

RANCANG BANGUN SISTEM INVENTARIS BARANG BERBASIS WEB MENGGUNAKAN ALGORITMA TF-IDF

R.Farras Yudritya Zikirrullah

Teknik Informatika, Universitas Muhammadiyah, Pontianak

E-mail: 201220066@unmuhpnk.ac.id

ABSTRAK

Pengelolaan inventaris barang yang masih dilakukan secara manual menimbulkan berbagai kendala, khususnya pada proses pencarian barang yang tidak efisien dan kurang akurat. Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem inventaris barang berbasis web dengan menerapkan algoritma Term Frequency-Inverse Document Frequency (TF-IDF) dan *Cosine Similarity* guna meningkatkan relevansi pencarian barang berdasarkan deskripsi teks. Data inventaris yang digunakan berasal dari Hotel Dangau dan terdiri dari deskripsi barang operasional hotel. Sistem dikembangkan menggunakan PHP dan MySQL serta diuji menggunakan metode *black box*. Hasil pengujian menunjukkan bahwa sistem mampu menampilkan hasil pencarian berdasarkan tingkat relevansi dengan waktu respons yang cepat, namun tingkat akurasi pencarian masih tergolong rendah. Rendahnya akurasi disebabkan oleh keterbatasan TF-IDF dalam menangani sinonim, makna kontekstual, serta ketidakteraturan deskripsi data. Meskipun demikian, sistem yang dikembangkan dapat menjadi solusi awal dalam pengelolaan inventaris berbasis web dan dapat dikembangkan lebih lanjut dengan pendekatan semantik.

Kata kunci

Sistem Inventaris, TF-IDF, Cosine Similarity, Information Retrieval, Web.

ABSTRACT

Inventory management systems that are still handled manually often cause inefficiencies, particularly in item retrieval processes. This study aims to design and implement a web-based inventory system by applying the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm and Cosine Similarity to improve search relevance based on textual descriptions. The inventory data were obtained from Hotel Dangau and consist of operational item descriptions. The system was developed using PHP and MySQL and tested using the black-box method. The results show that the system can retrieve relevant items efficiently with fast response time; however, the search accuracy remains relatively low. This limitation is mainly caused by TF-IDF's inability to handle synonyms, contextual meanings, and inconsistent item descriptions. Nevertheless, the proposed system can serve as an initial solution for web-based inventory management and may be enhanced further using semantic-based approaches.

Keywords

Inventory System, TF-IDF, Cosine Similarity, Information Retrieval, Web.

1. PENDAHULUAN

Pengelolaan inventaris barang merupakan aspek penting dalam menunjang kelancaran operasional hotel. Ketepatan pencatatan dan kemudahan pencarian data inventaris berpengaruh langsung terhadap efisiensi kerja dan pengambilan keputusan. Pada Hotel Dangau, pencatatan dan pencarian inventaris masih dilakukan secara manual, sehingga rentan terhadap kesalahan pencatatan, keterlambatan pencarian data, serta ketidaksesuaian informasi stok.

Proses pencarian barang yang masih bergantung pada kode inventaris menyebabkan kesulitan bagi staf gudang, terutama ketika jumlah data inventaris semakin meningkat dan deskripsi barang tidak terstandarisasi. Untuk mengatasi permasalahan

tersebut, diperlukan sistem pencarian berbasis teks yang mampu menampilkan barang berdasarkan relevansi deskripsi.

Salah satu metode yang umum digunakan dalam sistem temu kembali informasi adalah Term Frequency-Inverse Document Frequency (TF-IDF). Algoritma ini memberikan bobot pada kata kunci berdasarkan tingkat kepentingannya dalam dokumen. Dengan mengombinasikan TF-IDF dan Cosine Similarity, sistem dapat mengukur tingkat kemiripan antara query pengguna dan deskripsi barang. Penelitian ini bertujuan untuk mengimplementasikan algoritma TF-IDF pada sistem inventaris barang berbasis web serta menganalisis faktor-faktor yang memengaruhi tingkat akurasi pencarian. dan menemukan pola-pola tersembunyi. Algoritma FP-Growth merupakan salah satu algoritma data mining yang dapat digunakan untuk menemukan pola kombinasi datasets dalam data bencana kabut asap di Sumatera Selatan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan pengembangan sistem berbasis web dengan metode *Information Retrieval*. Objek penelitian adalah data inventaris barang Hotel Dangau yang terdiri dari nama barang, deskripsi, jumlah stok, dan tanggal masuk. Sistem dikembangkan menggunakan PHP dengan basis data MySQL.

2. 1 Tahapan Penelitian

Tahapan penelitian meliputi identifikasi masalah, studi literatur, analisis kebutuhan, perancangan sistem, implementasi, serta pengujian sistem. Sistem dikembangkan menggunakan bahasa pemrograman PHP dengan basis data MySQL, dan diimplementasikan dalam lingkungan web server XAMPP.

2. 2 Text Preprocessing

Sebelum dilakukan perhitungan TF-IDF, data deskripsi barang dan query pengguna melalui tahapan *text preprocessing* untuk meningkatkan akurasi pencarian. Tahapan yang dilakukan meliputi:

- Case Folding*, yaitu mengubah seluruh huruf menjadi huruf kecil.
- Tokenizing*, yaitu memecah teks menjadi kata-kata (token).
- Stopword Removal*, yaitu menghapus kata-kata umum yang tidak memiliki makna penting.
- Stemming*, yaitu mengubah kata ke bentuk dasarnya.

2. 3 Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF digunakan untuk memberikan bobot pada setiap kata dalam deskripsi barang. Nilai *Term Frequency (TF)* dihitung berdasarkan jumlah kemunculan kata dalam satu dokumen, sedangkan *Inverse Document Frequency (IDF)* mengukur tingkat keunikan kata dalam seluruh koleksi dokumen.

Nilai Term Frequency (TF) dihitung menggunakan Persamaan (1):

$$TF(t, d) = \frac{f(t, d)}{N(d)}$$

di mana $f(t, d)$ adalah jumlah kemunculan term t dalam dokumen d , dan $N(d)$ adalah jumlah total kata dalam dokumen tersebut.

Nilai Inverse Document Frequency (IDF) dihitung menggunakan Persamaan (2):

$$IDF(t) = \log_{10} \left(\frac{N}{df(t)} + 1 \right)$$

di mana N adalah jumlah seluruh dokumen dan $df(t)$ adalah jumlah dokumen yang mengandung term t . Bobot TF-IDF diperoleh dengan mengalikan nilai TF dan IDF, sebagaimana ditunjukkan pada Persamaan (3):

$$TF - IDF(t, d) = TF(t, d) \times IDF(t)$$

2. 4 Cosine Similarity

Untuk mengukur tingkat kemiripan antara query pengguna dan deskripsi barang, digunakan metode *Cosine Similarity*. Metode ini menghitung sudut kosinus antara dua vektor TF-IDF, sehingga menghasilkan nilai kemiripan antara 0 dan 1. Tingkat kemiripan antara query dan dokumen dihitung menggunakan *Cosine Similarity* seperti pada Persamaan (4):

$$Similarity = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Nilai kemiripan tertinggi menunjukkan bahwa dokumen tersebut paling relevan dengan query yang dimasukkan oleh pengguna.

2.5 Pengujian Sistem

Pengujian sistem dilakukan menggunakan metode *black box* untuk memastikan seluruh fungsi sistem berjalan sesuai dengan kebutuhan pengguna. Selain itu, dilakukan evaluasi hasil pencarian untuk menganalisis tingkat akurasi sistem dalam menampilkan barang yang relevan berdasarkan deskripsi.

3. HASIL DAN PEMBAHASAN

Sistem inventaris barang berbasis web berhasil diimplementasikan dengan fitur utama berupa pencarian barang berdasarkan deskripsi menggunakan algoritma TF-IDF dan *Cosine Similarity*. Data inventaris yang digunakan terdiri dari 85 dokumen deskripsi barang, yang mencakup berbagai kategori barang operasional Hotel Dangau. Sistem mampu menerima input query dari pengguna dan menampilkan daftar barang yang diurutkan berdasarkan tingkat relevansi tertinggi.

Table 1 Table Awal Data

No	Query	Deskripsi Barang (Dokumen)
1	alat penerangan kamar	Lampu LED untuk penerangan kamar
2	pendingin udara ruangan	AC split untuk kamar VIP
3	alat untuk membersihkan toilet	Sikat untuk membersihkan toilet
4	perangkat kerja staf kantor	Komputer LG – Perangkat kerja utama staf
5	media hiburan untuk tamu	Televisi untuk kamar tamu
6	alat kebersihan lantai	Pel untuk membersihkan lantai
7	perangkat dapur hotel	Panci masak untuk dapur hotel
8	alat untuk laundry hotel	Mesin cuci untuk laundry hotel
9	akses kontrol kamar	IPAQ HP (Kunci Kamar) – Sistem kontrol akses pintu kamar
10	perlengkapan makan tamu	Set peralatan makan untuk restoran hotel

Tahapan *case folding*, *tokenizing*, *stopword removal*, dan *stemming* menghasilkan kata-kata dasar yang digunakan pada proses pembobotan. Proses ini terbukti mengurangi noise data dan menyederhanakan variasi kata dalam deskripsi barang.

Tabel 2 Hasil case Folding

Query	lampu penerangan kamar
Dokumen 1	lampu led untuk penerangan kamar
Dokumen 2	ac split untuk kamar vip
Dokumen 3	sapu untuk membersihkan lantai
Dokumen 4	kipas angin dinding untuk kamar tamu
Dokumen 5	televisi untuk kamar tamu

Tabel 3 Hasil Tokenizing

Query	lampu penerangan kamar
Dokumen 1	lampu led untuk
Dokumen 2	ac split untuk
Dokumen 3	Sapu untuk membersihkan
Dokumen 4	Kipas angin dinding
Dokumen 5	Televisi untuk kamar

Tabel 4 Hasil Stopword removal

Query	lampu penerangan kamar
Dokumen 1	lampu led penerangan
Dokumen 2	Ac split kamar
Dokumen 3	sapu membersihkan lantai
Dokumen 4	kipas angin dinding
Dokumen 5	televisi kamar tamu

Tabel 5 Hasil Stemming

Query	lampu terang kamar
Dokumen 1	lampu led terang
Dokumen 2	ac split kamar
Dokumen 3	sapu bersih lantai
Dokumen 4	kipas angin dinding
Dokumen 5	televisi kamar tamu

Setelah preprocessing, dilakukan perhitungan Term Frequency (TF) untuk setiap term dalam dokumen. Tabel 6 menunjukkan hasil perhitungan TF untuk beberapa term utama.

Table 6 Hasil Perhitungan TF

Term (t)	TF => $\frac{n(t,d)}{n(d)}$					
	Q	D1	D2	D3	D4	D5
alat	1	0	0	0	0	0
terang	1	1	0	0	0	0
kamar	1	1	1	0	1	1
lampu	0	1	0	0	0	0
led	0	1	0	0	0	0
ac	0	0	1	0	0	0
split	0	0	1	0	0	0
vip	0	0	1	0	0	0
sapu	0	0	0	1	0	0
Total Kata :						
N(d)	4	4	4	3	3	2

Nilai IDF dihitung menggunakan jumlah dokumen sebanyak 85. Term yang jarang muncul memperoleh nilai IDF lebih tinggi, sedangkan term yang sering muncul di banyak dokumen memiliki bobot lebih rendah. Hasil perhitungan IDF ditunjukkan pada Tabel 7.

Tabel 7. Proses perhitungan nilai IDF

Term (t)	df	N	IDF(t) = $\log_{10} (N/df+1)$				
			D1	D2	D3	D4	D5
alat	1	85	1.4314	0	0	0	0
terang	1	85	1.4314	1.4314	0	0	0
kamar	5	85	1.2041	1.2041	1.2041	0	1.2041
lampu	1	85	1.4314	1.4314	0	0	0
led	1	85	1.4314	1.4314	0	0	0
ac	1	85	1.4314	0	1.4314	0	0
split	1	85	1.4314	0	1.4314	0	0
vip	1	85	1.4314	0	1.4314	0	0
sapu	1	85	1.4314	0	0	1.4314	0

Nilai TF dan IDF kemudian dikalikan untuk memperoleh bobot TF-IDF. Bobot ini merepresentasikan tingkat kepentingan suatu kata dalam dokumen terhadap keseluruhan koleksi.

Table 8 Hasil Pembobotan TF-IDF

Term (t)	$W = TF(t, d) \times IDF(t)$					
	Q	D1	D2	D3	D4	D5

alat	1.4314	0.3333	0.4771	0	0	0
terang	1.4314	0.3333	0.4771	0.25	0.3579	0
kamar	1.2041	0.3333	0.4013	0.25	0.3010	0.25
lampu	1.4314	0	0	0.25	0.3579	0
led	1.4314	0	0	0.25	0.3579	0
ac	1.4314	0	0	0	0	0.25
split	1.4314	0	0	0	0	0.25
vip	1.4314	0	0	0	0	0.25
sapu	1.4314	0	0	0	0	0
Dot Product vec 1 & vec 2 = $A_i \cdot B_i$						
Dot Product sum ($Q \cdot D$)	0.2914	0.1207	0.0000	0.1207	0.1207	

Setelah bobot TF-IDF diperoleh, dilakukan perhitungan *Cosine Similarity* antara query dan setiap dokumen. Nilai similarity digunakan untuk menentukan urutan hasil pencarian. Tabel 9 menunjukkan hasil perhitungan similarity.

Table 9 Hasil Perhitungan Similarity

<u>Kemiripan</u>	$Similarity = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$	<u>Persentase</u>
Sim (D1, Q)	$\frac{0.2914}{0.7850 \times 0.4676} = \frac{0.2914}{0.3669} = 0,7942$	79.42%
Sim (D2, Q)	$\frac{0.1207}{0.7850 \times 0.3009} = \frac{0.1207}{0.2363} = 0,5107$	51.07%
Sim (D3, Q)	$\frac{00.000}{0.7850 \times 0.4771} = \underline{\underline{0.0000}}$	0.00%
Sim (D4, Q)	$\frac{0.1207}{0.7850 \times 0.3009} = \frac{0.1207}{0.2363} = 0,5107$	51.07%
Sim (D5, Q)	$\frac{0.1207}{0.7850 \times 0.3009} = \frac{0.1207}{0.2363} = 0,5107$	51.07%

Berdasarkan hasil diatas dokumen D1 merupakan dokumen yang paling relevan terhadap query "alat penerangan kamar" karena mengandung dua term yang sama, yaitu "terang" dan "kamar", dengan tingkat similarity tertinggi sebesar 79,42%, sementara dokumen D2, D4, dan D5 memiliki relevansi sedang karena hanya memiliki kesamaan pada term "kamar", sedangkan dokumen D3 tidak relevan karena tidak mengandung satu pun term yang sesuai dengan query.

4. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, berikut adalah beberapa kesimpulan utama:

- a. Metode TF-IDF terbukti efektif dalam membantu pencarian barang pada sistem inventaris berbasis web, khususnya dalam lingkungan hotel yang memiliki banyak item dengan nama dan fungsi yang beragam.
- b. Proses text preprocessing seperti case folding, tokenizing, stopwords removal, dan stemming sangat berperan dalam meminimalkan noise pada data teks, sehingga meningkatkan akurasi pencarian berdasarkan deskripsi pengguna.
- c. Sistem mampu menghitung kemiripan (similarity) antara query pengguna dan deskripsi barang secara otomatis menggunakan metode cosine similarity, dan mengurutkan hasil pencarian berdasarkan nilai kemiripan tertinggi.
- d. Dari hasil pengujian black box, semua fungsi utama seperti pencarian barang, validasi input, dan penampilan hasil relevan berjalan sesuai harapan dan tanpa error.
- e. Sistem ini dapat membantu petugas inventaris dan staf hotel dalam mencari barang dengan lebih cepat dan efisien, meskipun tidak mengetahui nama persis barang, melainkan hanya berdasarkan deskripsi umum.

5. DAFTAR PUSTAKA

- A. Noviriandini, D. Ayu Ambarsari, and F. Aprian, "Information Retrieval System Pada Pencarian File Dokumen Berbasis Teks Dengan Metode Vector Space Model", [Online]. Available: <https://ejournal.bsi.ac.id/ejurnal/index.php/reputasi>
- M. Kharisma Dina Putri, dan Noora Qotrun Nada, G. Pusat Lantai, and J. Sidodadi Timur, *RANCANG BANGUN SISTEM INFORMASI INVENTARIS BARANG DI KANTOR BADAN PUSAT STATISTIK KABUPATEN PATI BERBASIS WEBSITE*, vol. 5. 2020.
- A. Hanani, "ALGORITMA TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF)." [Online]. Available: <https://www.researchgate.net/publication/376720501>
- A. Rawung, "Aplikasi Inventaris Barang Berbasis Web," no. 17013063, pp. 1–89, 2020.
- Y. Sari, A. Rizky Baskara, P. B. Prakoso, and N. Royani, "PERBANDINGAN METODE PEMBOBOTAN TF-RF DAN TF-IDF DENGAN DIKOMBINASIKAN DENGAN WEIGHTED TREE SIMILARITY UNTUK SISTEM REKOMENDASI BUKU", doi: 10.25126/jtiik.202295709.
- M. Aditya, N. Kalimantan, F. Agustin, U. P. Utama, and U. P. Utama, "Implentasi Information Retrieval System Menggunakan Metode Latent Semantic," *JID (Jurnal Info Digit.*, vol. 1, no. 2, pp. 744–755, 2023. "13650047".
- R. Al Rasyid, D. Handayani, and U. Ningsih, "Penerapan Algoritma TF-IDF dan Cosine Similarity untuk Query Pencarian Pada Dataset Destinasi Wisata," *J. Teknol. Inf. dan Komunikasi*, vol. 8, no. 1, p. 2024, 2024, doi: 10.35870/jti.
- N. Widyasari, Y. Palinggi, and M. Hermanto Penulis adalah Mahasiswa dan Dosen Pada Fakultas Ekonomi dan Bisnis Universitas Kutai Kartanegara, "ANALISIS PENERAPAN METODE FIRST IN FIRST OUT (FIFO."
- S. Mariko, "APLIKASI WEBSITE BERBASIS HTML DAN JAVASCRIPT UNTUK MENYELESAIKAN FUNGSI INTEGRAL PADA MATA KULIAH KALKULUS," *J. Inov. Teknol. Pendidik*, vol. 6, no. 1, pp. 80–91, 2019, doi: 10.21831/jitp.v6.1.22280.
- M. Raharjo, M. Napiah, and R. S. Anwar, "Perancangan Sistem Informasi Dengan PHP Dan MYSQL Untuk Pendaftaran Sekolah Di Masa Pandemi," 1045. [Online]. Available:

<http://jurnal.bsi.ac.id/index.php/co-science>

- A. Oktarini Sari and E. Nuari, "RANCANG BANGUN SISTEM INFORMASI PERSEDIAAN BARANG BERBASIS WEB DENGAN METODE FAST(FRAMEWORK FOR THE APPLICATIONS)," vol. 13, no. 2, p. 261, 2017.
- N. Alamsyah, M. Rasyidan, I. Kalimantan, M. A. Al, and B. Banjarmasin, "DETEKSI PLAGIARISME TINGKAT KEMIRIPAN JUDUL SKRIPSI PADA FAKULTAS TEKNOLOGI INFORMASI MENGGUNAKAN ALGORITMA WINNOWING," Oktober-Desember, 2019.
- Indra Gunawan, "Deteksi Plagiarisme Abstrak Skripsi dengan Menggunakan Algoritma Rabin Karp (Studi Kasus: Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang)," *Techno Xplore J. Ilmu Komput. dan Teknol. Inf.*, vol. 6, no. 2, pp. 75–81, 2021, doi: 10.36805/technoxplore.v6i2.1227.
- Nuzul Hikmah, Dyah Ariyanti, and Ferry Agus Pratama, "Implementasi Chatbot Sebagai Virtual Assistant di Universitas Panca Marga Probolinggo menggunakan Metode TF-IDF," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 2, pp. 133–148, Aug. 2022, doi: 10.35746/jtim.v4i2.225.
- D. Nurmalasari and W. Nengsih, "Sistem Rekomendasi Referensi Jurnal Ilmiah dengan Metode Frequent Itemset," *J. Appl. Informatics Comput.*, vol. 5, no. 2, pp. 176–181, 2021.
- A. M. Ahmad, Nuraini, Arienda, Addis Prasetyo, "Penerapan Information Retrieval Pada Search Engine," *Knowl. J. Inov. Has. Penelit. dan Pengemb.*, vol. 1, no. no.1, pp. 15–23, 2021.
- J. Teknik Ilmu dan Aplikasi, A. Datumaya Wahyudi Sumari, M. Afif Hendrawan, and A. Vira Bahrudin, "Sistem Temu Kembali Informasi Pesawat Udara Militer Menggunakan Jaringan Syaraf Tiruan Back Propagation Network".
- R. Kosasih and A. Alberto, "Analisis Sentimen Produk Permainan Menggunakan Metode TF-IDF Dan Algoritma K-Nearest Neighbor," vol. 6, no. 1, 2021, doi: 10.30743/infotekjar.v6i1.3893.
- R. Al Rasyid and D. H. U. Ningsih, "Penerapan Algoritma TF-IDF dan Cosine Similarity untuk Query Pencarian Pada Dataset Destinasi Wisata," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 8, no. 1, pp. 170–178, 2024, doi: 10.35870/jtik.v8i1.1416.
- M. S. Hidayatulloh, "Sistem Temu Kembali Informasi Dinas Pemuda Olahraga dan Pariwisata Kabupaten Purbalingga dengan Metode Vector Space Model," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 2, no. 2, pp. 1–10, 2020, doi: 10.20895/inista.v2i2.101.
- Y. A. Putri Gabriella, "Optimasi Penerimaan Siswa Baru Dengan Penerapan Algoritma Text Mining Dan Tf-Idf," *J. Comput. Informatics Res.*, vol. 2, no. 3, pp. 110–117, Jul. 2023, doi: 10.47065/comforch.v2i3.941.