

ANALISIS SENTIMEN ULASAN PRODUK SERUM WAJAH PADA BEAUTY BRAND SOMETHINC MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

Dinda Meisa Azzahra, Moch Hafid T, Syariful Alam

Program Studi Teknik Informatika S1, Stt Wastukancana

Jalan Cikopak No.53, Mulyamekar, Kec.Babakancikao, Kab. Purwakarta, Jawa Barat

dindameisa39@wastukancana.ac.id

ABSTRAK

Perkembangan jumlah pengguna media sosial didampingi dengan perkembangan bisnis produk perawatan kulit baik secara global maupun domestik di Indonesia. Pada penelitian ini untuk mengetahui review dari produk serum somethinc, dengan perbandingan Serum Niacinamide dan Serum AHA BAH. Yang terdapat berbagai ulasan positif dan negative dan neutral. Penelitian ini menggunakan metode text mining yang terdiri dari tahapan crawling data, labelling, cleaning, preprocessing (transformation, tokenization, filtering). Untuk algoritma yang digunakan yaitu naïve bayes karena mempunyai nilai probabilitas atau peluang tertinggi untuk pengklasifikasian data, untuk pembobotan menggunakan perhitungan TF-IDF, dan pengujian data menggunakan confusion matrix yang ada pada tools phyton. Dari hasil penelitian ini terdapat analisis sentimen dari ulasan produk serum Niacinamide dengan jumlah data sebanyak 720, Dan serum AHA BAH PAH dengan jumlah data sebanyak 824, Berdasarkan confusion matrix menggunakan tools phyton serta goggle colab. Dan setelah melakukan evaluasi menggunakan confusion matrix pada serum AHA BAH PAH didapatkan akurasi sebanyak (accuracy) 80%, presisi sebesar (Precision) 84%, serta (Recall) sebesar 94%. Pada Serum Niacinamid didapatkan akurasi sebanyak 64% presisi sebesar 67%, serta recall sebesar 69%, Disimpulkan bahwa hasil dari Penilaian diatas menunjukkan bahwa data produk serum Serum AHA BAH PAH memiliki hasil lebih baik dengan jumlah data sentimen Positif 675 dan nilai akurasi (accuracy) 80% dan Presisi (Precision) 84%, recall 94%. Yang tergolong positif.

Kata kunci: *serum, naïve Bayes, female daily, phyton, analisis, sentiment*

1. PENDAHULUAN

Perkembangan teknologi internet dan penggunaan media sosial telah berkontribusi pada pertumbuhan bisnis perawatan kulit, baik di tingkat global maupun domestik di Indonesia. Fenomena ini dapat diamati melalui kehadiran berbagai merek global yang memasuki pasar Indonesia, serta perkembangan bisnis lokal yang menciptakan merek perawatan kulit khas Indonesia. Terlihat bahwa pasar kosmetik dan perawatan kulit di Indonesia dibanjiri dengan produk dari merek lokal dan global, seperti The Body Shop, Wardah, Somethinc, Victoria's Secret, Avoskin, Loreal, dan sebagainya [1].

Salah satu merek lokal di Indonesia yang telah mengatasi kesulitan para kaum milenial dalam memilih produk kosmetik adalah Somethinc. Brand Somethinc ini didirikan oleh Irene Ursula pada bulan Maret 2019. Merek ini dikenal sebagai produsen kosmetik dan perawatan kulit lokal yang telah mendapatkan sertifikasi halal serta memenuhi standar kualitas internasional. Dalam kurun waktu 2 tahun sejak pendiriannya, akun Instagram resmi mereka, @somethincofficial, telah berhasil mengumpulkan lebih dari 995.000 pengikut. Pada tahun 2020, Somethinc meraih keberhasilan sebagai salah satu merek kosmetik yang masuk dalam daftar Top 50 Brand Indonesia [2]. Produk serum yang paling terkenal dari brand Somethinc adalah Serum Aha Bah Pah Peeling Solution, yang mendapatkan rating terendah sebesar 4.1 di Female Daily. Rating yang

rendah ini disertai dengan berbagai ulasan positif dan negatif, menunjukkan bahwa performa produk yang ditawarkan oleh Somethinc belum sepenuhnya memenuhi harapan dan kebutuhan pengguna perawatan kulit. Oleh karena itu, evaluasi perlu dilakukan oleh Somethinc untuk meningkatkan kualitas produk dan memperbaiki reputasi agar menjadi produk yang berkualitas tinggi dan diterima dengan baik oleh konsumen.

Ulasan atau tinjauan online dari pengguna dapat menjadi alat yang efektif dalam memperoleh informasi tentang produk atau layanan yang diberikan. Pemilihan platform Female Daily juga didukung oleh kredibilitas Female Daily sebagai komunitas media sosial pertama di Indonesia dalam bidang kecantikan yang menyediakan review terpercaya mengenai perawatan kulit, perawatan tubuh, perawatan rambut, dan kebutuhan wanita lainnya [3]. Dalam penerapan analisis sentiment digunakan metode algoritma, maka digunakan metode Naïve Bayes Classifier untuk mengetahui hasil ulasan dari pengguna produk Serum Aha Bah Pah Peeling Solution dari brand "Somethinc" yang memiliki rating terendah pada produk berbagai produk somethinc dapat dilihat pada website Female Daily. Yang terdapat berbagai ulasan positif dan negative, Dari banyak nya ulasan atau opini yang diberikan pengguna terhadap review produk Serum Aha Bah Pah Peeling Solution.

Dalam penelitian ini, dibahas mengenai metode pengklasifikasian ulasan produk menjadi opini positif

dan negatif menggunakan salah satu metode klasifikasi, yaitu Naïve Bayes. Penelitian yang dilakukan oleh Khalida dan Setiawati pada tahun 2020 ini bertujuan untuk mengevaluasi tingkat akurasi, presisi, dan recall dari klasifikasi Naïve Bayes Classifier dengan mengoptimalkan pembobotan Gain Ratio. Algoritma Naïve Bayes adalah teknik prediksi yang relatif sederhana, dimana prediksi kemungkinan kejadian di masa mendatang didasarkan pada data yang telah ada sebelumnya. [4]. Algoritma ini mampu mengolah jumlah data yang besar dengan tingkat akurasi yang tinggi. Metode yang sering digunakan adalah Naïve Bayes Classifier yang pertama kali dikemukakan oleh Revered Thomas Bayes. Dalam penelitian ini, aplikasi pengolahan data yang digunakan adalah Rapidminer, yang mendukung berbagai algoritma dan metode di dalamnya.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen adalah cabang penelitian yang mengkaji pandangan, perasaan, penilaian, pengevaluasian, sikap, dan emosi individu terhadap suatu topik, layanan, produk, individu, organisasi, atau kegiatan khusus. Tujuan dari analisis sentimen adalah untuk menentukan apakah pendapat atau komentar mengenai suatu masalah memiliki orientasi positif atau negatif, serta untuk menggunakannya sebagai panduan dalam meningkatkan layanan atau kualitas produk [5].

2.2. Brand Somethinc

Somethinc merupakan sebuah merek produk skincare yang didirikan oleh Irene Ursula pada bulan Maret 2019. Merek ini terkenal sebagai produsen produk make-up dan skincare lokal yang memiliki sertifikasi halal dan memenuhi standar internasional dalam hal kualitas. Dalam waktu 2 tahun sejak pendiriannya, akun resmi Instagram mereka dengan nama @somethincofficial telah berhasil mendapatkan lebih dari 995,000 pengikut. Pada tahun 2020, *Somethinc* berhasil meraih prestasi sebagai salah satu dari Top 50 Brand Indonesia di industri kosmetik [6].

2.3. Female Daily

Female Daily merupakan platform daring yang menyediakan wadah bagi pengguna untuk berbagi pengalaman mereka dalam menggunakan produk skincare. Di samping itu, forum di situs web Female Daily menjadi sumber utama untuk membaca ulasan produk kecantikan. Ulasan produk yang ditulis oleh para pengguna menjadi salah satu fitur yang paling diminati dan berperan penting dalam memberikan informasi kepada para pengunjung situs ini. Electronic Word of Mouth (E-WOM) yang terjadi di platform ini mencakup ulasan berdasarkan pengalaman pengguna dengan produk skincare atau kecantikan tertentu, yang kemudian digunakan sebagai referensi oleh pengguna lain sebelum memutuskan untuk membeli produk tersebut, meskipun tidak selalu sepenuhnya akurat [7].

2.4. Serum Wajah

Serum Niacinamide dan serum AHA BAH PAH Peeling Solution adalah produk yang dirancang khusus untuk mengatasi berbagai masalah kulit wajah, memberikan rasa tenang dalam merawat kulit. Kedua produk ini memiliki formula khusus yang bertujuan untuk membantu melindungi kulit dari berbagai masalah, mulai dari penuaan hingga ketidakmerataan warna wajah. Serum wajah memiliki konsentrasi bahan aktif yang tinggi, sehingga dapat dengan cepat diserap oleh kulit dan memberikan efek yang nyaman. Selain itu, karena viskositasnya yang rendah, serum mudah menyebar di permukaan kulit, membuat penggunaan serum lebih efisien. Serum dianggap sebagai jenis emulsi karena memiliki viskositas rendah dan memberikan performa maksimal dalam merawat kulit [8].

2.5. Naïve Bayes

Klasifikasi Naive Bayes merupakan klasifikasi yang bersifat supervised learning karena memiliki supervisor (manusia melakukan klasifikasi secara manual pada data yang digunakan dalam pelatihan) selaku pengajar dalam proses belajar atau learning. Selain itu, performansi Naive Bayes memiliki waktu klasifikasi yang singkat sehingga mempercepat proses sistem analisis sentimen. Pada penelitian ini pengujian digunakan dengan menggunakan 5 kelas (sangat negatif, negatif, netral, positif dan sangat positif) dan 2 kelas (negatif, dan positif) [9].

2.6. Text Mining

Secara umum, teks mining adalah konsep tentang memproses sejumlah besar dokumen yang ada dari waktu ke waktu dengan menggunakan berbagai analisis. Tujuan dari pengolahan teks ini adalah untuk mengidentifikasi dan mengekstrak informasi yang berharga dari sumber data tersebut, serta mengeksplorasi pola menarik dalam kasus text mining. Sumber data yang digunakan dalam proses ini biasanya berupa kumpulan atau koleksi dokumen yang tidak terstruktur, dan memerlukan pengelompokan untuk dapat diolah secara efisien [10].

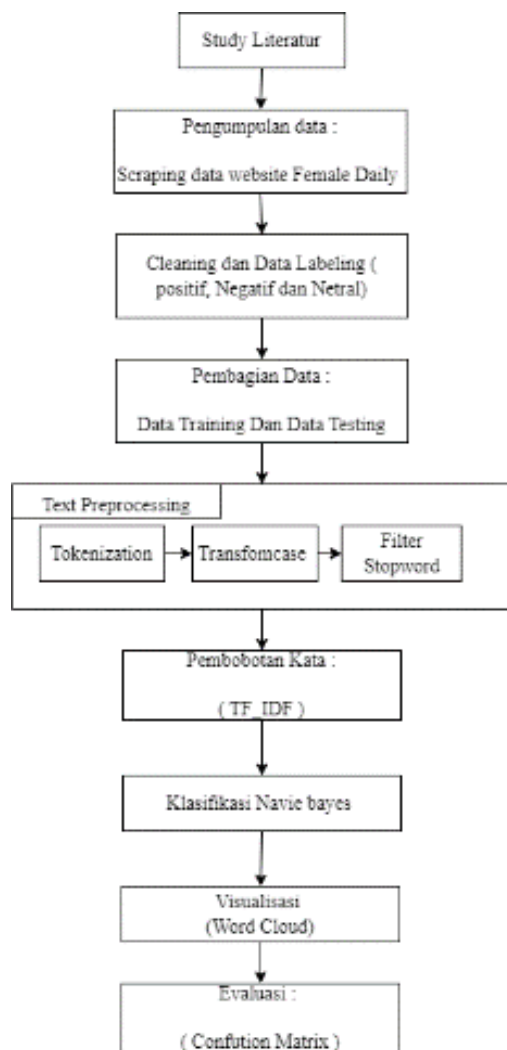
3. METODE PENELITIAN

Objek penelitian merupakan komponen yang penting dalam sebuah penelitian karena melalui objek penelitian tersebut, kita dapat mengidentifikasi variabel-variabel penting yang akan mendukung penelitian. Dalam penelitian yang akan dilakukan ini, objek penelitian yang akan difokuskan adalah ulasan pengguna produk Serum Niacinamide dan Serum AHA BAH PAH dari brand kecantikan "*Somethinc*" yang terdapat di platform Female Daily.

3.1. Kerangka Pemikiran

Kerangka pikir adalah sebuah sistem pemikiran yang dirancang berdasarkan aktivitas penelitian yang telah dilakukan. Kerangka berpikir merupakan suatu model konseptual yang menggambarkan bagaimana

teoriterhubung dengan berbagai faktor yang telah diidentifikasi sebagai masalah yang signifikan. [11]. Dapat dilihat pada gambar di bawah ini adalah kerangka pemikiran :



Gambar 1. Kerangka pemikiran

3.2. Study Literatur

Dalam penelitian ini, populasi yang dimaksud adalah semua ulasan atau review pengguna produk Serum Niacinamide "Somethinc" yang terdapat dalam database website Female Daily. Sebagai sampel, digunakan 720 ulasan produk Serum Niacinamide dan 824 ulasan produk Serum AHA BAH PAH Peeling Solution.

3.3. Scraping Data

Pada penelitian ini, metode pengumpulan data yang digunakan adalah metode web scraping dari Google Chrome. Pada alamat website : https://reviews.femaledaily.com/products/treatment/serum-essence-23/somethinc/niacinamide-moisture-beet-serum-19?cat=&cat_id=0&age_range=&skin_type=&skin_t

[one=&skin undertone=&hair texture=&hair type=&order=newest&page=1](https://reviews.femaledaily.com/products/treatment/serum-essence-23/somethinc/niacinamide-moisture-beet-serum-19?cat=&cat_id=0&age_range=&skin_type=&skin_t)

3.4. Labeling Dan Cleaning

Proses labeling adalah proses menentukan sentimen setiap ulasan tersebut, dalam penelitian ini dibagi menjadi 2 (Sentimen) positif dan negatif. Untuk ulasan yang diberikan label positif memberikan review yang "cocok, bagus dimuka", dan berkata "kurang cocok dimuka" diberikan label negative. Dengan menggunakan *tools python* dapat dilihat dibawah ini :

```

score = data["Compound"].values
sentiment = []
for i in score:
    if i >= 0.05:
        sentiment.append('positif')
    elif i <= -0.05:
        sentiment.append('negatif')
data["Sentiment"] = sentiment
data.head()
  
```

Gambar 2. Source Kode Python Labeling

Berikut adalah hasil dari proses labeling diatas, dapat dilihat tabel dibawah ini:

Tabel 1. Hasil Proses Labeling

Ulasan	Sentimen
klo di aku ga terlalu berasa ngefek sih, tp ga bikin Negatif BO jd krn sayang, kupake aja gtu afternya gamau nah pas bgt aku lg ganti skincare dan dia gacocok sm aha bha	Negatif
rekomen bagus buat pemula. krn ini udh ada AHA BHA PHA.	Positif

Proses cleaning adalah proses membersihkan ulasan dari kata-kata yang tidak dibutuhkan. Kata-kata yang disingkat adalah huruf HTML, kata kunci, symbol emosional, tagar.

```

def clean_text(text):
    text = re.sub('@[\^s]+', '', text)
    text = re.sub(r'(?://|http?://|www)\S+', '', text)
    text = text.translate(str.maketrans('', '', string.punctuation))
    text = re.sub('<.*?>', '', text)
    text = re.sub("\n", "", text)
    text = text.lower()
    text = re.sub(r"[a-zA-Z]\b", "", text)
    text = re.sub(r'^rt[\s]+', '', text)
    text = ' '.join(text.split())
    text = re.sub(r'\d+', '', text)
    emoticon_pattern = re.compile("[f")
  
```

Gambar 3. Source Kode Python Cleaning

Berikut adalah hasil dari cleaning, dapat dilihat tabel dibawah ini :

Tabel 2. Hasil Cleaning

Sebelum
Tekstur: Cair bening berwarna merah tua mudah menyerap dan tidak lengket dikulit. Efektivitas: serum ini gentle banget, awalnya sempet agak takut celkit-celkit ternyata nggak sama sekali. Suka banget sama hasilnya bikin kulit halus dan glowing. Tp setelah pakai serum ini rutin,

Tabel 3. Hasil Cleaning

Sesudah
tekstur cair bening berwarna merah tua mudah menyerap dan tidak lengket dikulit efektivitas serum ini gentle banget awalnya sempet agak takut celkitcelkit ternyata nggak sama sekali suka banget sama hasilnya bikin kulit halus dan glowing tp setelah pakai serum ini rutin

3.5. Pre-Processing Teks

Pre-processing adalah suatu langkah penting yang bertujuan untuk memilah data yang tidak relevan sehingga data menjadi lebih terstruktur dan informasi yang diperoleh menjadi lebih jelas. Dalam proses pre-processing, terdapat beberapa tahapan yang meliputi:

- a) **Case Folding:** tahap ini mengubah teks menjadi huruf kecil (lowercase), menghapus tanda baca, serta menghilangkan angka pada teks ulasan.

Tabel 4. Case Folding

Sebelum
Tekstur: Cair bening berwarna merah tua mudah menyerap dan tidak lengket dikulit. Efektivitas: serum ini gentle banget, awalnya sempet agak takut celkit-celkit ternyata nggak sama sekali. Suka banget sama hasilnya bikin kulit halus dan glowing. Tp setelah pakai serum ini rutin,

Tabel 5. Case Folding

Sesudah
tekstur cair bening berwarna merah tua mudah menyerap dan tidak lengket dikulit efektivitas serum ini gentle banget awalnya sempet agak takut celkitcelkit ternyata nggak sama sekali suka banget sama hasilnya bikin kulit halus dan glowing tp setelah pakai serum ini rutin

- b) **Tokenizing:** tahap ini berfungsi untuk membersihkan dan mengekstraksi kata-kata. Pada tahap ini, teks dipecah menjadi unit-unit terkecil yang disebut token, seperti kata-kata.

Tabel 6. Tokenizing

Sebelum
tekstur cair bening berwarna merah tua mudah menyerap dan tidak lengket dikulit efektivitas serum ini gentle banget awalnya sempet agak takut celkitcelkit ternyata nggak sama sekali suka banget sama hasilnya bikin kulit halus dan glowing tp setelah pakai serum ini rutin

Tabel 7. Tokenizing

Sesudah
['Tekstur', ':', 'Cair', 'bening', 'berwarna', 'merah', 'tua', 'mudah', 'menyerap', 'dan', 'tidak', 'lengket', 'dikulit', ':', 'Efektivitas', ':', 'serum', 'ini', 'gentle', 'banget', ':', 'awalnya', 'sempet', 'agak', 'takut', 'celkit-celkit', 'ternyata', 'nggak', 'sama', 'sekali', ':', 'Suka', 'banget', 'sama', 'hasilnya', 'bikin', 'kulit', 'halus', 'dan', 'glowing', ':', 'Tp', 'setelah', 'pakai', 'serum', 'ini', 'rutin', ':']

- c) **Filtering (Stop words):** tahap ini merupakan proses penyaringan kata-kata yang umumnya tidak memiliki makna signifikan dalam suatu dokumen.

Kata-kata tersebut sering disebut sebagai "stop words" dan dihilangkan dalam proses pre-processing.

Tabel 8. StopWord Removal

Sebelum
tekstur cair bening berwarna merah tua mudah menyerap dan tidak lengket dikulit efektivitas serum ini gentle banget awalnya sempet agak takut celkitcelkit ternyata nggak sama sekali suka banget sama hasilnya bikin kulit halus dan glowing tp setelah pakai serum ini rutin

Tabel 9. StopWord Removal

Sesudah
['Tekstur', ':', 'Cair', 'bening', 'berwarna', 'merah', 'tua', 'mudah', 'menyerap', 'dan', 'tidak', 'lengket', 'dikulit', ':', 'Efektivitas', ':', 'serum', 'ini', 'gentle', 'banget', ':', 'awalnya', 'sempet', 'agak', 'takut', 'celkit-celkit', 'ternyata', 'nggak', 'sama', 'sekali', ':', 'Suka', 'banget', 'sama', 'hasilnya', 'bikin', 'kulit', 'halus', 'dan', 'glowing', ':', 'Tp', 'setelah', 'pakai', 'serum', 'ini', 'rutin', ':']

3.6. TF_IDF

Setelah selesai melakukan preprocessing (transformasi, tokenisasi, penyaringan) pada dokumen, langkah selanjutnya adalah melakukan proses term weighting, di mana setiap term akan diberi bobot (nilai) yang menunjukkan tingkat pentingnya term tersebut dalam dokumen. Selanjutnya, dilakukan perhitungan bobot untuk setiap term yang dicari dalam setiap dokumen, dengan tujuan untuk mengetahui kegunaan dan kemiripan term tersebut dalam dokumen-dokumen tersebut. Semakin sering kata muncul dalam kumpulan dokumen, semakin tinggi nilai atau bobot kata tersebut. Berikut dibawah ini adalah tabel 4.13 sampel data ulasan Serum AHA BAH PAH dan Serum Niacinamide dapat dilihat di bawah ini :

Tabel 10. Sampel Data Ulasan Serum Aha Bah Pah

Ulasan	Dokumen
Tekstur cair bening berwarna merah tua mudah menyerap dan tidak lengket	D1
Pertama kali pakai udah langsung besoknya jerawat dimanamana	D2
Ini adaah produk peeling pertama. Tidak berespektasi banyak waktu beli	D3

Tabel 11. Sampel Data Ulasan Serum Niacinamide

Ulasan	Dokumen
awal cobain produk ini karena lagi jerawat dan review di sini pada bilang bagus	D1
cinta banget sama serum ini awal ga berekspektasi tinggi bisa pudar bekas jerawat	D2
Serum ini udah kuhabisin dua botol, lebih kentel teksturnya	D3

3.7. Term Frequency

Pada tahap ini semakin besar jumlah kemunculan istilah dalam dokumen, semakin besar bobotnya. Kata yang tidak muncul akan bernilai 0, dan jika ada maka

bernilai 1. Berikut hasil term frequency Menggunakan Excel dapat dilihat pada tabel 4.14 dibawah ini :

Tabel 12. Term Frequency Serum Niacinamide

TERM	TF		
	D1	D2	D3
awalnya	1	0	0
cobain	1	0	0
produk	1	0	0
ini	1	1	1
karena	1	0	0
lagi	1	0	0
jerawat	1	0	0
dan	1	0	0
review	1	0	0
di	1	0	0
sini	1	0	0
pada	1	0	0
bilang	1	0	0
bagus	1	0	0
cinta	0	1	0
banget	0	1	0
sama	0	1	0
serum	0	1	1
ga	0	1	0
berekspetasi	0	1	0

Tabel 13. Term Frequency Serum AHA BAH PAH

TERM	TF		
	D1	D2	D3
tekstur	1	0	0
cair	1	0	0
bening	1	0	0
warna	1	0	0
merah	1	0	0
tua	1	0	0
mudah	1	0	0
serap	1	0	0
dan	1	0	0
tidak	1	0	1
lengket	1	0	0
pertama	0	1	1
kali	0	1	0
pakai	0	1	0
udah	0	1	0
langsung	0	1	0
besok	0	1	0
jerawat	0	1	0
dimanamana	0	1	0
ini	0	0	1

3.8. Invers Dokumen Frequency

Pada tahap inverse document frequency setelah semua kata diberi bobot, selanjutnya menghitung document frequency (DF) atau berapa banyak kata yang ada dalam dokumen. Kemudian total dari seluruh dokumen dibagi dengan document frequency (DF) untuk menghasilkan nilai dokumen/DF.

Tabel 14. Invers Dokumen Frequency

TERM	DF	IDF
awalnya	1	0,477
cobain	1	0,477
produk	1	0,477
ini	2	0,176
karena	1	0,477
lagi	1	0,477
jerawat	1	0,477
dan	1	0,477
review	1	0,477
di	1	0,477
sini	1	0,477
pada	1	0,477
bilang	1	0,477
bagus	1	0,477
cinta	1	0,477
banget	1	0,477
sama	1	0,477
serum	2	0,176
ga	1	0,477
berekspetasi	1	0,477

Tabel 15. Invers Dokumen Frequency

TERM	DF	IDF
tekstur	1	0,477
cair	1	0,477
bening	1	0,477
warna	1	0,477
merah	1	0,477
tua	1	0,477
mudah	1	0,477
serap	1	0,477
dan	1	0,477
tidak	2	0,176
lengket	1	0,477
pertama	2	0,176
kali	1	0,477
pakai	1	0,477
udah	1	0,477
langsung	1	0,477
besok	1	0,477
jerawat	1	0,477
dimanamana	1	0,477
ini	1	0,477

3.9. Term Frequency Invers dokumen frequency

Pada tahap term frequency inverse document frequency untuk mendapatkan nilai tf-idf yaitu dengan cara mengalikan bobot tiap kata dari seluruh dokumen, dengan nilai dari idf kata tersebut.

Tabel 16. Term Frequency Invers Dokumen Frequency

	W = TF * DF	
D1	D2	D3
0,477	0	0
0,477	0	0
0,477	0	0
0,176	0	0,176
0,477	0	0
0,477	0	0

	W = TF * DF	
D1	D2	D3
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0	0,477	0
0	0,477	0
0	0,477	0
0	0,176	0,176
0	0,477	0
0	0,477	0
0	0,477	0
0	0,477	0
0	0,477	0
0	0,477	0
0	0	0,477
0	0	0,477
0	0	0,477
0	0	0,477
0	0	0,477

Tabel 17. Term Frequency Invers Dokumen

	W = TF * DF	
D1	D2	D3
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,477	0	0
0,176	0	0,176
0,477	0	0
0	0,176	0,176
0	0,477	0
0	0,477	0
0	0,477	0
0	0,477	0
0	0,477	0
0	0,477	0
0	0,477	0

4. HASIL DAN PEMBAHASAN

4.1. Klasifikasi Naïve Bayes

Dokumen ulasan yang sudah melewati proses sebelumnya bisa dikelompokkan dalam kategori tertentu berdasarkan kata-kata yang terdapat di dalamnya. Salah satu pendekatan yang digunakan untuk klasifikasi ini adalah metode Naïve Bayes, yang bertujuan untuk mencapai tingkat akurasi yang tinggi dan hasil yang optimal. Microsoft Excel merupakan salah satu alat yang digunakan untuk melakukan perhitungan klasifikasi Naïve Bayes, sedangkan

Python digunakan untuk melakukan pemodelan dan pengujian klasifikasi Naïve Bayes.

4.2. Model Klasifikasi Naïve bayes

```

prediksi benar: 67 data
prediksi salah: 16 data
Akurasi Algoritme: 80.72289156626506 %
TRUE NEGATIVE (TN): 2
FALSE NEGATIVE (FN): 4
TRUE POSITIVE (TP): 65
FALSE POSITIVE (FP): 12
PRECISION: 84.4155844155844 %
RECALL: 94.20289855072464 %
[[ 2 12]
 [ 4 65]]

```

	precision	recall	f1-score	support
negatif	0.33	0.14	0.20	14
positif	0.84	0.94	0.89	69
accuracy			0.81	83
macro avg	0.59	0.54	0.55	83
weighted avg	0.76	0.81	0.77	83

Gambar 4. Model Python Klasifikasi Naïve Bayes

4.3. Evaluasi (Confusion Matrix)

Pada tahap ini, dilakukan pengujian akurasi dengan menggunakan confusion matrix dan visualisasi. Evaluasi dilakukan untuk memeriksa kebenaran hasil klasifikasi. Dalam tahap ini, evaluasi menggunakan metode cross validation, di mana hasilnya akan diukur menggunakan Confusion Matrix untuk menentukan akurasi, presisi, dan recall. Berikut adalah tabel 3.1 Confusion Matrix dapat dilihat di bawah ini :

Tabel 18. Confusion Matrix

	Positif	Negatif
Negative	2	12
Positif	4	65

4.3.1. Confition Matrix Serum AHA BAH PAH

Akurasi : 80.72289%

$$\begin{aligned}
 \text{Akurasi} &= \frac{TN + TP}{TP + TN + FP + FN} \\
 &= \frac{2 + 65}{65 + 2 + 12 + 4} = \frac{67}{83} \\
 &= 0,8 = 80\%
 \end{aligned}$$

Rumus 1. Perhitungan Akurasi

Precision : 84.41558%

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FP} = \frac{65}{65 + 12} = \frac{65}{77} \\
 &= 0,84 = 84\%
 \end{aligned}$$

Rumus 2. Perhitungan Presisi

Recall : 94.202898%

$$Recall = \frac{TP}{TP + FN} = \frac{65}{65 + 4} = \frac{65}{69}$$

$$= 0,94 = 94\%$$

Rumus 3. Perhitungan Recall

4.3.2. Confution Matrix Serum Niacinamide

Akurasi : 80.72289%

$$\frac{TN + TP}{TP + TN + FP + FN} = \frac{38 + 55}{55 + 38 + 27 + 24}$$

$$= \frac{93}{144} = 0,64 = 64\%$$

Rumus 4. Perhitungan Akurasi

Precision : 84.41558%

$$Precision = \frac{TP}{TP + FP} = \frac{55}{55 + 27}$$

$$= \frac{55}{82} = 0,67$$

$$= 67\%$$

Rumus 5. Perhitungan Presisi

$$Recall = \frac{TP}{TP + FN} = \frac{55}{55 + 24} = \frac{55}{79} = 0,69$$

$$= 69\%$$

Rumus 6. Perhitungan Recall

4.4. Visualisasi (WordCloud)

- a. Hasil dari visualisasi word cloud Positif ditampilkan pada gambar 5. sebagai berikut:



Gambar 5. WordCloud Positif

Berdasarkan gambar 5. Word Cloud diatas terdapat kata – kata yang serong muncul pada data ulasan positif keseluruhan adalah “untuk”, “banget”, “yang”, “juga”, dan lain – lain.

- b. Hasil dari visualisasi word cloud negative ditampilkan pada gambar 4.16 sebagai berikut :



Gambar 6. WordCloud Negatif

Berdasarkan gambar 6. Word Cloud diatas terdapat kata – kata yang serong muncul pada data ulasan positif keseluruhan adalah “juga”, “tapi”, “yang”, “banget”, dan lain – lain.

5. KESIMPULAN DAN SARAN

Dari hasil penelitian diatas mengenai analisis sentiment ulasan Perkembangan jumlah pengguna media sosial didampingi dengan perkembangan bisnis produk perawatan kulit baik secara global maupun domestik di Indonesia. Pada penelitian ini untuk mengetahui riview dari produk serum somethinc, dengan perbandingan Serum Niacinamide dan Serum AHA BAH. Yang terdapat berbagai ulasan positif dan negative dan neutral. Penelitian ini menggunakan metode text mining yang terdiri dari tahapan crawling data, labelling, cleaning, preprocessing (transformation, tokenization, filtering). Untuk algoritma yang digunakan yaitu naïve bayes karena mempunyai nilai probabilitas atau peluang tertinggi untuk pengklasifikasian data, untuk pembobotan menggunakan perhitungan TF-IDF, dan pengujian data menggunakan confusion matrix yang ada pada tools phyton. Dari hasil penelitian ini terdapat analisis sentimen dari ulasan produk serum Niacinamide dengan jumlah data sebanyak 720, Dan serum AHA BAH PAH dengan jumlah data sebanyak 824, Berdasarkan confusion matrix menggunakan tools phyton serta goggle colaboryatory. Dan setelah melakukan evaluasi menggunakan confusion matrix pada serum AHA BAH PAH didapatkan akurasi sebanyak (accuracy) 80%, presisi sebesar (Precision) 84%, serta (Recall) sebesar 94%. Pada Serum Niacinamid didapatkan akurasi sebanyak 64% presisi sebesar 67%, serta recall sebesar 69%,. Disimpulkan bahwa hasil dari Penilaian diatas menunjukkan bahwa data produk serum Serum AHA BAH PAH memiliki hasil lebih baik dengan jumlah data sentimen Positif 675 dan nilai akurasi (accuracy) 80% dan Presisi (Precision) 84%, recall 94%. Yang tergolong positif.

produk serum wajah pada beauty brand something Serum Niacinamide yang berjumlah 720 dan Serum AHA BAH PAH Peeling Solution yang berjumlah 824 sudah melewati tahap preprocess text yang terdiri dari filtering, tokenization, transformation dan klasifikasi menggunakan algoritma naïve bayes serta evaluasi data dengan confusion matrix menggunakan tools phyton serta goggle colaboratory. Dan setelah melakukan evaluasi menggunakan confusion matrix pada serum AHA BAH PAH didapatkan akurasi sebanyak (accuracy) 80%, presisi sebesar (Precision) 84%, serta (Recall) sebesar 94%. Pada Serum Niacinamid didapatkan akurasi sebanyak 64% presisi sebesar 67%, serta recall sebesar 69%, Disimpulkan bahwa hasil dari Penilaian diatas menunjukan bahwa data produk serum Serum AHA BAH PAH memiliki hasil lebih baik dengan jumlah data sentimen Positif 675 dan nilai akurasi (accuracy) 80% dan Presisi (Precision) 84%, recall 94%. Yang tergolong positif. Berdasarkan nilai tersebut membuktikan sentiment mengenai ulasan produk Serum AHA BAH PAH pada beauty brand something terbukti lebih baik dalam mengklasifikasikan dokumen ulasan sentiment.

DAFTAR PUSTAKA

- [1] Asri Nabila. (n.d.). Asri Nabila, *ANALISIS SENTIMEN ULASAN PRODUK TONER PADA BEAUTY BRAND "THE BODY SHOP" MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE: STUDI KASUS DI FEMALE DAILY*.
- [2] Azizah, Y., Kurniawati, E., & Dyah, W. (n.d.). *KARAKTERISTIK SEDIAAN SERUM WAJAH DENGAN VARIASI KONSENTRASI SARI RIMPANG TEMU GIRING (Curcuma heyneana) TERFERMENTASI Lactobacillus bulgaricus CHARACTERISTICS OF FACIAL SERUM PREPARATION WITH VARIOUS CONCENTRATION OF TEMU GIRING (Curcuma heyneana) FERMENTED WITH Lactobacillus bulgaricus*.
- [3] Francia, C. (2022). *Pengaruh Electronic Word of Mouth Terhadap Minat Membeli Nature Republic* (Vol. 9, Issue 2).
- [4] Gunawan, B., Sasty, H., #2, P., Esyudha, E., & #3, P. (2018). *JEPIN (Jurnal Edukasi dan Penelitian Informatika) Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes*. 4(2), 17–29. www.femaledaily.com
- [5] Khalida, R., & Setiawati, S. (2020). Analisis Sentimen Sistem E-Tilang Menggunakan Algoritma Naive Bayes Dengan Optimalisasi Information Gain. *Journal of Information and Information Security (JIFORTY)*, 1(1). <http://ejurnal.ubharajaya.ac.id/index.php/jiforty>
- [6] Melyza, A., & Agus, R. M. (2021). SISWA TERHADAP PROSES PENERAPAN PEMBELAJARAN PENDIDIKAN JASMANI OLAHRAGA DAN KESEHATAN PADA PANDEMI COVID-19 DI SMA NEGERI 1 PADANG CERMIN. *Journal of Physical Education (JouPE)*, 2(1), 8–16.
- [7] Muchammad Shiddieqy Hadna, N., Insap Santosa, P., & Wahyu Winarno, W. (2016). STUDI LITERATUR TENTANG PERBANDINGAN METODE UNTUK PROSES ANALISIS SENTIMEN DI TWITTER. In *Seminar Nasional Teknologi Informasi dan Komunikasi*.
- [8] Nurjanah, W. E., Setya Perdana, R., & Fauzi, M. A. (2017). *Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet* (Vol. 1, Issue 12). <http://j-ptiik.ub.ac.id>
- [9] Sheren Yunanto, L. (2022). *ANALISIS PENGARUH KESADARAN HARGA DAN KUALITAS PRODUK TERHADAP KEPUASAN PELANGGAN DALAM MEMBENTUK LOYALITAS KONSUMEN PADA PELANGGAN SOMETHINC* (Vol. 10, Issue 1).
- [10] Wijaya, A. P., & Santoso, H. A. (2016). Naive Bayes Classification pada Klasifikasi Dokumen Untuk Identifikasi Konten E-Government Naive Bayes Classification on Document Classification to Identify E-Government Content. *Journal of Applied Intelligent System*, 1(1), 48–55.