

PEMANFAATAN TEXT MINING PADA SISTEM REKOMENDASI PEMILIHAN ALGORITMA ENKRIPSI TERBAIK UNTUK KEAMANAN DATA MENGGUNAKAN *CONTENT BASED FILTERING*

M. Iqbal ¹⁾, Mukti Qamal ²⁾, Yesy Afrilia ³⁾

^{1,2,3)} Teknik Informatika, Universitas Malikussaleh, Lhokseumawe, Indonesia
Email korespondensi: iqbal.210170038@mhs.unimal.ac.id ¹⁾

Abstract: The selection of an appropriate encryption algorithm is a crucial challenge in maintaining data security, yet many users face difficulties due to limited technical knowledge and the vast amount of literature required for analysis. This research aims to develop a recommendation system based on Content-Based Filtering (CBF) integrated with text mining techniques to facilitate faster, more accurate, and scientifically based encryption algorithm selection. The research methodology involved collecting 300 scientific articles from the Garuda portal of the Ministry of Education and Culture using web scraping techniques, followed by preprocessing steps such as tokenization, stopword removal, and case folding. The data was represented using the TF-IDF method, while the similarity between user requirements and the literature was calculated using Cosine Similarity. The results show that the most dominant algorithms discussed in the literature are RSA (52 articles), AES (40 articles), and RC4 (25 articles). The developed recommendation system proved effective because it was able to provide relevant suggestions tailored to specific needs, such as recommending AES as the primary algorithm for the "fast algorithm for sensitive data" criterion. The use of text mining and CBF in this system provides a more effective solution compared to manual methods because the system is able to perform literature analysis in depth and automatically, thereby minimizing the risk of selection which can have an impact on computational inefficiency and data security vulnerabilities.

Keywords: Encryption Algorithm, Cosine Similarity, Content-Based Filtering, Data Security, Text Mining, TF-IDF.

Abstrak: Pemilihan algoritma enkripsi yang tepat merupakan tantangan krusial dalam menjaga keamanan data, namun banyak pengguna menghadapi kesulitan karena keterbatasan pengetahuan teknis dan banyaknya literatur yang harus dianalisis. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi berbasis Content-Based Filtering (CBF) yang terintegrasi dengan teknik text mining untuk membantu proses pemilihan algoritma enkripsi secara lebih cepat, tepat, dan berbasis data ilmiah. Metodologi penelitian yang dilakukan meliputi pengumpulan 300 artikel ilmiah dari portal Garuda Kemdikbud melalui teknik web scraping, yang kemudian melalui tahapan preprocessing seperti tokenisasi, stopword removal, dan case folding. Data tersebut direpresentasikan menggunakan metode TF-IDF, sementara tingkat kemiripan antara kebutuhan pengguna dan literatur dihitung menggunakan Cosine Similarity. Hasil penelitian menunjukkan distribusi algoritma yang paling dominan dibahas dalam literatur adalah RSA (52 artikel), AES (40 artikel), dan RC4 (25 artikel). Sistem

rekomendasi yang dibangun terbukti efektif karena mampu memberikan saran yang relevan sesuai kebutuhan spesifik, seperti merekomendasikan AES sebagai algoritma utama untuk kriteria “algoritma cepat untuk data sensitif”. Pemanfaatan text mining dan CBF dalam sistem ini memberikan solusi yang lebih efektif dibandingkan metode manual karena sistem mampu melakukan analisis literatur secara mendalam dan otomatis, sehingga meminimalkan risiko kesalahan pemilihan yang dapat berdampak pada inefisiensi komputasi dan kerentanan keamanan data.

Kata kunci: Algoritma Enkripsi, Cosine Similarity, Content-Based Filtering, Keamanan Data, Text Mining, TF-IDF.

1. Pendahuluan

Era digital telah mentransformasi hampir seluruh aspek kehidupan, termasuk kebutuhan akan keamanan data yang semakin mendesak. Dengan meningkatnya volume data sensitif yang ditransmisikan dan disimpan secara elektronik, pemilihan algoritma enkripsi yang tepat menjadi fondasi kritis dalam melindungi informasi dari ancaman siber (Bhaskara et al., 2024). Namun, kompleksitas teknis dan banyaknya pilihan algoritma yang tersedia—masing-masing dengan karakteristik keamanan, kecepatan, dan efisiensi sumber daya yang berbeda—sering kali membingungkan bagi pengguna, administrator sistem, dan bahkan pengembang yang tidak memiliki spesialisasi di bidang kriptografi (Fahriani & Rosyid, 2019).

Permasalahan utama terletak pada keterbatasan waktu dan sumber daya untuk menganalisis literatur ilmiah yang terus bertambah. Pemilihan algoritma sering kali hanya mengandalkan popularitas atau rekomendasi umum, tanpa pertimbangan mendalam terhadap konteks spesifik data yang akan dilindungi. Pendekatan seperti ini berisiko menimbulkan celah keamanan, inefisiensi pemrosesan, atau pemborosan sumber daya komputasi. Oleh karena itu, diperlukan sebuah mekanisme yang dapat memberikan rekomendasi berbasis data yang obyektif, sistematis, dan dapat dipertanggungjawabkan.

Sistem rekomendasi berbasis konten (*Content-Based Filtering* atau CBF) menawarkan solusi dengan menganalisis karakteristik intrinsik dari item—dalam hal ini, algoritma enkripsi—dan mencocokkannya dengan profil kebutuhan pengguna (Ridhwanullah et al., 2024). Keunggulan CBF terletak pada kemampuannya mengatasi masalah *cold-start*, di mana pengguna baru tetap bisa mendapatkan rekomendasi relevan meski tanpa riwayat interaksi sebelumnya, karena rekomendasi didasarkan pada fitur item itu sendiri, bukan pada preferensi pengguna lain (Theo Santoso et al., 2024). Untuk mengisi basis pengetahuan sistem, teknik *text mining* diterapkan untuk mengekstrak dan mengolah informasi teknis dari kumpulan literatur ilmiah seperti jurnal, konferensi, dan artikel penelitian dalam jumlah besar (Afrillia et al., 2022). Kombinasi *text mining* dan CBF ini diharapkan dapat menghasilkan sistem yang mampu menganalisis lanskap penelitian terkini dan memberikan saran algoritma yang paling sesuai.

Penelitian ini bertujuan untuk: (1) Mengembangkan model implementasi *text mining* untuk mengekstrak informasi teknis dari literatur ilmiah secara otomatis; (2) Membangun sistem rekomendasi berbasis *Content-Based Filtering* yang mampu memberikan saran algoritma enkripsi berdasarkan masukan pengguna; dan (3) Menganalisis akurasi sistem melalui pengujian *Cosine Similarity* dan evaluasi metrik *Information Retrieval*. Dengan demikian, penelitian ini diharapkan dapat berkontribusi dalam mempermudah proses pengambilan keputusan teknis di bidang keamanan informasi, sekaligus menjadi studi aplikatif mengenai integrasi *text mining* dan sistem rekomendasi dalam domain yang bersifat teknis.

2. Kajian Kepustakaan

2.1. Sistem Rekomendasi dan Content-Based Filtering

Sistem rekomendasi adalah teknik komputasi yang dirancang untuk menyarankan item yang relevan kepada pengguna berdasarkan preferensi atau karakteristik tertentu (Qamal et al., 2021). Salah satu pendekatannya adalah *Content-Based Filtering* (CBF), yang bekerja dengan menganalisis atribut atau deskripsi item untuk menemukan kemiripan dengan preferensi pengguna yang diketahui (Ridhwanullah et al., 2024). Dalam konteks penelitian ini, item adalah algoritma enkripsi, dan atributnya diekstrak dari deskripsi teknis dalam literatur ilmiah. Metode ini cocok untuk domain dengan item yang memiliki deskripsi tekstual kaya, seperti literatur akademik, karena tidak bergantung pada data interaksi pengguna lain sehingga terhindar dari masalah *data sparsity* (Saputra et al., 2025). Dua komponen kunci dalam CBF adalah representasi fitur dan pengukuran kesamaan. Representasi fitur sering dilakukan menggunakan model *bag-of-words* yang diperkuat dengan pembobotan TF-IDF, sedangkan pengukuran kesamaan umumnya menggunakan *Cosine Similarity* untuk menghitung kedekatan vektor antara profil pengguna dan item (Pernanda & Hakiki, 2021).

2.2. Text Mining dan Pemrosesan Teks

Text mining mengacu pada proses mengekstrak pola dan pengetahuan bermakna dari data teks tidak terstruktur (G. R. Kumar, 2020). Dalam sistem rekomendasi, *text mining* berperan untuk mengubah dokumen teks (seperti abstrak jurnal) menjadi representasi terstruktur yang dapat diproses mesin. Tahapan utamanya meliputi *preprocessing* teks: *case folding* (penyeragaman huruf), *tokenization* (pemecahan teks menjadi kata), *stopword removal* (penghapusan kata umum), dan *stemming* (pengembalian kata ke bentuk dasar) (Afrillia et al., 2022). Hasil *preprocessing* kemudian direpresentasikan secara numerik, salah satunya dengan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), yang memberikan bobot tinggi pada kata yang sering muncul di dokumen tertentu tetapi jarang muncul di seluruh koleksi dokumen, sehingga mengidentifikasi kata kunci yang signifikan (Jain et al., 2024).

2.3. Algoritma Enkripsi dan Kriptografi

Kriptografi adalah ilmu dan seni menjaga kerahasiaan dan keutuhan informasi melalui teknik penyandian (Schneier, 2017). Algoritma enkripsi modern umumnya dikategorikan menjadi dua: simetris (menggunakan kunci yang sama untuk enkripsi dan dekripsi, contoh: AES, DES) dan asimetris (menggunakan pasangan kunci publik dan privat, contoh: RSA, ECC). Pemilihan algoritma bergantung pada berbagai faktor seperti jenis data, kebutuhan kecepatan, tingkat keamanan, dan lingkungan komputasi. Literatur ilmiah dari waktu ke waktu terus mengevaluasi dan membandingkan performa berbagai algoritma ini dalam skenario yang berbeda, yang menjadi sumber pengetahuan berharga bagi sistem rekomendasi yang dikembangkan.

2.4. Penelitian Terkait

Beberapa penelitian telah mengintegrasikan *text mining* dan CBF dalam berbagai domain. Asshiddiq dan Wulandhari (2024) berhasil menerapkannya untuk rekomendasi hotel, sementara Nouvalina dan Hati (2024) menggunakannya untuk rekomendasi produk *skincare*. Dalam domain pendidikan, Fahri (2020) memanfaatkan *text mining* untuk rekomendasi judul tugas akhir. Namun, penelitian yang secara spesifik memanfaatkan *text mining* dan CBF untuk rekomendasi algoritma enkripsi—dengan sumber data berupa literatur ilmiah—masih terbatas. Penelitian ini berupaya mengisi celah tersebut dengan fokus pada domain keamanan informasi yang bersifat teknis.

3. Metode Penelitian

3.1. Desain Penelitian dan Alur Kerja

Penelitian ini menggunakan pendekatan penelitian dan pengembangan (*R&D*) dengan tahapan sistematis yang dirancang untuk membangun sistem rekomendasi fungsional. Alur kerja utama penelitian digambarkan pada Gambar 1, yang mencakup: (1) Pengumpulan Data; (2) *Preprocessing* dan *Text Mining*; (3) Ekstraksi Fitur dan Pembobotan; (4) Pembangunan dan Implementasi Sistem Rekomendasi; serta (5) Evaluasi Sistem.

3.2. Pengumpulan Data (Data Collection)

Data primer penelitian berupa artikel ilmiah dikumpulkan melalui teknik *web scraping* dari portal *Garba Rujukan Digital* (Garuda) Kementerian Pendidikan dan Kebudayaan. Pencarian dilakukan dengan kata kunci seperti “algoritma enkripsi”, “keamanan data”, “AES”, “RSA”, dan “kriptografi”. Sebanyak 300 artikel yang memenuhi kriteria inklusi—berupa publikasi jurnal nasional terakreditasi dan prosiding konferensi dalam rentang tahun 2019–2024—berhasil dikumpulkan. Metadata setiap artikel (judul, abstrak, penulis, kata kunci) disimpan dalam format terstruktur (JSON) untuk diproses lebih lanjut.

3.3. Preprocessing Data dan Text Mining

Tahap *preprocessing* diterapkan pada teks abstrak artikel untuk membersihkan dan menstandarisasi data. Prosesnya meliputi:

1. *Case Folding*: Mengubah semua karakter menjadi huruf kecil.
2. *Tokenization*: Memisahkan teks menjadi unit kata (*tokens*) menggunakan *library* NLTK.
3. *Stopword Removal*: Menghapus kata penghubung dan kata umum yang tidak informatif (menggunakan daftar *stopwords* bahasa Indonesia dan Inggris).
4. *Stemming*: Mengurangi kata ke bentuk dasarnya menggunakan algoritma *Porter Stemmer*.

Teks yang telah dibersihkan kemudian digabungkan dengan informasi judul dan kata kunci artikel untuk membentuk korpus teks yang komprehensif.

3.4. Ekstraksi Fitur dengan TF-IDF dan Perhitungan Kesamaan

Korpus teks kemudian direpresentasikan ke dalam ruang vektor menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). Metode ini menghitung bobot setiap kata (*term*) dengan mempertimbangkan frekuensi kemunculannya dalam satu dokumen (TF) dan kelangkaannya di seluruh koleksi dokumen (IDF). Representasi ini menghasilkan matriks TF-IDF di mana setiap baris mewakili sebuah dokumen (artikel) dan setiap kolom mewakili sebuah kata dengan bobot tertentu.

Ketika pengguna memasukkan *query* (misalnya, “algoritma cepat untuk data sensitif”), *query* tersebut melalui proses *preprocessing* yang sama dan kemudian diubah menjadi vektor TF-IDF menggunakan *vocabulary* yang telah dibangun dari korpus. Tingkat kesamaan (*similarity*) antara vektor *query* dan vektor setiap dokumen dalam korpus dihitung menggunakan metrik *Cosine Similarity*. Rumusnya adalah:

$$\text{similarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|}$$

di mana A adalah vektor *query* dan B adalah vektor dokumen. Nilai kesamaan berkisar antara 0 (tidak mirip) hingga 1 (sangat mirip).

3.5. Implementasi Sistem Rekomendasi

Sistem rekomendasi dibangun dengan arsitektur berbasis web menggunakan *framework* Flask (Python) di sisi *backend* dan HTML/CSS/Bootstrap di sisi *frontend*. Alur kerja sistem adalah:

1. Pengguna memasukkan deskripsi kebutuhan keamanan data ke dalam antarmuka web.
2. Sistem memproses *input* tersebut menjadi vektor TF-IDF.
3. Sistem menghitung *Cosine Similarity* antara vektor *query* dan seluruh vektor dokumen dalam basis data.
4. Sistem mengurutkan dokumen (artikel) berdasarkan nilai kesamaan tertinggi.
5. Sistem mengidentifikasi algoritma enkripsi yang dibahas dalam artikel-artikel peringkat teratas.
6. Sistem menampilkan daftar rekomendasi algoritma, dilengkapi dengan nilai kesamaan, judul artikel pendukung, dan ringkasan kelebihan/kekurangan yang diekstrak.

3.6. Evaluasi Sistem

Kinerja sistem dievaluasi menggunakan metrik *Information Retrieval* (IR). Dua metrik utama yang digunakan adalah:

1. *Precision@k*: Proporsi item yang direkomendasikan dalam *k* teratas yang relevan. Dalam pengujian ini, digunakan $k = 3(Precision@3)$.
2. *Average Precision (AP)*: Rata-rata nilai *precision* pada setiap titik di mana item relevan ditemukan dalam daftar peringkat, memberikan gambaran kualitas peringkat secara keseluruhan. Evaluasi dilakukan dengan membuat sejumlah *query* uji yang merepresentasikan kebutuhan pengguna nyata. Relevansi setiap rekomendasi dinilai secara manual oleh peneliti berdasarkan kesesuaian algoritma yang direkomendasikan dengan konteks *query*.

4. Hasil dan Pembahasan

4.1. Hasil Pengumpulan dan Analisis Data Awal

Dari 300 artikel yang dikumpulkan, analisis distribusi topik menunjukkan bahwa algoritma RSA merupakan yang paling banyak dibahas (52 artikel), diikuti oleh AES (40 artikel), dan RC4 (25 artikel). Algoritma lain seperti Blowfish, ElGamal, dan DES juga muncul dengan frekuensi yang signifikan. Hasil ini menggambarkan tren penelitian di Indonesia yang masih berfokus pada algoritma-algoritma fundamental dan yang banyak diadopsi dalam standar.

4.2. Hasil Implementasi dan Pengujian Sistem

Sistem berhasil diimplementasikan dalam bentuk aplikasi web yang interaktif. Gambar 2 menunjukkan antarmuka awal sistem, tempat pengguna memasukkan *query*. Gambar 3 menunjukkan contoh tampilan hasil rekomendasi untuk *query*: “Saya butuh algoritma untuk mengamankan data folder”.

Tabel 1. Contoh Hasil Rekomendasi untuk Query Tertentu.

Peringkat	Algoritma Direkomendasikan	Nilai Cosine Similarity	Artikel Pendukung (Contoh)
1	Playfair Cipher	0.39	“Keamanan Data Folder dengan Algoritma Klasik...”
2	AES (Advanced Encryption Standard)	0.38	“Implementasi AES pada Sistem Keamanan Berkas...”
3	Kombinasi RSA & SHA-1	0.36	“Hybrid Cryptography untuk Integritas Data...”

Hasil di atas menunjukkan sistem mampu memetakan *query* bahasa natural yang relatif umum ke algoritma spesifik. Rekomendasi tidak hanya berupa nama algoritma, tetapi juga disertai bukti literatur yang mendukung.

4.3. Hasil Evaluasi Kinerja Sistem

Evaluasi dilakukan terhadap 10 *query* uji yang divariasikan. Hasil rata-rata evaluasi sistem adalah sebagai berikut:

- *Precision@3*: 1.00 (100%). Ini berarti untuk semua *query* uji, ketiga rekomendasi teratas yang diberikan sistem dinilai relevan oleh penilai.
- *Average Precision (AP)*: 0.71.

Nilai *Precision@3* yang sempurna mengindikasikan bahwa sistem sangat handal dalam memberikan rekomendasi awal yang tepat sasaran. Nilai AP sebesar 0.71 menunjukkan bahwa secara keseluruhan, sebagian besar dokumen relevan berhasil ditempatkan pada peringkat yang tinggi di dalam daftar hasil. Hasil ini membuktikan efektivitas pipeline *text mining* (TF-IDF) dan mekanisme pencocokan (*Cosine Similarity*) yang diterapkan.

4.4. Pembahasan

Keberhasilan sistem ini ditopang oleh beberapa faktor. Pertama, penggunaan *text mining* yang memadai memungkinkan ekstraksi fitur-fitur teknis kunci dari literatur. Kedua, metode CBF dengan TF-IDF dan *Cosine Similarity* terbukti efektif untuk domain di mana item dapat dideskripsikan secara tekstual. Sistem ini secara efektif menjembatani kesenjangan antara kebutuhan pengguna non-ahli dengan kompleksitas pengetahuan teknis di bidang kriptografi.

Namun, terdapat beberapa keterbatasan. *Query* yang terlalu umum atau ambigu (misalnya, “enkripsi yang aman”) terkadang menghasilkan rekomendasi yang kurang spesifik. Selain itu, sistem saat ini belum mempertimbangkan faktor performa real-time atau benchmark algoritma dalam rekomendasinya, karena sumber data hanya bersifat deskriptif dari literatur. Heterogenitas dalam penulisan abstrak jurnal juga dapat mempengaruhi konsistensi ekstraksi fitur.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan, dapat disimpulkan bahwa:

1. Integrasi teknik *text mining* (dengan TF-IDF) dan *Content-Based Filtering* (dengan *Cosine Similarity*) berhasil diterapkan untuk membangun sistem rekomendasi pemilihan algoritma enkripsi.
2. Sistem yang dikembangkan mampu memberikan rekomendasi yang relevan dan kontekstual berdasarkan analisis terhadap 300 artikel ilmiah, dengan kinerja evaluasi yang sangat baik ($Precision@3 = 1.00$, $AP = 0.71$).
3. Sistem ini berpotensi menjadi alat bantu yang efektif bagi pengguna dengan latar belakang teknis terbatas untuk mengambil keputusan awal dalam memilih algoritma enkripsi, sekaligus mempercepat proses *literature review* bagi peneliti atau mahasiswa.

5.2. Saran

Untuk pengembangan penelitian selanjutnya, disarankan:

1. Memperkaya basis data dengan sumber literatur internasional (seperti IEEE Xplore, SpringerLink) untuk mendapatkan cakupan algoritma dan analisis yang lebih global.
2. Mengintegrasikan teknik pemrosesan bahasa alami (NLP) yang lebih maju seperti word embeddings (Word2Vec, GloVe) atau model berbasis transformer (BERT) untuk meningkatkan pemahaman semantik terhadap query dan konten dokumen.
3. Mengembangkan sistem menjadi model hybrid yang menggabungkan Content-Based Filtering dengan Collaborative Filtering atau teknik knowledge-based, misalnya dengan memasukkan parameter teknis objektif (seperti ukuran kunci, kecepatan benchmark) ke dalam profil item.
4. Menambahkan fitur interaktif, seperti *feedback* pengguna terhadap rekomendasi, untuk menyempurnakan model secara bertahap (*relevance feedback*).

Daftar Kepustakaan

- Afrillia, Y., Rosnita, L., & Siska, D. (2022). Analisis sentimen ciutan Twitter terkait penerapan Permendikbudristek Nomor 30 Tahun 2021 menggunakan TextBlob dan *support vector machine*. *G-Tech: Jurnal Teknologi Terapan*, 6(2), 387–394. <https://doi.org/10.33379/gtech.v6i2.1778>
- Asshiddiq, A., & Wulandhari, L. A. (2024). Sistem rekomendasi hotel dengan ekstraksi fitur deskripsi menggunakan metode *text mining* dan *content-based filtering*. *Kesatria*, 5(4), 1636–1646.

<https://doi.org/10.30645/kesatria.v5i4.486>

- Bhaskara, I. M. A., Ariawan, M. P. A., Peling, I. B. A., & Prayudha, I. P. A. (2024). Studi literatur: Analisa perbandingan teori tentang tingkat keamanan antar algoritma simetris. *Jurnal Bangkit Indonesia*, 13(1), 40–45. <https://doi.org/10.52771/bangkitindonesia.v13i1.278>
- Fahri, M. U. (2020). Sistem rekomendasi pengambilan keputusan judul tugas akhir mahasiswa jurusan komputer menggunakan metode *text mining*. *Sentral Jurnal AMKI*, 1(1).
- Fahriani, N., & Rosyid, H. (2019). Implementasi teknik enkripsi dan dekripsi file video menggunakan algoritma Blowfish. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 6(6), 697. <https://doi.org/10.25126/jtiik.2019661465>
- Jain, S., Jain, S. K., & Vasal, S. (2024). An effective TF-IDF model to improve the text classification performance. Dalam *2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT)* (hlm. 1–4). IEEE. <https://doi.org/10.1109/CSNT60213.2024.10545818>
- Kumar, G. R. (2020). A summarization on *text mining* techniques for information extracting from applications and issues. *Journal of Mechanics of Continua and Mathematical Sciences*, spl5(1), 324–332. <https://doi.org/10.26782/jmcms.spl.5/2020.01.00026>
- Nouvalina, D. A., & Hati, K. (2024). Sistem rekomendasi produk *skin care* berdasarkan permasalahan kulit wajah dengan metode *content-based filtering*. *Jurnal Teknik Informatika STMIK Antar Bangsa*, 10(2), 60–67. <https://doi.org/10.51998/jti.v10i2.586>
- Pernanda, A. Y., & Hakiki, M. (2021). Penerapan *cosine similarity* sebagai metode pengukuran *similarity index* pada sistem pengajuan judul skripsi. *Jurnal Inovasi Pendidikan dan Teknologi Informasi*, 2(2), 93–99. <https://doi.org/10.52060/pti.v2i02.626>
- Qamal, M., Fajriana, F., & Mardhatillah, M. (2021). Metode Naïve Bayes untuk menentukan rekomendasi tempat wisata terbaik di Aceh. *TECHSI – Jurnal Teknik Informatika*, 13(1), 81. <https://doi.org/10.29103/techsi.v13i1.3132>
- Ridhwanullah, D., Kumarahadi, Y. K., & Raharja, B. D. (2024). *Content-based filtering* pada sistem rekomendasi buku informatika. *Jurnal Ilmiah SINUS*, 22(2), 57. <https://doi.org/10.30646/sinus.v22i2.840>
- Saputra, V. S., Ridwan, A., & Pratama, T. G. (2025). Rancang bangun sistem rekomendasi buku berbasis *item-based collaborative filtering* menggunakan algoritma *K-nearest neighbors*. *Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, 19(1), 40–47. <https://doi.org/10.47111/JTI>
- Schneier, B. (2017). *Applied cryptography: Protocols, algorithms, and source code in C* (20th anniversary ed.). Wiley.

Theo Santoso, D., Atina, V., & Hartanti, D. (2024). Prototipe sistem rekomendasi film Indonesia menggunakan pendekatan *content-based filtering* dan metode *vector space model*. *Infotek: Jurnal Informatika dan Teknologi*, 7(2), 444–455. <https://doi.org/10.29408/jit.v7i2.26083>