

## PENERAPAN METODE *CONTENT-BASED FILTERING* DALAM STRATEGI KOMUNIKASI PEMASARAN PADA *MARKETPLACE* TOKOPEDIA

Ahmad Syaifuddin<sup>1</sup>, Masnia Ningsih<sup>2</sup>

<sup>1</sup>Universitas Islam Majapahit Mojokerto  
e-mail: [syaifuddin@unim.ac.id](mailto:syaifuddin@unim.ac.id)

<sup>2</sup>Universitas Islam Majapahit Mojokerto  
\*e-mail korespondensi: [masnia\\_ningsih@unim.ac.id](mailto:masnia_ningsih@unim.ac.id)

### Abstrak

Strategi komunikasi pemasaran melibatkan rangkaian kegiatan interdisipliner yang memerlukan seperangkat keterampilan sistem dan koordinasi yang efektif antara semua unsur. Promosi produk menjadi strategi yang penting dalam membangun komunikasi pemasaran yang baik, salah satunya yaitu penggunaan fitur rekomendasi pada marketplace. Tujuan dari penelitian ini yaitu menerapkan *Integrated Marketing Communication* (IMC) dengan menggunakan *Content Base Filtering* untuk merekomendasikan produk dalam Bahasa Indonesia pada marketplace saat pengguna memilih suatu produk, menggunakan parameter nama dan deskripsi produk yang sama. Pengumpulan data dilakukan dengan cara *web scrapping* di Tokopedia, dengan dataset sebanyak 3233 catatan produk. Data kemudian dilakukan teknik pra-proses seperti penghapusan simbol, *stop-word removal*, terakhir dilakukan perhitungan kemiripan antar dokumen dengan TF-IDF dan *cosine similarity*. Fokus riset ini adalah menganalisis kualitas produk berdasarkan judul, kategori, dan deskripsi produk. Hasil menunjukkan bahwa *content base filtering* dapat memberikan rekomendasi produk dengan tingkat kepercayaan yang sesuai yaitu sebesar 97,39%. Penerapan *Integrated Marketing Communication* (IMC) dapat mengidentifikasi kemiripan dokumen yang termasuk nama dan deskripsi produk dengan menggunakan metode *content base filtering*.

**Kata Kunci:** Komunikasi Pemasaran, Tokopedia, *Content Base Filtering*

### Abstract

*Marketing communication strategy involves a series of interdisciplinary activities that require a set of system skills and effective coordination between all elements. Product promotion is an important strategy in building good marketing communications, one of which is the use of recommendation features on the marketplace. This research is to implement Integrated Marketing Communication (IMC) using Content Base Filtering in the marketplace to recommend products in Indonesian when users choose a product, using the same name and product description attributes. Data collection using web scrapping techniques at Tokopedia, with a dataset of 3,233 product records. The data is then carried out by pre-processing techniques such as symbol deletion, stop-word removal, finally the similarity calculation between documents is carried out using TF-IDF and cosine similarity. This study analyzes product quality based on title, category, and product description. The results show that content base filtering can provide product recommendations with an appropriate level of confidence that is equal to 97.39%. The application of Integrated Marketing Communication (IMC) can identify document similarities including product names and descriptions using the content base filtering method.*

**Keywords:** *Integrated Marketing Communication* (IMC), Tokopedia, *Content Base Filtering*

## 1. Pendahuluan

Berdasarkan survei BPS, sekitar 34,10% pelaku usaha terlibat dalam aktivitas *e-commerce* pada tahun 2022. Sebanyak 20% pelaku usaha di industri *e-commerce* menggunakan marketplace untuk promosi dan/atau pemberian online, sedangkan ada sebanyak besar pelaku usaha yang menggunakan situs web, email, pesan instan, dan media sosial mereka untuk memasarkan produknya. Menurut jenis barang/jasa yang dijual, barang yang dijual di marketplace yang berhubungan dengan makanan sekitar 41.50%, dan *fashion* sekitar 16,25% (BPS, 2022).

Di Indonesia sendiri terdapat beberapa marketplace antara lain Tokopedia, Shopee, Lazada, Bukalapak dan lain-lain. SEMRush mendapatkan data statistik dari marketplace terbesar yaitu Tokopedia, terdapat sekitar 236.300.000 pencarian dengan kata kunci Tokopedia, disusul Shopee 121.700.000 pencarian, selanjutnya ada marketplace lain seperti Lazada dan Bukalapak (Wisnuadi, 2022).

E-Commerce menggunakan berbagai fitur sebagai sarana komunikasi saat melakukan promosi, salah satunya adalah penggunaan rekomendasi produk. Setiap fitur rekomendasi berpotensi meningkatkan retensi pelanggan yang sangat penting bagi konsumen yang ingin membeli produk. Fitur rekomendasi dapat disebut sebagai bentuk dari strategi komunikasi pemasaran yang bertujuan untuk meningkatkan "*Brand Awareness*" pelanggan atau calon konsumen baru dengan lebih mudah. Apabila *Brand Awareness* pelanggan atau calon konsumen baru telah terbentuk, maka perusahaan akan lebih mudah dalam mempengaruhi mereka dan secara otomatis akan berdampak pada peningkatan penjualan.

*Integrated Marketing Communication (IMC)* adalah sistem untuk mengembangkan dan menerapkan program komunikasi pemasaran secara persiasif untuk pelanggan atau calon pelanggan secara terbuka (Terence, 2010). Berdasarkan pendapat Kotler & Keller, (2012) *Brand Awareness* dalam IMC ialah kemampuan pelanggan dalam memilih suatu barang pada kondisi yang berbeda-beda, seperti yang diperoleh dari penilaian pelanggan terhadap barang tentang baik atau buruknya kualitas layanan.

Selain itu ada istilah "*Brand Association*" yang mengungkapkan gagasan bahwa siapapun yang memiliki pengetahuan atau minat sekecil apapun terhadap merek

tertentu pada akhirnya akan mampu mempengaruhi persepsi konsumen terhadap merek tersebut atau "*Brand Image*" dalam pikiran mereka. Semakin sering konsumen yang memakai suatu merek, maka kesan terdapat merek tersebut juga akan meningkat (Priansyah, 2017).

Lebih jauh lagi mengenai *IMC* adalah bahwa tujuan komunikasi pemasaran tidak hanya untuk mempengaruhi kesadaran merek saja, akan tetapi lebih kepada usaha untuk menggerakkan perilaku konsumen kearah tujuan dilakukannya komunikasi pemasaran tersebut, yakni perilaku pembelian produk. *IMC* menggunakan berbagai cara atau jalan untuk bisa melakukan "Kontak" kepada konsumen dan calon konsumennya dengan maksud menjangkau khalayak potensialnya agar mereka selalu terhubung dengan merek atau perusahaan. Sehingga apabila sudah terjalin hubungan antara keduanya, maka bisa dikatakan itulah kunci keberhasilan dari sebuah pemasaran modern yang akan menimbulkan pembelian yang berulang, bahkan memunculkan loyalitas terhadap merek oleh konsumen

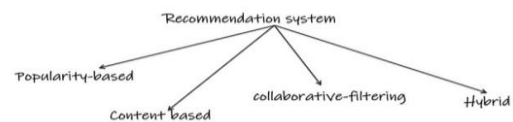
Sistem Rekomendasi menjadi bagian yang cukup penting dari kehidupan konsumen di internet, termasuk marketplace, sosial media, sampai pencarian online (Sadeli & Wibowo, 2022). Contohnya Amazon yang mendapatkan penghasilan dengan menerapkan sistem rekomendasi sebesar 35% dari total pembelian. Jadi, dapat diasumsikan bahwa Amazon tidak mengeluarkan banyak uang untuk pemasarannya, namun menggunakan rekomendasi hasil personalisasi tetap dapat memotivasi pelanggan. Ada berbagai macam produk di toko online, mulai dari peralatan kecil serta perangkat seluler. Karena itu, penjual pada toko online sangat perlu untuk mengidentifikasi karakteristik pelanggan dan mendorong pelanggan untuk terus melakukan pembelian pada aplikasi *e-commerce*. Inilah penjelasan mengapa sistem rekomendasi sangat berperan dalam strategi pemasaran pada *e-commerce* (Tanuar & Yosanny, 2010).

Sistem rekomendasi adalah aplikasi penelitian akademis paling populer sejauh ini. Sistem rekomendasi merupakan sistem berbasis data yang dipakai dalam merekomendasikan produk kepada konsumen dilihat dari pembelian sebelumnya atau perilaku lainnya. Selain itu, sistem menggunakan algoritma mesin pembelajaran

yang bertujuan untuk menentukan tingkat penggunaan pengguna sebagai entitas target (Negara et al., 2023).

Banyak perusahaan teknologi yang menggunakan sistem rekomendasi dalam meningkatkan retensi pelanggan, sebagai contoh Amazon yang menggunakan sistem rekomendasi untuk produnya, Youtube yang menggunakan sistem rekomendasi untuk video, Netflix menggunakan sistem rekomendasi untuk film untuk pengunjung. Sistem rekomendasi produk memberikan layanan penunjang bagi pengunjung serta dapat menguntungkan bagi pemilik usaha (Roy & Dutta, 2022)

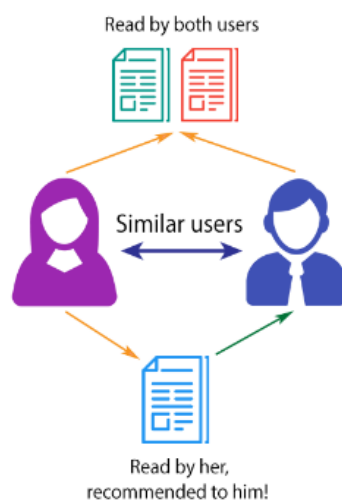
Ada beberapa jenis *recommendation system*, yaitu *Content Base Filtering* yang memanfaatkan parameter deskripsi produk yang sama dalam mendapatkan rekomendasi, serta *Collaborative Filtering* memanfaatkan parameter kemiripan pengguna atau pelanggan untuk mendapatkan rekomendasi produk (Hariri & Rochim, 2022).



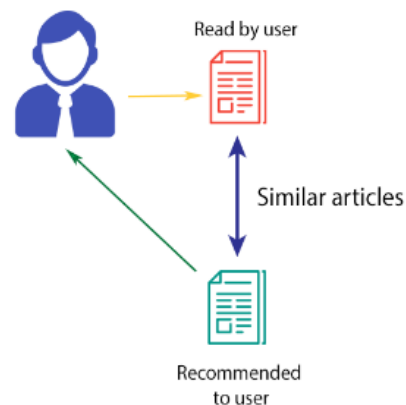
Gambar 1. Teknik Rekomendasi Sistem

- Popularity-based:** sistem rekomendasi yang mempromosikan produk ofensif dari situs web perusahaan ke setiap pengguna. Misalnya, IMDB mencantumkan 250 film yang trending di Youtube.
- Content-based:** sistem rekomendasi yang didasarkan pada kompatibilitas produk. Misalnya, pengunjung toko online yang ingin melihat produk lain yang mirip dengan produk yang dia pilih.
- Collaborative filtering:** sistem rekomendasi berdasarkan kemiripan data pengunjung. Misalnya, pengunjung X dan Y pernah melihat film Q, kemudian ada pengunjung X pernah melihat film P, maka sistem akan merekomendasikan pengunjung Y untuk melihat film P juga.
- Hybrid filtering:** sistem rekomendasi yang mengkombinasikan dua atau lebih metode diatas untuk membuat rekomendasi.

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



Gambar 2. Metode Rekomendasi Data Paling Populer  
Sumber: Gowtham S.R (2022)

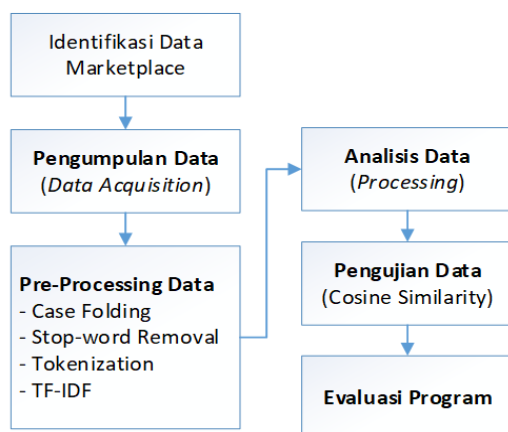
Pemilihan metode *content base filtering* didasarkan pada atribut yang tersedia pada dataset tokopedia yang terbatas ketika didapatkan melalui *web scrapping*, diantaranya adalah data judul produk, deskripsi, kategori, harga dan gambar produk. Penelitian Kuo & Cheng (2022) mendapatkan hasil penelitian bahwa metode *content-based filtering* adalah yang

metode terbaik dibandingkan dengan metode *vector space model* dan *Rocchio algorithm*.

Tujuan penelitian ini adalah untuk membuat aplikasi penerapan *Integrated Marketing Communication* (IMC) dengan menggunakan *Content Base Filtering* untuk merekomendasikan produk dalam Bahasa Indonesia ketika pengguna memilih suatu produk di marketplace.

## 2. Metode Penelitian

Penelitian ini dilakukan dengan beberapa prosedur dalam melakukan analisa *text mining*. Langkah-langkah ini termasuk pra-proses, penghitungan *TF-IDF* dan kemudian tahap pengujian/proses data, yang menerapkan algoritma *cosine similarity*. Tahap pra-proses yang dilakukan terdiri dari *case folding*, tokenisasi, *stop-word removal* berbahasa Indonesia, *TF-IDF* (Syarifuddin & Muslimin, 2022). Tahap proses/uji coba dilakukan dengan mengimplementasikan algoritme *cosine similarity* untuk mendapatkan nilai similaritas antar dokumen.



Gambar 3. Langkah Penelitian

Gambar 3 memberikan penjelasan tentang langkah-langkah penelitian, pertama dengan melakukan identifikasi data *marketplace*, kemudian dilakukan penarikan data dengan menggunakan teknik *scrapping*. Data kemudian dilakukan pra-proses data (*case folding*, *stopword removal*, *tokenization* dan *TF-IDF*) selanjutnya dilakukan analisa data yaitu menghitung *cosine similarity*, terakhir dilakukan evaluasi data.

Uji coba sistem dilakukan menggunakan *tools* Jupyter Notebook dengan paket python versi 3.9.12. Dalam penelitian ini dilakukan 3 kali tahap tes data untuk melihat, konsistensi dan rerata nilai similaritas yang dihasilkan dengan metode *Content Base Filtering*.

## 3. Hasil dan Pembahasan

### 3.1 Web Scrapping

Pengumpulan data dilakukan dengan mengambil data di website *e-commerce* Tokopedia dengan cara *scrapping*. Teknik *Scrapping* adalah teknik yang biasa digunakan oleh para *merchant* dan *dropshipper* untuk membuat konten produk di website *e-commerce* miliknya, setelah melakukan *scrapping*, *dropshipper* dapat mengubah dan menata produk sesuai yang mereka inginkan. Aplikasi yang paling sering dipakai serta dapat digunakan dengan mudah yaitu aplikasi *BigSeller*, yang memanfaatkan fitur *Scrape Products*.

Berikut langkah pengambilan data:

- 1) Langkah pertama, peneliti mengunjungi website Tokopedia, kemudian pilih kategori "Makanan dan minuman". Peneliti memilih kategori makanan dan minuman karena produk tersebut paling banyak dibeli berdasarkan hasil survei BPS 2022.
- 2) Langkah kedua, peneliti memilih lima toko besar dengan rating Produk diatas 4.5, lalu dilakukan penarikan data di lima toko tersebut, diperoleh data sebagai berikut:

Tabel 1. Sebaran Data

| Toko                                 | Produk      |
|--------------------------------------|-------------|
| Japfa Best Official                  | 466         |
| Oreo, Kraft & Cadbury Official Store | 165         |
| NutriMart                            | 560         |
| Pisang Goreng Madu Bu Nanik          | 282         |
| KedaiMart Official Store             | 1760        |
| <b>Total</b>                         | <b>3233</b> |

Berdasarkan penarikan data dari lima toko menghasilkan sekitar 3233 data makanan dan minuman. Berikut beberapa data yang diperoleh dengan melakukan *scrapping* dengan *BigSeller*.

Tabel 2. Contoh Data Produk

| Field            | Isian Data  |
|------------------|---|
| Product Name     | Paket Snack Oreo Cadbury Kraft Ritz Biskuat - Free Special Gift - B   |
| Long Description | Paket Snack Oreo Cadbury Kraft Ritz Biskuat - Gratis Hadiah Spesial<br><br>Isi paket :<br>3x Oreo Biskuit Blackpink Pink Cookie 123.5g<br>5x Cadbury Cokelat Dairy Milk 30g<br>1x Kraft Keju Cake Bolu Keju 16g [1 Box Isi 12 pcs]<br>3x Ritz Biskuit Sandwich Rasa Keju 118g<br>3x Biskuit Biskuit Coklat 50.4G - Snack Cemilan<br><br>*Free Goto Micha Panci Listrik Elektrik Mini / Heinz Sambal Mexico / Gaabor Timbangan Digital / Starco Laptop Table<br><br>Cadbury Dairy Milk merupakan coklat kombinasi susu yang lezat digabungkan menjadikan coklat batangan bercita rasa sempurna dan nikmat, cocok untuk temani momen spesial kamu agar lebih bermakna!<br><br>Ritz Cracker dengan rasa asin yang ringan di satu sisi. Rasa asin dari biskuit menjadi kombinasi yang lezat. Dengan tekstur |

| Field                 | Isian Data  |
|-----------------------|---|
|                       | yang renyah, diperkaya dengan kalsium dan zat besi. Ritz cocok untuk camilan bergizi!<br><br>Kraft Keju Cake Bolu Keju merupakan cake yang lembut dan ringan dengan isian keju tebal dan creamy yang meleleh di mulut. Dibuat dengan kebaikan susu dan keju asli yang membuat Keju Cake ini sangat lembut dan lezat dan akan membuat Anda ketagihan!<br><br>Biskuit terbuat dari perpaduan susu, coklat dan gandum yang menyatu dan menghasilkan sebuah biskuit yang enak, lezat, dan menyehatkan. Biskuit ini telah diperkaya dengan kandungan 9 vitamin dan 6 mineral yang baik bagi tubuh.<br>   |
| Short Description     | "Paket Snack Oreo Cadbury Kraft Ritz Biskuit - Gratis Hadiah Spesial<br>Isi paket :<br>3x Oreo Biskuit Blackpink Pink Cookie 123.5g<br>5x Cadbury Cokelat Dairy Milk 30g<br>1x Kraft Keju Cake Bolu Keju 16g [1 Box Isi 12 pcs]<br>3x Ritz Biskuit Sandwich Rasa Keju 118g<br>3x Biskuit Biskuit Coklat 50.4G - Snack Cemilan<br>*Free Goto Micha Panci Listrik Elektrik Mini / Heinz Sambal Mexico / Gaabor Timbangan Digital / Starco Laptop Table<br>Cadbury Dairy Milk merupakan coklat kombinasi susu yang lezat digabungkan menjadikan coklat batangan bercita rasa sempurna dan nikmat, cocok untuk temani momen spesial kamu agar lebih bermakna!<br>Ritz Cracker dengan rasa asin yang ringan di satu sisi. Rasa asin dari biskuit menjadi kombinasi yang lezat. Dengan tekstur yang renyah, diperkaya dengan kalsium dan zat besi. Ritz cocok untuk camilan bergizi!<br>Kraft Keju Cake Bolu Keju merupakan cake yang lembut dan ringan dengan isian keju tebal dan creamy yang meleleh di mulut. Dibuat dengan kebaikan susu dan keju asli yang membuat Keju Cake ini sangat lembut dan lezat dan akan membuat Anda ketagihan!<br>Biskuit terbuat dari perpaduan susu, coklat dan gandum yang menyatu dan menghasilkan sebuah biskuit yang enak, lezat, dan menyehatkan. Biskuit ini telah diperkaya dengan kandungan 9 vitamin dan 6 mineral yang baik bagi tubuh." |
| Supplier Source       | <a href="https://www.tokopedia.com/mondelez/paket-snack-oreo-cadbury-kraft-ritz-biskuit-free-special-gift-b-dcb64?extParam=whid%3D13056835">https://www.tokopedia.com/mondelez/paket-snack-oreo-cadbury-kraft-ritz-biskuit-free-special-gift-b-dcb64?extParam=whid%3D13056835</a>   |
| Variation Name1       | Type  |
| Option for Variation1 | A   |
| Original Price        | 274350  |
| Discount Price        | 0   |
| Currency              | IDR   |
| Stock                 | 110   |
| SKU                   | 9320806639  |
| Weight                | 1.51  |
| Item image1           | <a href="https://images.tokopedia.net/img/cache/700/VqbcmM/2023/5/15/f84fe8e7-a28d-4dd1-8c18-2c1b37bb903f.jpg">https://images.tokopedia.net/img/cache/700/VqbcmM/2023/5/15/f84fe8e7-a28d-4dd1-8c18-2c1b37bb903f.jpg</a>   |
| Item image2           | <a href="https://images.tokopedia.net/img/cache/700/VqbcmM/2023/5/15/8bd0bca5-3533-4e6d-8f6b-f4e893e50719.jpg">https://images.tokopedia.net/img/cache/700/VqbcmM/2023/5/15/8bd0bca5-3533-4e6d-8f6b-f4e893e50719.jpg</a>   |
| Variant Image1        | <a href="https://images.tokopedia.net/img/cache/700/VqbcmM/2023/5/15/f84fe8e7-a28d-4dd1-8c18-2c1b37bb903f.jpg">https://images.tokopedia.net/img/cache/700/VqbcmM/2023/5/15/f84fe8e7-a28d-4dd1-8c18-2c1b37bb903f.jpg</a>   |

Tabel 2 menunjukkan salah satu contoh data yang dapat diambil melalui *web scrapping*, ada beberapa field/parameter seperti nama produk, deskripsi produk versi kode, deskripsi produk versi tayang, alamat url toko, variasi produk, harga original produk, harga diskon, mata uang, stok, kode *barcode* produk, berat, url gambar produk

dan url gambar variasi produk.

Ada beberapa library python yang digunakan dalam penelitian ini, antara lain: Pandas, Numpy, Sklearn, NLTK stopword sastrawi. Library sastrawi digunakan karena teks yang proses pada penelitian ini adalah teks berbahasa Indonesia.

```
In [1]:
1 import pandas as pd
2 import numpy as np
3 from sklearn.metrics.pairwise import cosine_similarity
4 from sklearn.feature_extraction.text import TfidfVectorizer
5 from nltk.corpus import stopwords
6 import re
7 import random
8 from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory
9
10 df = pd.read_csv("scraped_product_food0.csv")
11 df.head()
```

Out[1]:

| ID | ProductName                                 | Description  | SupplierSource                                    | Variation1 | Original Price | DiscountPrice | Stock | SKU   | Weight |
|----|---|--|---|------------|----------------|---------------|-------|-------|--------|
| 0  | Knorr Cream of Chicken Soup 900 Gram / 1 Kg | Knorr Cream of Chicken Soup 900 Gram / 1 Kg<br>... | https://www.tokopedia.com/kedaimartstore/knorr... | NaN        | 125300         | 125300        | 35    | 06805 | 1.000  |
| 1  | Indocafe Cappuccino Bag Isi 50 x 25 Gram    | Indocafe Cappuccino Bag Isi 50 x 25 GramIndoca...  | https://www.tokopedia.com/kedaimartstore/indoc... | NaN        | 116900         | 116900        | 31    | 06806 | 1.250  |
| 2  | Milo Active-Go UHT 110 ml                   | Milo Active-Go UHT 110 mlMilo Active-Go UHT su...  | https://www.tokopedia.com/kedaimartstore/milo...  | NaN        | 3900           | 3900          | 900   | 06594 | 0.110  |
| 3  | Tiga Sapi Susu Kental Manis Putih 375 Gram  | Tiga Sapi Susu Kental Manis Putih 375 GramTiga...  | https://www.tokopedia.com/kedaimartstore/tiga...  | NaN        | 12200          | 12200         | 28    | 06407 | 0.375  |
| 4  | Yeos Sarang Burung 300 ml                   | Yeos Sarang Burung 300 mlYeos Sarang Burung mi...  | https://www.tokopedia.com/kedaimartstore/yeos...  | NaN        | 6500           | 6500          | 758   | 06406 | 0.300  |

Gambar 4. Proses Pengambilan Data pada Python

### 3.2 Pra-Proses

Langkah pra-proses terdiri dari beberapa langkah berikut: (Syaifuddin et al., 2020).

- Data cleansing* dilakukan dengan tahap menghapus URL dan @.
- Case folding* yaitu membuat semua kata dalam dokumen diubah menjadi huruf kecil.
- Stop-word removal* menggunakan *stop word removal* berbahasa Indonesia dari library NLTK.
- Tokenisasi data teks  
Proses text mining dapat bekerja dengan data teks, diperlukan tokenisasi.

Tokenisasi adalah cara memecah teks menjadi unit yang lebih kecil yang disebut token. Token dalam contoh ini adalah sebuah kata. Tfidf digunakan untuk mengisikan pada kolom matriks. *Term Frequency-Inverse Document Frequency* (TF-IDF) dipakai untuk menghitung relevansi kata dalam dokumen pada kumpulan dokumen (Mondi et al., 2020).

Adapun banyaknya data yang telah melalui pra-proses yaitu 3230 data. Berikut adalah gambar dari pra-proses data dengan menggunakan aplikasi python.

In [8]:

```

1 clean_spc1 = re.compile('[/(){}[\]\|@,;]')
2 clean_symbol = re.compile('[^0-9a-z #+_]')
3 stopwords = set(stopwords.words('indonesian'))
4
5 def clean_text(text):
6     text = text.replace('<br>', ' ') # replace new line into space
7     text = text.replace('NaN', '') # Replace NaN with an empty string
8     text = text.lower() # Lowercase text
9     text = clean_spc1.sub(' ', text)
10    text = clean_symbol.sub('', text)
11    text = ' '.join(word for word in text.split() if word not in stopwords)
12    return text
13
14 df['desc_clean'] = df['data1'].apply(clean_text)

```

In [9]:

df.head()

Out[9]:

| ID | ProductName                                 | Description  | SupplierSource                                    | Variation1 | Original Price | DiscountPrice | Stock | SKU   | Weight | data1   | desc_cle                           |
|----|---|--|---|------------|----------------|---------------|-------|-------|--------|---|------------------------------------|
| 0  | Knorr Cream of Chicken Soup 900 Gram / 1 Kg | Knorr Cream of Chicken Soup 900 Gram / 1 Kg<br>... | https://www.tokopedia.com/kedaimartstore/knorr... | nan        | 125300         | 125300        | 35    | 06805 | 1.000  | Knorr Cream of Chicken Soup 900 Gram / 1 Kg Kn... | knorr cre of chic soup 1 gram 1 km |
| 1  | Indocafe Cappuccino Bag Isi 50 x 25 Gram    | Indocafe Cappuccino Bag Isi 50 x 25 GramIndoca...  | https://www.tokopedia.com/kedaimartstore/indoc... | nan        | 116900         | 116900        | 31    | 06806 | 1.250  | Indocafe Cappuccino Bag Isi 50 x 25 Gram Indoc... | indoc cappucc bag isi 5 25 gr indc |
| 2  | Milo Active-Go UHT 110 ml                   | Milo Active-Go UHT 110 mlMilo Active-Go UHT su...  | https://www.tokopedia.com/kedaimartstore/milo...  | nan        | 3900           | 3900          | 900   | 06594 | 0.110  | Milo Active-Go UHT 110 ml Milo Active-Go UHT 1... | r activego 110 ml r activego 11    |
| 3  | Tiga Sapi Susu Kental Manis Putih 375 Gram  | Tiga Sapi Susu Kental Manis Putih 375 GramTiga...  | https://www.tokopedia.com/kedaimartstore/tiga...  | nan        | 12200          | 12200         | 28    | 06407 | 0.375  | Tiga Sapi Susu Kental Manis Putih 375 Gram Tig... | sapi s ke manis p 375 gr sapi st   |

Gambar 5. Pra-Proses Data

e. *Term Frequency Inverse Document Frequency* (TF-IDF)

TF-IDF untuk menggabungkan dua teori berdasarkan frekuensi kemunculan kata pada dokumen. Setiap *term* (kata) menunjukkan satu atau lebih kata kunci dalam dokumen. Frekuensi kemunculan data di dokumen menunjukkan beberapa kata yang sangat penting, sedangkan frekuensi di dokumen menunjukkan beberapa kata yang diterima secara umum. Berikut rumus TF-IDF:

$$tf = 0,5 + 0,5 \times \frac{tf}{\max(tf)} \quad (1)$$

$$idf_t = \log\left(\frac{D}{df_t}\right) \quad (2)$$

$$W_{d,t} = tf_{d,t} \times IDF_{d,t} \quad (3)$$

Keterangan:

tf = jumlah kata yang dicari pada sebuah dokumen

D = total dokumen

df = jumlah dokumen yang mengandung *term* t

idf = *Inversed Document Frequency* (log2)

d = dokumen ke-d

t = kata ke-t dari kata kunci

W = bobot dokumen ke-d terhadap kata ke-t

Berikut adalah hasil perhitungan TF-IDF dengan menggunakan aplikasi python.

```
In [12]: 1 df.set_index('ProductName', inplace=True)
2 tf = TfidfVectorizer(analyzer='word', ngram_range=(1, 3), min_df=0)
3 tfidf_matrix = tf.fit_transform(df['desc_clean'])
4 cos_sim = cosine_similarity(tfidf_matrix, tfidf_matrix)
5 cos_sim

Out[12]: array([[1.          , 0.00615663, 0.00270182, ..., 0.00462649, 0.00475556,
0.00482698],
[0.00615663, 1.          , 0.00206981, ..., 0.00498918, 0.00468615,
0.00273235],
[0.00270182, 0.00206981, 1.          , ..., 0.03156795, 0.0278124 ,
0.04176209],
...,
[0.00462649, 0.00498918, 0.03156795, ..., 1.          , 0.46730979,
0.51530113],
[0.00475556, 0.00468615, 0.0278124 , ..., 0.46730979, 1.          ,
0.49553339],
[0.00482698, 0.00273235, 0.04176209, ..., 0.51530113, 0.49553339,
1.          ]])
```

Gambar 6. Hasil percobaan pertama

f. *Analisis Content Base Filtering*

*Cosine Similarity* digunakan sebagai algoritme utama untuk melakukan analisa *content base filtering* yaitu. *Cosine similarity* dapat menghitung dan mengklasifikasikan dokumen berdasarkan tingkat kemiripan dokumen dengan cara mengambil nilai similaritas tertinggi dari dokumen kunci. Berikut ini adalah rumus *cosine similarity* yang digunakan.

$$\cos \alpha = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (4)$$

Keterangan:

A = vektor A

B = vektor B

A.B = dot product vektor A dan vektor B

|A| = panjang vektor A

|B| = panjang vektor B

|A||B| = cross product antara |A| dan |B|

Pada uji coba dalam implementasi metode *content based filtering*, fungsi rekomendasi produk didapatkan berdasarkan memiripan data dari hasil TF-IDF dan *cosine similarity* berdasarkan 10 produk dengan nilai *similarity* tertinggi. Selanjutnya diperoleh data hasil rekomendasi dengan nilai kemiripan dokumen tertinggi.

Berikut ini adalah gambar implementasi penerapan metode *Content Base Filtering*:

```

In [14]: 1 def recommendations(name, cos_sim = cos_sim):
2
3     recommended_product = []
4
5     # Mengambil ProductName hotel berdasarkan variabel indicies
6     idx = indices[indices == name].index[0]
7
8     # Membuat series berdasarkan skor kesamaan
9     score_series = pd.Series(cos_sim[idx]).sort_values(ascending = False)
10
11    # mengambil index dan dibuat 10 baris rekomendasi terbaik
12    top_10_indexes = list(score_series.iloc[1:11].index)
13
14    print("Recommending products similar to " + name)
15
16    for i in top_10_indexes:
17        recommended_product.append( list(df.index)[i]+ " (similar:" + str(score_series[i]) + ")")
18    return recommended_product
19
20    def MAE(recommended_product):
21        return
22    accuracy.mae(recommended_product, verbose=False)
23

```

```
In [15]: 1 recommendations('Susu Fresh Strawberry Greenfields 1L')
```

Recommending products similar to Susu Fresh Strawberry Greenfields 1L

```

Out[15]: ['Susu Fresh Strawberry Greenfields 200 ml - Daily Deals (similar:0.838099315664403)',
'Susu Fresh Strawberry Greenfields 200 ml (similar:0.754252862849736)',
'Susu Greenfields Fresh Full Cream 1 L (similar:0.6286300318934829)',
'Susu Fresh Full Cream Greenfields 200 ml (similar:0.5903978453617806)',
'Susu Fresh Choco Malt Greenfields 200 ml (similar:0.5638778133858797)',
'Greenfields Fresh Milk Choco Malt 1 L (similar:0.5633644138437607)',
'Susu Fresh Full Cream Greenfields 1.89 L (similar:0.563299331678872)',
'Greenfields Fresh Milk / Susu Pasteurisasi 500 ml (similar:0.5561167602731414)',
'Susu Greenfields UHT Strawberry 200 ml - Isi 32 pcs (similar:0.5450593494302487)',
'Susu Greenfields UHT Strawberry 200 ml - Isi 16 pcs (similar:0.5375994152160338)']

```

Gambar 7. Hasil percobaan pertama

```
In [16]: 1 recommendations('Paket Bundling Nugget Ayam Original So Good 400gr (Isi 2)')
```

Recommending products similar to Paket Bundling Nugget Ayam Original So Good 400gr (Isi 2)

```

Out[16]: ['Nugget Ayam So Good 400 gr (similar:0.7284585335779614)',
'Nugget Ayam Original So Good 400 gr - Payday Hemat (similar:0.7260585131688452)',
'Paket Bundling Nugget Ayam Dinobites So Good 400gr (Isi 2) (similar:0.6555035723218882)',
'So Good Chicken Nugget Original 400 gr "FREE MINIONS PEN" (similar:0.6073289939791935)',
'Nugget Stick Original So Good 200 gr - Daily Deals (similar:0.47762987632804077)',
'Nugget Ayam Animal Pizza So Good 400gr (similar:0.4722433536731828)',
'Nugget Stick Original So Good 200gr [Buy 1 Get 1 Free] (similar:0.46189588698014605)',
'Nugget Ayam Dinobites So Good 400 gr (similar:0.43231739305106515)',
'Nugget Ayam Stick Original So Good 200 gr (similar:0.4251852471496098)',
'Fish Nugget So Good 300 gr - Daily Deals (similar:0.423902812240374)']

```

Gambar 8. Hasil percobaan kedua

```
In [17]: 1 recommendations('Filma Minyak Goreng Refill Pouch 2L')
```

Recommending products similar to Filma Minyak Goreng Refill Pouch 2L

```

Out[17]: ['Filma Minyak Goreng Refill 2L (similar:0.9739300467299611)',
'Filma Minyak Goreng Refill Pouch 1L (similar:0.9197142531599796)',
'Filma Minyak Goreng Refill 1L (similar:0.9017820042400403)',
'Minyak Goreng Filma 1 L - TWIN PACK (similar:0.7849178585663193)',
'Minyak Goreng Refill Rose Brand 2L (similar:0.37350825260860143)',
'Greenfields Yogurt Drink Strawberry 250 ml (similar:0.325825010922238725)',
'Susu Greenfields Fresh Full Cream 1 L (similar:0.32321402148559275)',
'Greenfields Yogurt Drink Blueberry 250 ml (similar:0.32229413912802707)',
'Susu Greenfields UHT Full Cream 1 L - Paket Isi 3 (similar:0.3191026717812269)',
'Susu Greenfields UHT Full Cream [4 Pcs x 1L] (similar:0.31783218898487886)']

```

Gambar 9. Hasil percobaan ketiga

Percobaan pertama pada produk “Susu Fresh Strawberry Greenfields 1L”, didapatkan skor dengan nilai kemiripan sebesar 83,81%. Percobaan kedua pada produk “Paket Bundling Nugget Ayam Original So Good 400gr (Isi 2)”, didapatkan

skor dengan nilai kemiripan sebesar 72,84%. Percobaan ketiga pada produk “Filma Minyak Goreng Refill Pouch 2L”, didapatkan skor tertinggi dengan nilai kemiripan sebesar 97,39%.

Berdasarkan hasil tersebut, metode



*Content Base Filtering* dapat menjawab kebutuhan komunikasi pemasaran atau *Integrated Marketing Communication* (IMC), pada beberapa aspek, diantaranya adalah *brand awareness*. Konsumen dapat dengan mudah menjumpai dan mengidentifikasi produk berdasarkan hasil rekomendasi yang muncul.

Tabel 3. Hasil Pengujian

| Nama Produk yang dipilih                                  | Nilai Similaritas | Rerata Similaritas |
|---|-------------------|--------------------|
| Susu Fresh Strawberry Greenfields 1L                      | 0.84              | 0.61               |
| Paket Bundling Nugget Ayam Original So Good 400gr (Isi 2) | 0.72              | 0.54               |
| Filma Minyak Goreng Refill Pouch 2L                       | 0.97              | 0.56               |

Hasil percobaan tersebut sejalan dengan teori Kotler & Keller (2012) yang menjelaskan IMC (*Integrated Marketing Communication*) merupakan sistem komunikasi pemasaran secara persuasif kepada pelanggan ataupun calon pelanggan melalui proses pengembangan dan penerapan berbagai bentuk kegiatan secara berkesinambungan. Terdapat istilah *Brand Awareness* dalam IMC, yaitu cara pelanggan dalam membedakan merek dalam kondisi tertentu, misalnya tercermin dari pencapaian *brand recognition* atau *recall* mereka (Chitty et al., 2017).

Selain itu rekomendasi produk dapat memberikan aspek *brand association* dan *brand image* kepada konsumen. Yang sejalan dengan teori Kotler & Keller (2012) yang menjelaskan "*Brand Association*" merupakan suatu kesan yang muncul pada pelanggan dari segala ingatan dan pengalamannya mengenai suatu merek yang memiliki tingkat kekuatan tertentu sehingga pada akhirnya dapat membuat citra terhadap brand tersebut atau "*Brand Image*" di mata pelanggan. Semakin banyak pelanggan yang menggunakan produk maka semakin sering produk tersebut muncul pada daftar pencarian (Chitty et al., 2017).

Sistem rekomendasi dibuat dengan tujuan untuk mengurangi kemungkinan kerugian perusahaan dan meningkatkan *brand awareness* yang dimiliki pelanggan terhadap produk, khususnya produk yang dijual di *marketplace*. Karena semakin banyak pelanggan dapat mengenali produk, membuat *brand awareness* produk juga

semakin tinggi. Penggunaan gabungan indikator judul produk, deskripsi produk dan kategori produk dapat menjadi solusi dalam sistem rekomendasi produk.

#### 4. Kesimpulan

*Marketing Communication* merupakan strategi bisnis yang sangat penting karena tanpa komunikasi, konsumen atau pelanggan tidak akan mengetahui produk dan layanan yang tersedia. Penerapan *Integrated Marketing Communication* (IMC) dengan metode *content based filtering* dapat memberikan daftar nama rekomendasi produk yang memiliki kesamaan produk menggunakan parameter judul, deskripsi dan kategori produk. Fitur bisa dikategorikan sebagai fitur yang bertujuan untuk memberikan Efek kognitif di mana fitur ini dapat membentuk kesadaran informasi tertentu pada benak pelanggan yang pada akhirnya diharapkan dapat meningkatkan *brand awareness* dan berdampak pada peningkatan pembelian oleh konsumen.

#### Referensi

- BPS, B. P. S. (2022). *Statistik eCommerce 2022*. Badan Pusat Statistik. <https://www.bps.go.id/publication/2022/12/19/d215899e13b89e516caa7a44/statistik-e-commerce-2022.html>
- Chitty, B., Chitty, W., Luck, E., Barker, N., Sassenberg, A.-M., A. Shimp, T., & Andrews, J. C. (2017). *Integrated Marketing Communications with Online Study Tools 12 Months*. In *Cengage AU*. Cengage AU. [https://books.google.co.id/books?hl=id&lr=&id=1ENMDwAAQBAJ&oi=fnd&pg=PR1&dq=Terence+A.+Shimp+2010&ots=UO2MFliQEU&sig=BScBlr\\_zrtVkJUpTe04wqdipJ8&redir\\_esc=y#v=onepage&q=Terence A. Shimp 2010&f=false](https://books.google.co.id/books?hl=id&lr=&id=1ENMDwAAQBAJ&oi=fnd&pg=PR1&dq=Terence+A.+Shimp+2010&ots=UO2MFliQEU&sig=BScBlr_zrtVkJUpTe04wqdipJ8&redir_esc=y#v=onepage&q=Terence A. Shimp 2010&f=false)
- Gowtham S.R. (2022). *Practical Implementation of Content-Based Recommendation System | by Gowtham S R | Towards AI*. <https://pub.towardsai.net/how-to-build-a-content-based-recommendation-system-f7d881a53e9a>
- Hariri, F. R., & Rochim, L. W. (2022). Sistem Rekomendasi Produk Aplikasi Marketplace Berdasarkan Karakteristik Pembeli Menggunakan Metode User Based Collaborative Filtering. *Teknika*, 11(3), 208–217. <https://doi.org/10.34148/TEKNIKA.V11I3.538>

- Kotler, P., & Keller, K. L. (2012). *Manajemen Pemasaran*. Erlangga.
- Kuo, R. J., & Cheng, H. R. (2022). A content-based recommender system with consideration of repeat purchase behavior. *Applied Soft Computing*, 127, 109361. <https://doi.org/10.1016/J.ASOC.2022.109361>
- Mondi, R. H., Wijayanto, A., & Winarno, W. (2020). Recommendation System With Content-Based Filtering Method for Culinary Tourism in Mangan Application. *ITSMART: Jurnal Teknologi Dan Informasi*, 8(2), 65–72. <https://jurnal.uns.ac.id/itsmart/article/view/35008>
- Negara, E. S., Sulaiman, Andryani, R., Saksono, P. H., & Widyanti, Y. (2023). Recommendation System with Content-Based Filtering in NFT Marketplace. *Journal of Advances in Information Technology*, 14(3), 518–522. <https://doi.org/10.12720/JAIT.14.3.518-522>
- Priansyah, D. J. (2017). *Komunikasi Pemasaran Terpadu pada Era Media Sosial*. Pustaka Setia.
- Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, 9(1), 1–36. <https://doi.org/10.1186/S40537-022-00592-5/TABLES/3>
- Sadeli, S. P., & Wibowo, A. P. W. (2022). Sistem informasi Rekomendasi Produk Pada Website Ecommerce dengan Menggunakan Algoritma Apriori. *Jurnal Darma Agung*, 30(1), 675–682. <https://doi.org/10.46930/OJSUDA.V30I1.2586>
- Syaifuddin, A., Harianto, R. A., & Santoso, J. (2020). Analisis Trending Topik untuk Percakapan Media Sosial dengan Menggunakan Topic Modelling Berbasis Algoritme LDA. *INSYST: Journal of Intelligent System and Computation*, 2(1), 12–19. <https://doi.org/10.52985/INSYST.V2I1.150>
- Syaifuddin, A., & Muslimin, M. (2022). Analisis Sentimen pada Sosial Media Tentang Implementasi Kebijakan PSE Kominfo Menggunakan Algoritme Lexicon Based. *Seminar Nasional Fakultas Teknik*, 1(1), 7–14. <https://doi.org/10.36815/SEMASTEK.V1I1.2>
- Tanuar, E., & Yosanny, A. (2010). Analisis E-Bisnis Terhadap Amazon dan Aquarelle. *ComTech*, 1(2), 782–792.
- Terence, A. S. (2010). *Komunikasi Pemasaran Terpadu dalam Periklanan dan Promosi*. Salemba Empat.
- Wisnuadi, K. (2022). *[Data Statistik] Top 5 Marketplace yang Paling Digemari di Indonesia (Mei 2022)*. <https://dipstrategy.co.id/blog/data-statistik-top-5-marketplace-yang-paling-digemari-di-indonesia/>