JURNAL INFORMATIKA
Volume 4, Nomor 2, September 2025 ISSN 2809-3704

Rancang Bangun Aplikasi Pencarian Influencer Digital Berbasis Web dengan Metode TF-IDF

Feri Andana Putra[1]*, M. Ghofar Rohman[2], Siti Mujilahwati[3]

[1,2]Program Studi Teknik Informatika, Universitas Islam Lamongan, Indonesia

*Corresponding email: feriandanaputra@gmail.com

ABSTRAK – Perkembangan pemasaran digital mendorong meluasnya penggunaan influencer, namun proses pencariannya seringkali tidak efisien. Penelitian ini bertujuan merancang dan membangun aplikasi pencarian influencer digital berbasis web dengan menerapkan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk membantu pengguna menemukan influencer yang relevan. Sistem ini dikembangkan menggunakan PHP dan MySQL dengan model waterfall. Metode TF-IDF dan Cosine Similarity digunakan untuk menghitung dan mengurutkan relevansi antara kata kunci pencarian dengan biodata 100 influencer dari berbagai platform. Hasil implementasi menunjukkan semua fitur utama sistem, termasuk pencarian dan integrasi kontak, berjalan sesuai perancangan. Pengujian fungsionalitas dengan metode black box pada 18 skenario pengujian menunjukkan 100% keberhasilan. Survei kepuasan terhadap 20 responden menghasilkan skor rata-rata 3,87 dari 5, yang mengindikasikan penerimaan pengguna yang baik. Disimpulkan bahwa sistem berbasis TF-IDF ini mampu meningkatkan efisiensi dan akurasi pencarian influencer, serta dapat menjadi solusi efektif bagi pemilik merek untuk menemukan influencer yang tepat sasaran.

Kata Kunci: Influencer, Pencarian Informasi, TF-IDF, Cosine Similarity, Pemasaran Digital.

ABSTRACT – *The growth of digital marketing has led to the widespread use of influencers, yet the search process for them is often inefficient. This research aims to design and build a web-based digital influencer search application by implementing the Term Frequency–Inverse Document Frequency (TF-IDF) method to help users find relevant influencers. The system was developed using PHP and MySQL with a waterfall model. The TF-IDF and Cosine Similarity methods were used to calculate and rank the relevance between search keywords and the biodata of 100 influencers from various platforms. Implementation results show that all main system features, including search and contact integration, function as designed. Black box testing on 18 test scenarios resulted in 100% functional success. A satisfaction survey of 20 respondents yielded an average score of 3.87 out of 5, indicating good user acceptance. It is concluded that this TF-IDF-based system can improve the efficiency and accuracy of influencer searches, serving as an effective solution for brand owners to find targeted influencers.*

***Keywords*:** *Influencer, Information Retrieval, TF-IDF, Cosine Similarity, Digital Marketing.*

PENDAHULUAN

Era digital telah mengubah strategi pemasaran secara fundamental, dengan digital marketing menjadi pendekatan utama untuk menjangkau audiens. Salah satu strategi yang paling efektif dalam ekosistem ini adalah influencer marketing, yang memanfaatkan figur publik digital untuk mempromosikan produk melalui berbagai platform media sosial [1]. Influencer terbukti memiliki pengaruh signifikan terhadap keputusan pembelian konsumen karena hubungan kepercayaan yang mereka bangun dengan pengikutnya, bahkan seringkali lebih dipercaya daripada iklan konvensional [2].

Meskipun potensi influencer marketing sangat besar, proses pencarian influencer yang tepat sasaran masih menjadi tantangan utama. Banyak pemilik merek (brand) melakukan pencarian secara manual, sebuah proses yang tidak sistematis, memakan waktu, dan sangat subjektif [3]. Ketergantungan pada proses manual dan perantara merupakan masalah umum yang menghambat pertumbuhan

usaha, seperti yang dialami oleh pengrajin Tenun Ikat di Lamongan yang pemasarannya terbatas pada pengumpul [4]. Kelemahan metode manual ini seringkali menghasilkan pilihan influencer yang tidak relevan dengan citra merek atau target audiens kampanye [5], sehingga efektivitas pemasaran menurun. Celah ini menunjukkan adanya kebutuhan mendesak akan sebuah sistem terstruktur berbasis data yang dapat mengotomatisasi dan mengoptimalkan proses pencarian influencer.

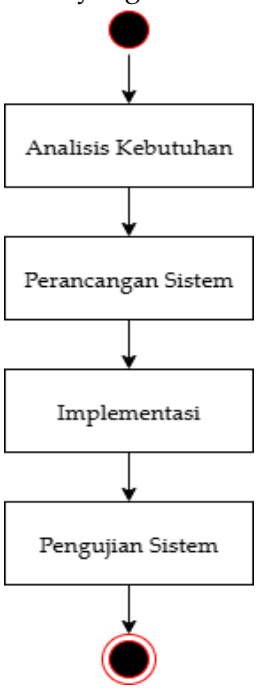
Berbagai penelitian sebelumnya telah membuktikan efektivitas metode TF-IDF dan Cosine Similarity dalam sistem pencarian informasi berbasis teks [6]. Metode ini telah berhasil diterapkan dalam beragam domain, seperti pencarian destinasi wisata [7], rekomendasi artikel berita [8], deteksi similaritas judul tugas akhir [9], sistem rekomendasi lowongan kerja berdasarkan keahlian [10], optimasi pencarian dari hasil web scraping [11], hingga rekomendasi produk pada e-commerce [12]. Studi-studi tersebut secara konsisten menunjukkan bahwa kombinasi kedua metode mampu memberikan hasil pencarian yang akurat. Meskipun demikian, penerapan metode ini dalam konteks spesifik pencarian influencer digital untuk kebutuhan pemasaran masih terbatas, sehingga menjadi celah penelitian (research gap) yang diisi oleh penelitian ini.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan solusi berupa perancangan dan pembangunan aplikasi pencarian influencer digital berbasis web. Untuk mengukur relevansi antara profil influencer dengan kueri pencarian pengguna, penelitian ini menerapkan metode Term Frequency–Inverse Document Frequency (TF-IDF). Oleh karena itu, tujuan utama penelitian ini adalah untuk merancang, membangun, dan mengevaluasi fungsionalitas sebuah sistem pencarian influencer yang efisien dan relevan, dengan mengimplementasikan metode TF-IDF sebagai inti dari mesin pencarinya.

METODE

Penelitian ini menggunakan metode rekayasa perangkat lunak (*software engineering*) untuk merancang, membangun, dan mengevaluasi sebuah aplikasi pencarian *influencer* digital berbasis web. Proses pengembangan sistem mengadopsi model *waterfall*, yang dipilih karena alurnya terstruktur dan berurutan, sesuai untuk proyek dengan ruang lingkup yang telah terdefinisi dengan jelas di awal [13].

Alur penelitian dilaksanakan melalui beberapa tahapan sistematis yang bertujuan untuk membangun sebuah sistem informasi pendukung keputusan [14]. Tahap pertama adalah analisis kebutuhan, yang mencakup identifikasi masalah pada proses pencarian *influencer* manual dan studi literatur untuk menentukan pendekatan solusi. Tahap kedua adalah perancangan sistem, di mana arsitektur aplikasi, struktur basis data, dan antarmuka pengguna dimodelkan. Tahap ketiga adalah implementasi, yang meliputi pengkodean aplikasi menggunakan PHP dan MySQL, serta integrasi algoritma pencarian. Tahap terakhir adalah pengujian sistem, yang terdiri dari pengujian fungsionalitas dan evaluasi kepuasan pengguna untuk mengukur efektivitas aplikasi yang telah dibangun [15].



Gambar 1. Diagram Alur Penelitian

2.1. Pengumpulan Data

Sistem ini menggunakan dataset simulasi yang terdiri dari 100 data profil *influencer* digital. Dataset ini dibuat secara manual untuk merepresentasikan informasi kunci seperti biodata singkat, kategori konten (misalnya, Kecantikan, Teknologi, Kuliner), dan jumlah pengikut di berbagai platform media sosial (Instagram, TikTok, YouTube, X). Pendekatan ini dipilih karena belum tersedianya sumber data publik yang terstruktur dan seragam untuk profil *influencer*. Setiap biodata dalam dataset berfungsi sebagai dokumen teks yang menjadi korpus untuk dianalisis oleh algoritma pencarian.

2.2. Implementasi Algoritma Pencarian

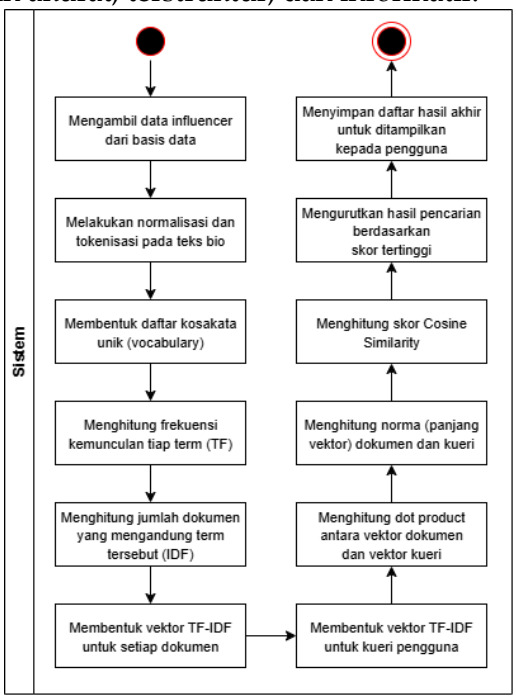
Inti dari fungsionalitas pencarian pada sistem ini adalah penerapan gabungan metode *Term Frequency–Inverse Document Frequency* (TF-IDF) dan *Cosine Similarity*, yang merupakan pendekatan standar dalam bidang *information retrieval* (IR) untuk menemukan dan memeringkat relevansi dokumen teks [16]. Alur kerja algoritma ini adalah sebagai berikut:

1. Input Kueri: Pengguna memasukkan satu atau lebih kata kunci pencarian.
2. Preprocessing Data: Baik kueri dari pengguna maupun teks biodata dari 100 *influencer* dalam basis data akan melalui tahap *preprocessing* sederhana yang mencakup *case folding* (mengubah semua teks menjadi huruf kecil) dan *tokenization* (memecah teks menjadi kata-kata individual atau *term*).
3. Pembobotan TF-IDF: Sistem menghitung bobot TF-IDF untuk setiap *term* dalam setiap dokumen (biodata *influencer*). Perhitungan ini terdiri dari tiga bagian:
   a. *Term Frequency* (TF), yang mengukur frekuensi kemunculan sebuah *term* t dalam sebuah dokumen d.
   $$TF(t,d) = \frac{f_{t,d}}{\Sigma_k f_{k,d}}$$
   b. *Inverse Document Frequency* (IDF), yang mengukur keunikan sebuah *term* t di seluruh koleksi dokumen N.
   $$IDF(t) = \log\left(\frac{N}{n_t}\right)$$
   c. Skor akhir TF-IDF adalah perkalian antara TF dan IDF.
   $$TFIDF(t,d) = TF(t,d) \times IDF(t)$$
4. Perhitungan Kemiripan: Setelah setiap dokumen dan kueri direpresentasikan sebagai vektor bobot TF-IDF, tingkat kemiripan antara vektor kueri A dan vektor dokumen B dihitung menggunakan *Cosine Similarity*.
   $$\text{CosineSimilarity}(A,B) = \frac{A.B}{\|A\| \times \|B\|}$$
   Nilai yang mendekati 1 menandakan tingkat relevansi yang tinggi.
5. Pemeringkatan Hasil: Sistem akan mengurutkan daftar *influencer* berdasarkan skor *Cosine Similarity* tertinggi, kemudian menampilkannya kepada pengguna.

Proses ini dimulai ketika pengguna memasukkan kata kunci pencarian melalui form yang tersedia di halaman utama sistem. Setelah menerima input, sistem mengambil seluruh data *influencer* dengan status akun "Aktif" dari basis data. Data bio dari masing-masing *influencer* kemudian diproses melalui tahapan normalisasi dan tokenisasi untuk membentuk representasi kata-kata yang akan dianalisis. Kata kunci dari pengguna juga diproses dengan cara yang sama agar dapat dibandingkan secara sepadan dengan data bio.

Selanjutnya, sistem menghitung *Term Frequency* (TF) dari setiap kata dalam bio, dan *Inverse Document Frequency* (IDF) berdasarkan distribusi kemunculan kata di seluruh dokumen. Hasil TF dan IDF ini dikalikan untuk membentuk bobot TF-IDF per term pada masing-masing dokumen dan pada kata kunci kueri. Setelah vektor TF-IDF terbentuk, sistem menghitung skor kemiripan menggunakan *Cosine Similarity* untuk mengukur seberapa dekat vektor kata kunci dengan vektor bio *influencer*. Skor ini mencerminkan tingkat relevansi antara kata kunci dan deskripsi bio yang tersedia. Hasil pencarian kemudian

diurutkan berdasarkan skor kemiripan tertinggi, dan ditampilkan kepada pengguna dalam bentuk daftar *influencer* yang paling relevan. Proses ini memungkinkan pencarian dilakukan tidak hanya dengan mencocokkan kata secara literal, tetapi juga dengan mempertimbangkan pentingnya kata dalam konteks keseluruhan data yang tersedia, sehingga menghasilkan hasil pencarian yang lebih akurat, terstruktur, dan informatif.



Gambar 2. *Activity Diagram* TF-IDF dan *Cosine Similarity*

2.3. Prosedur Pengujian Sistem

Evaluasi sistem dilakukan melalui dua pendekatan:

1. Pengujian Fungsional: Menggunakan metode *black box testing* untuk memverifikasi apakah semua fitur aplikasi berjalan sesuai dengan yang diharapkan. Pendekatan pengujian fungsional ini juga digunakan secara efektif dalam pengembangan sistem lain untuk memastikan aplikasi berjalan sesuai kebutuhan sebelum diimplementasikan [17]. Pengujian ini mencakup 18 skenario, meliputi fungsi registrasi, login, pencarian, hingga fitur kontak via WhatsApp
2. Pengujian Pengguna: Melibatkan survei kepuasan pengguna yang disebar kepada 20 responden. Kuesioner ini menggunakan skala Likert untuk mengukur persepsi pengguna terhadap kemudahan penggunaan, efektivitas fitur pencarian, dan tampilan antarmuka sistem.

HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dari implementasi dan pengujian sistem pencarian *influencer*, diikuti dengan pembahasan yang menginterpretasikan temuan tersebut dalam konteks tujuan penelitian.

3.1. Hasil Pengujian Fungsional

Pengujian fungsionalitas dilakukan dengan metode *black box testing* terhadap 18 skenario pengujian yang mencakup seluruh alur kerja sistem. Pengujian ini memverifikasi fungsi-fungsi inti mulai dari autentikasi pengguna, fitur pencarian, manajemen profil, hingga fitur administrasi. Hasil pengujian menunjukkan bahwa 100% skenario berhasil dan semua fitur dapat berjalan sesuai dengan yang diharapkan tanpa ditemukan galat (*error*) fungsional. Ringkasan hasil pengujian fungsionalitas disajikan pada Tabel 1.

Tabel 1. Ringkasan Hasil Pengujian

No.	Fitur Yang Diuji	Contoh Skenario	Hasil
1	Autentikasi dan Manajemen Akun	Registrasi pengguna baru, Login, Edit Profil.	Berhasil
2	Fitur Pencarian (Inti)	Pencarian dengan kata kunci, penggunaan filter.	Berhasil
3	Interaksi dan Fitur Tambahan	Kontak via WhatsApp, pelaporan akun.	Berhasil
4	Manajemen Admin	Kelola pengguna, kelola laporan, lihat survei.	Berhasil

3.2. Kinerja Algoritma dan Simulasi Pencarian

Untuk mendemonstrasikan kinerja algoritma, dilakukan simulasi pencarian dengan kueri "makeup luar negeri". Sistem berhasil memproses kueri dengan alur TF-IDF dan *Cosine Similarity* untuk menemukan profil *influencer* yang paling relevan. Sebagai contoh, sebuah profil dengan bio yang mengandung kata "makeup" (4 kali), "luar" (1 kali), dan "negeri" (1 kali) dianalisis. Dengan total korpus 97 dokumen (*influencer* aktif), sistem menghitung bobot TF-IDF untuk setiap kata dan menghasilkan skor *Cosine Similarity* sebesar 0.344 untuk profil tersebut. Nilai ini menunjukkan tingkat relevansi yang signifikan, membuktikan bahwa sistem tidak hanya mencocokkan kata kunci secara literal tetapi juga mempertimbangkan bobot dan keunikan kata dalam keseluruhan data. Langkah perhitungan yang dilakukan sistem meliputi:

1. Tokenisasi & Normalisasi
   Sistem mengubah seluruh teks menjadi huruf kecil, lalu memecahnya menjadi kata-kata unik (token).
   Query terms:
   ["makeup", "luar", "negeri"]
   Sampel salah satu bio *influencer* setelah diproses:
   ["saya", "senang", "mengeksplorasi", "berbagai", "tren", "makeup", "dari", "dalam", "dan", "luar", "negeri", "dari", "warna", "lipstik", "terbaru", "hingga", "gaya", "riasan", "mata", "ala", "korea", "saya", "selalu", "tertarik", "mencoba", "dan", "membagikan", "hasilnya", "saya", "juga", "membuat", "konten", "tutorial", "makeup", "natural", "untuk", "pemula", "yang", "mudah", "diikuti", "penonton", "saya", "rata-rata", "adalah", "remaja", "dan", "dewasa", "muda", "yang", "sedang", "belajar", "mengekspresikan", "diri", "melalui", "makeup", "saya", "percaya", "bahwa", "makeup", "bukan", "untuk", "menutupi", "tapi", "untuk", "memperkuat", "kepercayaan", "diri", "setiap", "minggu", "saya", "membagikan", "video", "singkat", "yang", "menunjukkan", "tips", "praktis", "dan", "inspiratif"]
   Jumlah total token = 50 kata
2. Menghitung TF (*Term Frequency*)
   Term Frequency dihitung sebagai jumlah kemunculan kata dalam dokumen. Berdasarkan dokumen di atas:
   - "makeup" muncul sebanyak 4 kali
   - "luar" muncul sebanyak 1 kali
   - "negeri" muncul sebanyak 1 kali

   Rata-rata TF:
   $$\text{TF Rata} - \text{rata} = \frac{4+1+1}{3} = 2$$
3. Menghitung *Inverse Document Frequency* (IDF)
   Berdasarkan isi tabel pengguna di database, jumlah dokumen N = 97 bio *influencer* aktif. Setelah dilakukan pemindaian token pada semua bio aktif, terdapat 5 dokumen dengan *term* makeup, terdapat 2 dokumen dengan *term* luar, dan terdapat 2 dokumen dengan *term* negeri.
   - IDF (makeup) = log (97 / 5) = 2,965
   - IDF (luar) = log (97 / 2) = 3,882
   - IDF (negeri) = log (97 / 2) = 3,882

   Rata-rata IDF:
   $$\text{IDF Rata} - \text{rata} = \frac{2{,}965 + 3{,}882 + 3{,}882}{3} = 3{,}576$$
   Jadi, ketiga term memiliki IDF yang bervariasi sesuai dengan seberapa jarang kemunculannya dalam korpus. Kata "makeup" relatif lebih umum, sementara "luar" dan "negeri" lebih jarang, sehingga nilainya lebih tinggi.
4. Menghitung Vektorisasi dan *Cosine Similarity*
   Kata kueri diubah ke dalam vektor bobot TF-IDF.
   Vektor dokumen:
   V_d = [TF-IDF("makeup"), TF-IDF("luar"), TF-IDF("negeri")]
   = [4×IDF, 1×IDF, 1×IDF]
   = [15,228, 3,807, 3,807]
   Vektor kueri (query TF = 1 untuk semua term):
   V_q = [1×IDF, 1×IDF, 1×IDF]

= [3,807, 3,807, 3,807]

Lalu menghitung:

a. Dot Product:
∑(Vd[i]×Vq[i])
= (15,228×3,807) + (3,807×3,807) + (3,807×3,807)
= 58,03+14,5+14,5
= 87,03

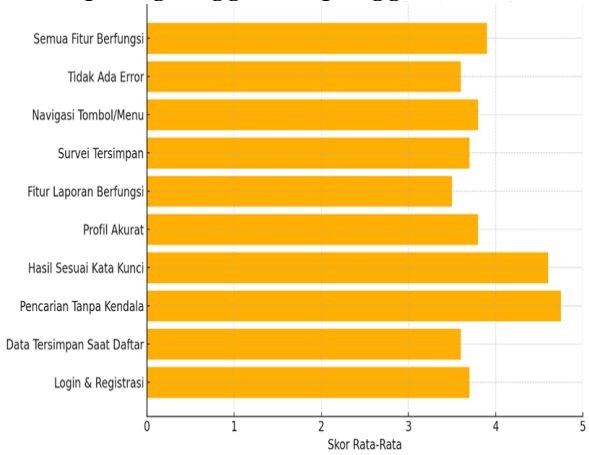
b. Norm Vektor:
$\|V_d\| = \sqrt{15{,}228^2 + 3{,}807^2 + 3{,}807^2} = 16{,}28$
$\|V_q\| = \sqrt{3{,}807^2 + 3{,}807^2 + 3{,}807^2} = 6{,}597$

c. *Cosine Similarity*:
$\cos(\theta) = \frac{87{,}03}{16{,}28 \times 6{,}597} = 0{,}344$

Nilai *cosine similarity* sebesar 0.344 menunjukkan bahwa bio ini cukup relevan dengan kueri yang diberikan, karena memuat semua kata kunci dan memiliki TF-IDF tinggi untuk kata dominan seperti "makeup". Nilai ini juga menegaskan bahwa metode ini tidak hanya mengandalkan keberadaan kata, melainkan memperhitungkan distribusi dan dominasi kata kunci terhadap keseluruhan isi dokumen.

3.3. Hasil Evaluasi Pengguna

Evaluasi dari sisi pengguna dilakukan melalui survei kepuasan yang diisi oleh 20 responden. Hasil survei menunjukkan skor rata-rata kepuasan keseluruhan sebesar 3,87 dari skala 5, yang mengindikasikan penerimaan yang baik dari pengguna. Secara spesifik, fitur pencarian sebagai fungsi utama sistem mendapatkan penilaian tertinggi. Pernyataan "Saya dapat menggunakan fitur pencarian influencer tanpa kendala" memperoleh skor rata-rata 4,75, dan pernyataan "Hasil pencarian yang muncul sesuai dengan kata kunci yang saya masukkan" memperoleh skor 4,60. Hasil ini divalidasi lebih lanjut melalui visualisasi pada Gambar 1, yang menunjukkan bahwa fungsionalitas pencarian (diwakili oleh pertanyaan-pertanyaan awal) secara konsisten dinilai paling tinggi oleh pengguna.



Gambar 3. Grafik Rata-Rata Hasil Survei Pengguna

Pembahasan

Hasil pengujian fungsional dan evaluasi pengguna secara konsisten menunjukkan bahwa aplikasi pencarian *influencer* yang dikembangkan telah berhasil mencapai tujuannya. Keberhasilan 100% pada pengujian *black box* mengonfirmasi bahwa sistem stabil secara teknis dan semua fitur yang dirancang dapat beroperasi dengan baik. Temuan ini diperkuat oleh skor kepuasan pengguna yang tinggi (3,87 dari 5), yang menandakan bahwa aplikasi tidak hanya berfungsi, tetapi juga diterima dengan baik oleh pengguna akhir.

Penilaian yang sangat positif terhadap fitur pencarian (skor rata-rata >4,60) secara langsung memvalidasi efektivitas penerapan metode TF-IDF dan *Cosine Similarity*. Hasil ini sejalan dengan penelitian-penelitian sebelumnya [6], [7], [8] yang juga menunjukkan keberhasilan TF-IDF dalam domain pencarian teks yang berbeda. Penelitian ini membuktikan bahwa metode tersebut juga sangat relevan dan akurat ketika diterapkan pada domain baru, yaitu pencarian profil *influencer* untuk kebutuhan pemasaran digital.

Lebih lanjut, sistem ini secara signifikan meningkatkan efisiensi proses pencarian dibandingkan dengan metode manual. Jika pencarian manual melalui berbagai platform media sosial dapat memakan waktu beberapa

menit per profil dan bersifat subjektif, sistem ini mampu menyajikan daftar kandidat yang relevan dalam hitungan detik. Transformasi dari proses manual ke sistematis ini merupakan kontribusi praktis utama dari penelitian ini, karena dapat menghemat waktu dan beban kognitif pengguna secara drastis.

Meskipun hasil yang diperoleh sangat positif, penelitian ini memiliki keterbatasan utama pada penggunaan dataset simulasi. Untuk pengembangan selanjutnya, sistem dapat ditingkatkan dengan mengintegrasikan teknik *web scraping* untuk mengambil data *influencer* secara *real-time* dari media sosial. Langkah ini akan membawa sistem menuju implementasi teknologi data yang lebih besar, namun juga akan dihadapkan pada tantangan nyata seperti isu ketersediaan data, standarisasi data dari berbagai platform, serta masalah privasi data pengguna [18]. Selain itu, fungsionalitas sistem dapat diperluas dari sekadar pencarian relevansi menjadi deteksi pola untuk verifikasi kredibilitas. Dengan menerapkan teknik *data mining* seperti *Clustering* (misalnya *K-Means*) dan *Classification* (misalnya J48), sistem dapat dikembangkan untuk mengidentifikasi pola perilaku tidak otentik dari seorang *influencer* [19]. Pendekatan ini serupa dengan yang digunakan untuk mendeteksi dan memprediksi kejahatan siber di sektor *marketplace*, sehingga dapat menambah nilai signifikan bagi pemilik merek untuk memastikan mereka bekerja sama dengan *influencer* yang tidak hanya relevan, tetapi juga kredibel [19].

KESIMPULAN

Penelitian ini telah berhasil merancang dan membangun aplikasi pencarian influencer digital berbasis web dengan mengimplementasikan metode TF-IDF dan Cosine Similarity secara efektif. Keberhasilan ini dibuktikan melalui hasil pengujian fungsional black box pada 18 skenario yang menunjukkan 100% fungsionalitas sistem berjalan sesuai rancangan, serta melalui evaluasi pengguna yang menghasilkan skor kepuasan rata-rata 3,87 dari 5, dengan penilaian tertinggi pada aspek relevansi hasil pencarian. Dengan demikian, dapat disimpulkan bahwa aplikasi yang dikembangkan mampu menjawab seluruh tujuan penelitian dengan menyediakan solusi yang efisien dan terstruktur, yang secara signifikan mempercepat proses penemuan influencer yang relevan dibandingkan dengan metode manual.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah memberikan bantuan dan dukungan selama proses penelitian, sehingga artikel ini dapat diselesaikan dengan baik.

REFERENSI

[1] C. Lou and S. Yuan, “Influencer Marketing: How Message Value and Credibility Affect Consumer Trust of Branded Content on Social Media,” *J. Interact. Advert.*, vol. 19, no. 1, pp. 58–73, Jan. 2019.

[2] M. De Veirman, V. Cauberghe, and L. Hudders, “Marketing through Instagram influencers: the impact of number of followers and product-influencer match on brand attitude and purchase intentions,” *Int. J. Advert.*, vol. 36, no. 5, pp. 798–828, Aug. 2017.

[3] Sudha and K. Sheena, “Impact of Influencers in Consumer Decision Process: The Fashion Industry,” *SCMS J. Indian Manag.*, vol. 14, no. 3, pp. 14–30, Jul. 2017.

[4] M. R. Zamroni, M. Sholihin, S. Mujilahwati, A. A. Bettaliyah, R. Wardhani, D. E. Anggraini, and M. Z. Qomaruddin, “Peningkatan Penjualan pada Pengrajin Ikat Tenun di Desa Parengan, Kecamatan Maduran, Kabupaten Lamongan dengan E-Commerce Berbasis Web,” in *Prosiding*

Seminar Nasional Sains dan Teknologi, vol. 12, no. 1, Nov. 2022, pp. 438-443.

[5] S. Khamis, L. Ang, and R. Welling, "Self-branding, 'micro-celebrity' and the rise of Social Media Influencers," *Celebrity Stud.*, vol. 8, no. 2, pp. 191–208, Apr. 2017.

[6] R. A. Rasyid dan D. H. U. Ningsih, "Penerapan Algoritma TF-IDF dan Cosine Similarity untuk Query Pencarian Pada Dataset Destinasi Wisata," *J. Teknol. Inf. dan Komun.*, vol. 8, no. 1, pp. 1–10, 2024.

[7] N. Silalahi dan G. L. Ginting, "Rekomendasi Berita Berkaitan dengan Menerapkan Algoritma Text Mining dan TF-IDF," *Bull. Comput. Sci. Res.*, vol. 3, no. 4, pp. 276–282, Jun. 2023.

[8] A. Amrulloh dan I. F. Adam, "Sistem Pencarian Similaritas Judul Tugas Akhir Menggunakan Metode TF-IDF," *J. CoreIT*, vol. 7, no. 2, pp. 74–81, Dec. 2021.

[9] S. A. Wibowo, H. A. Santoso, and I. P. Windasari, "Implementation of TF-IDF and Cosine Similarity for Job Recommendation System based on Skills," *J. RESTI (Rekayasa Sist. dan Teknol. Inf.)*, vol. 7, no. 4, pp. 834–841, Aug. 2023.

[10] A. F. Hidayatullah, I. S. Supono, and A. Lukito, "Hoax News Detection on Social Media using TF-IDF and Naive Bayes Classifier," *J. Ilm. Tek. Elektro Komput. dan Inform. (JITEKI)*, vol. 9, no. 3, pp. 582–591, Sep. 2023.

[11] D. P. Lestari, R. Sarno, and D. A. Kurniady, "Product recommendation system for e-commerce using TF-IDF and user-based collaborative filtering," *IJEECS (Indones. J. Electr. Eng. Comput. Sci.)*, vol. 25, no. 1, pp. 449–458, Jan. 2022.

[12] R. S. Pressman and B. R. Maxim, *Software Engineering: A Practitioner's Approach*, 9th ed. New York: McGraw-Hill Education, 2020.

[13] K. C. Laudon and J. P. Laudon, *Management Information Systems: Managing the Digital Firm*, 16th ed. Harlow: Pearson, 2020.

[14] I. Sommerville, *Software Engineering*, 10th ed. Harlow: Pearson, 2016.

[15] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge: Cambridge University Press, 2008.

[16] A. Rohman, R. Y. Perkasa, A. S. Hidaytullah, and M. G. Rohman, "Implementasi Metode Waterfall Pada Rancang Bangun Sistem Pengarsipan Surat Berbasis Web," *Generation Journal*, vol. 6, no. 2, pp. 134-143, Jul. 2022.

[17] N. Fajriyah, W. Setiawan, E. Dewi, and T. Duha, "Implementasi Teknologi Big Data di Era Digital," *JURNAL INFORMATIKA*, vol. 1, no. 1, pp. 1-7, Mar. 2022.

[18] J. Waruwu, W. Hadinata, S. Febriyani, and R. Wijayanti, "Data Mining Techniques in Detecting and Predicting Cyber Crimes In Marketplace Sector," *JURNAL INFORMATIKA*, vol. 1, no. 1, pp. 8-11, Mar. 2022.

[19] S. Mujilahwati and L. D. Windasari, "Implementasi Metode k-Nearest Neighbor (k-NN) untuk Memprediksi Penjualan Buah di Indonesia berbasis Website," in *PROSIDING SEMINAR NASIONAL TEKNOLOGI DAN SAINS TAHUN 2024*, vol. 3, Kediri, 2024, pp. 7-14.