

## Rekomendasi Pemilihan Baju Wanita dengan Metode Naive Bayes

Ari Avianto<sup>1\*</sup>, Nur Nafi'iyah<sup>1</sup>, Nur Qomariyah Nawafilah<sup>3</sup>

<sup>1,3</sup>Program Studi Teknik Informatika, Fakultas Teknik, Universitas Islam Lamongan  
Jalan Veteran Nomor 53A Lamongan, 62211  
e-mail: ariavianto01@gmail.com

**Abstrak**— Algoritma Naive Bayes dapat digunakan untuk rekomendasi baju wanita. Metode ini akan diterapkan untuk memberikan rekomendasi pemilihan baju wanita sesuai kelas baju karakter konsumen. Tujuan penelitian, yaitu membuat sistem yang dapat digunakan untuk memberikan rekomendasi baju dengan algoritma Naive Bayes. Kategori baju yang direkomendasikan adalah *Blouses*, *Dresses*, *Jackets* dan *Jeans*. Jumlah dataset yang digunakan sebanyak 500 data yaitu 490 data *training* dan 10 data *testing* yang terbagi kedalam 7 atribut yaitu *Age*, *Rating*, *Recommended* IND, *Positive Feedback Count*, *Division Name*, *Departement Name* dan *Class Name*, aplikasi penerapan metode *naive bayes* ini dapat memberikan Sistem Pendukung Keputusan (SPK) Rekomendasi Pemilihan Baju Wanita dengan akurasi 80 %.

Kata kunci: *rekomendasi baju wanita, naive bayes*.

### 1. Pendahuluan

Pembelian dan penjualan secara online menjadi suatu perkembangan dalam dunia perdagangan. Karena masyarakat dapat membeli baju atau menjual baju tanpa harus bertatap muka. Dengan banyaknya website atau e-commerce yang menyediakan produk baju ataupun lainnya dalam jual beli. Maka dari transaksi tersebut akan terekap dalam dataset. Kumpulan dataset tersebut dapat digunakan sebagai informasi atau pengetahuan. Transaksi jual beli yang sudah terekap dan tersimpan dalam database dapat diolah. Data mining adalah proses mengolah dan menggali informasi dan pengetahuan dari kumpulan dataset. Proses menggali akan dapat digunakan untuk memberikan rekomendasi pada pembeli baru. Algoritma yang dapat digunakan untuk menggali data, salah satunya algoritma naive bayes. Di mana proses mendapatkan pengetahuan berdasarkan dataset yang dihitung nilai probabilitasnya[1]. Dalam penelitian ratih menyebutkan bahwa pembeli akan kebingungan dengan banyaknya pilihan baju yang ditawarkan dalam e-commerce. Dan pembeli akan lebih dimudahkan jika sistem atau e-commerce memberikan rekomendasi baju yang sesuai tipe tubuh, bahan, model, warna dan harga berdasarkan data transaksi jual beli di e-commerce. Kategori yang dipakai untuk memberikan rekomendasi berdasarkan style, price, rating, size, season, neckline, sleevelength, waiseline, material, febrictype, decoration, patterntype, recommendation[1].

Penelitian lainnya terkait algoritma Naive Bayes yang digunakan pada website e-commerce adalah Zia Ayu. Di mana algoritma naive bayes digunakan untuk mengklasifikasi ulasan atau komentara para pembeli di aplikasi Shopee. Proses klasifikasi ulasan atau komentar mengenai pelayanan dan produk, seperti dari hal ongkir, buruk, susah, gambar,

pengiriman, dan penjual. Kelas klasifikasi terhadap komentar adalah puas atau tidak puas, atau positif dan negatif[2]. Klasifikasi dapat menggunakan berbagai algoritma, dan algoritma naive bayes memberikan nilai akurasi yang tinggi. Pada penelitian Zia Ayu, proses klasifikasi komentar pembeli pada aplikasi Shopee akurasi 97,4%[2]. Pada penelitian Nur Nafi'iyah, dan Mochammad Ainun dengan menggunakan algoritma Naive Bayes secara berturut-turut nilai akurasi 80%[5], dan 74,5%[6].

Dari beberapa penelitian sebelumnya maka peneliti ingin mengimplementasikan algoritma naive bayes dalam memberikan rekomendasi baju wanita berdasarkan kriteria age, rating, recommended, positive feedback count, devision name, departement name. Dan kelas rekomendasi Blouses, Dresses, Jackets, dan Jeans.

### 2. Tinjauan Pustaka

#### 2.1. Naive Bayes

Naive bayes suatu algoritma yang menerapkan nilai probabilitas dari suatu kemungkinan antara data query terhadap dataset. Di mana data query akan dihitung nilai probabilitasnya berdasarkan dataset. Proses menghitung nilai probabilitas dengan Persamaan 1.

$$p(H | E) = \frac{p(E | H) * p(H)}{p(E)}$$

Di mana  $p(H|E)$  adalah nilai probabilitas hipotesis terhadap evidence/variabel. Dan  $p(E|H)$  adalah nilai probabilitas evidence query terhadap semua kelas hipotesis. Dan  $p(H)$  adalah nilai

probabilitas semua kelas terhadap semua dataset. P(E) adalah nilai probabilitas evidence (semua variabel input). Dan untuk menghitung probabilitas hipotesis dalam Persamaan 2.

$$p(H) = \frac{N_{kelas}}{N_{seluruh\_dataset}} \quad 2$$

Algoritma dari data mining yang juga dapat digunakan dalam jual beli online atau e-commerce adalah association Rule. Di mana algoritma association rule yang digunakan dalam penelitian Arif Ismail dan kawan-kawan adalah Apriori. Algoritma Apriori digunakan untuk strategi menyusun etalase penjualan pakaian. Etalase berjalan secara online perlu disusun secara strategis agar dapat meningkatkan penjualan[3]. Hasil dari penelitian Arif Ismail mengatakan bahwa adanya hubungan antara baju yang dibeli bersamaan terhadap penyusunan etalase[3]. Market Analisis juga salah satu algoritma data mining association rule, di mana market basket analisis dapat digunakan untuk menyusun tata letak jualan pada Minimarket Aira. Salah satu contoh penyusunan tata letak jaluannya, yaitu antara menyusun snack kacang dengan minuman kopi. Ditunjukkan dari hasil market basket analisis konsumen yang membeli snack kacang maka 97,2% akan membeli snack lainnya, dan 95% akan membeli kopi[4].

### 3. Metode Penelitian

Sistem rekomendasi pemilihan baju wanita menggunakan metode *Naïve bayes* mempunyai beberapa tahapan. Langkah pertama yang dilakukan dalam metode *Naïve bayes* yaitu menghitung nilai probabilitas P(H) dengan cara menghitung jumlah item berdasarkan *class name* kemudian dibagi dengan jumlah dataset keseluruhan seperti dalam Persamaan 2. Langkah kedua menghitung data item query yang diinputkan dibagi dengan nilai jumlah tiap kelas tersebut P(E|H). Kemudian mengkalikan nilai P(E|H) berdasarkan *class name* nya masing-masing, langkah terakhir yaitu mengkalikan hasil nilai P(E|H) dengan nilai probabilitas setiap kelas P(H). Sehingga akan didapatkan hasil dari item rekomendasi untuk konsumen masuk kedalam *class name* yang nilai probabilitas tertinggi. Dalam proses perhitungan naive bayes dibutuhkan data training, dalam kasus ini penulis menggunakan data pemilihan baju wanita sebagai data training dalam Tabel 1.

Berdasarkan Tabel 1 dapat dihitung klasifikasi data baju wanita apabila diinputkan Data testing: X = (Age = 54, Rating = 5, Recommended IND = 1, Positif feedback count = 0, Division name = General,

Departement name = Dresses, Class name = Dresses) menggunakan metode *Naïve Bayes*, maka klasifikasi data baju wanita dapat ditentukan melalui langkah berikut:

#### P (Ci)

P (Class name = "Dresses") = 42/114 =

0,3684210526

P (Class name = "Blouses") = 61/114 =

0,5350877193

P (Class name = "Jackets") = 5/114 = 0,0438596491

P (Class name = "Jeans") = 6/114 = 0,0526315789

#### P (X|Ci)

P (Age = "54" | Class name = "Dresses") = 1/42 =

0,0238095238

P (Age = "54" | Class name = "Blouses") = 1/61 =

0,0163934426

P (Age = "54" | Class name = "Jackets") = 0/5 = 0

P (Age = "54" | Class name = "Jeans") = 0/6 = 0

P (Rating = "5" | Class name = "Dresses") = 19/42 =

0,4523809524

P (Rating = "5" | Class name = "Blouses") = 28/61 =

0,4590163934

P (Rating = "5" | Class name = "Jackets") = 3/5 = 0,6

P (Rating = "5" | Class name = "Jeans") = 4/6 =

0,6666666667

P (Recommended IND = "1" | Class name =

"Dresses") = 35/42 = 0,8333333333

P (Recommended IND = "1" | Class name =

"Blouses") = 47/61 = 0,7704918033

P (Recommended IND = "1" | Class name =

"Jackets") = 5/5 = 1

P (Recommended IND = "1" | Class name =

"Jeans") = 6/6 = 1

P (Positif feedback count = "0" | Class name =

"Dresses") = 20/42 = 0,4761904762

P (Positif feedback count = "0" | Class name =

"Blouses") = 23/61 = 0,3770491803

P (Positif feedback count = "0" | Class name =

"Jackets") = 4/5 = 0,8

P (Positif feedback count = "0" | Class name =

"Jeans") = 2/6 = 0,3333333333

P (Division name = "General" | Class name =

"Dresses") = 27/42 = 0,6428571429

P (Division name = "General" | Class name =

"Blouses") = 43/61 = 0,7049180328

P (Division name = "Generale" | Class name =

"Jackets") = 4/5 = 0,8

P (Division name = "General" | Class name =

"Jeans") = 6/6 = 1

P (Departement name = "Dresses" | Class name =

“Dresses”) = 42/42 = 1  
 P (Departement name = “Dresses” | Class name = “Blouses”) = 0/61 = 0  
 P (Departement name = “Dresses” | Class name = “Jackets”) = 0/5 = 0  
 P (Departement name = “Dresses” | Class name = “Jeans”) = 0/6 = 0

$P(X| \text{Class name} = \text{“Dresses”}) = 0,0238095238 * 0,4523809524 * 0,8333333333 * 0,4761904762 * 0,6428571429 * 1 = 0,0027476976$   
 $P(X| \text{Class name} = \text{“Blouses”}) = 0,0163934426 * 0,4590163934 * 0,7704918033 * 0,3770491803 * 0,7049180328 * 0 = 0$   
 $P(X| \text{Class name} = \text{“Jackets”}) = 0 * 0,6 * 1 * 0,8 * 0,8 * 0 = 0$   
 $P(X| \text{Class name} = \text{“Jeans”}) = 0 * 0,6666666667 * 1 * 0,3333333333 * 1 * 0 = 0$

**P (X|Ci) \* P (Ci)**  
 P (Class name = “Dresses”) \* P(Class name = “Dresses”) = 0,0027476976 \* 0,3684210526 = 0,0010122807  
 P (Class name = “Blouses”) \* P(Class name = “Blouses”) = 0 \* 0,5350877193 = 0  
 P (Class name = “Jackets”) \* P(Class name = “Jackets”) = 0 \* 0,0438596491 = 0  
 P (Class name = “Jeans”) \* P(Class name = “Jeans”) = 0 \* 0,0526315789 = 0

Jadi, untuk Age = 54, Rating = 5, Recommended IND = 1, Positif feedback count = 0, Division name = General, Departement name = Dresses, Class name = Dresses masukke Class name “Dresses”.

10	50	3	1	1	General	Dresses	Dresses
11	41	5	1	0	General	Tops	Blouses
12	47	5	1	0	General	Dresses	Dresses
13	33	4	1	2	General	Tops	Blouses
14	55	4	1	14	General	Dresses	Dresses
15	31	2	0	7	General	Dresses	Dresses
16	34	3	1	0	General	Dresses	Dresses
17	55	5	1	0	General	Tops	Blouses
18	52	5	1	0	General	Tops	Blouses
19	39	3	1	3	General	Dresses	Dresses
20	36	4	1	0	General Petite	Dresses	Dresses
21	41	5	1	0	General Petite	Dresses	Dresses
22	56	3	0	1	General Petite	Dresses	Dresses
23	38	4	1	10	General Petite	Tops	Blouses
24	36	2	0	0	General	Tops	Blouses
25	27	5	1	4	General	Tops	Blouses
26	59	5	1	1	General Petite	Dresses	Dresses
27	48	5	1	5	General Petite	Tops	Blouses
28	31	4	1	0	General	Dresses	Dresses
29	28	2	0	0	General Petite	Tops	Blouses
30	47	5	1	1	General	Tops	Blouses
31	62	4	1	2	General Petite	Dresses	Dresses
32	68	5	1	1	General	Dresses	Dresses
33	68	5	1	3	General	Tops	Blouses
34	39	4	1	0	General Petite	Tops	Blouses
35	39	4	1	9	General Petite	Dresses	Dresses
36	38	2	0	4	General Petite	Tops	Blouses
37	24	4	1	3	General	Tops	Blouses
38	51	4	1	0	General Petite	Dresses	Dresses
39	29	5	1	5	General Petite	Tops	Blouses
40	31	5	1	0	General	Dresses	Dresses
41	23	4	1	3	General Petite	Tops	Blouses
42	44	1	0	0	General Petite	Tops	Blouses
43	69	5	1	2	General	Tops	Blouses
44	23	2	0	9	General	Tops	Blouses
45	66	5	1	0	General	Tops	Blouses
46	47	4	1	11	General	Tops	Blouses
47	66	4	1	0	General	Tops	Blouses
48	23	5	1	3	General	Tops	Blouses
49	47	5	1	0	General	Dresses	Dresses
50	44	3	1	4	General	Dresses	Dresses
51	50	5	1	24	General	Tops	Blouses
52	61	5	1	3	General	Dresses	Dresses
53	32	2	0	0	General	Dresses	Dresses
54	51	4	1	1	General	Dresses	Dresses

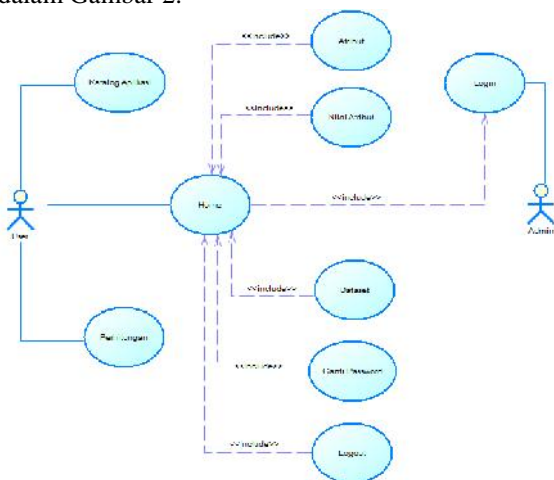
No	Age	Rating	Recommended IND	Positive Feedback Count	Division Name	Department Name	Class Name
1	34	5	1	4	General	Dresses	Dresses
2	60	3	0	0	General	Dresses	Dresses
3	47	5	1	6	General	Tops	Blouses
4	49	2	0	4	General	Dresses	Dresses
5	24	5	1	0	General	Dresses	Dresses
6	34	5	1	0	General	Dresses	Dresses
7	53	3	0	14	General	Dresses	Dresses
8	39	5	1	2	General Petite	Dresses	Dresses
9	53	5	1	2	General Petite	Dresses	Dresses



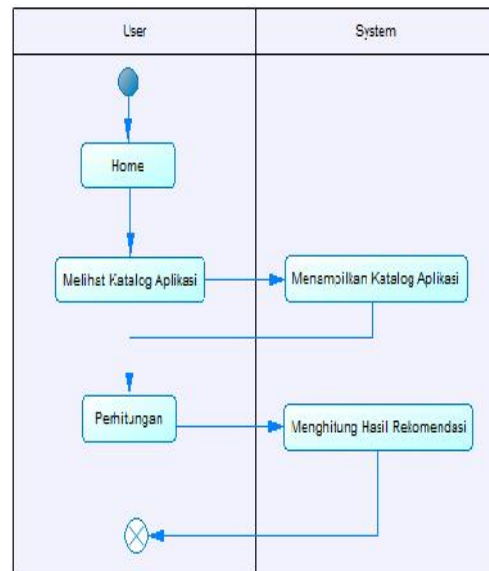
55	32	5	1	2	General Petite	Dresses	Dresses
56	37	2	0	0	General	Tops	Blouses
57	54	5	1	4	General	Tops	Blouses
58	63	5	1	0	General	Jackets	Jackets
59	29	5	1	0	General Petite	Dresses	Dresses
60	24	5	1	0	General	Tops	Blouses
61	34	3	1	1	General	Dresses	Dresses
62	52	4	1	0	General	Jackets	Jackets
63	62	4	1	3	General Petite	Tops	Blouses
64	39	3	1	0	General	Dresses	Dresses
65	36	5	1	0	General	Jackets	Jackets
66	65	5	1	12	General Petite	Tops	Blouses
67	31	4	1	0	General	Jackets	Jackets
68	52	4	1	0	General	Tops	Blouses
69	36	5	1	0	General	Tops	Blouses
70	39	4	1	2	General	Tops	Blouses
71	39	3	0	9	General	Tops	Blouses
72	39	3	1	2	General	Tops	Blouses
73	53	3	0	0	General	Tops	Blouse
74	36	3	1	1	General	Tops	Blouses
75	44	3	1	0	General Petite	Tops	Blouses
76	48	5	1	0	General Petite	Tops	Blouses
77	53	4	1	7	General Petite	Tops	Blouses
78	33	4	1	0	General Petite	Tops	Blouses
79	52	1	0	8	General Petite	Tops	Blouses
80	64	4	1	2	General Petite	Tops	Blouses
114	25	5	1	6	General	Dresses	Dresses

Tabel 1. Data Training

Alur sistem yang akan dibangun seperti dalam Gambar 1. Dan alur dari sistem untuk pengguna dalam Gambar 2.



Gambar 1. Usecase Diagram



Gambar 2. Alur Kegiatan Pengguna Sistem

#### 4. Hasil dan Pembahasan

Sistem ini dibangun menggunakan bahasa PHP dan berbasis web offline. Di mana terdapat beberapa tabel yang berelasi. Diantaranya ada Tabel data training, dan ada Tabel data testing. Tabel data training berisikan dataset yang digunakan sebagai acuan menghitung nilai probabilitas. Dan Tabel data testing sebagai tempat menyimpan data ujicoba dan sekaligus sebagai proses perhitungan nilai akurasi. Sistem ini mempunyai 4 kelas dapat memberikan rekomendasi ke pembeli, yaitu dalam tabel 2. Dan atribut atau variabel yang digunakan, yaitu dalam tabel 3.

No	Nilai Atribut
1	Blouses
2	Dresses
3	Jackets
4	Jeans

Tabel 2. Kelas Sistem Rekomendasi Baju Wanita

No	Atribut
1	Age
2	Rating
3	Recommended IND
4	Positive Feedback
5	Devision Name
6	Departement Name

Tabel 3. Atribut atau Variabel Input

Di mana setiap value dari inputan variabel dalam karakter atau number. Misalnya inputan Age dalam number, inputan rating dalam angka akan tetapi mempunyai interval [1 5]. Variabel yrecommened IND juga berupa number dan mempunyai interval pilihan 0 dan 1. Variabel positive feedback juga berupa number yang berkisar antara 1-100. Sedangkan variabel devision name dan departement name berupa karakter dan kategori. Misalnya devision name kategorinya dalam Tabel 4, dan departement name mempunyai kategori dalam Tabel 5.

No	Value Kategori
1	General
2	General Petite

Tabel 4. Nilai Kategori dari Varibel Devision Name

No	Value Kategori
1	Dresses
2	Tops
3	Jackets
4	Bottoms

Tabel 5. Nilai Kategori dari Variabel Departement Name

## 5. Kesimpulan

Sistem ini dibangun menggunakan metode *naive bayes* sehingga dapat membantu menyelesaikan permasalahan pemilik toko saat merekomendasikan baju wanita kepada pembeli. Sistem ini disajikan dengan pengklasifikasian baju wanita dengan metode *naive bayes*, sehingga mampu memberi rekomendasi pada pemilik toko berdasarkan *class* "Dresses", "Blouses", "Jackets" dan "Jeans". Hasil dari perhitungan program berdasarkan data testing akurasi adalah 80%.

## 6. Daftar Pustaka

- [1] R. Y. Hayuningtyas, "Penerapan Algoritma Naive Bayes untuk Rekomendasi Pakaian Wanita," *Jurnal Informatika*, vol. 6, no. 1, pp. 18-22, 2019.
- [2] Z. A. N. Gumilang, "Implementasi Naive Bayes Classifier dan Asosiasi untuk Analisis Sentimen Data Ulasan Aplikasi E-Commerce Shopee pada Situs Google Play," Skripsi-Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia Yogyakarta, 2018.
- [3] Arif Ismail Husin, Farida Mulyaningsih, "Penerapan Metode Data Mining Analisis Terhadap Data Penjualan Pakaian dengan Algoritma Apriori," in *SNIPTEK STMIK Nusa Mandiri*, Jakarta, 2015.
- [4] J. R. Putri, "Perancangan Ulang Tata Letak Berdasarkan Market Basket Analisis," Skripsi-Sekolah Tinggi Ilmu Ekonomi, Bandung, 2015.
- [5] Nur Nafi'iyah, Siti Mujilawati, "Analisis Algoritma Backpropagation dan Naive Bayes," in *SENASIF*, Malang, 2018.
- [6] Mochammad Ainun Rozaq, Nur Nafi'iyah, Masruroh, "Sistem Pendukung Keputusan Pemilihan Kualitas Songkok Berdasarkan Bahan Baku Menggunakan Metode Naive Bayes," *JMAI*, vol. 3, no. 2, pp. 68-72, 2019.