

Analisis Sentimen Kategori Aspek Pada Ulasan Produk Menggunakan Metode KNN Dengan Seleksi Fitur *Mutual Information*

1st Alex Wira Wilantapoera
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
alexwira@students.telkomuniversi
ty.ac.id

2nd Widi Astuti
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
widwdu@telkomuniversity.ac.id

3rd Mahendra Dwifebri
Purbolaksono
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
mahendradwifebri@telkomuniversi
ty.ac.id

Abstrak

Analisis sentimen adalah bidang yang cukup populer untuk menganalisis opini, sikap, dan emosi terhadap suatu subjek dari banyak orang. Dalam hal ini teks ulasan menjadi alat untuk menilai, menimbang, dan mengkritik sebuah produk yang diulas. Produk adalah suatu hal yang dapat memuaskan seorang konsumen dalam bentuk yang beragam seperti barang, jasa, dan sebagainya. Pada penelitian ini dilakukan sebuah implementasi sebuah klasifikasi pada ulasan produk kecantikan dari situs *Female Daily* dengan menggunakan metode *k-Nearest Neighbor* (kNN). Metode kNN merupakan metode yang umum digunakan untuk klasifikasi. Lalu menggunakan *Mutual Information* (MI) sebagai metode seleksi fiturnya. Pada penelitian ini dihasilkan nilai akurasi 91,59% pada aspek price, 90,33% pada aspek packaging, 50,05% pada aspek product, dan 85,89% pada aspek aroma.

Kata kunci : *k-Nearest Neighbour*, Ulasan Produk, Ulasan, Produk, *Mutual Information*, Analisis sentimen.

I. PENDAHULUAN

1.1. Latar Belakang

Saat ini jual beli produk secara *online* sangat diminati oleh banyak orang, sehingga banyak situs jual beli *online* yang bermunculan dan menyediakan forum ulasan sebagai wadah untuk menampung *feedback* dari para pembelinya. Salah situs jual beli *online* adalah *Female Daily* yang menjual berbagai jenis produk kecantikan. Pada penelitian dilakukan klasifikasi teks otomatis yang sering disebut juga sebagai analisis sentimen [1]. Analisis ini dilakukan pada ulasan produk kecantikan yang terdapat pada forum *Female Daily*.

Pada penelitian ini dilakukan pengklasifikasian menggunakan metode *k-Nearest Neighbor* (KNN) karena berdasarkan penelitian yang dilakukan oleh Novelty dan Adiwijaya [2] pada penelitiannya yang

Abstract

Sentimen analysis is a field that is quite populer for analyzing opinions, attitudes, and emotions towards a subject from many people. In this case, the review text becomes a tool for assessing, weighing, and criticizing a product being reviewed. Product is something that can satisfy a consumer in various forms such as goods, services, and so on. In this study, an implementation of a classification model for beauty product reviews from the *Female Daily* website was carried out using the *k-Nearest Neighbor* (kNN) method. The kNN method is a commonly used method for classification. Then use *Mutual Information* (MI) as the feature selection method. In this research, the accuracy value is 91,59% on the price aspect, 90,33% on the packaging aspect, 50,05% on the product aspect, and 85,89% on the aroma aspect.

Keywords: *k-Nearest Neighbor*, Product Reviews, Reviews, Products, *Mutual Information*, Sentimen analysis.

berjudul *Sentimen Analysis on Movie Reviews Using Information Gain and k-Nearest Neighbor* didapatkan hasil bahwa pengujian menggunakan kNN pada analisis sentimen menghasilkan akurasi lebih tinggi jika dibandingkan dengan metode *Naive Bayes*, *Support Vector Machine*, dan *Random Forrest* dengan akurasi sebesar 96,8%.

Penelitian ini juga menggunakan seleksi fitur untuk mereduksi kompleksitas model, meningkatkan efisiensi komputasi pada model, dan mengurangi kesalahan generalisasi yang disebabkan oleh *noise* akibat fitur yang tidak relevan [3].

Seleksi fitur memiliki dua kategori, yaitu *classifier-dependent* yang terdiri dari *wrapper* dan *embedded methods*, lalu *classifier-independent* yang terdiri dari *filter methods* [4]. Pada penelitian ini digunakan seleksi fitur dengan jenis *filter methods*, yaitu *Mutual Information* (MI), seleksi fitur ini dipilih

karena MI adalah salah satu metode seleksi fitur yang paling populer di antara seleksi fitur yang lain [4].

Adapun metode seleksi fitur yang digunakan adalah *Mutual Information*, metode seleksi fitur ini dipilih karena pada penelitian seperti yang dilakukan oleh Fahmi et. al. [5] pada penelitiannya yang berjudul *News Topic Classification Using Mutual Information And Bayesian Network* menghasilkan nilai akurasi yang lebih tinggi saat menggunakan MI dengan kenaikan 29,39%. Sedangkan, jika dibandingkan dengan penelitian yang dilakukan oleh Ekky et. al. [6] pada penelitiannya yang berjudul *Analisis Sentimen Pada Ulasan Produk Kecantikan Menggunakan k-Nearest Neighbor Dan Information Gain* menghasilkan akurasi yang tidak maksimal setelah menggunakan seleksi fitur *Information Gain*.

1.2. Topik dan Batasannya

Penelitian ini dilakukan pada dataset yang didapatkan dari forum ulasan pada situs *Female Daily*, dataset yang digunakan sudah dalam bentuk format CSV, dan juga sudah diberikan label dan aspek. Dataset yang didapatkan berjumlah 3960 ulasan. Namun, aktualnya dataset yang digunakan pada penelitian ini dipisahkan berdasarkan aspeknya dengan masing-masing datanya sebanyak 3665 ulasan.

1.3. Tujuan

Tujuan penelitian ini adalah untuk membangun sebuah model klasifikasi menggunakan metode kNN untuk mengklasifikasi teks ulasan produk kecantikan pada situs jual beli *online Female Daily* dan menggunakan metode seleksi fitur MI untuk melihat pengaruhnya pada hasil klasifikasi.

1.4. Organisasi Tulisan

Sebagai gambaran tentang penulisan pada penelitian ini, berikut ini akan disajikan garis besar pembahasan dari isi laporan ini pada setiap bab sebagai berikut.

1. Pendahuluan yang terdiri dari latar belakang, topik dan batasan penelitian, tujuan penelitian, dan organisasi penulisan.
2. Studi Terkait yang terdiri dari penelitian terkait yang dijadikan dasar teori penelitian ini, pembahasan teoritis terkait analisis sentimen, metode kNN, dan seleksi fitur menggunakan *Mutual information*.
3. Implementasi Klasifikasi membahas tentang gambaran umum rancangan system yang akan dibangun, pembahasan tentang *preprocessing*, lalu pembahasan seleksi fitur menggunakan *Mutual Information*, klasifikasi, dan evaluasi hasil klasifikasi.
4. Skenario Pengujian yang terdiri dari pembahasan dari pengujian nilai k, pengujian *majority voting* dan *weighted*

kNN, pengujian pengaruh seleksi fitur menggunakan *mutual information*, dan evaluasi proses klasifikasi menggunakan metode *k-fold cross validation*.

5. Kesimpulan hasil penelitian.
6. Daftar Pustaka yang terdiri dari paper, jurnal, dan makalah yang menjadi acuan teori penelitian.

II. KAJIAN TEORI

Penelitian terkait yang menggunakan metode klasifikasi kNN untuk analisis sentimen. Diantaranya adalah penelitian yang dilakukan oleh Daeli dan Adiwijaya [2] dalam pengklasifikasian menggunakan metode kNN dan *Information Gain* menunjukkan hasil yang paling baik jika dibandingkan dengan metode *Naïve Bayes*, *Support Vector Machine*, dan *Random Forest* dengan akurasi mencapai 96,8%. Lalu, penelitian yang dilakukan oleh Wisnu et. al. [7] pada penelitiannya yang berjudul *Sentimen Analysis on Customer Satisfaction of Digital Payment in Indoensia: A Comparative Study Using kNN and Naive Bayes* mengatakan bahwa penggunaan kNN untuk klasifikasi lebih baik dari pada *Naive Bayes* dengan akurasi mencapai 91%.

Selain itu, ada juga penelitian lainnya seperti yang dilakukan oleh Pascal et. al. [8] pada penelitiannya yang berjudul *A Simple kNN Alogrithm for Text Classification* penggunaan kNN untuk pengklasifikasian menghasilkan akurasi sebesar 95,5%. Lalu pada penelitian yang dilakukan oleh Febrianti [11] pada penelitiannya yang berjudul *Analisis Sentimen Pada Ulasan Produk Lazada Berbahasa Indonesia Menggunakan K-Nearest Neighbor (K-NN) Dan Perbaikan Kata Menggunakan Jaro Winkler Distance*, didapatkan hasil bahwa penggunaan metode kNN untuk mengklasifikasi teks menghasilkan akurasi yang cukup baik dengan akurasi sebesar 74%. Dan juga pada penelitian yang dilakukan oleh Surohman et. al. [12] pada penelitiannya yang berjudul *Analisis Sentimen Terhadap Review Fintech Dengan Metode Naive Bayes Classifier dan K-Nearest Neighbor* penggunaan metode kNN untuk klasifikasi teks pada analisis sentimen menghasilkan akurasi sebesar 82,92%.

Lalu analisis sentimen yang dilakukan oleh Fauzi et. al. [13] pada penelitiannya yang berjudul *Twitter Sentimen Analysis on 2013 Curriculum Using Ensemble Features and k-Nearest Neighbor*, didapatkan hasil bahwa penggunaan metode kNN untuk klasifikasi teks pada analisis sentimen menghasilkan akurasi sebesar 96%. Selanjutnya penelitian yang dilakukan oleh Sumandee et. al. [14] pada penelitiannya yang berjudul *Sentimen Analysis Approach Based on N-Gram and kNN Classifier*, didapatkan hasil bahwa penggunaan metode kNN untuk klasifikasi teks pada analisis sentimen menghasilkan akurasi sebesar 86%.

Selain penggunaan metode kNN untuk pengklasifikasiannya, pada penelitian ini juga menggunakan seleksi fitur MI. Penelitian terkait penggunaan seleksi fitur MI ini seperti yang dilakukan oleh Fahmi et. al. [5] pada penelitiannya menggunakan MI menghasilkan akurasi 75,34% sedangkan tanpa MI diperoleh akurasi 45,95%. Lalu pada penelitian yang dilakukan oleh Ritter et. al. [9] pada penggunaan MI menghasilkan nilai maksimum dengan akurasi 95%. Lalu pada penelitian yang dilakukan oleh Karunia et. al. [15] pada penelitiannya yang berjudul *Online News Classification Using Naive Bayes Classifier with Mutual Information for Feature Selection*, didapatkan hasil bahwa penggunaan MI untuk seleksi fitur dapat meningkatkan akurasi dari pengklasifikasian yang telah dilakukan dimana akurasinya saat menggunakan seleksi fitur MI mencapai 70%.

2.1. Analisis Sentimen

Analisis sentimen adalah sebuah cara yang digunakan untuk mengolah komentar yang diberikan oleh pemesan atau pelanggan melalui berbagai media, mengenai sebuah produk, jasa ataupun sebuah instansi [1]. Sistem analisis sentimen merupakan sistem yang digunakan untuk melakukan proses analisis otomatis pada ulasan produk *online* berbahasa Indonesia untuk memperoleh informasi meliputi sentimen yang merupakan bagian dari ulasan *online* [18].

2.2. Metode kNN

K-Nearest Neighbor atau biasa disingkat kNN merupakan pengklasifikasi berbasis jarak [8]. Untuk pemecahan masalah pengklasifikasian, metode inilah yang paling sederhana diantara metode klasifikasi lainnya [11]. Metode pada penelitian ini menggunakan metode kNN dan menunjukkan hasil dengan akurasi sebesar 82,92% [12]. Dari pernyataan

di atas dapat disimpulkan bahwa penggunaan metode kNN pada sebuah teks sangat cocok.

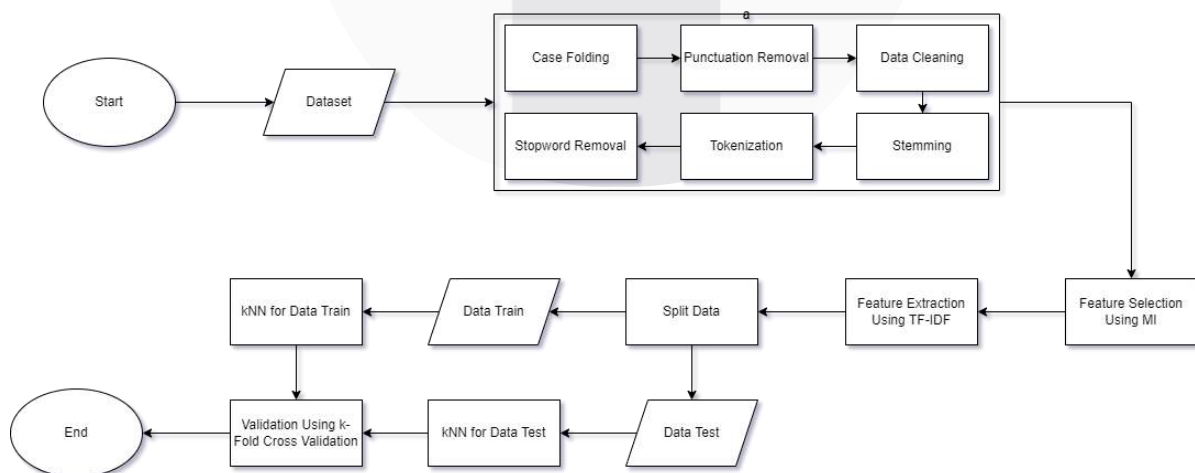
2.3. Seleksi Fitur Menggunakan *Mutual Information*

Seleksi fitur adalah salah satu teknik *data mining* yang umum digunakan pada tahapan *preprocessing* [19]. Salah satu metode seleksi fitur yang sering digunakan untuk menghitung bobot dari fitur adalah *Mutual information*. MI menunjukkan seberapa banyak informasi ada atau tidaknya sebuah *term* memberikan kontribusi dalam membuat keputusan klasifikasi secara benar [15].

III. METODE

3.1. Gambaran Umum Sistem

Sistem yang digunakan dalam penelitian ini dibangun menggunakan bahasa pemrograman *Python*. Pada sistem ini yang pertama kali dilakukan adalah membagi dataset menjadi 4 buah file CSV, dataset mentah dibagi berdasarkan aspeknya. Jadi, ada 4 buah dataset yaitu dataset untuk aspek *price*, dataset untuk aspek *packaging*, dataset untuk aspek *product*, dan dataset untuk aspek *aroma*. Dimana masing-masing dataset berisikan 3665 buah ulasan. Selanjutnya, masing-masing dataset dimasukkan kedalam tahap *preprocessing* yang terdiri dari *case folding*, *punctuation removal*, *data cleaning*, *stopword removal*, *tokenizing*, *stemming*. Selanjutnya dilakukan seleksi fitur menggunakan MI untuk mengurangi *noise* [3]. Setelah itu, dilakukan ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Selanjutnya dilakukan *data splitting* untuk membagi data menjadi data latih dan data uji. Setelah tahap-tahap tersebut dilakukan, selanjutnya melakukan klasifikasi menggunakan kNN. Dan tahap terakhir dilakukan validasi menggunakan *k-fold cross validation*. Untuk lebih jelasnya dapat dilihat pada gambar 1 berikut ini.



Gambar 1. Gambaran Umum Sistem yang Dibangun

3.2. Dataset

Pada tahap awal yang dikerjakan adalah membaca dataset yang dilakukan untuk melakukan penelitian tugas akhir ini. Dataset yang digunakan merupakan ulasan produk kecantikan dalam bahasa

Indonesia dari situs jual beli *online Female Daily*. Berikut disajikan tabel 1 yang menampilkan contoh dataset ulasan produk:

Tabel 1. Contoh Raw Dataset

Text Review	Price	Packaging	Product	Aroma
sunscreen termahal yang pernah gue beli ini kayanya. but it's worth it sih and will definitely buy again. sukanya sama sunscreen ini: - high spf - nggak meninggalkan white cast. perfectly blends into the skin - nggak membuat muka berminyak - very light - doesn't clog pores produk ini berhasil membuat gue jadi mau pakai sunscreen :)	-1	0	1	0

3.3. Preprocessing

Data pada ulasan prduk merupakan data teks yang tidak terstruktur, oleh karena itu diperlukan tahapan *preprocessing* untuk memastikan data di dalam dataset sudah terstruktur. Adapun beberapa proses yang dilakukan dalam tahap ini, diantaranya sebagai berikut.

perlindungan dr matahari baguuss tp emg cpt bkin berminyak wajahnya hehe tp klo ud pkai skinaqua rasany tenang krna wajah bebas sunburn hihi	perlindungan dr matahari baguuss tp emg cpt bkin berminyak wajahnya hehe tp klo ud pkai skinaqua rasany tenang krna wajah bebas sunburn hihi
--	--

3.3.1. Punctuation Removal

Pada tahap ini dilakukan penghapusan simbol-simbol yang tidak dipakai. Adapun contoh dari proses ini dapat dilihat pada tabel berikut.

Tabel 2. Contoh Proses Punctition Removal

Input	Output
ini andalanku klo mau keluar2 rumah, teksturnya ringan, ngga lengket, cepat menyerap, dan yg psti perlindungan dr matahari baguuss.. tp emg cpt bkin berminyak wajahnya hehe.. tp klo ud pkai skinaqua rasany tenang krna wajah bebas sunburn hihi	ini andalanku klo mau keluar2 rumah teksturnya ringan ngga lengket cepat menyerap dan yg psti perlindungan dr matahari baguuss tp emg cpt bkin berminyak wajahnya hehe tp klo ud pkai skinaqua rasany tenang krna wajah bebas sunburn hihi

3.3.3. Normalization

Tahap ini dilakukan untuk mengubah kata-kata yang tidak standar ke dalam bentuk standarnya dengan menggunakan kamus yang berisi kata-kata tidak standar dan kata-kata standarnya. Berikut ini contoh dari proses *normalization*.

Tabel 4. Contoh Proses Normalization

Input	Output
ini andalanku klo mau keluar2 rumah teksturnya ringan ngga lengket cepat menyerap dan yg psti perlindungan dr matahari baguuss tp emg cpt bkin berminyak wajahnya hehe tp klo ud pkai skinaqua rasany tenang krna wajah bebas sunburn hihi	ini andalanku kalau mau keluar2 rumah teksturnya ringan tidak lengket cepat menyerap dan yang pasti perlindungan dari matahari bagus tapi memang cepat membuat berminyak wajahnya hehe tapi kalau sudah pakai skinaqua rasanya tenang karena wajah bebas sunburn hihi

3.3.2. Case Folding

Pada Tahap ini dilakukan untuk mengolah huruf kapital menjadi huruf kecil. Berikut ini contoh dari proses *case folding*.

Tabel 3. Contoh Proses Case Folding

Input	Output
ini andalanku klo mau keluar2 rumah teksturnya ringan ngga lengket cepat menyerap dan yg psti	ini andalanku klo mau keluar2 rumah teksturnya ringan ngga lengket cepat menyerap dan yg psti

3.3.4. Data Cleaning

Tahap ini dilakukan untuk menghapus simbol-simbol dan angka yang terdapat pada *text review*. Berikut ini contoh dari proses *data cleaning*.

Tabel 5. Contoh Proses Data Cleaning

Input	Ouput
-------	-------

ini andalanku kalau mau keluar2 rumah teksturnya ringan tidak lengket cepat menyerap dan yang pasti perlindungan dari matahari bagus tapi memang cepat membuat berminyak wajahnya hehe tapi kalau sudah pakai skinaqua rasanya tenang karena wajah bebas sunburn hihi	ini andalanku kalau mau keluar rumah teksturnya ringan tidak lengket cepat menyerap dan yang pasti perlindungan dari matahari bagus tapi memang cepat membuat berminyak wajahnya hehe tapi kalau sudah pakai skinaqua rasanya tenang karena wajah bebas sunburn hihi
---	--

rasanya tenang karena wajah bebas sunburn hihi	“berminyak”, “wajahnya”, “heeh”, “kalua”, “sudah”, “pakai”, “skinaqua”, “rasanya”, “tenang”, “karena”, “wajah”, “bebas”, “sunburn”, “hihi”
--	--

3.3.5. Stopword Removal

Tahap ini dilakukan untuk menghapus kata yang tidak memiliki keterkaitan antara *text review* dengan labelnya. Berikut ini contoh proses *stopword removal*.

Tabel 6. Contoh Proses *Stopword Removal*

Input	Output
ini andalanku kalau mau keluar rumah teksturnya ringan tidak lengket cepat menyerap dan yang pasti perlindungan dari matahari bagus tapi memang cepat membuat berminyak wajahnya hehe tapi kalau sudah pakai skinaqua rasanya tenang karena wajah bebas sunburn hihi	ini andalanku kalau mau keluar rumah teksturnya ringan tidak lengket cepat menyerap yang pasti perlindungan matahari bagus memang cepat membuat berminyak wajahnya hehe kalau sudah pakai skinaqua rasanya tenang karena wajah bebas sunburn hihi

3.3.6. Tokenization

Tahap ini memisahkan kata-kata dalam kumpulan data menjadi token. Berikut ini contoh dari proses *tokenizing*.

Tabel 7. Contoh Proses *Tokenization*

Input	Output
ini andalanku kalau mau keluar rumah teksturnya ringan tidak lengket cepat menyerap yang pasti perlindungan matahari bagus memang cepat membuat berminyak wajahnya hehe kalau sudah pakai skinaqua	“ini”, “andalanku”, “kalua”, “mau”, “keluar”, “rumah”, “teksturnya”, “ringan”, “tidak”, “membuat”, “lengket”, “cepat”, “menyerap”, “yang”, “pasti”, “perlindungan”, “matahari”, “bagus”, “cepat”, “membuat”,

3.3.7. Stemming

Tahap ini dilakukan untuk mengubah kata yang memiliki imbuhan menjadi kata bakunya dengan menggunakan *library* Sastrawi pada bahasa pemrograman *Python*. Berikut ini contoh dari proses *stemming*.

Tabel 8. Contoh Proses *Stemming*

Input	Output
“ini”, “andalanku”, “kalua”, “mau”, “keluar”, “rumah”, “teksturnya”, “ringan”, “tidak”, “membuat”, “lengket”, “cepat”, “menyerap”, “yang”, “pasti”, “perlindungan”, “matahari”, “bagus”, “cepat”, “membuat”, “berminyak”, “wajahnya”, “hehe”, “kalua”, “sudah”, “pakai”, “skinaqua”, “rasanya”, “tenang”, “karena”, “wajah”, “bebas”, “sunburn”, “hihi”	“ini”, “andalanku”, “kalua”, “mau”, “keluar”, “rumah”, “tekstur”, “ringan”, “tidak”, “buat”, “lengket”, “cepat”, “serap”, “yang”, “pasti”, “lindung”, “matahari”, “bagus”, “cepat”, “buat”, “minyak”, “wajah”, “hehe”, “kalua”, “sudah”, “pakai”, “skinaqua”, “rasa”, “tenang”, “karena”, “wajah”, “bebas”, “sunburn”, “hihi”

3.4. Seleksi Fitur Menggunakan *Mutual Information*

Pada tahap selanjutnya, dataset yang telah melalui tahap *preprocessing* dimasukkan ke dalam tahap seleksi fitur. Seleksi fitur mempunyai tujuan yaitu melakukan pemilihan subset pada fitur yang relevan dari fitur aslinya dengan cara menghapus fitur yang tidak cocok, berlebihan, atau fitur yang terlalu banyak. Antara suku *t* dan kategori *c*, A menunjukkan berapa kali *t* dan *c* terjadi bersama-sama dan B menunjukkan berapa kali *t* muncul tanpa *c*, C menunjukkan berapa kali *c* terjadi tanpa *t* dan jumlah total dokumen dilambangkan dengan *N* [16].

MI mengukur seberapa banyaknya informasi ada atau tidaknya sebuah istilah berkontribusi untuk membuat klasifikasi yang benar pada sebuah kelas. Secara formal, rumusnya dituliskan sebagai persamaan (1).

$$I(U; C) = \frac{N_{11}}{N} \log_2 \frac{NN_{11}}{N_1.N_1} + \frac{N_{01}}{N} \log_2 \frac{NN_{01}}{N_0.N_1} + \frac{N_{10}}{N} \log_2 \frac{NN_{10}}{N_1.N_0} + \frac{N_{00}}{N} \log_2 \frac{NN_{00}}{N_0.N_0} \quad (1)$$

dimana N adalah variabel acak yang memiliki nilai $e_t = 1$ (dokumen yang mengandung *term t*) dan $e_t = 0$ (dokumen yang tidak mengandung *term t*), dan C adalah variabel acak yang memiliki nilai $e_c = 1$ (dokumen pada kelas c) dan $e_c = 0$ (dokumen tidak pada kelas c).

Misalnya N_{10} adalah jumlah dokumen yang mengandung *term t* dan tidak ada pada kelas c . $N_{11} = N_{10} + N_{11}$ di mana kita menghitung jumlah dokumen yang mengandung *term t* dengan mengabaikan dokumen tersebut berada pada kelas apa.

Lalu N adalah jumlah seluruh dokumen yang mengandung *term t* pada kelas c (N_{11}), dokumen yang mengandung *term t* tidak pada kelas c (N_{10}), dokumen yang tidak mengandung *term t* pada kelas c (N_{01}), dan dokumen yang tidak mengandung *term t* tidak pada kelas c (N_{00}). Dan dirumuskan sebagai persamaan (2).

$$N = N_{00} + N_{01} + N_{10} + N_{11}$$

3.5. Klasifikasi Menggunakan kNN

Pada tahap selanjutnya, yaitu melakukan klasifikasi pada dataset yang telah melalui tahap *preprocessing* dan seleksi fitur. Metode yang digunakan pada tahap klasifikasi adalah kNN. Pada algoritma kNN terdapat penghitungan jarak antara dua titik, yaitu titik berupa data latih dan titik berupa data uji yang dapat di definisikan dengan rumus *Euclidean* sebagai persamaan (3).

$$d(x, y) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

3.6. Evaluasi

Evaluasi ini dilakukan untuk mengukur performa dari sistem yang telah dibangun. Performa klasifikasi yang dibangun ini dapat diukur dengan menggunakan beberapa metode, pada penelitian ini digunakan metode *confusion matrix*. *Confusion matrix* adalah metode evaluasi yang memiliki cara kerja membandingkan data yang diprediksi dengan data aktualnya.

Tabel 9. Evaluasi Performa Klasifikasi KNN Menggunakan *Confussion Matrix*

Nilai Prediksi	Positive	Neutral	Negative
Positive	TP	FP	FP
Neutral	FNeu	TNeu	FNeu
Negative	FN	FN	TN

Keterangan:

TP = data yang diprediksi positif dan data yang sebenarnya positif

FP = data yang diprediksi positif tapi data yang sebenarnya tidak positif

TNeu = data yang diprediksi netral dan data yang sebenarnya netral

FNeu = data yang diprediksi netral tapi data yang sebenarnya tidak netral

TN = data yang diprediksi negatif dan data yang sebenarnya negatif

FN = data yang diprediksi negatif tapi data yang sebenarnya tidak negatif

IV. HASIL DAN PEMBAHASAN

Pada penelitian ini terdapat empat skenario pengujian. Pengujian pertama, yaitu menguji nilai k untuk mendapatkan nilai k yang paling optimal. Pengujian kedua, yaitu membandingkan *majority voting* dan *weighted kNN*. Pengujian ketiga, yaitu membandingkan pengaruh seleksi fitur menggunakan *Mutual Information*. Dan pengujian keempat, yaitu melakukan evaluasi pada klasifikasi dengan menggunakan metode *k-fold cross validation*.

4.1. Pengujian Nilai k Pada kNN

Pengujian pertama adalah mencari nilai k yang menghasilkan nilai akurasi yang optimal. Nilai k yang diujikan adalah $k=1$ sampai dengan $k=10$ dengan menggunakan *weighted kNN*. Pengujian ini dilakukan untuk menganalisa nilai k yang paling optimal dan akan digunakan pada seluruh pengujian. Berikut ini disajikan tabel yang menampilkan hasil pengujian k serta akurasinya pada masing-masing aspek.

Tabel 10. Hasil Pengujian Nilai k Pada kNN

K	Akurasi			
	Price	Packaging	Product	Aroma
1	94,59%	94,04%	48,65%	91,10%
2	95,28%	94,90%	40,35%	91,21%
3	95,28%	94,30%	51,34%	91,07%
4	94,07%	93,81%	51,07%	91,59%
5	94,07%	94,00%	52,99%	90,96%
6	93,90%	93,97%	51,84%	91,24%
7	93,90%	94,00%	51,40%	90,72%
8	94,07%	94,04%	51,56%	91,35%
9	94,07%	94,04%	52,11%	91,49%
10	94,07%	94,04%	52,28%	91,59%

Hasilnya menunjukkan bahwa nilai k yang optimal pada aspek price adalah 2 dan 3 dengan akurasi sebesar 95,28%, aspek packaging nilai k yang optimal adalah 2 dengan akurasi 94,90%, aspek product nilai k yang optimal adalah 5 dengan akurasi 52,99%, dan aspek aroma nilai k yang optimal adalah 4 dan 10 dengan akurasi 91,59%.

4.2. Pengujian *Majority Voting* dan *Weighted kNN*

Pada penelitian ini, digunakan *euclidean distance* untuk menghitung jarak antara data latih dengan data uji. Setelah mendapatkan *k* tetangga terdekat dari pengujian pertama, selanjutnya menentukan label pada data baru dengan salah satu dari dua cara yang ada, yaitu *majority voting* dan *weighted kNN*. Pertama *majority voting*, cara ini tidak mempertimbangkan jarak data uji terhadap data latih terdekatnya, namun pelabelan pada data uji dilakukan berdasarkan label data latih terdekatnya yang terbanyak. Kedua *weighted kNN*, cara ini mempertimbangkan jarak data baru dengan *k* data latih terdekatnya, sehingga semakin dekat data uji dengan data latih tersebut, maka semakin besar kemungkinan data uji dilabeli dengan label yang sama.

Adapun klasifikasi yang dibangun dalam penelitian ini menggunakan *library scikit-learn* dengan modul *KNeighborsClassifier*. Pada modul tersebut untuk melakukan *majority voting* caranya adalah dengan men-set parameter “*weights*” dengan “*uniform*”. Sedangkan untuk *weighted kNN* caranya adalah dengan men-set parameter “*weights*” dengan “*distance*”.

Pengujian ini dilakukan dengan membandingkan akurasi tertinggi yang didapatkan dengan nilai *k* yang telah didapatkan dari pengujian pertama. Di bawah ini disajikan tabel hasil pengujian tersebut. Dapat dilihat bahwa ada perbedaan akurasi pada aspek *product* dan *aroma*, di mana pada aspek *product* memiliki akurasi yang lebih tinggi saat menggunakan *majority voting* dengan nilai 52,99%, sedangkan pada aspek *aroma* memiliki nilai akurasi tertinggi pada saat menggunakan *weighted kNN* dengan nilai 91,66%.

Tabel 11. Hasil Perbandingan Akurasi Penggunaan *Majority Voting* dan *Weighted kNN*

Aspek	Akurasi (Majority Voting)	Akurasi (Weighted)
Price	95,28%	95,28%
Packaging	94,90%	94,90%
Product	52,99%	52,06%
Aroma	91,59%	91,66%

Dari pengujian ini dapat disimpulkan bahwa perbedaan hasil penggunaan *majority voting* dan *weighted kNN* tidak signifikan.

4.3. Pengujian Pengaruh Seleksi Fitur Menggunakan *Mutual Information*

Selanjutnya adalah pengujian pengaruh seleksi fitur menggunakan MI. Pada pengujian ini, digunakan metode *weighted kNN* berdasarkan pengujian kedua, dan juga menggunakan *k* yang telah didapatkan dari pengujian pertama. Pengujian ini dilakukan untuk menganalisis pengaruh seleksi fitur menggunakan MI pada pengklasifikasian. Di bawah ini disajikan tabel yang menampilkan akurasi hasil

klasifikasi setelah menggunakan MI dan tanpa MI. dari tabel ini bisa dilihat pada saat menggunakan MI akurasi yang didapatkan pada aspek *price*, *packaging*, dan *aroma* memiliki akurasi yang lebih tinggi jika dibandingkan dengan saat tanpa MI. Pada aspek *price* akurasinya naik 39,77%, lalu pada aspek *packaging* akurasinya naik 7,31%, dan pada aspek *aroma* akurasinya naik 9,18%. Sedangkan pada aspek *product* akurasinya turun 11,29%

Tabel 12. Hasil Pengujian Seleksi Fitur Menggunakan MI dan Tanpa MI

Aspek	Akurasi (menggunakan MI)	Akurasi (tanpa MI)
Price	91,59%	51,82%
Packaging	90,33%	83,02%
Product	50,05%	61,34%
Aroma	85,89%	76,71%

Hal tersebut dapat terjadi karena pengaruh dari MI yang memiliki cara kerja mengkalkulasi seberapa besar pengaruh suatu kata (*term*) terhadap aspeknya. Misal, jika dilihat pada tabel 13 kata “mahal” pada aspek *price* sangat berpengaruh karena kemunculannya pada kelas negatif mencapai 513 kali, sedangkan pada kelas positif dan netral kemunculannya berturut-turut adalah 3 dan 2 kali. Sedangkan pada aspek *product* terjadi penurunan akurasi, kemungkinan hal ini diakibatkan oleh reduksi fitur yang dilakukan MI.

Tabel 13. Hasil Seleksi Fitur Menggunakan MI

	Term/fitur	Negatif	Netral	Positif	MI Negatif	MI Netral	MI Positif
Price	Mahal	513	2	3	0.43251	0.00002	0.10003
					0.51327	0.00001	0.08138
					0.00007	0.00004	0.00007
Packaging	Packagingnya	53	26	130	0.00012	0.00008	0.00015
					0.00024	0.00012	0.00015
					0.00001	0.00004	0.00007
Product	Coklat	231	103	444	0.00017	0.00005	0.00011
					0.00027	0.00007	0.00014
					0.00006	0.00009	0.00004
Aroma	Wangi	47	115	428	0.00001	0.00002	0.00002
					0.00001	0.00003	0.00002
					0.00049	0.00006	0.00004
			0.00000	0.00009	0.00007		

Dari pengujian ini dapat disimpulkan bahwa pengaruh penggunaan seleksi fitur MI relatif baik karena dapat meningkatkan akurasi dari hasil klasifikasi pada beberapa aspek.

4.4. Evaluasi Proses Klasifikasi Menggunakan *k-fold Cross Validation*

Setelah melakukan serangkaian pengujian, selanjutnya melakukan validasi model klasifikasi yang telah dibangun menggunakan metode *k-fold cross validation*. Pada pengujian ini digunakan hasil klasifikasi dengan seleksi fitur menggunakan MI, lalu menggunakan hasil *weighted kNN*, dan menggunakan nilai *k* yang telah didapatkan dari pengujian pertama.

Validasi klasifikasi ini menggunakan nilai *k=5* pada *k-fold*. Dari validasi ini didapatkan akurasi pada masing-masing aspek. Berikut ini disajikan tabel 14 yang menampilkan akurasi setelah dilakukan validasi menggunakan *k-fold cross validation*.

Tabel 14. Hasil Validasi Menggunakan *k-fold Cross Validation*

K	Akurasi			
	Price	Packaging	Product	Aroma
1	92,89%	81,39%	48,78	80,74%
2	89,05%	75,81%	64,55%	75,09%
3	89,94%	69,76%	54,63%	72,49%
4	84,91%	71,02%	66,82%	59,10%
5	88,42%	72,89%	64,82%	63,56%

Hasilnya pada aspek *price*, *packaging*, dan *aroma* memiliki nilai akurasi tertinggi pada nilai *k=1*, di mana nilai akurasi untuk aspek *price* sebesar 92,89%, *packaging* 81,39%, dan *aroma* 80,74%. Sedangkan pada aspek *product* nilai tertingginya didapatkan pada nilai *k=4* dengan nilai akurasi sebesar 66,82%.

V. KESIMPULAN

Kesimpulannya, analisis sentimen menggunakan metode *kNN* dengan seleksi fitur menggunakan MI pada penelitian ini dengan dataset yang berupa ulasan produk kecantikan dari forum ulasan pada situs *Female Daily* memiliki akurasi yang relatif baik. Berdasarkan hasil serangkaian pengujian yang telah dilakukan menunjukkan hasil yang cukup baik, dimana pengklasifikasian menggunakan metode *kNN* pada aspek *price* menghasilkan akurasi maksimal pada nilai *k=2* dengan akurasi 95,28%, lalu pada aspek *packaging* menghasilkan akurasi maksimal pada nilai *k=2* dengan akurasi 94,90%, lalu pada aspek *product* menghasilkan akurasi maksimal pada nilai *k=5* dengan akurasi 52,99%, dan pada aspek *aroma* menghasilkan akurasi maksimal pada nilai *k=4* dengan akurasi 91,59%. Selanjutnya pada pengujian *majority voting* dan *weighted kNN* menunjukkan bahwa perbedaan akurasi tidak signifikan, di mana nilai akurasi pada aspek *product* dengan *majority voting* adalah 52,99% dan dengan menggunakan *weighted kNN* didapatkan akurasi 52,06%. Lalu pada aspek *aroma* dengan menggunakan *majority voting* didapatkan akurasi 91,59%, dan dengan menggunakan *weighted kNN* didapatkan akurasi 91,66%. Sedangkan pada aspek *price* dan *packaging* tidak ada perbedaan sama sekali saat menggunakan

majority voting ataupun *weighted kNN* dimana akurasinya adalah 95,28% pada aspek *price*, dan 95,90% pada aspek *packaging*. Namun, penggunaan MI untuk seleksi fitur menghasilkan pengaruh yang relatif baik, di mana pada aspek *price*, *packaging*, dan *aroma* akurasinya meningkat. Adapun akurasi yang didapatkan pada aspek *price* setelah menggunakan MI meningkat 39,77%, lalu pada aspek *packaging* meningkat 7,31%, dan pada aspek *aroma* meningkat 9,18%. Lalu pada hasil pengujian evaluasi menggunakan *k-fold cross validation* dengan jumlah *k=5* menghasilkan akurasi terbaik pada nilai *k=1* untuk aspek *price* dengan akurasi sebesar 92,89%, lalu hasil terbaik pada aspek *packaging* didapatkan pada nilai *k=1* dengan akurasi 81,39%, lalu hasil terbaik pada aspek *product* didapatkan pada nilai *k=4* dengan akurasi 66,82%, dan pada aspek *aroma* hasil terbaiknya didapatkan pada nilai *k=1* dengan akurasi sebesar 80,74%.

REFERENSI

- [1] Na, J. C., Thet, T. T., & Khoo, C. S. (2010). Comparing sentiment expression in movie reviews from four online genres. *Online Information Review*.
- [2] Daeli, N. O. F., & Adiwijaya, A. (2020). Sentimen analysis on movie reviews using Information gain and K-nearest neighbor. *Journal of Data Science and Its Applications*, 3(1), 1-7.
- [3] Kumar, Ajitesh. 2021. Machine Learning – Feature Selection vs Feature Extraction. [Online] Availabel at: <https://vitalflux.com/machine-learning-feature-selection-feature>
- [4] Macedo, F., Oliveira, M. R., Pacheco, A., & Valadas, R. (2019). Theoretical foundations of forward feature selection methods based on mutual information. *Neurocomputing*, 325, 67-89.
- [5] Nurfikri, F. S., & Mubarak, M. S. (2018, May). News topic classification using mutual information and bayesian network. In *2018 6th International Conference on Information and Communication Technology (ICoICT)* (pp. 162-166). IEEE.
- [6] Al Faraby, S. (2021). Analisis Sentimen Pada Ulasan Produk Kecantikan Menggunakan K-nearest Neighbor Dan Information Gain. *eProceedings of Engineering*, 8(5).
- [7] Wisnu, H., Afif, M., & Ruldevyani, Y. (2020). Sentimen analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naïve Bayes. In *Journal of Physics: Conference Series* (Vol. 1444, No. 1, p. 012034). IOP Publishing.
- [8] Soucy, P., & Mineau, G. W. (2001, November). A simple KNN algorithm for text categorization. In *Proceedings 2001 IEEE international conference on data mining* (pp. 647-648). IEEE.

- [9] Ritter, G. L., Lowry, S. R., Woodruff, H. B., & Isenhour, T. L. (1976). Relation between mutual information and classification. *Analytical Chemistry*, 48(7), 1027-1031.
- [10] Jiang, S., Pang, G., Wu, M., & Kuang, L. (2012). An improved K-nearest-neighbor algorithm for text categorization. *Expert Systems with Applications*, 39(1), 1503-1509.
- [11] Febrianti, Y. M. (2018). *Analisis Sentimen Pada Ulasan "Lazada" Berbahasa Indonesia Menggunakan K-Nearest Neighbor (K-NN) Dengan Perbaikan Kata Menggunakan Jaro Winkler Distance* (Doctoral dissertation, Universitas Brawijaya).
- [12] Surohman, S., Aji, S., Rousyati, R., & Wati, F. F. (2020). Analisa Sentimen Terhadap Review Fintech Dengan Metode Naive Bayes Classifier Dan K-Nearest Neighbor. *EVOLUSI: Jurnal Sains dan Manajemen*, 8(1).
- [13] Irfan, M. R., Fauzi, M. A., Tibyani, T., & Mentari, N. D. (2018). Twitter sentimen analysis on 2013 curriculum using ensemble features and k-nearest neighbor. *International Journal of Electrical and Computer Engineering (IJECE)*, 8(6), 5409-5414.
- [14] Kaur, S., Sikka, G., & Awasthi, L. K. (2018, December). Sentimen analysis approach based on N-gram and KNN classifier. In *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)* (pp. 1-4). IEEE.
- [15] Karunia, S. A. (2017). *Online News Classification Using Naive Bayes Classifier with Mutual Information for Feature Selection*. *ITSMART: Jurnal Teknologi dan Informasi*, 6(1), 11-15.
- [16] Khan, F. H., Qamar, U., & Bashir, S. (2016). SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentimen polarity detection. *Applied Soft Computing*, 39, 140-153.
- [17] Paul, A. K., & Shill, P. C. (2016, December). Sentimen mining from bangla data using mutual information. In *2016 2nd international conference on electrical, computer & telecommunication engineering (ICECTE)* (pp. 1-4). IEEE.
- [18] Putra, S. J., Khalil, I., Gunawan, M. N., Amin, R. I., & Sutabri, T. (2018, November). A hybrid model for social media sentimen analysis for Indonesian text. In *Proceedings of the 20th International Conference on Information Integration and Web-Based Applications & Services* (pp. 297-301).
- [19] Adnyana, I. M. B. (2019). Penerapan Feature Selection untuk Prediksi Lama Studi Mahasiswa. *Jurnal Sistem dan Informatika (JSI)*, 13(2), 72-76.