

Sistem Rekomendasi Produk Kebutuhan Pada E-Commerce Menggunakan Metode K-Nearest Neighbor

William Saputra
Program Studi Informatika, Fakultas MIPA, Universitas Udayana
Jalan. Kampus Bukit Jimbaran, Badung, Bali, Indonesia
Williamsyt2005@gmail.com

Abstract

The growing number of product information available on the internet brings challenges to both customer and online businesses in the e-commerce environment. Customer often have difficulty when looking for products on the internet because of the number of products sold on the internet. In addition, online businessman often experience difficulties because they has much data about products, customers and transactions, thus causing online businessman have difficulty to promote the right product to a particular customer target. A recommendation system was developed to address those problem with various methods such as Collaborative Filtering, ContentBased, and Hybrid. Collaborative filtering method uses customer's rating data, content based using product content such as title or description, and hybrid using both as the basis of the recommendation. In this research, the k-nearest neighbor algorithm is used to determine the top-n product recommendations for each buyer. The result of this research method Content Based outperforms other methods because the sparse data, that is the condition where the number of rating given by the customers is relatively little compared the number of products available in e-commerce.

Keyword: KKN, K-Nearset Neighbor, E-commerce, Rekomendasi

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam dunia perdagangan. *E-commerce* atau perdagangan elektronik telah menjadi salah satu pilar utama dalam aktivitas jual beli di era digital. Kemudahan akses, kenyamanan, dan variasi produk yang ditawarkan oleh platform e-commerce membuatnya semakin diminati oleh konsumen. Namun, dengan banyaknya pilihan produk yang tersedia, sering kali konsumen merasa kesulitan untuk menemukan produk yang sesuai dengan kebutuhan dan preferensinya.

Sistem rekomendasi produk muncul sebagai solusi untuk mengatasi permasalahan tersebut. Sistem ini bertujuan untuk membantu konsumen menemukan produk yang relevan berdasarkan data historis dan preferensi pribadi mereka. Salah satu metode yang efektif untuk mengimplementasikan sistem rekomendasi adalah metode K-Nearest Neighbor (K-NN). Metode K-NN adalah salah satu teknik pembelajaran mesin yang bekerja dengan cara mengidentifikasi item-item yang mirip berdasarkan kedekatan fitur-fitur yang dimiliki.

Konsep sistem rekomendasi telah digunakan oleh berbagai bisnis online seperti amazon.com dan ebay.com sebagai alat bisnis. Sistem rekomendasi dilaporkan telah meningkatkan penjualan produk dan membangun loyalitas pembeli (Mobasher, 2007). Penelitian ini dilakukan dengan tujuan untuk mengembangkan algoritma K-NN untuk sistem rekomendasi produk pada e-commerce, menguji keakuratan dan efektivitas rekomendasi yang dihasilkan oleh sistem, serta menganalisis dampak penerapan sistem rekomendasi terhadap kepuasan pengguna dan peningkatan penjualan pada platform e-commerce.

2. Metode Penelitian

Algoritma KNN adalah salah satu algoritma yang sering digunakan untuk melakukan klasifikasi. Algoritma termasuk dalam algoritma *lazy learning* yang mudah untuk diimplementasikan (Alkhatib et al., 2013). Dalam penggunaan algoritma KNN dibagi menjadi dua bagian yaitu data latih dan data uji. Data latih digunakan untuk melakukan dasar prediksi, sedangkan data uji akan dilakukan untuk

menampilkan nilai-nilai yang diprediksi dalam algoritma (Imandoust dan Bolandraftar, 2013). Adapun langkah-langkah dari algoritma K-Nearest Neighbor sendiri adalah sebagai berikut:

1. Menentukan parameter k (jumlah tetangga terdekat)
2. Hitung jarak data latih dengan semua data uji
3. Urutkan jarak tersebut berdasarkan nilai yang terkecil sejumlah k.
4. Tentukan kelompok data uji berdasarkan label mayoritas pada k.

Untuk menghitung jarak antara dua titik pada algoritma KNN digunakan metode *Euclidean Distance* yang dapat digunakan pada *1-dimensional space*, *2-dimensional space*, atau *multi-dimensional space*.

1-dimensional space berarti perhitungan jarak hanya menggunakan satu variabel bebas (*independent variable*), *2-dimensional-space* berarti ada dua variabel bebas, dan *multi-dimensional space* berarti ada lebih dari dua variabel.

Secara formula, perumusan dari *Euclidean distance* pada *1-dimensional space* adalah sebagai persamaan (1) berikut.

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2} \quad (1)$$

Lalu untuk penghitungan lebih dari satu maka penjumlahan dari persamaannya akan sebagai persamaan (2) berikut.

$$dis = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2 + (y_{1i} - y_{2i})^2 + \dots} \quad (2)$$

3. Hasil dan Pembahasan

Pada algoritma KNN (K-Nearest Neighbor) sendiri, kita akan simulasi sebuah masalah dan membuat pemrograman tersebut dengan implementasi bahasa java.

Semisal disediakan sebuah data tabel sebagai acuan kebutuhan yang penting dari perlengkapan yang dibutuhkan oleh sang pengguna. Dimana disini permasalahannya adalah pengguna nantinya akan memasukkan jumlah data sesuai yang diperlukan untuk mengetahui urutan rekomendasi kebutuhan apa yang diperlukan terlebih dahulu

Table 1. Data Kebutuhan

Age	Income
29	350
51	430
33	290

Pengguna memiliki data rentang *age* bernilai 34 dan nilai *income* sebanyak 390. Sesuai dengan perhitungan pada rumus sebelumnya maka, dapat ditulis sebagai berikut.

1. Data 1

$$dis = \sqrt{\sum_{i=0}^n (34 - 29)^2 + (390 - 350)^2} = 40,31 \quad (3)$$

2. Data 2

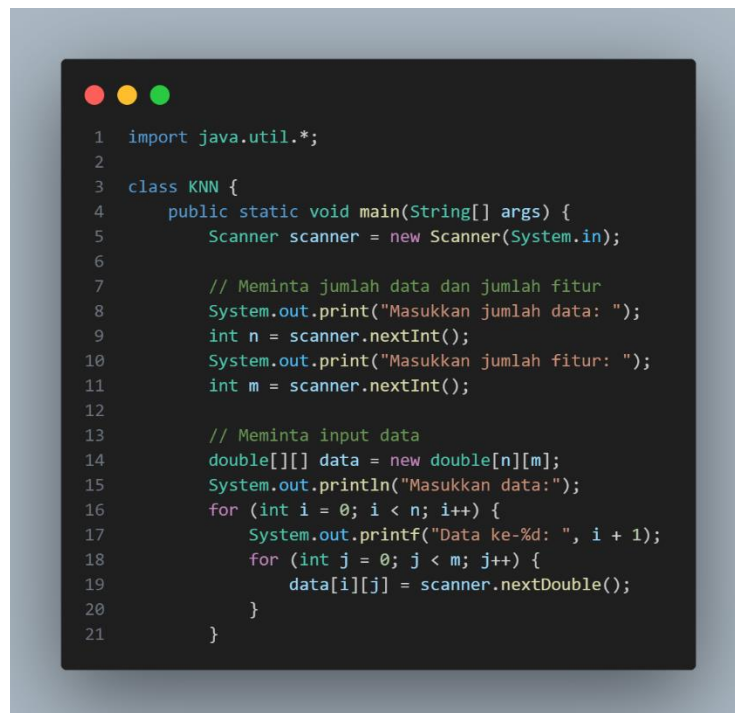
$$dis = \sqrt{\sum_{i=0}^n (34 - 51)^2 + (390 - 430)^2} = 43,46 \quad (4)$$

3. Data 3

$$dis = \sqrt{\sum_{i=0}^n (34 - 33)^2 + (390 - 290)^2} = 100,01 \quad (5)$$

Dari hasil perhitungan oleh data yang sudah ada, maka pengguna akan mengetahui apa yang akan direkomendasikan oleh system untuk keperluan yang dibutuhkan. Mulai dari data pertama, lalu kedua, dan diakhiri oleh data ke tiga. Hal ini akan memudahkan pengguna untuk mengatur dan memenuhi kebutuhan sesuai apa yang direkomendasikan dari metode K-Nearest Neighbor.

Adapun penggunaan implementasi dari pemrograman bahasa Java yang telah di buat.



```
1  import java.util.*;
2
3  class KNN {
4      public static void main(String[] args) {
5          Scanner scanner = new Scanner(System.in);
6
7          // Meminta jumlah data dan jumlah fitur
8          System.out.print("Masukkan jumlah data: ");
9          int n = scanner.nextInt();
10         System.out.print("Masukkan jumlah fitur: ");
11         int m = scanner.nextInt();
12
13         // Meminta input data
14         double[][] data = new double[n][m];
15         System.out.println("Masukkan data:");
16         for (int i = 0; i < n; i++) {
17             System.out.printf("Data ke-%d: ", i + 1);
18             for (int j = 0; j < m; j++) {
19                 data[i][j] = scanner.nextDouble();
20             }
21         }
22     }
23 }
```

Gambar 1. Java 1

```

1 // Meminta input data yang akan diprediksi
2 double[] newData = new double[m];
3 System.out.println("\nMasukkan data baru yang akan diprediksi:");
4 for (int i = 0; i < m; i++) {
5     System.out.printf("Fitur ke-%d: ", i + 1);
6     newData[i] = scanner.nextDouble();
7 }
8
9 // Meminta input nilai k
10 System.out.print("\nMasukkan nilai k: ");
11 int k = scanner.nextInt();
12
13 // Menghitung jarak Euclidean dan menemukan k tetangga terdekat
14 double[][] distances = new double[n][2];
15 for (int i = 0; i < n; i++) {
16     double dist = 0.0;
17     for (int j = 0; j < m; j++) {
18         dist += Math.pow(data[i][j] - newData[j], 2);
19     }
20     distances[i][0] = Math.sqrt(dist); // Jarak Euclidean
21     distances[i][1] = i; // Index data
22 }
23

```

Gambar 2. Java 2

```

1 // Mengurutkan berdasarkan jarak (kolom pertama)
2 Arrays.sort(distances, Comparator.comparingDouble(a -> a[0]));
3
4 // Menampilkan k tetangga terdekat
5 System.out.printf("\n%d Tetangga terdekat:\n", k);
6 for (int i = 0; i < k; i++) {
7     int index = (int) distances[i][1];
8     System.out.printf("Data ke-%d dengan jarak %.2f\n", index + 1, distances[i][0]);
9 }
10 scanner.close();
11 }
12 }
13 }
14

```

Gambar 3. Java 3

Dengan hasil output sebagai berikut

```

Masukkan jumlah data: 3
Masukkan jumlah fitur: 2
Masukkan data:
Data ke-1: 29 350
Data ke-2: 51 430
Data ke-3: 33 290

Masukkan data baru yang akan diprediksi:
Fitur ke-1: 34
Fitur ke-2: 390

Masukkan nilai k: 3
3 Tetangga terdekat:
Data ke-1 dengan jarak 40.31
Data ke-2 dengan jarak 43.46
Data ke-3 dengan jarak 100.00

=== Code Execution Successful ===

```

Gambar 4. Output

4. Kesimpulan

Implementasi teori metode *K-Nearest Distance* dalam sistem rekomendasi e-commerce memberikan keuntungan signifikan dalam hal akurasi dan relevansi. Penilaian dengan menggunakan perhitungan *euclidean distance* membantu penuh dalam penerapan dalam penggunaan dalam bahasa *java*. Namun penggunaan ini masih perlu dioptimalkan dengan bantuan *cosine similarity* dan gagasan *confusion matrix* untuk mendapatkan nilai *precision*, *recall*, dan *f-measure*. Penelitian lebih lanjut diperlukan untuk mengoptimalkan algoritma dan mengeksplorasi potensi integrasi dengan metode *machine learning* lainnya.

References

- [1] Alkhatib, K., Najadat, H., Hmeidi, I., & Shatnawi, M. K. A. (2013). Stock price prediction using k-nearest neighbor (kNN) algorithm. *International Journal of Business, Humanities and Technology*, 3(3), 32-44.
- [2] Imandoust, S. B., & Bolandraftar, M. (2013). Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background. *International journal of engineering research and applications*, 3(5), 605-610.
- [3] Prasetya, C. S. D. (2017). Sistem Rekomendasi Pada E-Commerce Menggunakan K-Nearest Neighbor. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)* p-ISSN, 2355, 7699.
- [4] Aisha, D., & Kusumawati, R. (2022). IMPLEMENTASI METODE ALGORITMA COLLABORATIVE FILTERING DAN K-NEAREST NEIGHBOR PADA SISTEM REKOMENDASI E-COMMERCE. *Jurnal ilmiah Sistem Informasi dan Ilmu Komputer*, 2(3), 25-38.
- [5] Dharmawan, H., Hilabi, S. S., & Karniawulan, I. (2023). Sistem Rekomendasi Buku dengan Metode K-Nearest Neighbor (K-NN) pada Gramedia. *ZONAsi: Jurnal Sistem Informasi*, 5(1), 16-25.
- [6] Rahardja, C. A., Juardi, T., & Agung, H. (2019). Implementasi algoritma K-Nearest Neighbor pada website rekomendasi laptop. *Jurnal Buana Informatika*, 10(1), 75-84.
- [7] Sebastian, D. (2019). Implementasi Algoritma K-Nearest Neighbor untuk Melakukan Klasifikasi Produk dari beberapa E-marketplace. *Jurnal Teknik Informatika dan Sistem Informasi*, 5(1).
- [8] Muhammad, D. I., Ermatita, E., & Fali, N. (2021). Penggunaan K-Nearest Neighbor (KNN) untuk Mengklasifikasi Citra Belimbing Berdasarkan Fitur Warna. *Informatik: Jurnal Ilmu Komputer*, 17(1), 9-16.
- [9] Afifah, L., "Algoritma K-Nearest Neighbor (KNN) untuk Klasifikasi", 23 November 2020. [Online]. Available: <https://ilmudatapy.com/algoritma-k-nearest-neighbor-knn-untuk-klasifikasi/>. Diakses pada 9 Juli 2024.