

ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP PRODUK KECANTIKAN MENGGUNAKAN *NEIGHBOR WEIGHTED K-NEAREST NEIGHBOR*

Fauzan Lukmanul Hakim¹, Kania Evita Dewi²

Program Studi Teknik Informatika, Universitas Komputer Indonesia^{1,2}
Jl. Dipatiukur No. 112-116 Bandung
E-mail : kania.evita.dewi@email.unikom.ac.id²

Abstrak

Penelitian ini bertujuan untuk menguji performansi *Neighbor Weighted K-Nearest Neighbor* (NWKNN) dalam menangani dataset yang tidak seimbang dalam kasus analisis sentiment berbasis aspek. Data yang digunakan didalam penelitian ini adalah ulasan produk kecantikan yang berasal dari situs kaggle. Diperoleh data sebanyak 2.449 ulasan. Setiap ulasan produk sebelum masuk ketahapan klasifikasi, melalui preprocessing. Dalam penelitian ini tahapan preprocessing terdiri dari proses casefolding, cleaning, tokenisasi, normalisasi, stemming, convert negasi, dan stopword removal. Agar hasil preprocessing dapat diolah oleh algoritma klasifikasi maka setiap ulasan yang sudah dipreprocessing masuk kedalam ekstraksi fitur. Metode ekstraksi fitur yang digunakan dalam penelitian ini adalah TF-IDF. Hasil ekstraksi fitur lah yang masuk kedalam proses klasifikasi. Didalam penelitian ini setiap ulasan melalui proses klasifikasi beberapa kali. Karena dalam penelitian ini dalam penanganan multilabel menggunakan teknik binary relevance. Setiap klasifikasi menggunakan NWKNN. Pengklasifikasian dilakukan sebanyak empat kali sesuai dengan aspek yang digunakan didalam penelitian ini, yaitu: harga, kemasan, efektifitas dan aroma. Sehingga setiap klasifikasi menghasilkan polaritas untuk setiap aspek, yaitu: positif, negative, atau non sentiment. Hasil pengujian performansi dengan *Confusion Matrix* dihasilkan performansi NWKNN lebih tinggi dibandingkan KNN untuk masing-masing aspek, dalam f1-score. Dimana nilai e dan k yang optimal untuk metode NWKNN yaitu nilai k=40 dan e=2. Ini menunjukkan bahwa NWKNN terbukti lebih baik bekerja jika dataset tidak seimbang dibandingkan KNN.

Kata kunci: Analisis Sentimen, NWKNN, *Euclidean Distance*, ulasan produk

Abstract

This research aims to test the performance of Neighbor Weighted K-Nearest Neighbor (NWKNN) in handling imbalanced datasets in the case of aspect-based sentiment analysis. The data used in this research are beauty product reviews from the Kaggle site. Data obtained from 2,449 reviews. Every product review before entering the classification stage, goes through preprocessing. In this research, the preprocessing stages consist of casefolding, cleaning, tokenization, normalization, stemming, convert negation, and stopword removal processes. So that the preprocessing results can be processed by a classification algorithm, each review that has been preprocessed is included in feature extraction. The feature extraction method used in this research is TF-IDF. The results of feature extraction are included in the classification process. In this research, each review went through a classification process several times. Because in this research, multilabel handling uses binary relevance techniques. Each classification uses NWKNN. Classification was carried out four times according to the aspects used in this research, namely: price, packaging, effectiveness and aroma. So each classification produces a polarity for each aspect, namely: positive, negative, or non-sentimental. The results of performance testing with Confusion Matrix showed that NWKNN's performance was higher than KNN's for each aspect, in terms of f1-score. Where the optimal e and k values for the NWKNN method are k=40 and e=2. This shows that NWKNN proves to perform better when the dataset is imbalanced compared to KNN.

Keywords: Sentiment Analysis, NWKNN, *Euclidean Distance*, Review Product.

1. PENDAHULUAN

Analisis sentimen, juga disebut penggalan opini, merupakan teknik pemrosesan bahasa alami yang digunakan untuk mengidentifikasi dan menganalisis sentimen atau opini dari sebuah teks tentang sebuah

entitas dan atributnya lalu membedakan opini tersebut menjadi positif, netral dan negative. Entitas yang dinilai dapat berupa produk, layanan, organisasi, individu, acara, masalah, atau topik[1]. Dimana analisis sentimen dapat diterapkan pada berbagai jenis dokumen, seperti *tweet*, ulasan produk, dan artikel berita. Dimana untuk mendapatkan penilaian yang lebih detail dapat dilakukan dengan analisis sentimen berbasis aspek[1]. Salah satu jenis produk yang dapat terbantu dengan ulasan yaitu produk kecantikan. dimana ulasan dapat membantu pembeli dalam menentukan produk yang akan dibeli untuk mengetahui pendapat masyarakat yaitu positif dan negatif mengenai aspek-nya, seperti harga, packaging, efektifitas, dan aroma[2].

Salah satu metode yang sering digunakan untuk klasifikasi analisis sentimen adalah metode *K-Nearest Neighbor* (K-NN), salah satu penelitian yang menggunakan metode K-NN dan *information gain* dalam melakukan analisis sentimen berbasis aspek pada ulasan produk kecantikan menghasilkan akurasi rata-rata sebesar 74.21%[2], namun metode K-NN memiliki kekurangan dalam penggunaan data yang tidak seimbang[3], dimana rata-rata performansi *F1-Score* yang didapatkan sebesar 45.19% hal ini dikarenakan distribusi polaritas untuk masing-masing aspek tidak seimbang[2].

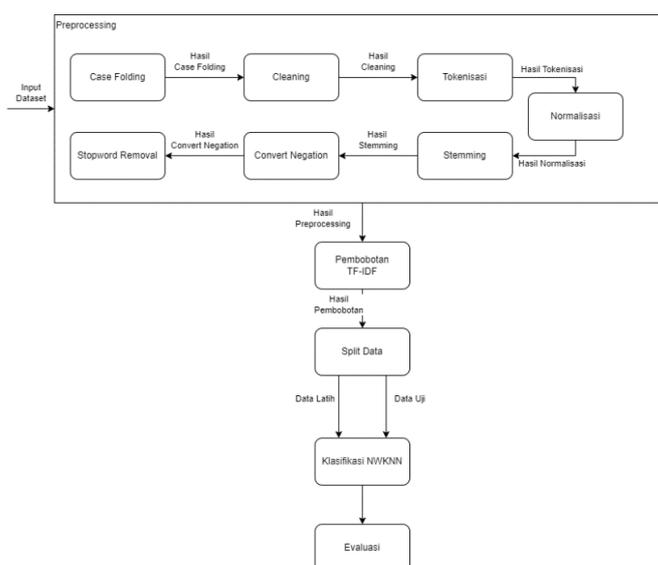
Salah satu metode yang dapat menangani permasalahan penyebaran data yang tidak seimbang yaitu Metode *Neighbor Weighted K-Nearest Neighbor* (NWKNN), metode NWKNN merupakan pengembangan untuk metode K-NN yang dapat meningkatkan performansi dari metode K-NN[4]. Dalam penelitian analisis sentimen dengan kasus *review* aplikasi *mobile* menggunakan metode NWKNN[5] dengan menggunakan data seimbang, dimana jumlah data positif yang digunakan berjumlah 100 dan data negatif berjumlah 100 menghasilkan *F1-Score* tertinggi 90% untuk kedua metode dengan nilai K=15, sedangkan untuk jumlah data tidak seimbang menggunakan metode K-NN mendapatkan *F1-Score* 48% sedangkan untuk metode NWKNN mendapatkan *F1-Score* 80% untuk K=45. Namun hasil dalam analisis sentiment berbasis aspek dengan menggunakan metode NWKNN belum diketahui.

Didalam penelitian ini, dilakukan klasifikasi ASBA pada data ulasan produk kecantikan. Penelitian ini menunjukkan performa NWKNN yang lebih baik dibanding KNN dalam klasifikasi ASBA, walaupun dengan data yang imbalance. Dengan menggunakan binary relevance dalam menangani masalah multilabel, NWKNN dapat mengklasifikasikan empat aspek dan 3 polaritas yang terdapat dalam data produk kecantikan tersebut.

2. METODOLOGI

2.1. Alur Penelitian

Alur penelitian adalah gambaran proses awal penelitian hingga penelitian berakhir. Proses penelitian diawali dengan memasukkan dataset yang diambil dari situs Kaggle.com, setelah dataset terbaca dilakukan tahap preprocessing, hasil dari preprocessing akan dilakukan tahap pembobotan kata dengan menggunakan TF-IDF, proses pembagian data latih dan data uji, klasifikasi, menggunakan NWKNN dan terakhir evaluasi performansi dari hasil klasifikasi. Alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur penelitian

2.2. Data

Data yang digunakan pada penelitian berasal dari ulasan produk kecantikan pada situs “Female Daily” yang diambil dari kaggle.com dengan url <https://www.kaggle.com/datasets/hafidahmusthaanah/skincare-review?select=00.+Review.csv>. Data yang digunakan berjumlah 2449 data, dimana aspek yang akan digunakan yaitu harga, aroma, efektivitas dan kemasan[2], serta akan diberikan label secara manual untuk masing-masing aspek. Sedangkan setiap aspek memiliki polaritas, yaitu: positif, negative, dan non sentiment. Didalam penelitian ini non sentiment artinya, didalam ulasan yang sedang diklasifikasi tidak ada aspek yang ditanyakan sehingga tidak memiliki polaritas.

2.2. Multilabel

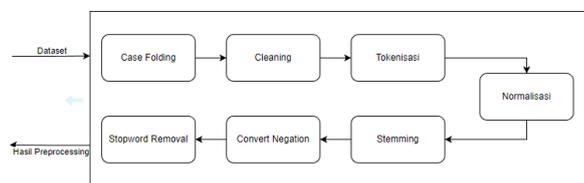
Multilabel classification adalah sebuah masalah klasifikasi dimana setiap sampel data dapat diklasifikasikan ke dalam lebih dari satu kelas pada saat yang sama. Sebagai contoh, dalam klasifikasi genre film, satu film dapat termasuk dalam beberapa genre seperti komedi, romantis, dan drama sekaligus. Dimana terdapat beberapa cara untuk menyelesaikan masalah *multilabel classification* yaitu *binary relevance*, *classifier chains*, *label powerset*[6]. Didalam penelitian ini menggunakan *binary relevance*.

2.3. Binary Relevance

Binary Relevance merupakan salah satu cara dalam klasifikasi *multilabel*. Cara kerja dari metode *binary relevance*, yaitu memisahkan setiap label untuk per kelas atau per aspek, lalu pada masing-masing aspek tersebut akan dijadikan sebuah model menggunakan metode yang akan digunakan, setelah proses training dilakukan akan dilakukan pengujian dengan data baru untuk masing-masing aspek menggunakan metode yang akan digunakan, lalu akan didapatkan hasil prediksi.[6]

2.4. Preprocessing

Tahap *preprocessing* dilakukan sebelum dilakukannya proses klasifikasi, tujuan dari tahap *Preprocessing* ini dilakukan untuk memastikan data sudah siap digunakan dan dan menyeragamkan bentuk kata sehingga dapat mengurangi volume kata, sehingga memungkinkan analisis atau pemrosesan data yang lebih akurat dan efektif. Karena sebelum dilakukan tahap *preprocessing dataset* memungkinkan memiliki *noise* yang tinggi yang dapat memengaruhi proses klasifikasi[2][7].



Gambar 2. Alur *preprocessing*

Adapun tahap *preprocessing* yang dilakukan pada penelitian ini adalah sebagai berikut:

A. Case Folding

Case Folding merupakan tahapan praproses dimana setiap kata pada *dataset* diubah atau diseragamkan. Setiap ulasan produk didalam peneleitian ini diubah menjadi huruf kecil [7]

B. Cleaning

Cleaning merupakan tahapan pada praproses dimana menghapus tanda baca, angka, dan simbol-simbol pada *dataset* [7].

C. Normalization

Normalization atau normalisasi merupakan proses memperbaiki kata-kata yang salah eja atau disingkat dan mengubah kata tidak baku menjadi kata baku berdasarkan kamus [7]. Kamus didapat dari github dengan sumber <https://raw.githubusercontent.com/haryoa/indo-collex/main/dict/informal-formal-Indonesian-dictionary.tsv> serta terdapat penambahan dari peneliti.

D. Stopword Removal

Stopword Removal merupakan sebuah proses penghapusan kata tidak penting seperti kata penghubung atau kata yang tidak memiliki makna. Pada tahap ini digunakan *stoplist*, dimana *stoplist* berisi sekumpulan kata yang tidak memiliki makna atau biasa disebut dengan *stopword* [7]. *Stoplist* yang digunakan dalam penelitian ini adalah *stoplist* yang digunakan library nltk sastrawi dan juga beberapa kata seperti wkwwk, hehehe, hahaha, yang ditambahkan oleh peneliti kedalam *stoplist*.

E. *Tokenization*

Tokenization atau tokenisasi adalah tahap memisahkan atau memotong sebuah kalimat menjadi kata per kata dengan sebuah pemisah yaitu blank space atau spasi[7].

F. *Stemming*

Stemming, merupakan tahapan praproses yang bertujuan untuk menghapus imbuhan yang terdapat pada kata atau mengubah kata menjadi bentuk dasarnya[7]. Proses stemming yang digunakan dalam penelitian ini juga menggunakan library nltk sastrawi.

G. *Convert Negation*

Convert Negation adalah tahapan menggabungkan kata negasi dengan kata selanjutnya, karena kata negasi dapat mengubah makna sebuah sentimen dalam sebuah ulasan[8]. Kata-kata negasi yang dimaksud dalam penelitian ini seperti: tidak, bukan, belum, jangan, kurang dan lainnya.

2.5. Pembobotan *Term Frequency-Inverse Document Frequency (TF-IDF)*

Metode TF-IDF digunakan untuk mengukur tingkat kepentingan kata terhadap sebuah dokumen, dengan memberikan bobot terhadap kata, penentuan nilai bobot dilakukan dengan cara menghitung frekuensi kemunculan kata dalam dokumen[9]. Perhitungan TF (*Term Frequency*) mengukur seberapa sering suatu kata muncul dalam sebuah dokumen atau teks, sedangkan DF (*Document Frequency*) merupakan jumlah dari dokumen/teks yang terdapat suatu kata/term[10].

Perhitungan TF [11] dilihat pada persamaan (1),

$$TF_{ij} = \frac{f_i(d_j)}{\sum_{i=1}^k f_i(d_j)} \tag{1}$$

Dimana :

- $TF_{i,j}$: Term Frequency ke-i pada dokumen ke-j
- $f_i(d_j)$: Frekuensi kemunculan term ke-i pada dokumen ke-j
- $\sum_{i=1}^k f_i(d_j)$: Total term pada dokumen j

Perhitungan IDF [12] dilihat pada persamaan (2)

$$IDF = \log \left(\frac{1+D}{1+DF_j} \right) + 1 \tag{2}$$

Dimana :

- IDF : *Inverse Document Frequency*
- D : Jumlah Ddokumen
- DF_j : Jumlah dokumen pada term j

Perhitungan TF-IDF dilihat pada persamaan (3)

$$W(i, j) = tf_{(i,j)} * IDF \tag{3}$$

Dimana :

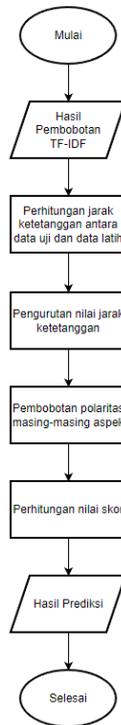
- $W(i, j)$: bobot kata j terhadap dokumen i
- $tf_{(i,j)}$: Term Frequency ke-i pada dokumen ke-j
- IDF : *Inverse Document Frequency*

Perhitungan TF-IDF dengan menggunakan *Sklearn* [12] dilihat pada persamaan (4)

$$W(i, j) = \frac{tf_{(i,j)} * IDF}{\sqrt{\sum_{s=1}^k (tf_{(i,j)} * IDF)^2}} \tag{4}$$

2.6. *Neighbor Weighted K-Nearest Neighbor (NWKNN)*

Metode NWKNN merupakan sebuah pengembangan dari metode KNN yang memiliki keunggulan dalam mengatasi masalah data yang tidak seimbang[5]. Pada NWKNN, terdapat penambahan beberapa tahapan, seperti menghitung bobot yang akan digunakan pada tahap menghitung skor. Secara lebih rinci, tahapan dalam proses klasifikasi NWKNN dapat dilihat pada Gambar 3



Gambar 3. Alur penelitian

Perhitungan jarak menggunakan *euclidean distance* [13] dilihat pada persamaan 5

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{5}$$

Dimana :

- d : Jarak antara data uji dan data latih
- n : jumlah data
- x_i : fitur ke- i pada data uji
- y_i : fitur ke- i pada data latih

Perhitungan bobot [5] dihitung persamaan 6

$$Weight_i = \frac{1}{\left(\frac{num(C_i^d)}{\min\{Num(C_j^d)\}_{j=1, \dots, k^*}}\right)^{1/E}} \tag{6}$$

Dimana :

- $num(C_i^d)$: jumlah data latih d pada kelas i
- $num(C_j^d)$: jumlah data latih d pada kelas j, dimana j terdapat dalam himpunan k tetangga terdekat
- E : bilangan lebih dari 1

Perhitungan nilai skor[14] dihitung persamaan 7

$$Score(q, C_i) = Weight_i * \left(\sum_{d_j \in KNN(X)} (\sqrt{\sum_{i=1}^n (x_i - y_i)^2}) * \delta(d_j, C_i)\right) \tag{7}$$

Dimana :

- $Weight_i$: bobot kelas i
- $\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$: Jarak antara data latih dan data u
- $\delta(d_j, C_i)$: akan bernilai 1 jika nilai jarak $\in C_i$
akan bernilai 0 jika nilai jarak $\notin C_i$

C_i : kelas atau kategori i

2.8. Confusion Matrix

Confusion Matrix merupakan salah satu alat yang dapat mengukur performa dari sebuah model klasifikasi, *confusion matrix* berbentuk seperti tabel yang mencakup matrix dua dimensi, dimana dimensi yang pertama diisi oleh kelas sebenarnya dari suatu objek dan di dimensi yang lain diisi oleh kelas yang dihasilkan oleh *classifier* atau hasil prediksi[15]. Istilah yang biasa ada dalam confusion matrix adalah TP, TN, FP, dan FN. TP adalah kondisi kelas actual dan hasil prediksi sama-sama melabeli data dengan label positif. TN adalah kondisi kelas aktuan dan hasil prediksi sama-sama menyatakan data tersebut berlabel negatif. FP adalah kondisi kelas aktual mengatakan data tersebut berlabel negative sedangkan hasil prediksi berlabelnya negative. FN adalah kondisi kelas aktualnya memberikan label positif tetapi hasil prediksinya positif. Misalkan dalam matriks confusion multilabel akan memiliki dimensi lebih besar dibandingkan confusion matrix kelas biner. Confusion matrix multikelas memiliki dimensi sebesar $n \times n$, dimana n menunjukkan banyak kelas yang berbeda C_i , dimana $i = \{1,2, \dots, n\}$. *Confusion matrix* menggunakan *multi-label*[16] dapat dilihat pada Tabel 1 berikut.

Tabel 1. Confusion Matrix

		Kelas Prediksi			
		C_1	C_2	...	C_n
Kelas Aktual	C_1	$C_{1,1}$	FP	...	$C_{1,n}$
	C_2	FN	TP	...	FN

	C_n	$C_{n,1}$	FP	...	$C_{n,n}$

Precision adalah proporsi dari pelabelan yang teridentifikasi dengan benar[16], rumus untuk mencari *precision*[16] dapat dihitung menggunakan persamaan 8 sebagai berikut:

$$Precision(C_i) = \frac{TP(C_i)}{TP(C_i)+FP(C_i)} \tag{8}$$

Recall merupakan proporsi dari informasi yang dapat ditemukan dari label[16], rumus yang dapat digunakan untuk mencari *recall*[16] dapat dihitung menggunakan persamaan 9 sebagai berikut:

$$Recall(C_i) = \frac{TP(C_i)}{TP(C_i)+FN(C_i)} \tag{9}$$

F1-Score merupakan pengujian perfomansi untuk mengetahui ketepatan model dalam penggunaan data[2], rumus untuk mencari *F1-Score*[16] dapat dihitung menggunakan persamaan 10 sebagai berikut:

$$F1\ Score(C_i) = 2 \times \frac{Precision(C_i) \times Recall(C_i)}{Precision(C_i) + Recall(C_i)} \tag{10}$$

3. HASIL DAN PEMBAHASAN

Penggunaan data pada pengujian terdiri dari empat aspek yaitu harga, kemasan, efektivitas dan aroma. Jumlah data yang digunakan berjumlah 2.449, dengan pembagian polaritas data untuk aspek harga adalah 339 data positif, 117 data negative dan 2.095 data non sentimen, untuk aspek kemasan adalah 134 data positif, 28 data negative dan 2.389 data non sentimen, untuk aspek efektivitas adalah 1.771 data positif, 685 data negative dan 95 data non sentimen dan aspek aroma adalah 502 data positif, 130 data negative dan 1.919 data non sentimen.

Pengujian dilakukan sebanyak tiga kali, pengujian pertama yang dilakukan pada pembagian rasio data dengan empat skenario. Setelah didapatkan performa terbaik pada pengujian pertama, maka rasio pembagian data akan digunakan pada pengujian kedua, dimana pengujian kedua akan melakukan analisis sentimen untuk masing-masing aspek menggunakan NWKNN dengan mengubah nilai parameter e dan k . Pada pengujian ketiga akan menggunakan rasio pembagian data pada pengujian pertama dan akan dilakukan analisis sentimen untuk masing-masing aspek dengan menggunakan metode KNN dengan mengubah nilai parameter k . Hasil pengujian perfomansi akan dihitung dengan menggunakan metode *confusion matrix*.

3.1 Pengujian Rasio Data

Tujuan dari pengujian ini adalah untuk mengevaluasi kinerja sistem terhadap variasi jumlah data latih dan data uji dengan mengubah rasio pembagian data. Rasio pembagian data latih yang digunakan yaitu 90:10, 80:20, 70:30 dan 60:40. dimana data akan dibagi secara acak menjadi data latih dan data uji. K dan E yang digunakan dalam pengujian ini adalah K=10 [17] dan E=3 [18]. Hasil pengujian rasio pembagian data dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pengujian Rasio Data

Rasio	F1-Score
60:40	0.48
70:30	0.49
80:20	0.48
90:10	0.46

3.2 Pengujian NWKNN

Penelitian ini bertujuan untuk mengamati pengaruh variasi nilai k dan e terhadap perfomansi metode klasifikasi NWKNN. Sebanyak 2449 data digunakan untuk pengujian, dan data tersebut dibagi secara acak dengan rasio 70:30 berdasarkan hasil terbaik dari pengujian pertama. Nilai k yang diuji terdiri dari 2, 4, 8, 10, 15, 20, 25, 30, 35 dan 40[17]. Nilai e yang diuji adalah 2, 3, 4, 5, 6 [18].

Hasil pengujian untuk aspek harga dilihat pada Tabel 3.

Tabel 3. Hasil Pengujian NWKNN Aspek Harga

K	E				
	2	3	4	5	6
2	0.46	0.46	0.46	0.46	0.46
4	0.44	0.44	0.44	0.42	0.42
8	0.54	0.54	0.48	0.44	0.44
10	0.56	0.49	0.48	0.44	0.43
15	0.58	0.56	0.50	0.45	0.43
20	0.60	0.51	0.49	0.46	0.44
25	0.59	0.54	0.47	0.44	0.38
30	0.61	0.63	0.46	0.43	0.42
35	0.63	0.52	0.44	0.42	0.38
40	0.66	0.52	0.44	0.39	0.39

Hasil pengujian untuk aspek kemasan dilihat pada Tabel 4.

Tabel 4. Hasil Pengujian NWKNN Aspek Kemasan

K	E				
	2	3	4	5	6
2	0.46	0.46	0.46	0.46	0.46
4	0.44	0.44	0.44	0.42	0.42
8	0.54	0.54	0.48	0.44	0.44
10	0.56	0.49	0.48	0.44	0.43
15	0.58	0.56	0.50	0.45	0.43
20	0.60	0.51	0.49	0.46	0.44
25	0.59	0.54	0.47	0.44	0.38
30	0.61	0.63	0.46	0.43	0.42
35	0.63	0.52	0.44	0.42	0.38
40	0.66	0.52	0.44	0.39	0.39

Hasil pengujian untuk aspek efektivitas dilihat pada Tabel 5.

Tabel 5. Hasil Pengujian NWKNN Aspek Efektifitas

K	E				
	2	3	4	5	6
2	0.46	0.46	0.46	0.46	0.46
4	0.44	0.44	0.44	0.42	0.42
8	0.54	0.54	0.48	0.44	0.44
10	0.56	0.49	0.48	0.44	0.43
15	0.58	0.56	0.50	0.45	0.43
20	0.60	0.51	0.49	0.46	0.44
25	0.59	0.54	0.47	0.44	0.38
30	0.61	0.63	0.46	0.43	0.42
35	0.63	0.52	0.44	0.42	0.38
40	0.66	0.52	0.44	0.39	0.39

Hasil pengujian untuk aspek aroma dilihat pada Tabel 6.

Tabel 6. Hasil Pengujian NWKNN Aroma

K	E				
	2	3	4	5	6
2	0.46	0.46	0.46	0.46	0.46
4	0.44	0.44	0.44	0.42	0.42
8	0.54	0.54	0.48	0.44	0.44
10	0.56	0.49	0.48	0.44	0.43
15	0.58	0.56	0.50	0.45	0.43
20	0.60	0.51	0.49	0.46	0.44
25	0.59	0.54	0.47	0.44	0.38
30	0.61	0.63	0.46	0.43	0.42
35	0.63	0.52	0.44	0.42	0.38
40	0.66	0.52	0.44	0.39	0.39

3.3 Pengujian KNN

Penelitian ini bertujuan untuk mengamati pengaruh nilai k terhadap performansi metode klasifikasi KNN. Sebanyak 2449 data digunakan untuk pengujian, dan data tersebut dibagi secara acak dengan rasio 70:30 berdasarkan hasil terbaik dari pengujian pertama. Nilai k yang diuji antara lain 2, 4, 8, 10, 15, 20, 25, 30, 35, dan 40[17].

Hasil pengujian untuk KNN dilihat pada Tabel 7.

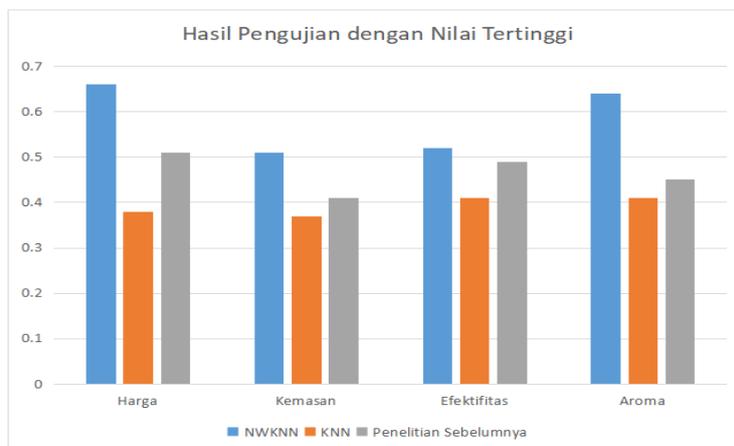
Tabel 7. Hasil Pengujian KNN

K	F1-Score Aroma	F1-Score Kemasan	F1-Score Efektivitas	F1-Score Aroma
2	0.37	0.37	0.04	0.30
4	0.38	0.45	0.12	0.35
8	0.34	0.37	0.46	0.36
10	0.34	0.34	0.47	0.37
15	0.33	0.32	0.41	0.41
20	0.33	0.32	0.41	0.39
25	0.32	0.32	0.40	0.40
30	0.33	0.32	0.39	0.41
35	0.31	0.32	0.39	0.41
40	0.31	0.32	0.40	0.41

3.4 Kesimpulan Pengujian

Pada pengujian NWKNN didapatkan F1-Score tertinggi untuk aspek harga sebesar 0,66, untuk aspek kemasan sebesar,0,66, untuk aspek efektivitas sebesar 0,66, dan untuk aspek aroma sebesar 0,66. Pada

pengujian KNN didapatkan F1-Score tertinggi untuk aspek harga sebesar 0,38, untuk aspek kemasan sebesar 0,45 untuk aspek efektifitas sebesar 0,47 dan untuk aspek aroma sebesar 0,41. Pada pengujian sebelumnya dengan menggunakan *information gain* dan KNN [2] didapatkan F1-Score tertinggi untuk aspek harga sebesar , untuk aspek kemasan sebesar, untuk aspek efektifitas sebesar dan untuk aspek aroma sebesar.



Gambar 4. Perbandingan Hasil Pengujian

Dari hasil pengujian menggunakan nilai F1-Score tertinggi untuk masing-masing aspek, dapat disimpulkan bahwa metode NWKNN memberikan hasil yang lebih baik dibandingkan dengan KNN serta penelitian sebelumnya yang menggunakan KNN dan *Information Gain*. Untuk menentukan nilai k dan e yang optimal secara keseluruhan, dilakukan perhitungan rata-rata F1-Score dengan menggunakan hasil pengujian pada Tabel 3, 4, 5, dan 6. Rata-rata F1-Score dari seluruh pengujian dapat dilihat pada Tabel 8.

Tabel 8. Rata-Rata F1-Score

K	E				
	2	3	4	5	6
2	0.315	0.315	0.315	0.315	0.315
4	0.335	0.3475	0.345	0.34	0.335
8	0.4125	0.4725	0.46	0.4525	0.4375
10	0.47	0.4825	0.455	0.4425	0.4425
15	0.5175	0.495	0.4725	0.4525	0.4225
20	0.545	0.4825	0.4625	0.44	0.425
25	0.55	0.4875	0.4425	0.425	0.405
30	0.55	0.5025	0.45	0.4275	0.42
35	0.5675	0.48	0.43	0.4175	0.4075
40	0.5725	0.4725	0.44	0.41	0.405

4. PENUTUP

Berdasarkan hasil pengujian, metode NWKNN dapat digunakan untuk melakukan analisis sentimen berbasis aspek pada produk kecantikan. NWKNN ini mampu mengatasi masalah ketidakseimbangan data dengan baik, di mana rata-rata *F1-Score* NWKNN mencapai 0,48, sementara rata-rata *F1-Score* KNN hanya mencapai 0,35. Hasil *F1-Score* tertinggi untuk setiap aspek dapat dicapai dengan menggunakan nilai e=2, sedangkan nilai k dan e yang optimal untuk semua aspek adalah k=40 dan e=2.

Untuk meningkatkan hasil penelitian ini, dapat dilakukan pengembangan lebih lanjut dengan menggunakan teknik *preprocessing* yang berbeda untuk mengetahui pengaruhnya terhadap akurasi hasil. Hal ini dapat membantu meningkatkan kinerja metode NWKNN dalam analisis sentimen berbasis aspek terhadap produk kecantikan.

DAFTAR PUSTAKA

- [1] B. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*. Chicago: 2015. doi: 10.1017/CBO9781139084789.
- [2] E. Yulianti Prastika S, S. Al Faraby, and M. Dwifabri P, "Analisis Sentimen pada Ulasan Produk Kecantikan Menggunakan K-Nearest Neighbor dan Information Gain," vol. 8, no. 5, pp. 10091–10105, 2021, doi: 10.34818/eoe.v8i5.15729.
- [3] R. Siringoringo, "Klasifikasi Data Tidak Seimbang Menggunakan Algoritma SMOTE dan k-Nearest Neighbor," *J. ISD*, vol. 3, no. 1, pp. 44–49, 2018.
- [4] A. Rivaldi, P. P. Adikara, and S. Adinugroho, "Klasifikasi Penyimpangan Tumbuh Kembang pada Anak Menggunakan Metode Neighbor Weighted K-Nearest Neighbor (NWKNN)," *Teknol. Inf. dan Ilmu*, vol. 2, no. 7, 2017, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1698>
- [5] I. Indriati and A. Ridok, "Sentiment Analysis for Review Mobile Applications Using Neighbor Method Weighted K-Nearest Neighbor (Nwknn)," *J. Environmental Eng. Sustain. Technol.*, vol. 3, no. 1, pp. 23–32, 2016, doi: 10.21776/ub.jeest.2016.003.01.4.
- [6] G. Tsoumakas and I. Katakis, *Multi-Label Classification*, vol. 3, no. 3. 2007. doi: 10.4018/jdwm.2007070101.
- [7] R. M. Juwita, E. Haerani, S. K. Gusti, and S. Ramadhani, "Klasifikasi Berita Menggunakan Metode K-Nearest Neighbor," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 2, pp. 259–268, Apr. 2022, doi: 10.32672/jnkti.v5i2.4192.
- [8] R. Amalia, M. A. Bijaksana, and D. Darmantoro, "Negation handling in sentiment classification using rule-based adapted from Indonesian language syntactic for Indonesian text in Twitter," *J. Phys. Conf. Ser.*, vol. 971, no. 1, 2018, doi: 10.1088/1742-6596/971/1/012039.
- [9] N. I. Widiastuti, E. Rainarli, and K. E. Dewi, "Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen," *J. Infotel*, vol. 9, no. 4, p. 416, 2017, doi: 10.20895/infotel.v9i4.312.
- [10] D. C. Hidayati, S. Al Faraby, and A. Adiwijaya, "Klasifikasi Topik Multi Label pada Hadis Shahih Bukhari Menggunakan K-Nearest Neighbor dan Latent Semantic Analysis," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 140, 2020, doi: 10.30865/jurikom.v7i1.2013.
- [11] M. N. Saadah, R. W. Atmagi, D. S. Rahayu, and A. Z. Arifin, "Information Retrieval Of Text Document With Weighting TF-IDF and LCS," *J. Ilmu Komput. dan Inf.*, vol. 6, no. 1, p. 34, Oct. 2013, doi: 10.21609/jiki.v6i1.216.
- [12] scikit-learn developers, "scikit-learn user guide Release 0.18.2 scikit-learn developers," pp. 1–2050, 2017, [Online]. Available: https://scikit-learn.org/0.18/_downloads/scikit-learn-docs.pdf diakses: 18 April 2023
- [13] B. L. Yudha, L. Muflikhah, and R. C. Wihandika, "Klasifikasi Risiko Hipertensi Menggunakan Metode Neighbor Weighted K- Nearest Neighbor (NWKNN)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 2, pp. 897–904, 2018.
- [14] M. N. Saadah, R. W. Atmagi, D. S. Rahayu, and A. Z. Arifin, "Identifikasi Penyakit Gagal Ginjal Menggunakan Metode Neighbor (NWKNN)," *J. Ilmu Komput. dan Inf.*, vol. 6, no. 1, p. 34, Oct. 2013, doi: 10.21609/jiki.v6i1.216.
- [15] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021, doi: <http://dx.doi.org/10.30645/j-sakti.v5i2.369>.
- [16] I. Markoulidakis, G. Kopsiaftis, I. Rallis, and I. Georgoulas, "Multi-Class Confusion Matrix Reduction method and its application on Net Promoter Score classification problem," *ACM Int. Conf. Proceeding Ser.*, pp. 412–419, 2021, doi: 10.1145/3453892.3461323.
- [17] P. N. Fadila, . I., and D. E. Ratnawati, "Identifikasi Jenis Attention Deficit Hyperactivity Disorder (ADHD) Pada Anak Usia Dini Menggunakan Metode Neighbor Weighted K-Nearest Neighbor (NWKNN)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, p. 193, Sep. 2016, doi: 10.25126/jtiik.201633195.
- [18] A. Satria, Marji, and D. E. Ratnawati, "Klasifikasi Jenis Kanker Berdasarkan Struktur Protein Menggunakan Metode Neighbor Weighted K-Nearest Neighbor (NWKNN)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 4, pp. 3617–3624, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>