support, Feedback)*

Agar sistem dapat terus berjalan dan berkembang sesuai dengan perannya, Tahap *Deployment* adalah saat implementasi *software* ke klien, pemeliharaan *software* secara berkala, perbaikan *software*, evaluasi *software*, dan pengembangan *software* berdasarkan umpan balik yang diberikan. (Pressman, 2015:17).

Prosedur penelitian dalam bentuk alur algoritma levenshtein distance seperti pada gambar



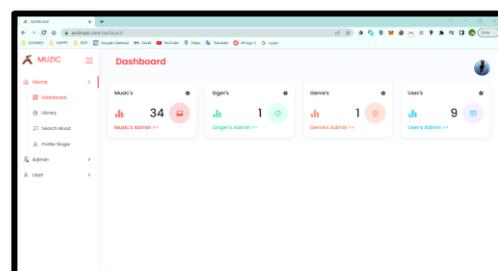
Gambar 3.2 Alur Algoritma *Levenshtein Distance*

IV PEMBAHASAN

4.1 Hasil dan Pembahasan

4.1.1 Implementasi Interface Website

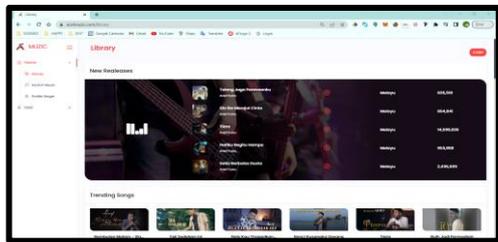
Algoritma *Levenshtein Distance* yang diterapkan pada *website* Lagu Melayu yang telah di hosting dengan link <https://acelmuzic.com/>. Berikut hasil dari penerapan algoritma *Levenshtein Distance* pada *Website* Lagu Melayu:



Gambar 4.1 Interface Search Music

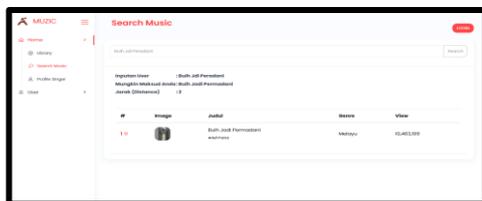
Pada *Interface Dashboard* hanya dapat di akses oleh *Admin* yang memberikan informasi seputar banyaknya data dalam aplikasi *website* dan memiliki tombol detail yang terletak di bawah informasi, dimana jika di klik *text* tersebut akan

langsung masuk ke halaman admin dari *Music's Admin*, *Singer's Admin*, *Genre's Admin*, dan *User's Admin*.



Gambar 4.2 Interface Library

Pada *Interface Library* baik *Admin* ataupun *User* dapat mengaksesnya. Dimana tampilan *Library* ini memiliki beberapa *cardbox* dimana *cardbox* tersebut berisi informasi dari *New Songs*, *Top 8 Trending*, dan *Songs* yang dimana jika di klik salah satu dari *cardbox* tersebut kita akan langsung masuk ke *menu detail* lagu.



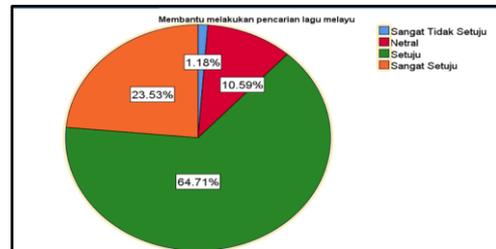
Gambar 4.3 Interface Search Music

Pada *Interface Search Music*, Pengguna dapat mengakses halaman awal ini, di halaman ini terdapat 1(satu) *Text Box* yang di gunakan untuk memasukan *Keyword* yang akan di lakukan proses pencarian menggunakan algoritma *Levenshtein Distance* jika menekan tombol *Search* yang ada di samping *textbox*.

4.2 Analisis Statistik Deskriptif

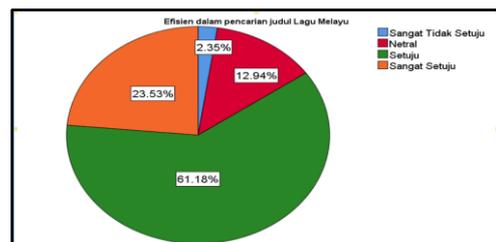
Pengujian ini menggunakan statistik deskriptif. Data yang digunakan untuk pengujian ini diambil dari kuesioner yang diisi oleh orang *random* dari usia > 15 tahun

sampai > 50 tahun. Berikut ini adalah hasil analisis dari beberapa pertanyaan yang didapat dari data responden :



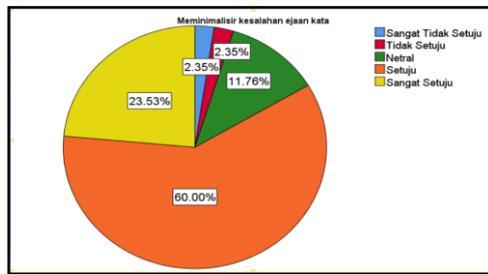
Gambar 4.4 Membantu Melakukan Pencarian Judul Lagu Melayu

Berdasarkan hasil pengelolaan data kuesioner tentang website membantu melakukan pencarian judul lagu melayu diketahui bahwa responden menjawab setuju dengan presentase 65.71% dan sangat setuju dengan 23.53% Hal ini dapat dinyatakan bahwa hampir seluruh responden menjawab setuju bahwa website ini sangat membantu melakukan pencarian lagu melayu karya Arif Putra.



Gambar 4.5 Efisien dalam Pencarian Judul Lagu Melayu

Berdasarkan hasil pengelolaan data kuesioner diketahui bahwa 61.18% dan 23.53% responden menjawab setuju ataupun sangat setuju. Hal ini dapat dinyatakan bahwa hampir seluruh responden menjawab setuju bahwa website lebih efisien dalam pencarian judul lagu melayu karya Arif Putra.



Gambar 4.6 Meminimalisir Kesalahan Ejaan Kata

Berdasarkan hasil pengelolaan data kuesioner diketahui bahwa 60% menjawab setuju dan 23,53% menjawab sangat setuju. Hal ini dapat dinyatakan bahwa hampir seluruh responden menjawab setuju bahwa website membantu dalam meminimalisir kesalahan ejaan kata.

V Kesimpulan

Berdasarkan Penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan algoritma Levenshtein Distance Untuk Misspelled Word Pada Pencarian Lagu Melayu Berbasis Website berhasil diimplementasikan dan terbukti dari hasil responden dengan jumlah 85 orang menyatakan 23.53% sangat setuju dan 60% setuju bahwa *algoritma Levenshtein Distance* yang diterapkan dapat meminimalisir kesalahan ejaan kata pada *website* pencarian Lagu Melayu.

VI DAFTAR PUSTAKA

- Andriyan, Wendy, Sarwan Septiawan Septiawan, and Annisa Aulya. 2020. "Perancangan Website Sebagai Media Informasi Dan Peningkatan Citra Pada SMK Dewi Sartika Tangerang." *Jurnal Teknologi Terpadu* 6(2):79–88. doi: 10.54914/jtt.v6i2.289.
- Baeza Yates, Ricardo, and Berthier Ribeiro-Neto. 1999. *Modern Information Retrieval*. Vol. 1. New York: Addison Wesley.
- Banoë, Pono. 2003. *Kamus Musik*. Yogyakarta: Kanisius.
- Budiharto, Widodo. 2016. "Knowledge-Information-Retrieval.Pdf." 155.
- Darnita, Yulia, and Muntahanah Muntahanah. 2019. "Aplikasi Sistem Pencarian Lagu Menggunakan Algoritma Levenshtein Distance." *Journal of Technopreneurship and Information System (JTIS)* 1(3):61–67. doi: 10.36085/jtis.v1i3.68.
- Jamalus. 1988. *Panduan Pengajaran Buku Pengajaran Musik Melalui Pengalaman Musik*. Jakarta: Proyek Pengembangan Lembaga Pendidikan.
- Manning, D., Prabhakar Raghavan, and Hinrich Schütze. 2009. *Introduction to Modern Information Retrieval (2nd Edition)*. 2nd ed.
- Pratama, B. P. S. A. Pamungkas. 2016. "Analisis Kinerja Algoritma Levenshtein Distance." *Logika* 6(2):131–43.
- Pratama, Kevin Rizky. 2022. "Kenapa Sering Typo Saat Mengetik?" *Kompas.Com*.
- Pratiwi, Ryan Sarah. 2022. "Cross-Genre Bakal Tetap Jadi Tren Musik Di Tahun 2022." *Kompas.Com*.
- Pressman. 2015. *Software Engineering: A Practitioner's Approach (Eight Edit)*. McGraw-Hill Education. Vol. 59. 8th ed. New York: Raghu Srinivasan.
- Randy, Erha. 2019. "Peluang Musik Dangdut Di Era Digital, Ini Kata Rahayu Kertawiguna." *Nalar.Id*.

- Ritonga, Timbun. 2008. "Orkes Musik El Suraya Di Kota Medan (1977- 1990)."
- Wahyudi, David, and Aurino R. A. Djamaris. 2018. *Metode Statistik Untuk Ilmu Dan Teknologi Pangan*.
- Widiarko, Iqbal. 2023. "Sejarah Musik Dangdut Hingga Perkembangannya Di Indonesia." *Celebrities.Id*. Retrieved (<https://www.celebrities.id/read/sejarah-musik-dangdut-60mlio>).
- Yulianto, Muhamad Maulana, Riza Arifudin, and Alamsyah Alamsyah. 2018. "Autocomplete and Spell Checking Levenshtein Distance Algorithm To Getting Text Suggest Error Data Searching In Library." *Scientific Journal of Informatics* 5(1):67–75. doi: 10.15294/sji.v5i1.14148.