SNESTIK

Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika

https://ejurnal.itats.ac.id/snestik dan https://snestik.itats.ac.id

Informasi Pelaksanaan :
SNESTIK V - Surabaya, 26 April 2025
Fakultas Teknik Elektro dan Teknologi Informasi, Institut Teknologi Adhi Tama Surabaya

Informasi Artikel:
DOI : 10.31284/p.snestik.2025.7581

Prosiding ISSN 2775-5126

Fakultas Teknik Elektro dan Teknologi Informasi-Institut Teknologi Adhi Tama Surabaya
Gedung A-ITATS, Jl. Arief Rachman Hakim 100 Surabaya 60117 Telp. (031) 5945043
Email : snestik@itats.ac.id

SISTEM REKOMENDASI PADA PROUK *MARKETPLACE* "LAISYAH" MENGGUNAKAN METODE COSINE SIMILARITY

Chrisna Adrian Dwiputra Haryono[1], Septiyawan Rosetya W[2], dan Rani Rotul Muhima[3]

Institut Teknologi Adhi Tama Surabaya

e-mail: adrianharyono77@gmail.com

ABSTRACT – Font 10

The development of the internet has profoundly transformed people's lifestyles, particularly in shopping habits. Today, sellers and buyers no longer need to meet face-to-face to complete transactions. Consumers increasingly prefer shopping through marketplaces, reflecting a shift from traditional in-store shopping to more convenient online experiences. Marketplaces like "Laisyah" provide this convenience, enabling customers to purchase or order products without the need for direct interaction with the tailors. Furthermore, transaction security is ensured through payment gateways, which serve as third parties to facilitate payments. To enhance this marketplace, we have implemented a recommendation system based on search keyword history and product descriptions. This system assists customers in more easily discovering products that align with their preferences. According to the survey results, the data from 7 individuals was used for testing. The overall accuracy of this system reached 44.5%. While this accuracy level yielded relevant results for some user searches, there are still limitations, primarily due to insufficient data. The system lacks specific and pertinent information about products that meet user needs, leading to suboptimal outcomes. Therefore, improvements and adjustments to the data are necessary for better system performance. Continued research and development are crucial to enhance accuracy and fulfill user expectations.

Keywords: TF-IDF, Cosine Similarity, Recommendation System

ABSTRAK

Perkembangan internet telah membawa perubahan besar dalam pola hidup masyarakat, termasuk dalam cara berbelanja. Saat ini, penjual dan pembeli tidak perlu bertatap muka untuk melakukan transaksi. Masyarakat lebih memilih berbelanja melalui marketplace karena gaya hidup yang berubah, dari belanja langsung ke toko fisik menjadi belanja online yang lebih nyaman. Marketplace seperti "Laisyah" menyediakan kemudahan ini, di mana pelanggan dapat membeli atau memesan produk tanpa harus bertemu langsung dengan penjahit. Selain itu, keamanan transaksi juga terjamin dengan adanya payment gateway yang berperan sebagai pihak ketiga untuk mengelola pembayaran. Dalam upaya mengembangkan marketplace ini, kami untuk menerapkan sistem rekomendasi yang didasarkan pada riwayat kata kunci pencarian dan deskripsi produk. Dengan sistem ini, pelanggan dapat menemukan produk yang sesuai dengan keinginan mereka dengan lebih mudah. Metode yang digunakan untuk sistem rekomendasi adalah metode *cosine similarity*. Metode ini dapat menentukan kesamaan dua dokumen. Hal ini yang menjadi landasan untuk menentukan rekomendasi berdasarkan kedekatan histori kata kunci. Pengujian pada penelitian dilakukan oleh 7 *surveyor* yang mencoba sistem rekomendasi ini. Berdasarkan hasil pengujian, diperoleh rata-rata akurasi sebesar 61,43%. Hasil akurasi yang belum maksimal disebabkan karena keterbatasan data dari pihak *market place* yang belum memiliki data atau barang yang cukup spesifik dan relevan dengan kebutuhan percobaan.
Kata Kunci: TF-IDF, *Cosine Similarity*, Sistem Rekomendasi

PENDAHULUAN

Perkembangan internet telah menyebabkan perubahan paradigma baru dalam pola hidup, termasuk mempengaruhi tempat penjualan dimana penjual dan pembeli tidak harus bertatap muka. Proses jual beli secara tradisional membutuhkan pertemuan tatap muka antara penjual dan pembeli, tetapi proses jual beli saat ini hanya dengan 'klik' di perangkat atau komputer dengan koneksi internet. Umumnya masyarakat lebih memilih berbelanja melalui *marketplace* karena adanya perubahan gaya hidup dari pembelian langsung ke toko online, kenyamanan dan informasi terkait kebutuhan. [1]

Dengan adanya *marketplace* ini, pelanggan dapat membeli atau memesan sesuai keinginannya tanpa harus menemui penjahit secara langsung. Aplikasi *marketplace* ini juga dapat memastikan keamanan transaksi karena menggunakan *gateway payment* menjadi pihak ketiga sebagai rekening bersama [2]. Salah satu perkembangan teknologi dalam sistem *marketplace* adalah personalisasi. Personalisasi *web* adalah proses mengidentifikasi atau mengumpulkan kecenderungan pengguna untuk berbelanja *online* menggunakan *marketplace*. Salah satu manfaat personalisasi adalah dengan adanya sistem yang memberikan rekomendasi bagi pengguna yang menggunakan sistem ini untuk mencari produk.

Seperti dalam kajian [3] tentang rekomendasi produk untuk Emina Cosmetics. Tujuan dari penelitian ini adalah untuk mengembangkan sebuah aplikasi yang dapat memberikan rekomendasi produk dari Emina Cosmetics kepada pengguna berdasarkan produk yang mereka cari. Maka dari itu, konsumen akan kesulitan untuk mencari barang persis seperti apa yang mereka inginkan. Termasuk salah satunya *marketplace* yang berbasis dari "laisyah". *Marketplace* ini masih dalam ke tahap pengembangan yaitu rekomendasi yang akan ditampilkan melalui beranda yang berdasarkan history kata kunci, agar mempermudah untuk apa yang diperlukan dan diinginkan oleh pengguna.

Tingkat perbedaan antara dua *item* ditentukan dengan menggunakan ukuran jarak. Semakin besar perbedaan antara keduanya, semakin besar nilai jaraknya. Konsep jarak dan perbedaan sering digunakan secara bergantian, meskipun jarak mengacu pada bagian dari perbedaan. Jarak adalah ukuran tingkat ketidakrataan atau kerataan antara dua *item* sedangkan disparitas adalah ukuran kontras atau perbedaan antara dua objek berdasarkan banyak kualitas. Singkatnya, pengukuran jarak digunakan untuk menentukan tingkat pemisahan antara dua obyek.

Akibatnya selisih pasangan benda yang berbeda akan menjadi besar dan nilai pengukuran jarak akan jauh, sedangkan selisih pasangan benda yang lebih mirip akan semakin kecil dan nilai pengukuran jarak akan dekat [4].

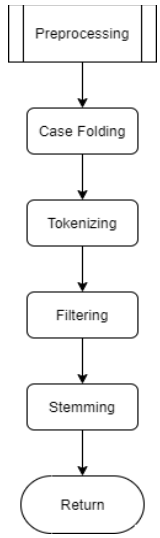
Metode *Cosine Similarity* dipilih untuk digunakan pendekatan *history* kata kunci dalam sistem rekomendasi ini. Seperti pada penelitian sebelumnya [5], [6]metode ini dapat digunakan untuk menunjukkan tingkat kesamaan antara dua dokumen. Untuk menentukan kesamaan antara dokumen, nilai dalam rentang 0 hingga 1.diperlukan, dengan nilai 0 menunjukan bahwa tidak ada kesamaan sedangkan nilai 1 menunjukan bahwa dokumen tersebut identik. Dalam hal konteks ini, vector digunakan untuk merepresantikan produk dan prefresi pengguna; dan dengan menggunakan metode ini, sistem mampu mengidentifikasi produk yang paling mirip dengan preferensi pengguna. [7]

`*Cosine Similarity* juga digunakan untuk menghitung sudut yang terbentuk. Kelebihannya meskipun ada data yang mirip namun memiliki jarak yang jauh, maka sudut yang terbentuk akan kecil dan perhitungannya jauh lebih baik daripada menggunakan metode *Euclidean Distance* dan lainnya. Dalam menentukan pembobotan kata pada penelitian ini digunakan pembobotan TF-IDF. Teknik ini digunakan untuk menghilangkan istilah yang paling umum dan mengekstrak istilah yang relevan dari korpus [8]. TF-IDF digunakan karena salah satu metode *text-processing* yang simpel dan relevan. [9]

METODE

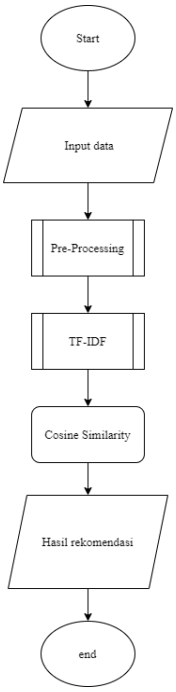
Penulis menjelaskan analisis dan perancangan sistem yang dapat memberikan rekomendasi berupa preferensi berupa menampilkan *item*/produk menggunakan *cosine similarity*. *Marketplace* ‘Laisyah’ merupakan sumber dataset yang akan diambil. Dengan adanya pengembangan sistem rekomendasi.

Terdapat tahapan-tahapan dalam melakukan pengembangan sistem rekomendasi untuk *item*/produk yaitu : *input* data, *input*an ini merupakan pengambilan data dari histori *query*; pada tahap *pre-processing* ini bertujuan untuk mengubah data mentah yang tidak terstruktur menjadi data yang terstruktur dari dataset peneliti.lalu di olah dengan TF-IDF metode untuk menghitung bobot suatu kata dalam dokumen berdasarkan frekuensinya dan seberapa jarang kata tersebut muncul di dokumen lain. Lalu akhir dari prosses yaitu *Cosine similarity* yang sering digunakan dalam mengukur tingkat kesamaan antara dua dokumen seperti digambar 1:



Gambar 1. Alur pre-preproccecing

Lalu tahap selanjutnya ialah diproses dengan dengan dengan metode nya yaitu cosine similarity, sseperti gambar 2



Gambar 2 Diagram Alir *cosine similarity*

TF-IDF merupakan metode yan diterapkan untuk membeikan bobot pada hubungan antara suatu term dengan dokumen yang relevan. Prinsip kerja TF-IDF melibatkan pemberian bobot pada suatu term yang berdasarkan frekuensi kemunuculan term tersebut dalam dokumen dan mengevaluasi term tersebut dalam keseluruhan koleksi dokumen. Untuk melakukan perhitungan term frequency (TF) bisa menggunakan rumus pada persamaan (1)

$$tf_{i.f} = f_{i.f} \quad 1$$

TF(Term Frequency) didefinisikan sebagai jumlah kemunculan suatu term ti dalam sebuah dokumen $d\ j$. Oleh karena itu, untuk mengitung nilai TF kita cukup menghitung berapa kali suatu term muncul dalam dokumen tersebut [10].sedangkan untuk perhitungan inverse document frequency menggunaka n persamaan (2)

$$idf_{i.f} = log\ (\frac{N}{df_i}) \quad 2$$

Untuk menghitunnnnng inverse document frequency (idf_i) dengan memperhitungkan jumlah dokumen dalam koleksi yang term (t_i) muncul di dalamnya. Sedangkan jumlah total dokumen terhadap pada sistem merupakan N. perhitungan idf_i digunakan untuk mengetahui jumlah term yang dicari (dfi) yang muncul dalam dokumen lain yang ada pada database.

Dalam menghitung bobot suatu term dalam sebuah dokumen, digunakan perkalian antara nilai tf dan idf. Hal ini menunjukan bahwa dokumen yang paling baik dijelaskan adalah dokunen yang memiliki term yang sering muncul di dokumen tersebut dan jarang muncul di dokumen lainnya. Berikut adalah rumus perhitungan bobot Term Frequency-Inverse Document Frequency (TF-IDF) pada persamaan 3

$$W_{i/f} = tf_{i.f} \log\left(\frac{N}{df_i}\right) \quad 3$$

Cosine Similiraty adalah metode yang sering digunakan dalam mengukur tingkat kesamaan antara dua dokumen [11]Metode ini mengukur cosinus sudut antara dua vektor, yang mencerminkan arah dan bukan hanya besarnya vektor. Nilai cosine similarity berkisar antara -1 hingga 1, di mana 1 menunjukkan bahwa kedua vektor identik, 0 menunjukkan tidak ada kesamaan, dan -1 menunjukkan bahwa kedua vektor berlawanan arah. Berikut adalah contoh persamaannya 4.

Rumus Cosine Similarity

Cosine similarity antara dua vektor q dan dj dihitung dengan rumus berikut:

$$Sim(q, d_j) = \frac{\text{q}.d_j}{|\text{q}|.|d_j|} = \frac{\sum_{i=1}^{n} w_{i.q} X w_{i.j}}{\sqrt{\sum_i^t (w_{i.q})^2} \; X \sqrt{\sum_i^t (w_{i.j})^2}} \quad 4$$

Dimana:

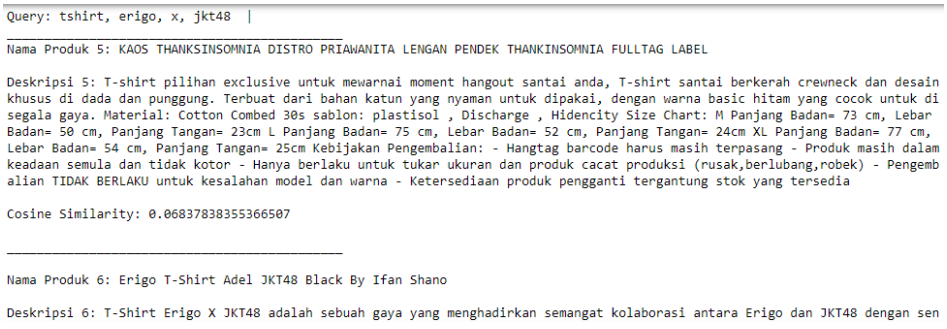
- q adalah vektor query.
- d

j adalah vektor dokumen ke-j.

- $wi.q$ adalah bobot elemen ke-i. dalam vektor query q.
- $wi.j$ adalah bobot elemen ke -i dalam vektor dokumen dj.
- n adalah jumlah elemen dalam vektor.
- t adalah jumlah elemen dalam vektor (dapat sama dengan n).

HASIL DAN PEMBAHASAN

Metode *cosine similarity* digunakan untuk menentukan rekomendasi. Rekomendasi ditentukan berdasrakan perhitungan kedekatan dengan *query*. Salah satu contoh hasil rekomendasi dari penelitian ini menggunakan cosine similarity ditunjukkan pada Gambar 9.



Gambar 3 Hasil salah satu *cosine similarity*

Sistem rekomendasi yang dibangun pada penelitian ini dilakukan pengujian oleh 7 orang *surveyor*. Berdasarkan hasil pengujian, diperoleh rata-rata akurasi berdasarkan kemiripan hasil rekomendasi dengan query sebesar 61,43 %. Berdasarkan hasil akurasi tersebut, sistem yang dibangun dapat memberikan hasil rekomendasi yang sesuai, namun perlu ditingkatkan tingkat akurasinya. Hal ini disebabkan keterbatasan data. Dari pihak *market place* belum memiliki data atau barang yang cukup spesifik dan relevan dengan kebutuhan percobaan, sehingga hasil yang diperoleh belum maksimal.

KESIMPULAN

Berdasarkan hasil penelitian sistem rekomendasi pada produk *marketplace* “laisyah” dengan menggunakan metode cosine similarity dapat diambil kesimpulan bahwa algoritma *cosine*

similarity telah terbukti mampu secara efektif mengukur kesamaan antar histori pencarian berdasarkan atribut tertentu, seperti deskripsi produk. Hal ini memungkinkan sistem untuk memberikan rekomendasi produk yang relevan dan sesuai dengan preferensi pengguna. Berdasarkan hasil pengujian oleh 7 *surveyor* yang mencoba sistem rekomendasi ini diperoleh rata-rata akurasi sebesar 61,43%. Hasil akurasi yang belum maksimal disebabkan karena keterbatasan data dari pihak *market place* yang belum memiliki data atau barang yang cukup spesifik dan relevan dengan kebutuhan percobaan.

Daftar Pustaka

[1] S. K. Sari and S. Andarini, "Pengaruh Kualitas Pelayanan Tokopedia terhadap Kepuasan Konsumen pada Pengguna Tokopedia di Surabaya," *Religion Education Social Laa Roiba Journal,* vol. 4 no 5, 2022.

[2] M. R. Aly, Rancang Bangun Aplikasi Marketplace Fashion Berbasis Web Menggunakan Metode Prototype, 2022.

[3] F. B. A. Larasati and H. Februariyanti, "Sistem Rekomendasi Product Emina Cometics dengan Menggunakan Metode Content," *Manajemen Informatika & Sistem Informasi,* vol. 4, pp. 45-54, Januari 2021.

[4] M. Kirisci, "New cosine similarity and distance measures for Fermatean," *Knowledge and Information Systems ,* vol. 65, pp. 855-868, 2022.

[5] D. Gunawan, C. A. Sembring and M. A. Budiman, "The Implementation of Cosine Similarity to Calculate Text," *Jurnal of Physics: Conference Series,* pp. 1-6, 2018.

[6] A. R. Lahitani, A. E. Permanasari and N. A. Setiawan, "Cosine Similarity to Determine Similarity Measure:," *IEEE.*

[7] K. Park, J. S. Hong and W. Kim, "A Methodology Combining Cosine Similarity with Classifier for Text Classification," *Applied Artificial Intelligence,* pp. 396-411, 2020.

[8] P. Bafna, D. Pramod and A. Vaidya, "Document Clustering: TF-IDF approach," *International Conference on Electrical, Electronics, and Optimization Techniques ,* pp. 61-66, 2016.

[9] D. E. &. P. I. Cahyani, "Performance comparison of TF-IDF and Word2Vec models for emotion text classification," *Bulletin of Electrical Engineering and Informatics,* vol. 10(5), pp. 2780-2788, 2021.

[10] R. F. Oeyliawan and D. Gunawan, "Aplikasi Rekomendasi Buku pada Katalog Perpustakaan Universitas Multimedia Nusantara Menggunakan Vector Space Model," *ULTIMATICS,,* vol. IX, no. 2085-4552, pp. 97-105, 2017.

[11] R. T. Wahyuni, D. Prastiyanto and E. Supraptono, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi," *Jurnal Teknik Elektro,* vol. 9 no 1, pp. 18-23, 2017.

Halaman ini sengaja dikosongkan