

ANALISA SENTIMEN TERHADAP REVIEW PRODUK KECANTIKAN MENGUNAKAN METODE NAIVE BAYES CLASSIFIER

Dita Novita Sari¹, Firda Adelia², Fita Rosdiana³, Belasana Butar Butar⁴, Muhadi Hariyanto⁵

^{1,2,3,5} Program Studi Sistem Informatika, STMIK Nusa Mandiri

Jl. Ciledug Raya No.108, Cipulir, Kec. Kby. Lama. Kota Jakarta Selatan

⁴ Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Bina Sarana Informatika

Jl. Kramat Raya No.98, Senen, Jakarta Pusat.

Email: ditano11162296@nusamandiri.ac.id¹, firdaa11162250@nusamandiri.ac.id², fitaro11162403@nusamandiri.ac.id³,
belsana.bbb@bsi.ac.id⁴, muhadi.mho@nusamandiri.ac.id⁵

Abstract

Nowadays cosmetic products have become the main needs of women who are targets in terms of type and brand of a cosmetic. But not all cosmetics have good quality in accordance with the needs of consumers, therefore in order not to choose the wrong product, provided an opinion website that has tried the product, and so as not to take long to see reviews of various opinions on the website, therefore sentiment analysis is one solution to solve the problem to group opinions or cosmetic reviews by classifying the review into positive or negative opinions automatically. The review data to be tested are emoticons and adjectives that have been manually labeled. Using the Naïve Bayes method has the advantage of being simple, fast and has high accuracy. The resulting accuracy value will be the benchmark for finding the best tester model for sentiment classification cases. By using RapidMiner it is known that 110 negative data classified 85 data was declared negative as predicted and 25 data predicted positive but the result was negative, as well as 110 classified positive data 92 as predicted and 18 data predicted negative but the result was positive. Based on the calculation of accuracy value obtained by 80.45%.

Article history

Received Jul 28, 2020

Revised Okt 13, 2020

Accepted Nov 02, 2020

Available Online Nov 28, 2020

Keywords

Naïve Bayes, Review, Sentimen, Clasification, RapidMiner.

Abstrak

- Saat ini produk kosmetik sudah menjadi kebutuhan utama kaum wanita yang merupakan target dari segi jenis maupun brand sebuah kosmetik. Namun tidak semua kosmetik memiliki kualitas yang baik sesuai dengan kebutuhan konsumen, maka dari itu agar tidak salah memilih produk, disediakan website yang berbentuk opini yang sudah mencoba produk tersebut, dan agar tidak memakan waktu lama untuk melihat review dari berbagai opini yang ada di website, oleh karena itu analisis sentimen merupakan salah satu solusi mengatasi masalah untuk mengelompokkan opini atau review kosmetik dengan cara pengklasifikasian review tersebut menjadi opini positif atau negatif secara otomatis. Data review yang akan diuji merupakan emoticon dan kata sifat yang sudah diberi label secara manual. Dengan menggunakan metode Naïve Bayes memiliki kelebihan yaitu sederhana, cepat dan memiliki akurasi yang tinggi. Nilai akurasi yang dihasilkan akan menjadi tolak ukur untuk mencari model pengujian yang terbaik untuk kasus klasifikasi sentimen. Dengan menggunakan RapidMiner diketahui bahwa 110 data negatif terklasifikasi 85 data dinyatakan negatif sesuai prediksi dan 25 data diprediksi positif tetapi hasilnya negatif, begitu juga 110 data positif terklasifikasi 92 sesuai prediksi dan 18 data diprediksi negatif tetapi hasilnya positif. Berdasarkan hasil perhitungan nilai akurasi yang didapatkan sebesar 80.45%.

Riwayat

Diterima 28 Jul 2020

Revisi 13 Okt 2020

Disetujui 02 Nov 2020

Terbit 28 Nov 2020

Kata Kunci

Naïve Bayes, Review, Sentimen, Klasifikasi, RapidMiner.

PENDAHULUAN

Produk kosmetik yang beredar di pasaran saat ini sangat beragam, Namun tidak semua kosmetik memiliki kualitas yang baik sesuai kebutuhan konsumen dan hal ini yang harus diperhatikan oleh para konsumen. Sebelum konsumen memutuskan untuk membeli kosmetik sebaiknya konsumen mengetahui dengan detail produk yang akan dibeli, hal ini dapat dipelajari dari testimoni dan opini atau hasil review dari konsumen yang sudah membeli dan menggunakan produk tersebut sebelumnya.

Beberapa review tentang produk kosmetik dapat membantu konsumen dalam mengetahui kualitas brand kosmetik tersebut layak atau tidak untuk digunakan[1]. Banyaknya website untuk review menyebabkan ketersediaan data review sangat banyak. Saat ini konsumen yang menulis opini dan pengalaman secara online semakin meningkat.

Review produk dapat memberikan manfaat besar bagi konsumen karena dengan membaca review, konsumen mendapat informasi yaitu dapat mengetahui kualitas suatu produk dari review konsumen lain yang membagikan pengalamannya tentang produk terkait[2].

Di era saat ini sentimen dapat menjadi potensi besar bagi perusahaan industri kosmetik yang ingin mengetahui umpan balik dari masyarakat terhadap brand kosmetik mereka, selain itu masyarakat yang baru ingin mencoba berbagai produk kosmetik tersebut dapat melihat review produk mana yang baik dan tidak dari klasifikasi review produk kecantikan. Klasifikasi sentimen bertujuan untuk mengatasi masalah ini dengan secara otomatis mengelompokkan review pengguna menjadi opini positif atau negatif[1].

Ketersediaan data review inilah yang menjadi alasan penelitian ini menggunakan sumber data berupa review produk untuk Analisa Sentimen Terhadap Review Produk Kecantikan Menggunakan Metode Naïve Bayes Classifier dan tools yang digunakan adalah aplikasi RapidMiner.

LANDASAN TEORI

A. Review

Review adalah teks yang mengapresiasi suatu karya seni atau sastra. Tujuan review adalah untuk mengevaluasi suatu karya dan menyatakan kelebihan dan kekurangan dari suatu karya yang bertujuan untuk memberikan rekomendasi bagi pembaca atas karya tersebut[3]. Review produk kecantikan ini ditunggu-tunggu karena banyak perempuan-perempuan yang ingin membeli produk kecantikan, namun meragukan kualitasnya.

Menemukan bahwa review secara online sumber informasi paling penting bagi pelanggan untuk menghasilkan niat pembelian pelanggan[3].

B. Data Mining

Data Mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data[4].

Konsep data mining semakin dikenal sebagai tools penting dalam manajemen informasi karena jumlah informasi yang semakin besar jumlahnya[5].

C. Text Mining

Text Mining merupakan penerapan konsep dan teknik data mining untuk mencari pola dalam teks, yaitu proses penganalisis teks guna mencari informasi yang bermanfaat untuk tujuan tertentu[6].

Penambahan teks dapat dianggap sebagai proses dua tahap yang diawali dengan penerapan terstruktur terhadap sumber data teks, dan dilanjutkan dengan ekstraksi informasi dan pengetahuan yang relevan dari data terstruktur ini dengan menggunakan teknik dan alat yang sama dengan penambahan data.

D. Analisa Sentiment

Sentiment analysis merupakan sebuah metode yang digunakan untuk memahami, mengekstrak data opini, dan mengolah data tekstual secara otomatis untuk mendapatkan sebuah sentiment yang terkandung dalam sebuah opini[7].

E. Klasifikasi Naive Bayes

Klasifikasi Naive Bayes merupakan salah satu metode dari machine learning. Naive Bayes merupakan metode klasifikasi berdasarkan teorema Bayes yang banyak digunakan. Metode ini mengasumsikan bahwa pengaruh nilai atribut pada suatu kelas adalah independen terhadap nilai dari atribut-atribut lain[8]

F. Pre-Processing

Pre-processing adalah proses pembersihan dan mempersiapkan teks untuk klasifikasi[9]. Teks online biasanya mengandung banyak noise dan bagian tidak informatif pada tingkat kata-kata.

Pre-processing yang digunakan dalam penelitian ini adalah transform cases dan tokenizing. Transform Cases berfungsi mengubah karakter huruf besar menjadi kecil pada dokumen, sedangkan To-kenize berfungsi membuat token pada dokumen.

Adapun tujuan pada penelitian ini adalah untuk:

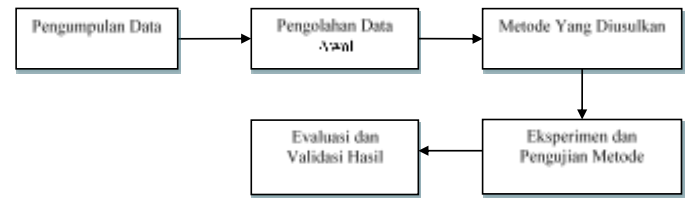
1. Memproleh gambaran mengenai opini atau review masyarakat tentang kosmetik yang baik dan tidak sebelum membelinya.
2. Dapat melakukan klasifikasi opini positif dan negatif terhadap review produk kecantikan yang ditujukan kepada femaledaily.com.
3. Untuk mengetahui akurasi klasifikasi menggunakan metode Naive Bayes

METODE PENELITIAN

Metode Penelitian secara umum sebagai kegiatan ilmiah yang dilakukan secara bertahap dimulai dengan penentuan topik, pengumpulan data dan menganalisis data, sehingga nantinya diperoleh suatu pemahaman dan pengertian atas topik, gejala atau isu tertentu[9].

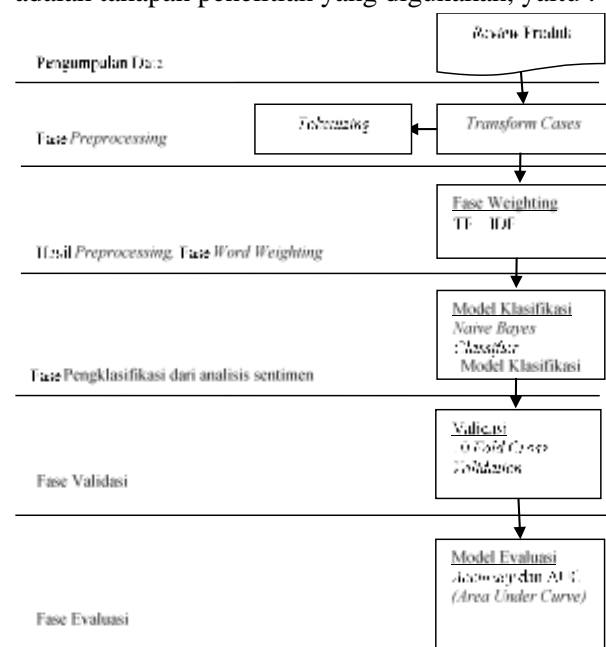
Classifier. Temuan Hasil Penelitian

Dalam bab ini penulis akan membahas tentang implementasi terhadap sistem yang sesuai dengan metode pengumpulan data, metode pengembangan sistem yang digunakan dan temuan hasil penelitiannya. Dalam mengembangkan sistem ini, penulis menggunakan metode pengembangan sistem prototype yang terdiri dari :



Gambar 1 Metode Penelitian[10]

Dalam penelitian ini analisis sentimen review produk kecantikan menggunakan pendekatan algoritma Naive Bayes Classifier. Review data yang telah dikumpulkan sebanyak 220 data dibagi menjadi 110 data positif dan 110 data negatif akan melalui tahapan pre-processing untuk mendapatkan kata-kata yang relevan untuk diklasifikasikan. Kemudian proses validasi akan dilakukan menggunakan 10 fold cross validation dengan evaluasi menggunakan accuracy dan AUC (area under curve). Berikut adalah tahapan penelitian yang digunakan, yaitu :



Gambar 2 Tahapan Penelitian[10]

A. Pengumpulan Data

Data yang didapat kemudian dipisah menjadi 2 kategori yaitu positif dan negatif yang disimpan dalam folder. Berikut merupakan contoh review kata sentimen yang telah dikategorikan:

Tabel 1 : Review Sentimen

Review Sentimen	
Positif	Negatif
satu satunya aloe vera yang bekerja sangat baik dikulitku	jerawatan gede cuyyy besokannya! Dan kalau dipakai disekitar hidung rada perih
pertama kali nyoba aloe vera ini dan udh sekitar sebulan lebih makenya dikulit aku	ternyata zonk bgt di muka aku wkwk bikin large pores keliatannnn dan jerawat banyak di pipi waktu aku jadiin sleeping mask baunya juga gaenak
aloe vera gel ini juga wanginya enak, beneran menenangkan wajah bgt	Baunya kimia banget dan teksturnya lebih cair dari gel yg seharusnya
ini sangat aman, menenangkan, melembabkan, gak ada reaksi negatif	jerawat kecil 2 muncul, dan 1 gede jg ikutan nimbrung
make di muka dan di diemin semalem beruntusan jadi lebih reda	bikin BERUNTUSAN BANGETTT dan teksturnya sangat lengket, abis make ini muka kaya berminyak kerlap kerlip dan lama nyerepnya

B. Fase Pre – Processing

Pada fase ini data akan diproses dalam 2 tahapan pre-processing yaitu tokenizing berfungsi untuk menghilangkan tanda baca atau simbol[11]. Transform Cases berfungsi untuk mengubah huruf kapital (upper case) menjadi huruf kecil (lower case) seluruhnya[12].

C. Fase Word Weighting

Hasil dari fase pre-processing nantinya akan melalui word weighting atau pembobotan kata menggunakan TF – IDF.

D. Fase Pengklasifikasi

Pada fase ini kata yang sudah melalui preprocessing dan sudah diberikan pembobotan nilai selanjutnya akan melalui proses klasifikasi dengan metode Naive

Bayes Classifier. Adapun langkah NBC sebagai berikut[13] :

- Langkah pertama yaitu melakukan pelatihan terhadap dokumen agar sesuai kategorinya atau data training.
- Langkah kedua yaitu proses klasifikasi dokumen untuk mengetahui keakuratan dokumen atau data testing.
- Langkah ketiga yaitu mencari nilai probabilitas bagi setiap kategori berdasarkan dokumen. Berikut rumusnya :

$$P(c | x) = \frac{p(x | c)p(c)}{p(x)}$$

Keterangan :

- X : Data dengan class yang belum diketahui
- C : Hipotesis data merupakan suatu class spesifik
- P(c|x) : Probabilitas hipotesis berdasarkan kondisi
- P(c) : Probabilitas hipotesis
- P(x|c) : Probabilitas berdasarkan kondisi pada hipotesis
- P(x) : Probabilitas c

E. Fase Validasi

Pada fase ini data akan dibagi menjadi 10 bagian secara acak dengan menggunakan formasi dari model yang nantinya akan dibuat dalam RapidMiner dan dari model tersebut nanti akan menghasilkan akurasi yang diperoleh dari hasil perhitungan dengan melihat seberapa banyak data yang telah berhasil diklasifikasi dengan benar.

F. Fase Evaluasi

Pada fase ini hasil dari pengujian model validasi menggunakan cross validation serta evaluasi akan menggunakan accuracy dan AUC (area under curve).

HASIL DAN PEMBAHASAN

Data yang digunakan dalam pengklasifikasian teks berjumlah 220 review yang dibagi menjadi 110 review dengan sentimen negatif dan 110 review dengan sentimen positif dari website <https://femaledaily.com>.

Data review sentimen disatukan ke dalam folder dengan nama negatif dan positif. Setiap dokumen memiliki format.txt yang dapat dibuka menggunakan notepad.



Gambar 3 Data Review Sentimen

A. Pre-Processing Teks Review

a) Tokenizing

Pada tahap tokenize ini akan menghilangkan atau menghapus angka, tanda baca dan simbol dalam data sentimen.

Tabel 2 : Perbandingan Sebelum dan Sesudah Tokenize

Sebelum	Sesudah
Satu satunya aloe vera yang bekerja sangat baik dikulitku	Satu satunya aloe vera yang bekerja sangat baik dikulitku
Lembab tapi udahnya langsung jerawat dan kusam	Lembab tapi udahnya langsung jerawat dan kusam
Aloe vera gel yang cocok dimuka, akhirnya menemukan aloe vera gel ini yang aman dimuka	Aloe vera gel yang cocok dimuka, akhirnya menemukan aloe vera gel ini yang aman dimuka
Aku ga suka sama aloe vera yg ini, karena lengket terus bau alcohol	Aku ga suka sama aloe vera yg ini karena lengket terus bau alcohol
Hasilnya memuaskan. Bruntusan aku dibagian pipi dan jidat hilang. Sukaaa banget. Dan yang utama adalah harganya murah banget	Hasilnya memuaskan. Bruntusan aku dibagian pipi dan jidat hilang. Sukaaa banget. Dan yang utama adalah harganya murah banget

Karena baunya bau kimia banget.	Karena baunya bau kimia banget
Ini sangat aman, melembabkan, gak ada reaksi negatif.	Ini sangat aman, melembabkan, gak ada reaksi negatif
Setelah aku beli beberapa hari dia malah bikin kulit aku makin kering.	Setelah aku beli beberapa hari dia malah bikin kulit aku makin kering

b) Transform Cases

Pada tahap ini transform cases akan mengubah huruf besar menjadi huruf kecil.

Tabel 3 : Perbandingan Sebelum dan Sesudah Transform Cases

Sebelum	Sesudah
S atu satunya aloe vera yang bekerja sangat baik dikulitku	s atu satunya aloe vera yang bekerja sangat baik dikulitku
L embab tapi udahnya langsung jerawat dan kusam	l embab tapi udahnya langsung jerawat dan kusam
A loe vera gel yang cocok dimuka, akhirnya menemukan aloe vera gel ini yang aman dimuka	a loe vera gel yang cocok dimuka, akhirnya menemukan aloe vera gel ini yang aman dimuka
A ku ga suka sama aloe vera yg ini, karena lengket terus bau alcohol	a ku ga suka sama aloe vera yg ini, karena lengket terus bau alcohol
H asilnya memuaskan. Bruntusan aku dibagian pipi dan jidat hilang. Sukaaa banget.	h asilnya memuaskan. bruntusan aku dibagian pipi dan jidat. s ukaaa banget.
B runtsuan aku dibagian pipi dan jidat hilang. Sukaaa banget.	memuaskan b runtsuan aku dibagian pipi dan jidat. d an yang utama adalah

D an yang utama adalah harganya murah harganya murah banget	harganya murah banget
K arena baunya bau kimia banget.	k arena baunya bau kimia banget.
S etelah aku beli beberapa hari dia malah bikin kulit aku makin kering.	s etelah aku beli beberapa hari dia malah bikin kulit aku makin kering.
I ni sangat aman, melembabkan, gak ada reaksi negatif.	i ni sangat aman melembabkan gak ada reaksi negatif.

RowNo	label	text	metabola_0	metabola_1	metabola_2	tf	idf	tfidf
123	positif	in premium	11M	D:ISTAK No.	Nov 28, 2019	0	0	0
124	positif	Diagn sajer	191M	D:ISTAK No.	Jun 24, 2020	0	0.338	0
125	positif	panjang	12M	D:ISTAK No.	Nov 28, 2019	0	0	0
126	positif	Siapa khal	13M	D:ISTAK No.	Nov 28, 2019	0	0	0.282
127	positif	Siapa khal	14M	D:ISTAK No.	Nov 28, 2019	0	0	0
128	positif	Pakanya k	15M	D:ISTAK No.	Nov 28, 2019	0	0	0
129	positif	Siapa khal	16M	D:ISTAK No.	Nov 28, 2019	0	0	0
130	positif	panjang	17M	D:ISTAK No.	Nov 28, 2019	0	0.196	0
131	positif	panjang	18M	D:ISTAK No.	Nov 28, 2019	0	0	0
132	positif	Siapa khal	19M	D:ISTAK No.	Nov 28, 2019	0	0	0
133	positif	Siapa khal	20M	D:ISTAK No.	Nov 28, 2019	0	0	0
134	positif	Siapa khal	21M	D:ISTAK No.	Nov 28, 2019	0	0	0
135	positif	Siapa khal	22M	D:ISTAK No.	Nov 28, 2019	0	0	0

A. Gambar 4 Hasil dari Bobot Nilai Menggunakan Algoritma TF-IDF.

B. Hasil Word Weighting

Word weighting atau pemberian bobot nilai. Berikut hasil dari perhitungan proses pemberian bobot yang menggunakan algoritma TF-IDF.

Dari hasil preprocessing teks terdapat kata umum yang mengandung makna diantara, baik, bau, cocok, lengket, kimia, kering, menenangkan, lembab, murah, aman. Langkah awal awal untuk menentukan frekuensi kemunculan kata pada suatu dokumen yaitu tentukan term frequency (tf) seperti dilihat pada tabel berikut:

Tabel 4 : Term Frequency

Term	Term Frequency (tf)							
	D1		D2		D3		D4	
	Positif	Negatif	Positif	Negatif	Positif	Negatif	Positif	Negatif
Baik	1	0	0	0	0	0	0	0
Bau	0	0	0	1	0	0	0	1
Cocok	0	0	0	0	1	0	0	0
Lengket	0	0	0	1	0	0	0	1
Kimia	0	0	0	1	0	0	0	0
Kering	0	0	0	0	0	0	0	1
Memuaskan	1	0	0	0	0	0	0	0
Lembab	0	0	0	1	1	0	0	0
Murah	1	0	0	0	0	0	0	0
Aman	0	0	0	0	2	0	0	0

Setelah menentukan term frequency (tf) pada dokumen. Selanjutnya menentukan inverse document frequency (idf) atau banyak dokumen dari suatu term (t) muncul. Dapat dilihat pada table berikut :

Tabel 5 : Inverse Document Frequency

Term	Inverse Document Frequency	
	Positif	Negatif
Baik	1	0
Bau	0	2
Cocok	1	0
Lengket	0	1
Kimia	0	1
Kering	0	1
Memuaskan	1	0
Lembab	1	1
Murah	1	0
Aman	2	0

Tabel tersebut merupakan kemunculan kata dari dokumen teks pada kelas positif dan negatif pada sentiment produk. Selanjutnya mencari opini tertinggi pada dokumen. Diketahui jumlah document frequency (df) yaitu 13 yang dimana komentar positif 7 dan negatif 6.

C. Klasifikasi Manual Naïve Bayes Classifier

Tahapan Naïve Bayes dengan memproses klasifikasi teks untuk menentukan probabilitas kelas opini tertinggi. Berikut tahapan :

- a) Tahap pertama mengetahui kategori teks yaitu baik, bau, cocok, lengket, kimia, kering, suka, lembab murah, dan aman.
- b) Tahap kedua terdapat pada table berikut :

Tabel 6 : Tahap Klasifikasi

Teks	Klasifikasi		
	Positif	Negatif	
Baik	1/7	0/6	1/13
Bau	0/7	2/6	2/13
Cocok	1/7	0/6	1/13
Lengket	0/7	1/6	1/13
Kimia	0/7	1/6	1/13
Kering	0/7	1/6	1/13
Memuaskan	1/7	0/6	1/13

Lembab	1/7	1/6	1/13
Murah	1/7	0/6	1/13
Aman	2/7	0/6	1/13
	P(positif) = 7/13	P(negatif) = 6/13	

Dari hasil klasifikasi manual pada dokumen diatas, selanjutnya dengan mencari nilai probabilitas yang tinggi pada kelas positif dan negatif.

- c) Dan tahap ketiga dengan mencari nilai probabilitas tertinggi. Sebagai berikut :

$$P(c|x)=P(\text{positif|baik})= ((1/7*7/13))/((1/13)) = (0,142*0,538)/0,076 = 1,005$$

$$P(c|x)=P(\text{positif|cocok})= ((1/7*7/13))/((1/13)) = (0,142*0,538)/0,076 = 1,005$$

$$P(c|x)=P(\text{positif|memuaskan})= ((1/7*7/13))/((1/13)) = (0,142*0,538)/0,076 = 1,005$$

$$P(c|x)=P(\text{positif|lembab})= ((1/7*7/13))/((1/13)) = (0,142*0,538)/0,076 = 1,005$$

$$P(c|x)=P(\text{positif|murah})= ((1/7*7/13))/((1/13)) = (0,142*0,538)/0,076 = 1,005$$

$$P(c|x)=P(\text{positif|aman})= ((2/7*7/13))/((2/13)) = (0,285*0,538)/0,153 = 1,002$$

$$P(|\text{positif})=1,005*1,005*1,005*1,005*1,005*1002 = 1,027$$

$$\text{Probabilitas} = P(\text{positif})*P(P|\text{positif})=0,538*1,027= 0,552$$

$$P(c|x)=P(\text{negatif|bau})= ((2/6*6/13))/((2/13)) = (0,333*0,461)/0,153= 1.003$$

$$P(c|x)=P(\text{negatif|lengket})= ((1/6*6/13))/((1/13)) = (0,166*0,461)/0,076 = 1.006$$

$$P(c|x)=P(\text{negatif|kimia})= ((1/6*6/13))/((1/13)) = (0,166*0,461)/0,076 = 1.006$$

$$P(c|x)=P(\text{negatif|kering})= ((1/6*6/13))/((1/13)) = (0,166*0,461)/0,076 = 1.006$$

$$P(|\text{negatif})=1,003*1,006*1,006*1,006 = 1,021$$

$$\text{Probabilitas} = P(\text{negatif})*P(P|\text{negatif})=0,461*1,021 = 0,470$$

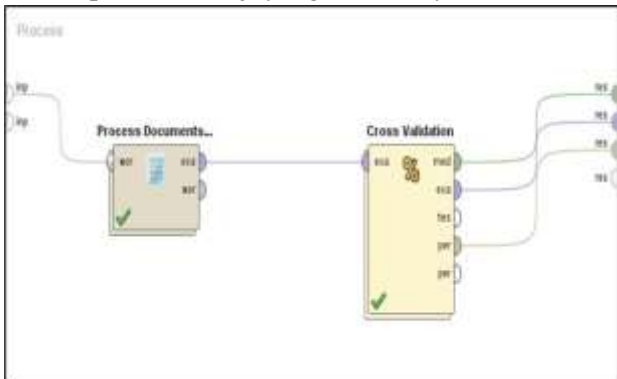
Hasil perhitungan probabilitas pada kategori kelas positif dan negatif pada dokumen tersebut. Diperoleh probabilitas tertinggi pada kelas positif dengan jumlah 0,552. Dari hasil perhitungan probabilitas tersebut

dapat dilihat bahwa review produk mendapatkan komentar atau opini positif.

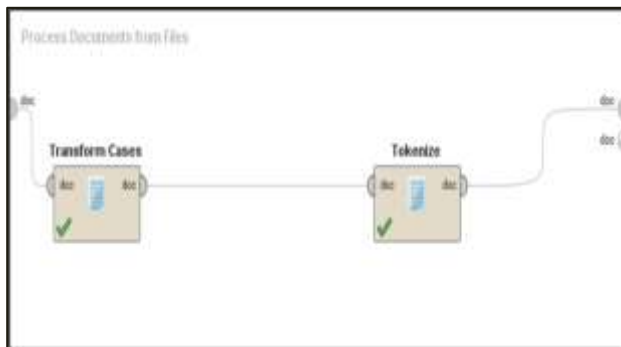
D. Klasifikasi Sentiment Naïve Bayes Classifier dengan Rapid Miner

Klasifikasi sentiment ini menggunakan metode Naïve Bayes Classifier dengan software RapidMiner 9.3 sebagai gambar berikut :

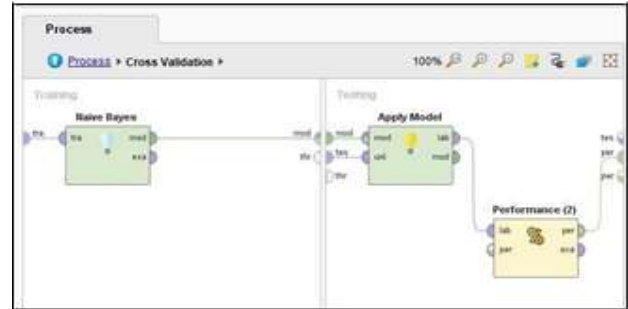
Cross Validation memiliki dua subproses, subproses pelatihan dan subproses pengujian. Subproses Pelatihan digunakan untuk melatih model. Model yang terlatih kemudian diterapkan dalam subproses Pengujian. Kinerja model diukur selama fase Pengujian. Cross Validation akan melakukan pengulangan data sebanyak 10 kali sehingga menampilkan data uji yang sebenarnya.



Gambar 5 Model Klasifikasi Naïve Bayes Classifier dengan RapidMiner 9.3



Gambar 6 Tahap Pre-Processing



Gambar 7 Desain Model Klasifikasi Naïve Bayes Classifier Menggunakan RapidMiner 9.3

E. Hasil Akurasi

Hasil dari klasifikasi berupa accuracy dan AUC, dengan metode Naïve Bayes Classifier.

accuracy: 80.45% +/- 10.29% (micro average: 80.45%)			
	true negatif	true positif	class precision
pred negatif	85	18	82.52%
pred positif	25	92	78.63%
class recall	77.27%	83.64%	

Gambar 8 Hasil Akurasi Naïve Bayes Classifier.

Berdasarkan gambar IV.5 menunjukkan bahwa akurasi sebesar 80.45% dari 110 data negatif terklasifikasi 85 data dinyatakan negatif sesuai prediksi. Dan 25 data diprediksi positif tetapi hasilnya negatif, begitu dengan 110 data positif terklasifikasi 92 sesuai prediksi dan 18 data diprediksi negatif tetapi hasilnya positif.

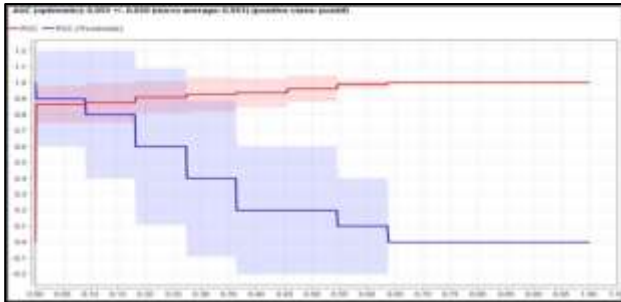
B. Hasil akurasi dapat dihitung kembali dengan rumus manual untuk mencocokkan hasil dari RapidMiner dengan perhitungan :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{92+85}{92+85+18+25} = \frac{177}{220} = 0,8045$$

- C. Keterangan :
- D. TP : True Positif
- E. TN : True Negatif
- F. FP : False Positif
- G. FN : False Negatif

Berdasarkan hasil perhitungan manual menggunakan rumus tersebut, yang didapatkan sebesar 0.8045 jika di jadikan persen maka 80.45% yang dihasilkan dengan RapidMiner dan Perhitungan rumus manual bernilai sama.

Dengan ini perhitungan divisualisasikan dengan AUC (Area Under Curve) yang dimana terdapat kesimpulan analisis sentiment kelas positif. Dilihat pada Gambar berikut :



Gambar 9 Kurva AUC

Dari gambar 8 terdapat grafik ROC dengan nilai AUC (Area Under Curve) sebesar 0.951. Hal tersebut yang dimana hasilnya excellent classification.

Tabel 7 : Perbandingan Klasifikasi Nilai AUC

0.90 - 1.00	<i>Excellent Classification</i>
0.80 - 0.90	<i>Good Classification</i>
0.70 - 0.80	<i>Fair Classification</i>
0.60 - 0.70	<i>Poor Classification</i>
0.50 - 0.60	<i>Failure</i>

KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan pada penelitian ini bahwa analisis sentiment pada review produk kecantikan menggunakan Algoritma Naïve Bayes Classifier sangat efektif untuk digunakan sebagai proses klasifikasi website yang dibutuhkan dalam sistem analisis sentiment ini yang menunjukkan bahwa 110 data negatif terklasifikasi 85 data dinyatakan negatif sesuai prediksi. Dan 25 data diprediksi positif tetapi hasilnya negatif, begitu

dengan 110 data positif terklasifikasi 92 sesuai prediksi dan 18 data diprediksi negatif tetapi hasilnya positif. Berdasarkan hasil perhitungan manual menggunakan rumus tersebut, nilai akurasi yang didapatkan sebesar 0.8045 jika di jadikan persen maka 80.45%.

berhasil dibangun dan dapat berfungsi sesuai tujuan yaitu menyisipkan pesan rahasia ke dalam gambar. Sistem ini juga telah berhasil mengembalikan pesan rahasia dengan menggunakan kunci yang sama sewaktu melakukan enkripsi.

REFERENSI

- D. A. Kristiyanti, “Analisis Sentimen Review Produk Kosmetik Menggunakan Algoritma Support Vector Machine Dan Particle Swarm Optimization Sebagai,” Semin. Nas. Inov. Tren 2015 “Peluang dan Tantangan Indones. Dalam Menyikapi Afta 2015,” pp. 134–141, 2015.
- J. Speed and S. P. Engineering, “Journal Speed – Sentra Penelitian Engineering dan Edukasi – Volume 10 No 4 – November - 2018,” vol. 10, no. 4, pp. 92–97, 2018.
- L. Arista and H. Lasmana, “Pengaruh Review Oleh Sarah Ayu Pada Produk Kecantikan Di Youtube Dan Brand Awareness Terhadap Keputusan Menggunakan Produk,” *Scriptura*, vol. 9, no. 1, pp. 26–34, 2019.
- D. Aprilia, D. Aji Baskoro, L. Ambarwati, and I. W. S. Wicaksana, “Belajar Data Mining Dengan Rapid Minner,” p. 139, 2013.
- J. Oliver, *Hilos Tensados*, vol. 1, no., pp. 1–476, 2019.
- Fatmawati, “2013,” pp. 27–42, 2009.
- E. M. Sipayung, H. Maharani, I. Zefanya, and D. S. Informasi, “No Title,” vol. 8, no. 1, pp. 958–965, 2016.
- B. Jeremy, Albert; Christanti, Viny; Mulyawan, “Opinion Mining Untuk Ulasan Produk Dengan,” pp. 9–16, 2003.
- C. R. Semiawan, “Metode Penelitian Kualitatif: Jenis, Karakteristik dan Keunggulannya.”
- N. Normah, “Naïve Bayes Algorithm For Sentiment Analysis Windows Phone Store Application Reviews,” *SinkrOn*, vol. 3, no.

2, p. 13, 2019.

- I. Lingga, *Pemodelan Deteksi Body Shaming Di Media Sosial Twitter Menggunakan Algoritma Naïve Bayes*, vol. 8, no. 2. 2019.
- E. Sihite, R. D. Ramadhani, M. Zidny, and R. Adhitama, "Text Processing Clustering dalam Menentukan Profesi Berdasarkan Data Twitter," pp. 103–108, 2018.
- A. Pandhu and H. Agus, "Naive Bayes Classification pada Klasifikasi Dokumen Untuk Identifikasi Konten E-Government," *J. Appl. Intell. Syst.*, vol. 1, no. 1, pp. 48-55–55, 2016.