
Sistem Rekomendasi Produk Menggunakan Metode *User-Based Collaborative Filtering* Pada *Digital Marketing*

Satia Suhada¹, Saeful Bahri², Setyo Bagus Nugraha³, Taufik Hidayatulloh⁴, Dede Wintana^{5*}

^{1,4}*Program Studi Sistem Informasi, Universitas Bina Sarana Informatika, Sukabumi, Indonesia*

^{2,3,5}*Program Studi Ilmu Komputer, Universitas Bina Sarana Informatika, Sukabumi, Indonesia*

Informasi Artikel

Diterima: 28-05-2023

Direvisi: 26-06-20223

Diterbitkan: 13-07-2023

Kata Kunci

Informasi; Sistem
Rekomendasi, *User-Based
Colaborative Filtering*, Digital
Marketing, RMSE, Cosine
Similarity.

***Email Korespondensi:**

Dede.dwe@bsi.ac.id

Abstrak

Sistem rekomendasi telah diterapkan dalam digital marketing yang digunakan dalam pemasaran produk dan jasa. Sistem rekomendasi digunakan untuk memberikan penawaran barang maupun jasa sesuai dengan kebiasaan dan minat pelanggan terhadap produk dan jasa yang diusulkan, namun pada prakteknya penawaran produk yang tepat untuk pelanggan mengarah pada ide mengembangkan sistem rekomendasi produk. Data pembelian yang didapat dari pelanggan dapat digunakan untuk menganalisis kebutuhan pelanggan dan preferensi produk. Pada sistem rekomendasi, Collaborative Filtering merupakan salah satu algoritma yang paling umum digunakan. Tujuan dari penelitian ini adalah untuk mengetahui seberapa akurat sistem rekomendasi berdasarkan pembelian barang yang serupa antar konsumen menggunakan User-based Collaborative Filtering. Berdasarkan hasil penelitian, User-based Collaborative Filtering menggunakan perhitungan *Cosine Similarity* dapat diterapkan dan menghasilkan 10 rekomendasi produk dengan nilai *RMSE* sebesar 0.9.

Abstract

The recommendation system has been implemented in digital marketing used in marketing products and services. The recommendation system is used to provide offers of goods and services in accordance with customer habits and interests in the proposed products and services, but in practice the right product offering for customers leads to the idea of developing a product recommendation system. Purchase data obtained from customers can be used to analyze customer needs and product preferences. In the recommendation system, Collaborative Filtering is one of the most commonly used algorithms. The purpose of this study is to find out how accurate the recommendation system is based on the purchase of similar goods between consumers using User-based Collaborative Filtering. Based on the results of the study, User-based Collaborative Filtering using Cosine Similarity calculations can be applied and produce 10 product recommendations with an RMSE value of 0.9.

1. Pendahuluan

Dewasa ini, perkembangan teknologi informasi dan komunikasi di Indonesia telah mempengaruhi banyak kalangan, baik masyarakat umum, pemerintah maupun dunia usaha. (Putra et al., 2022), Sistem rekomendasi merupakan salah satu bentuk pemanfaatan teknologi informasi dalam melakukan pemasaran produk dan atau jasa, yang dilakukan secara cepat dan tepat kepada target pasar, sistem rekomendasi telah banyak diterapkan seperti dalam *e-commerce* dan beberapa Media digital seperti, Whatsapp, Facebook, Youtube dan Instagram merupakan platform yang biasa digunakan dalam pemasaran digital. Menurut data dari lembaga pemasaran digital AS: *We Are Social, platform* media sosial yang paling banyak digunakan di Indonesia pada Januari 2017 adalah YouTube (49%) dan Facebook (48%). Tempat berikutnya adalah Instagram (39%), Twitter (38%), Whatsapp (38%) dan Google (36%). Sisanya berbasis FB Messenger, Line, LinkedIn, BBM, Pinterest dan Wechat secara berurutan. (Badan Pusat Statistik, n.d.)(Wahyuni Windasari, 2021) .

Pesatnya perkembangan teknologi tentu mengubah cara memasarkan produk yang sebelumnya konvensional (*offline*) beralih ke digital (*online*). *Digital marketing* muncul sebagai strategi yang lebih prospektif karena dapat memungkinkan calon pelanggan untuk mencari tahu informasi mengenai produk yang akan dibeli serta bertransaksi secara digital. *Digital Marketing* adalah kiat memasarkan produk dan jasa yang dilakukan melalui media digital(Rahmadi, 2020).

Semakin berkembangnya pasar maka produk yang ditawarkan kepada pelanggan menjadi bervariasi dan ini mengakibatkan proses penawaran produk yang sesuai dengan kemauan pelanggan menjadi tantangan tersendiri(Feng et al., 2021), Permasalahan dalam menawarkan produk yang tepat untuk pelanggan mengarah pada ide mengembangkan sistem rekomendasi produk. Data pembelian yang didapat dari pelanggan dapat digunakan untuk menganalisis kebutuhan pelanggan dan preferensi produk. Salah satu algoritma yang paling umum digunakan dalam sistem rekomendasi adalah *Collaborative Filtering* (Alhijawi & Kilani, 2020).

Algoritma *Collaborative Filtering* dapat memberikan rekomendasi berdasarkan pelanggan yang memiliki ketertarikan yang sama dengan pelanggan lainnya. *Collaborative Filtering* (CF) terbagi atas dua jenis yaitu *User-based CF* dan *Item-based CF*. Penelitian ini *User-based Collaborative Filtering* digunakan untuk merekomendasikan produk kepada pelanggan yang mempunyai histori pembelian yang sama dengan *user* lainnya. Sistem rekomendasi ini diharapkan dapat membantu perusahaan dalam menawarkan produk yang tepat untuk pelanggan.

Pendekatan *Item Based Collaborative Filtering* memberikan rekomendasi berdasarkan kesamaan antar item, sedangkan pada pendekatan *User Based Collaborative Filtering* sistem memberikan rekomendasi kepada *User* berdasarkan *Item-Item* yang disukai atau dinilai oleh *User* lain yang memiliki kesamaan. Kelebihan dari pendekatan *User Based Collaborative Filtering* adalah dapat menghasilkan rekomendasi yang berkualitas lebih baik.(Khusna et al., 2021)

2. Metode Penelitian

2.1 Analisis Masalah

Digital Marketing adalah kiat memasarkan produk dan jasa yang dilakukan melalui media digital. Dengan digital marketing, perusahaan dapat dengan mudah menjangkau para calon konsumen. Masalah timbul ketika perusahaan ingin merekomendasikan produk mereka kepada konsumen lama dengan produk yang berbeda. Dengan perilaku konsumen yang berbeda-beda, perusahaan kebingungan dalam merekomendasikan produk yang tepat untuk konsumen lama mereka, data perilaku yang dimiliki hanya data pembelian produk saja dikarenakan data yang didapatkan berasal dari sosial media tempat perusahaan mengiklankan produk mereka.

Untuk mengatasi masalah tersebut, diperlukan suatu sistem rekomendasi produk berdasarkan data pembelian konsumen dan salah satu metode yang cocok untuk mengatasi masalah tersebut yaitu menggunakan metode *User-based Collaborative Filtering* untuk merekomendasikan produk berdasarkan konsumen yang mempunyai kemiripan dalam pembelian produk dengan konsumen lainnya. Dengan adanya sistem rekomendasi ini diharapkan dapat memudahkan perusahaan dalam menargetkan produk mereka kepada konsumen yang mempunyai perilaku pembelian produk yang berbeda-beda secara otomatis dan diharapkan konsumen dapat lebih mudah untuk menemukan produk yang diminati.

2.2 Analisis Kebutuhan

Analisis kebutuhan perangkat lunak berguna untuk mengukur kebutuhan perangkat lunak dan diperlukan untuk membangun sistem rekomendasi. Dalam tahap ini kebutuhan perangkat lunak yang dibutuhkan dalam penelitian ini diantaranya Python 3.8 lalu menggunakan *Windows Subsystem for Linux* (WSL) dengan sistem operasi Ubuntu 20.4 kemudian tools *text editor* dalam penelitian ini menggunakan *Visual Studio Code* dan menggunakan extension *Python Extension Pack* yang tersedia pada *text editor* untuk menjalankan program dan *Microsoft Word* digunakan untuk mengisi laporan penelitian.

2.3 Kebutuhan Sistem

Pada analisis kebutuhan untuk penelitian dibutuhkan data penjualan yang berasal dari salah satu perusahaan yang bergerak di bidang perniagaan digital dan menggunakan bahasa python untuk melakukan *preprocessing* dan pengujian metode pada penelitian ini. Selain itu dibutuhkan juga sebuah *library* yang digunakan untuk memfasilitasi penelitian yang sedang berlangsung. Library yang digunakan dalam penelitian ini adalah *turicreate*, *pandas*, *numpy*, *scikit-learn* dan *flask*. *Turicreate* adalah pustaka yang menyediakan fungsi-fungsi yang dibutuhkan untuk melakukan perhitungan menggunakan metode machine learning. *Pandas* adalah pustaka yang digunakan untuk mengolah data sebelum digunakan pada *Turicreate*. *Numpy*, digunakan untuk mengolah dan menyimpan data primitif. *Scikit-learn* digunakan untuk memisahkan *data training* dan *data testing*. *Flask* digunakan untuk mengimplementasikan model yang telah disimpan ke dalam website.

2.4 Pengumpulan Data

Proses pengumpulan data yang digunakan adalah data sekunder yang di ambil berdasarkan data yang ditambahkan oleh konsumen yang meliputi data konsumen serta produk yang dibeli di market palece dengan jumlah data yang didapatkan sebanyak 180 ribu data penjualan .

2.5 Preprocessing Data

Pada tahap ini, ada beberapa proses yang dilakukan secara berurutan untuk mendapatkan data training, kemudian memecah data menjadi dua, yaitu data pengguna dan data transaksi yang dilakukan oleh pengguna tersebut. Contoh data pembelian terlihat pada gambar 1.

	customerId	product_id	name	price	date	quantity
91330	+628*****	9	ZURIAT 500GR	157000	23/3/2021 09:35:00	2
152831	+628*****	124	RAVENS EGSY - COKLAT MUDA,	153000	21/3/2022 10:36:32	1
45083	+628*****	9	ZURIAT 500GR	157000	3/1/2021 07:49:00	2
144799	+628*****	187	Premium Black Garlic 250GR	300000	4/2/2022 08:29:48	2
67681	+628*****	9	ZURIAT 500GR	87000	8/2/2021 15:25:00	1

Gambar 1. Data Pembelian

Setelah data dibagi menjadi menjadi data dua yakni data pengguna dan transaksi kemudian pada data pembelian produk dipisah dengan membuat baris baru untuk setiap produk id yang berbeda lalu menghitung jumlah produk yang dibeli dan dapat diliaht pada gambar 2.

	customerId	productId	purchase_count
0	+628*****	9	1
1	+628*****	4	1
2	+628*****	9	1
3	+628*****	9	1
4	+628*****	4	1

Gambar 2. Data Transaksi dengan kolom purchase count

Kemudian membuat kolom baru dengan nama *purchase_dummy* dan memberikan nilai 1 untuk setiap baris sebagai tanda produk tersebut telah dibeli oleh konsumen untuk memudahkan proses komputasi yang dapat dilihat pada gambar 3.

	customerid	productid	purchase_count	purchase_dummy
0	+628*****	9	1	1
1	+628*****	4	1	1
2	+628*****	9	1	1
3	+628*****	9	1	1
4	+628*****	4	1	1
...
164979	+628*****	9	1	1
164980	+628*****	122	1	1
164981	+628*****	124	1	1
164982	+628*****	122	1	1
164983	+628*****	9	1	1

Gambar 3. Data Transaksi dengan menambahkan kolom *purchase dummy*

2.6 Pembagian Data

Data dipecah menjadi dua bagian antara lain data training dan data test. Proses pembagian data sangat penting dilakukan sebelum melakukan evaluasi model *User-based Collaborative Filtering*. Umumnya rasio yang digunakan untuk data training lebih besar daripada data testing. Rasio *train-test* yang digunakan untuk model *User-based Collaborative Filtering* adalah 80:20. Pemilihan rasio 80:20 ini didasarkan pada penelitian terdahulu yang menyatakan bahwa untuk rasio data training sebesar 80% akan memberikan performa yang jauh lebih baik, (Bagaskara et al., 2022) terlebih pada dataset yang jumlahnya banyak sehingga terdapat 131987 data training dan 32997 data test. Data training akan digunakan untuk membangun modal prediksi dan *data testing* digunakan untuk mengevaluasi model.

Data *training* dan *data testing* perlu dikonversikan terlebih dahulu ke dalam *SFrame (scalable dataframe)* yang merupakan sebuah tabular objek *dataframe* yang dapat menampung data dengan skala besar yang digunakan untuk memproses data pada *library turicreate*. Konversi data ditunjukkan pada gambar 4.

```
train_data = tc.SFrame(train)
test_data = tc.SFrame(test)
```

Gambar 4. Konversi data training dan data testing

2.7 Membangun Model Dasar

Tahapan selanjutnya adalah membangun model dasar berdasarkan popularitas produk dengan jumlah pembelian yang paling banyak dengan tujuan untuk dijadikan sebagai acuan dasar sistem rekomendasi dengan pendekatan yang lebih mudah. Model dasar juga dapat digunakan untuk konsumen baru yang mengakses website atau aplikasi e-commerce.

Model dasar dibangun menggunakan method *popularity_recommender.create()* lalu menambahkan data training dan fitur yang digunakan yaitu id produk dan nomor konsumen. fitur atau kolom yang digunakan untuk membangun model dasar adalah *purchase_count* yang menghitung banyaknya produk yang dibeli oleh setiap konsumen. Selanjutnya model perlu memanggil method *recommend()* dan parameter yang digunakan adalah sebuah list nomor konsumen dan jumlah barang yang ingin direkomendasikan. Lalu memanggil method *print_rows()* untuk menampilkan hasil proses dan menambahkan jumlah baris yang ingin ditampilkan. Gambar 5 menunjukkan hasil proses rekomendasi berdasarkan popularitas produk.

Dari hasil proses model dasar pada Gambar 5 menampilkan beberapa nilai diantaranya id konsumen, id produk, score dan rank. Score didapatkan dengan cara menghitung rata-rata produk yang dibeli oleh semua konsumen dan rank menampilkan 10 score tertinggi produk untuk direkomendasikan ke konsumen.

customerId	productId	score	rank
(+62)	140	2.0	1
(+62)	205	2.0	2
(+62)	133	2.0	3
(+62)	134	2.0	4
(+62)	136	2.0	5
(+62)	37	1.4736842105263157	6
(+62)	78	1.3333333333333333	7
(+62)	82	1.3333333333333333	8
(+62)	33	1.2857142857142858	9
(+62)	227	1.2857142857142858	10
(+62)	140	2.0	1
(+62)	205	2.0	2
(+62)	133	2.0	3
(+62)	134	2.0	4
(+62)	136	2.0	5
(+62)	37	1.4736842105263157	6
(+62)	78	1.3333333333333333	7
(+62)	82	1.3333333333333333	8
(+62)	33	1.2857142857142858	9
(+62)	227	1.2857142857142858	10
(+62)	140	2.0	1
(+62)	205	2.0	2
..			
(+62)	227	1.2857142857142858	10

Gambar 5. Hasil proses rekomendasi berdasarkan popularitas produk

3. Hasil dan Pembahasan

3.1 Implementasi *Model User-based Collaborative Filtering*

Data yang digunakan oleh model *User-based Collaborative Filtering* sedikit berbeda dengan model dasar terutama pada fitur yang digunakan. Sebelumnya model dasar menggunakan *purchase_count* untuk menghitung nilai rata-rata produk yang dibeli oleh seluruh konsumen. Fitur yang digunakan untuk membangun model collaborative adalah *purchase_dummy* yang menyatakan transaksi atau pembelian yang dilakukan konsumen untuk memudahkan dalam proses komputasi pada model.

3.2 Membangun *Model User-based Collaborative Filtering*

Tahapan selanjutnya adalah membangun model *User-based Collaborative Filtering* menggunakan *method item_similarity_recommender.create()*. Pada tahapan ini dibutuhkan beberapa parameter yaitu, *data training*, id konsumen, id produk, target atau fitur yang digunakan dan *cosine* untuk *similarity_type* yang merupakan metode perhitungan yang dipakai untuk model *User-based Collaborative Filtering*.

3.3 Proses Rekomendasi Model

Setelah model dibangun dan disimpan, lalu memanggil *method recommend()* dan menambahkan parameter daftar nomor konsumen dan jumlah produk yang ingin direkomendasikan. Kemudian memanggil *method print_rows()* untuk menampilkan hasil model *User-based Collaborative Filtering* dan menambahkan nilai parameter untuk mengatur jumlah baris yang ingin ditampilkan. Gambar 6 menunjukkan hasil proses model *User-based Collaborative Filtering*.

customerId	productId	score	rank
+62	3	0.1719377636909485	1
+62	5	0.1502666473388672	2
+62	8	0.05874055624008179	3
+62	58	0.05754208564758301	4
+62	59	0.033487677574157715	5
+62	16	0.018985629081726074	6
+62	7	0.016689300537109375	7
+62	10	0.01659083366394043	8
+62	187	0.013733983039855957	9
+62	14	0.013240575790405273	10
+62	3	0.0	1
+62	124	0.0	2
+62	254	0.0	3
+62	5	0.0	4
+62	187	0.0	5
+62	127	0.0	6
+62	4	0.0	7
+62	122	0.0	8
+62	9	0.0	9
+62	125	0.0	10
+62	3	0.1719377636909485	1
+62	5	0.1502666473388672	2
..			
+62	14	0.013240575790405273	10

Gambar 6. Hasil proses model User-based Collaborative Filtering

3.4 Tahapan Implementasi

3.4.1 Halaman Home

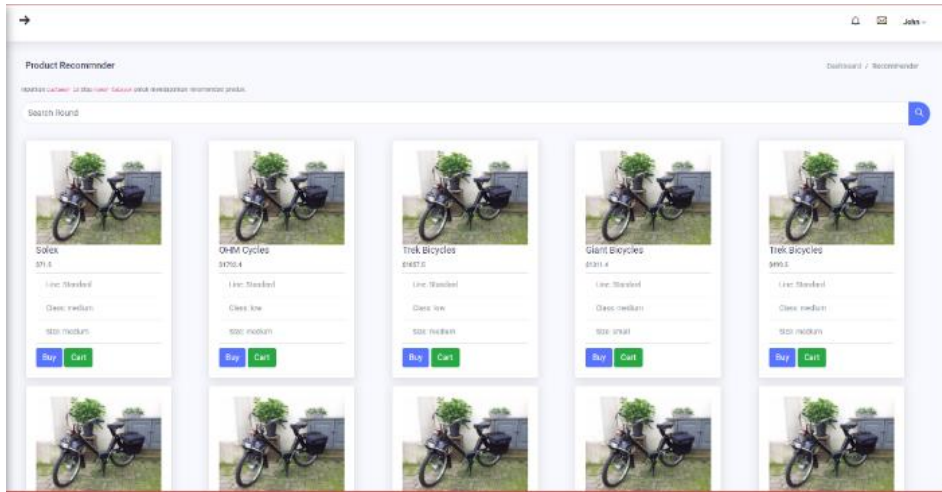
Pada halaman ini terdapat daftar transaksi yang ditampilkan dalam bentuk tabel beberapa variabel yang diantaranya total transaksi, total konsumen dan total produk yang dibungkus menggunakan card. Gambar 7 menunjukkan halaman home.

Name	Position	Office	Salary
Aki Satou	Accountant	Tokyo	\$142,700
Angelica Ramoz	Chief Executive Officer (CEO)	London	\$1,200,000
Ashton Cox	Junior Technical Author	San Francisco	\$80,000
Bradley Errol	Software Engineer	London	\$137,000
Brendan Wagner	Software Engineer	San Francisco	\$200,850
Brett Wilkinson	Integration Specialist	New York	\$177,000
Bruno Nash	Software Engineer	London	\$165,800
Cameron Vanon	Front-end Support	New York	\$106,150
Data Observer	Sales Assistant	New York	\$145,800

Gambar 7. Halaman Home

3.4.2 Halaman Recommender

Halaman *recommender* merupakan halaman yang berfungsi untuk mencari dan menampilkan produk yang direkomendasikan kepada konsumen. Pada halaman ini konsumen menginput nomor mereka ke kolom pencarian kemudian nomor konsumen akan dicocokkan dengan output dari model *collaborative filtering*. Gambar 8 menunjukkan halaman *recommender*.



Gambar 8. Halaman Recommender

3.5 Hasil Pengujian

Pengujian dilakukan terhadap model dasar, model Collaborative Filtering dan website, hal tersebut bertujuan untuk dijadikan acuan mengenai kemampuan sistem yang telah dibuat. Dalam melakukan testing pada model Collaborative Filtering digunakan data transaksi yang berasal dari data testing dengan total data sebanyak 2323 transaksi yang terdiri dari 23 konsumen yang masing-masing akan mendapatkan 10 rekomendasi produk. Untuk menguji kedua model sistem rekomendasi, maka digunakan konsep precision dan recall pada nilai k. Untuk mendapatkan nilai mean precision dan recall maka digunakan rumus.

$$Precision(k) = \frac{\text{Jumlah produk relevan } k \text{ yang direkomendasikan}}{\text{Jumlah produk } k \text{ yang direkomendasikan}}$$

$$Recall(k) = \frac{\text{Jumlah produk relevan } k \text{ yang direkomendasikan}}{\text{Total produk relevan}}$$

Dimana nilai k merupakan banyaknya produk yang disarankan dan produk relevan merupakan semua produk yang ada didataset. Kemudian menghitung Root Mean Squared Error (RMSE) untuk mengetahui besarnya nilai error yang didapat pada kedua model. Untuk mendapatkan nilai RMSE maka digunakan rumus.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

Dimana y_i sama dengan data sebenarnya (data awal), \hat{y}_i sama dengan data hasil estimasi (data akhir), n sama dengan jumlah data. Seberapa akurat hasil pengukuran estimasi diindikasikan dengan hasil RMSE yang mempunyai nilai kecil mendekati nol (Sutawinaya et al., 2017). Untuk pengujian aplikasi digunakan nomor konsumen yang kemudian diinputkan ke dalam kolom pencarian lalu menekan tombol cari untuk mengetahui produk yang direkomendasikan oleh salah satu model yang terdiri dari 10 produk.

3.6 Pengujian Model

Pengujian pada kedua model dilakukan terhadap konsumen menggunakan data testing yang berjumlah 3674 data transaksi. Tujuan dari pengujian ini adalah untuk mengetahui nilai precision-recall and nilai error yang didapat.

Sutawinaya, I. P., Astawa, I. N. G. A., & Hariyanti, N. K. D. (2017). Perbandingan Metode Jaringan Saraf Tiruan Pada Peramalan Curah Hujan. *Prosiding SNSebatik 2017*.

Wahyuni Windasari, T. Z. (2021). Pemetaan dan Analisis Faktor yang Mempengaruhi Persentase Usaha E-Commerce di Indonesia. *Akuntansi, Ekonomi Dan Manajemen Bisnis*, 9(2), 130–135.