

Opini Media Sosial *Facebook* Terhadap Produk Hijab Menggunakan Metode Text Mining

Agus Heryanto ^{1,*}, Rully Pramudita ²

¹ Teknik Informatika; Sekolah Tinggi Teknologi Bandung; Jl. Soekarno-Hatta No.378, Kb. Lega, Kec. Bojongloa Kidul, Kota Bandung, (022) 5224000; e-mail: agusheryanto@sttb.ac.id, agus.heryanto.1992@gmail.com.

² Manajemen Informatika; Universitas Bina Insani; Jl. Siliwangi No. 6 Rawa Panjang, Bekasi Barat, telp 02188958130; e-mail: rullypramudita@binainsani.ac.id

* Korespondensi: e-mail: agus.heryanto.1992@gmail.com

Diterima: 20 April 2020; Review: 06 Mei 2020; Disetujui: 15 Mei 2020

Cara sitasi: Heryanto A, Pramudita R. 2020. Opini Media Sosial *Facebook* Terhadap Produk Hijab Menggunakan Metode *Text Mining*. *Information System for Educators and Professionals*. 4 (2): 168 – 177.

Abstrak: Analisis opini bertujuan untuk melihat opini masyarakat atau kelompok mengenai entitas tertentu. Analisis sentimen digunakan untuk menilai sebuah produk, apakah memberikan review yang baik dan disukai masyarakat ataupun sebaliknya. Penelitian dilakukan untuk mengklasifikasikan opini atau sentimen masyarakat terhadap produk hijab. Hijab yang sekarang menjadi trend dikalangan remaja muslim, sebagian besar remaja muslim merupakan para pengguna media sosial terutama *facebook*. Sehingga pandangan para pengguna hijab ini akan tertuang dalam sebuah opini, baik itu opini positif maupun negatif. Oleh karena itu, penelitian ini mengambil opini masyarakat terhadap produk hijab Zoya dan El-zatta melalui dua *fanspage facebook* dari Zoya Lovers dan El-zatta Hijab. Metode penghitungan dan pengklasifikasian yang dipakai yaitu dengan menggunakan algoritma *naive bayes*. Hasil penelitian menunjukkan bahwa sentiment positif terbentuk paling banyak yaitu 432 komentar pada Zoya, dan 395 komentar pada EL-zatta. Jika dibandingkan dengan sentimen negatif yang hanya 103 komentar pada El-zatta dan 63 komentar pada Zoya. Jadi secara keseluruhan, persepsi masyarakat terhadap dua produk hijab yang tertuang pada komentar di *fanspage facebook* memberikan penilaian yang positif.

Kata kunci: analisis opini, *facebook*, hijab, *naive bayes*, *text mining*.

Abstract: Opinion analysis aims to see the opinions of the public or group regarding certain entities. Sentiment analysis is used to assess a product, whether to give a good review and accepted by the public or the opposite. Studies are conducted to classify public opinion or sentiment toward hijab products. The current hijab is a trend among muslim youths, most of muslim teenagers are social media users, mainly facebook. That the opinion of the hijab users will be featured in an opinion, whether positive or negative. Therefore, the study takes public opinion on the products of hijab Zoya and El-zatta through two facebook fans page from Zoya Lovers and El-zatta Hijab. The counter-intelligence and classification methods were adopted using a naive boyes algorithm. Research indicates that the most positive sentiment was 432 comments on Zoya and 395 comments on El-zatta. Compared to the negative sentiment which was just 103 comments on El-zatta and 63 comments on Zoya. So on the whole, the public's conception of two hijab products, featured on comments of facebook fans page has a positive assessment.

Keywords: *facebook*, hijab, *naive bayes*, *sentiment analysis*.

1. Pendahuluan

Alat bertukar informasi dan komunikasi modern saat ini yang berkembang pesat di masyarakat adalah media sosial. Bahkan sosial media ini bagian hidup masyarakat menengah ke atas. Sehingga media sosial, menjadi wadah dalam berkomunikasi yang menyediakan ruang interaksi tanpa batas tempat dan waktu. Oleh karena itu, informasi yang didapat pun akan selalu terbaru dan berasal dari manapun yang memang itu dibutuhkan oleh masyarakat. Disisi lain, remaja perempuan di Indonesia yang sebagian besarnya beragama islam sekarang ini sudah banyak yang berhijrah. Ini ditandai dengan banyaknya perempuan muslimah yang berhijab. Bahkan sekarang hijab sudah menjadi *trend* masa kini yang desain dan tampilannya banyak diminati oleh para remaja. Dimana sebagian besar remaja itu adalah pengguna media sosial termasuk *facebook*. Beberapa produsen hijab diindonesia memanfaatkan sosial media untuk memperkenalkan dan memasarkan produknya kepada masyarakat, penelitian dilakukan terhadap dua produk hijab yaitu : Zoya dan El-Zatta.

Banyaknya pertukaran informasi di media sosial *facebook* baik itu postingan berupa status, komentar ataupun percakapan dan yang lainnya, cenderung bersifat jujur dan spontan diungkapkan. Hal tersebut, dapat dimanfaatkan untuk mengetahui tingkat popularitas suatu produk, kepuasan pelanggan dari sebuah produk, dan apa yang pelanggan harapkan dari sebuah produk. Itu semua menunjukkan persepsi masyarakat dari sebuah produk. Oleh karena itu, pengusaha harus pandai menentukan segmen mana yang akan dimasuki serta juga harus pandai menetapkan pasar sasarannya sehingga tepat[1]. Untuk itu perlu suatu analisis opini ataupun opini masyarakat terhadap produk hijab, yang dalam penelitian ini produk hijabnya adalah Zoya dan El-zatta. Analisis opini yang dihasilkan dari persepsi masyarakat, bisa opini positif ataupun negatif

2. Metode Penelitian

Analisis Sentimen

Analisis sentimen merupakan suatu bidang interdisipliner ilmu, yang melakukan pemecahan masalah secara relevan dan terpadu dengan tinjauan dari berbagai sudut pandang ilmu yang serumpun. Analisis sentimen terdiri dari pemrosesan bahasa alami, analisis teks dan komputasi linguistik untuk mengidentifikasi sentiment atau opini dari sebuah data atau dokumen[2]. Nama lain dari Analisis Sentimen adalah *Opinion Mining* yang merupakan bidang dalam *Natural Language Processing (NLP)* yang membangun sistem yang mencoba mengidentifikasi dan mengekstrak pendapat dalam teks. Saat ini, analisis sentimen adalah topik yang banyak diteliti dan berkembang sangat pesat, sebab memiliki banyak *tools* untuk pengolahan datanya[3].

Dalam penelitian mengikuti beberapa teknik penelitian sebelumnya dengan menggunakan kata – kata bermakna sentimen sebagai penanda sentimen pada komentar dari data yang diambil. Namun dalam penelitian ini tidak memasukan *emoticon* sebagai penanda sentimen, hanya kata – kata yang bermakna sentimen berbahasa indonesia. Berikut ini beberapa daftar kata yang digunakan sebagai kata kunci (*Keyword*).

Tabel 1. *Keyword* Sentimen

| Kata Positif | Kata Negatif |
|---|---|
| bagus, cantik, ok, mau, berapa, pengen, ingin, suka, makasih, keren, kapan, diskon, minta, warna, beli, turun, harga, terima kasih, pesen, cocok, borong, alhamdulillah | jelek, enggak, kurang, kosong, kehabisan, habis, tutup, jangan, cuma, abis, jauh, bohong, hadeh, haduh, |

Sumber: Hasil Penelitian (2020)

Text Mining

Text mining merupakan proses *mining* atau menambang suatu informasi dari data yang tersaji dalam jumlah besar, dalam hal ini adalah teks. Proses ini dilakukan dalam rangka penggalian, pengolahan, serta pengaturan pada informasi dengan menganalisa keterkaitan antara informasi satu dengan yang lainnya[4]. Proses *text mining* ini sendiri merupakan hal yang berdasarkan kombinasi yang baik serta kata, dan frase yang sesuai, serta melibatkan proses ekstraksi fitur yang menjadi kunci efektif utama dalam tahap pelatihan atau yang sering disebut dengan *training*[5].

Pre-Processing

Tahap pre-processing atau tahap persiapan data adalah proses untuk mempersiapkan data awal sebelum dilakukan proses – proses yang lain. Pada umumnya, tahap persiapan data ini dilakukan dengan cara mengeliminasi data yang tidak sesuai atau mengubah data menjadi bentuk yang lebih mudah diproses oleh sistem. Tahap persiapan data ini sangat penting dalam proses analisis sentimen, terutama pada kalimat yang tidak formal dan tidak terstruktur serta memiliki noise yang besar pada data dari media sosial [6].

Tokenizing

Proses yang dilakukan pada *tokenizing* merupakan awal yang sangat penting, salah satunya adalah *text to lowercase* [7]. Proses selanjutnya adalah penguraian, proses penguraian yang dimaksud adalah membagi teks menjadi kumpulan kata tanpa memperhatikan keterhubungan diantara kata satu dengan yang lain serta peran dan posisinya pada kalimat, karakter diterima dalam kumpulan kata menurut abjad[8]

Stopword Filtering

Proses selanjutnya adalah memeriksa stop word list, stopword list adalah daftar kata-kata yang semestinya dihilangkan, jika kata pada dataset terdapat pada stop word list maka kata akan dihilangkan. Tetapi jika tidak terdapat di dalamnya maka proses akan berlanjut tanpa menghilangkan kata pada dokumen [7].

Data *stopword* pada penelitian ini diambil dari *static.hikaruyuuki.com* yang datanya berjumlah 758 kata. Serta diambil juga dari beberapa penelitian sebelumnya. Dibawah ini beberapa contoh *stopword* yang dimasukkan.

Tabel 2 *Data Stopword*

| No | Kata Stopword |
|----|---------------|
| 1 | ada |
| 2 | dari |
| 3 | dan |
| 4 | di |
| 5 | dulu |
| 6 | ini |
| 7 | karena |
| 8 | ke |
| 9 | yang |

Sumber: Hasil Penelitian (2020)

Stemming

Dalam teks berbahasa Indonesia pada proses stemming akan berbeda dengan dengan teks yang berbahasa Inggris. Jika pada teks yang berbahasa Inggris, hanya diperlukan proses untuk menghilangkan sufiks. Sedangkan pada teks yang berbahasa Indonesia, selain sufiks yang dihilangkan, ada prefiks dan konfiks yang juga perlu untuk dihilangkan[9].

Transform Case

Merupakan proses yang dilakukan guna merubah seluruh karakter dengan huruf kapital menjadi huruf standar guna mendukung klasifikasi hal tersebut dilakukan karena komentar sebagian besar merupakan kata-kata sentimen yang mengandung huruf kapital[10] .

Metode Naïve Bayes

Metode Naïve Bayes adalah metode klasifikasi yang digunakan untuk menentukan kelompok kelas. Metode ini dapat mengelola data dalam jumlah besar dengan hasil akurasi yang tinggi[11].

Teorema Bayes merupakan metode pendekatan statistik pada persoalan klasifikasi untuk melakukan inferensi induksi. Hal pertama yang dibahas adalah tentang konsep dasar dan definisi pada Teorema Bayes, Lalu menggunakan teorema bayes ini untuk melakukan proses klasifikasi dalam Data Mining[12]. Teorema Bayes memiliki bentuk umum seperti pada rumus 1:

$$P(H|X) = \frac{P(X|H) P(H)}{P(X)} \dots\dots\dots(1)$$

Keterangan :

- X = Data dengan class yang belum diketahui
- H = Hipotesis data X merupakan suatu class spesifik
- P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X
- P(H) = Probabilitas hipotesis H
- P(X|H) = Probabilitas X berdasarkan kondisi tersebut
- P(X) = Probabilitas dari X

3. Hasil dan Pembahasan

Pada pembahasan penelitian ini yaitu melakukan pengolahan data komentar dari *Fanspage* Zoya dan El-Zatta yang dilakukan secara urut sesuai dengan urutan metode *naïve bayes* sehingga diperoleh sentimen atau opini positif maupun negatif.

Persiapan Data

Sumber data yang akan diolah pada penelitian ini berasal dari *fanspage* Zoya Lovers dan *Fanspage* El-Zatta Hijab. Data berasal dari komentar – komentar terkait produk dan pelayanan dari pengunjung *fanspage* yang diambil dengan menggunakan aplikasi yang disediakan oleh *facebook* yaitu *netvizz*. Aplikasi *netvizz* merupakan aplikasi yang disediakan oleh *facebook* untuk mengambil data *facebook* secara gratis. Data yang dapat diambil bermacam – macam mulai dari *group data*, *page data*, *page like network*, *search dan link status*. Namun yang akan diambil pada penelitian kali ini adalah hanya data yang ada atau data yang tercakup di *page data*. Sesuai dengan studi kasus, organisasi yang akan diteliti adalah Zoya dan El-Zatta, maka *fanspage* yang diambil adalah *fanspage* dari Zoya dengan nama *fanspage* Zoya Lovers dan untuk El-Zatta yaitu *fanspage* El-Zatta Hijab. Data yang diambil dan yang diolah adalah data komentar dari kedua *fanspage* tersebut yang berbahasa indonesia. Data komentar dianggap sebagai ungkapan dari pengguna ataupun pelanggan yang langsung disampaikan, sehingga data komentar bisa dianggap sebagai sebuah sentimen ataupun opini publik. Opini yang didapat tersebut akan dipisahkan menjadi opini yang negatif dan opini yang positif. Opini yang negatif dan positif ditandai dengan kata – kata yang tertuang pada komentar para pelanggan yang memberikan komentarnya.

Tahap Data Training

Penelitian ini mengikuti beberapa teknik penelitian sebelumnya dengan menggunakan kata – kata bermakna sentimen sebagai penanda sentimen pada komentar dari data yang diambil. Namun, dalam penelitian ini tidak memasukan *emoticon* sebagai penanda sentimen, hanya kata – kata yang bermakna sentimen berbahasa indonesia. Kata – kata yang bermakna *sentimen* tersebut dibuatkan *data training*. *Data training* yang digunakan sebesar 1000 data. Data yang diambil tersebut mengandung data sentimen atau opini baik negatif sebanyak 500 data dan yang opini positif pun 500 data.

Untuk kebutuhan *training*, data yang berhasil dikumpulkan tersebut akan dikategorikan secara manual dan menilai sentimen yang terkandung dalam komentar tersebut kedalam dua kategori yaitu, positif dan negatif. Pada tabel 2 adalah contoh komentar yang dikategorikan sebagai sentimen positif atau negatif.

Tabel 3. *Data Training*

| No | Teks Komentar | <i>Fanspage</i> | Sentimen |
|----|--|-----------------|----------|
| 1 | Warna nya cantik cantik jadi pengen beli lagi dan lagi | Zoya | Positif |
| 2 | Saya cari warna merah susah | Zoya | Negatif |
| 3 | Bagus banget warna nya | El Zatta | Positif |
| 4 | Tapi sayang ngga ada yang polos nya | El Zatta | Negatif |

Sumber: Hasil Penelitian (2020)

Tahap Pre-Processing

Dalam tahapan *pre-processing* dibagi kedalam beberapa tahap lagi, yaitu tahap *tokenizing*, penghilangan *stopwords* dan *stemming*. Sehingga dari tahap tersebut akan dilanjutkan menjadi tahap *transformation*.

Tokenizing

Selanjutnya adalah *tokenizing*, dalam tahap ini dari data yang berupa komentar yang berbentuk kalimat kemudian dilakukan penghilangan delimeter dan penguraian kalimat. Misalkan contoh kalimat seperti ini: “oke.. saya minta tolong sis, bisa gak di upload gambar motif - motif terbarunya..”

Maka kalimat akan menjadi:

Oke
Saya
Minta
Tolong
Sis
Bisa
Gak
Di
Upload
Gambar
Motif
Terbarunya

Stopword Removal

Langkah – langkah untuk proses *Stopword Removal*: 1) Kata - kata penyusun kalimat hasil dari tahap *text pre-processing* dijadikan sebagai masukan; 2) Kemudian dibandingkan dengan kata - kata yang ada di *database stopwords*; 3) Jika kata yang dimasukkan sama dengan kata di *database stopwords* maka kata yang dimasukkan akan dihapus. Namun jika kata yang di masukkan tidak sama dengan kata yang ada di *database stopwords* maka tersebut tidak dihapus.

Maka Data dari setelah proses tokenizing dan sekarang dibandingkan dengan stopwords menjadi:

Oke
Saya
Minta
Tolong
Sis
~~Bisa~~
~~Gak~~
Di
Upload
Gambar
Motif
Terbarunya

Stemming

Langkah- langkah proses *stemming* adalah: Langkah pertama, kata yang belum di-*stemming* dibandingkan ke dalam database kamus kata dasar. Jika di temukan, maka kata tersebut diasumsikan sebagai kata dasar dan algoritma berhenti. Jika kata tidak sesuai dengan kata dalam kamus. Langkah kedua, Jika kata di - input memiliki pasangan awalan- akhiran “be - lah”, “be - an”, “me - i”, “di - i”, “pe - i”, atau “te - i” maka langkah *stemming* selanjutnya adalah 5, 3, 4, 5, 6, tetapi jika kata yang di- input tidak memiliki pasangan awalan- akhiran tersebut, langkah *stemming* berjalan normal yaitu 3, 4, 5, 6, 7. Langkah ketiga hilangkan partikel dan kata ganti kepunyaan. Pertama hilangkan partikel (“ - lah”, “ - kah”, “ - tah”, “ - pun”). Setelah itu hilangkan juga kata ganti kepunyaan (“ - ku”, “ - mu”, atau “ - nya”). Contoh : kata “komputerkulah”, proses *stemming* pertama menjadi “komputerku” dan proses *stemming* selanjutnya menjadi “komputer”. Apabila kata “komputer” ada di dalam kamus maka algoritma pun akan berhenti. Sesuai dengan model imbuhan, menjadi [[[AW+]AW+]AW+] Kata Dasar [+AK];. Langkah keempat hilangkan juga akhiran (“ - i”, “ - an”, dan “ - kan”), sesuai dengan model imbuhan, maka menjadi [[[AW+]AW+]AW+] kata dasar. Langkah kelima penghilangan awalan (“be - “, “di - “, “ke - “, “me - “, “pe - “, “se - “, dan “te - “) mengikuti langkah - langkah a) Algoritma akan berhenti jika: Awalan diidentifikasi bentuk sepasang imbuhan yang tidak

diperbolehkan dengan akhiran (berdasarkan tabel 2.1) yang dihapus pada langkah ketiga, diidentifikasi awalan yang sekarang identik dengan awalan yang telah dihapus sebelumnya atau, kata tersebut sudah tidak memiliki awalan; b) Identifikasi jenis awalan dan peluruhannya bila diperlukan jenis awalan ditentukan dengan aturan : Jika awalan dari kata adalah “di- “, “ke - “, atau “se- “ maka awalan dapat langsung dihilangkan, hapus awalan “te- “, “be - “, “me- “, atau “pe- “ yang menggunakan aturan peluruhan yang dijelaskan sebelumnya. Sebagai contoh kata “menulis”, setelah menghilangkan awalan “me - “ maka kata yang didapat adalah “nulis”. Karena kata “nulis” tidak ditemukan dalam database kata dasar maka karakter “n” diganti dengan karakter “t” sehingga dihasilkan kata “tulis” dan kata “tulis” merupakan kata yang sesuai dengan kata yang terdapat di database, sehingga algoritma pun akan berhenti. Langkah keenam jika semua langkah gagal, maka kata yang diuji pada algoritma ini dianggap sebagai kata dasar.

Penggunaan Naïve Bayes

Penghitungan dan klasifikasi *Naïve Bayes Classifier* dengan menampilkan enam komentar yang ada pada *fanspage Zoya Lovers* berupa hasil dari proses - proses *text mining*. Empat dokumen kometar adalah data *training*, sedangkan satu dokumen komentar akan digunakan sebagai dokumen *testing* yang akan diklasifikasi.

Tabel 4. Tabel Hasil Klasifikasi

| Dokumen | Isi Komentar | Kategori |
|---------|--|----------|
| 1 | Bagus banget warna nya | Positif |
| 2 | Mau donk, berapa harga nya | Positif |
| 3 | cantik - cantik motif nya saya suka | Positif |
| 4 | tapi sayang yang turun harga cuma segiempat saja...yang lain tidak | Negatif |

Sumber: Hasil Penelitian (2020)

Berdasarkan datatersebut, terdapat 24 kata yang unik {bagus, banget, warna, nya, mau, donk, berapa, harga, cantik, motif, saya, suka, tapi, sayang, yang, turun, cuma, segiempat, saja, lain, tidak, ngga, ada, polos}. Setiap kata yang ada pada *data training* dipetakan ke dalam tabel frekuensi seperti tabel 5, untuk di hitung jumlah frekuensi kemunculan dari setiap kata.

Tabel 5. Frekuensi Kemunculan Data

| | Dokumen | | | | |
|-----------|---------|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Bagus | 1 | | | | |
| Banget | 1 | | | | |
| Warna | 1 | | | | |
| Nya | 1 | 1 | 1 | | 1 |
| Mau | | 1 | | | |
| Donk | | 1 | | | |
| Berapa | | 1 | | | |
| Harga | | 1 | | | |
| Cantik | | | 2 | | |
| Motif | | | 1 | | |
| Saya | | | 1 | | |
| Suka | | | 1 | | |
| Tapi | | | | 1 | 1 |
| Sayang | | | | 1 | 1 |
| Yang | | | | 2 | |
| Turun | | | | 1 | |
| Cuma | | | | 1 | |
| Segiempat | | | | 1 | |
| Saja | | | | 1 | |
| Lain | | | | 1 | |
| Tidak | | | | 1 | |
| Ngga | | | | | 1 |
| Ada | | | | | 1 |

| | Dokumen | | | | |
|----------|---------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| Polos | | | | | 1 |
| Kategori | Positif | Positif | Positif | Negatif | Negatif |

Sumber: Hasil Penelitian (2020)

Hasil pengolahan berdasarkan tabel 5 dapat diketahui bahwa: Jumlah frekuensi n positif = 15; Jumlah frekuensi n negatif = 18; Jumlah kata ($|kosakata|$) = 24. Berdasarkan nilai yang telah diolah dapat dicari probabilitas setiap kata $P(X_i|V_j)$ dengan menggunakan rumus (1).

Pada Sentimen Positif :

$$\begin{aligned}
 P(\text{bagus}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{warna}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{nya}|\text{positif}) &= (3+1)/(15+24) = 0,1026 \\
 P(\text{mau}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{donk}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{berapa}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{harga}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{cantik}|\text{positif}) &= (2+1)/(15+24) = 0,0769 \\
 P(\text{motif}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{saya}|\text{positif}) &= (1+1)/(15+24) = 0,0513 \\
 P(\text{suka}|\text{positif}) &= (1+1)/(15+24) = 0,0513
 \end{aligned}$$

Pada Sentimen Negatif

$$\begin{aligned}
 P(\text{tapi}|\text{negatif}) &= (2+1)/(18+24) = 0,0714 \\
 P(\text{sayang}|\text{negatif}) &= (2+1)/(18+24) = 0,0714 \\
 P(\text{yang}|\text{negatif}) &= (2+1)/(18+24) = 0,0714 \\
 P(\text{turun}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{cuma}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{segiempat}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{saja}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{lain}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{tidak}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{ngga}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{ada}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{polos}|\text{negatif}) &= (1+1)/(18+24) = 0,0476 \\
 P(\text{nya}|\text{negatif}) &= (1+1)/(18+24) = 0,0476
 \end{aligned}$$

Pada komentar yang akan di klasifikasi adalah : “saya suka warna nya”

Diketahui :

Nilai $P(V_j)$ untuk dokumen positif = $3/5 = 0,6$

Nilai $P(V_j)$ untuk dokumen negatif = $2/5 = 0,4$

Jika $V_j =$ positif :

$$\begin{aligned}
 &P(\text{saya}|\text{positif}) * P(\text{suka}|\text{positif}) * P(\text{warna}|\text{positif}) * P(\text{nya}|\text{positif}) * P(\text{positif}) \\
 V_{\text{map}} &= (1+1)/(15+24) * (1+1)/(15+24) * (1+1)/(15+24) * (1+1)/(15+24) * 0,6 \\
 &= 4,1554E-6
 \end{aligned}$$

Jika $V_j =$ negatif :

$$\begin{aligned}
 &P(\text{saya}|\text{negatif}) * P(\text{suka}|\text{negatif}) * P(\text{warna}|\text{negatif}) * P(\text{nya}|\text{negatif}) * P(\text{negatif}) \\
 V_{\text{map}} &= (1+1)/(18+24) * (0+1)/(18+24) * (0+1)/(18+24) * (1+1)/(18+24) * 0,4 \\
 &= 0
 \end{aligned}$$

Pada hasil perhitungan tersebut didapat bahwa nilai V_{map} untuk kategori sentimen positif memiliki nilai lebih tinggi bahkan mutlak dibandingkan kategori sentimen negatif. Sehingga dapat disimpulkan bahwa contoh data/komentar tersebut adalah data/komentar yang memiliki nilai sentimen positif.

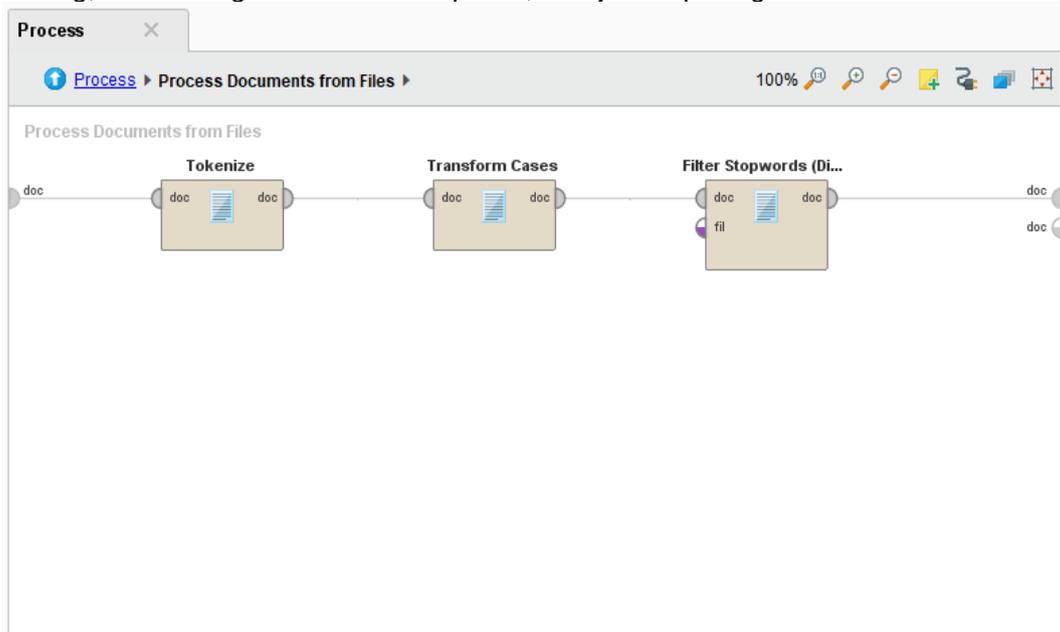
Pengujian Menggunakan Rapid Miner

Data yang diambil sebanyak 500 data komentar dari kedua *fanspage*, selama kurun waktu dari 1 Januari 2020 sampai dengan 31 Maret 2020. Namun data tersebut menyusut

ketika melalui proses *training*. Data training yang digunakan sebanyak 500 data training positif dan 500 data training negatif.

Untuk melakukan penelitian dari data yang sudah dapatkan, baik data training dan data testing berupa komentar dari *fanspage facebook* menggunakan aplikasi RapidMiner 7.6.1. Pertama adalah proses *training* dengan menggunakan 1000 data atau komentar, yang telah diklasifikasikan menjadi 500 data positif dan 500 data negatif. Selanjutnya data dimasukkan kedalam aplikasi RapidMiner.

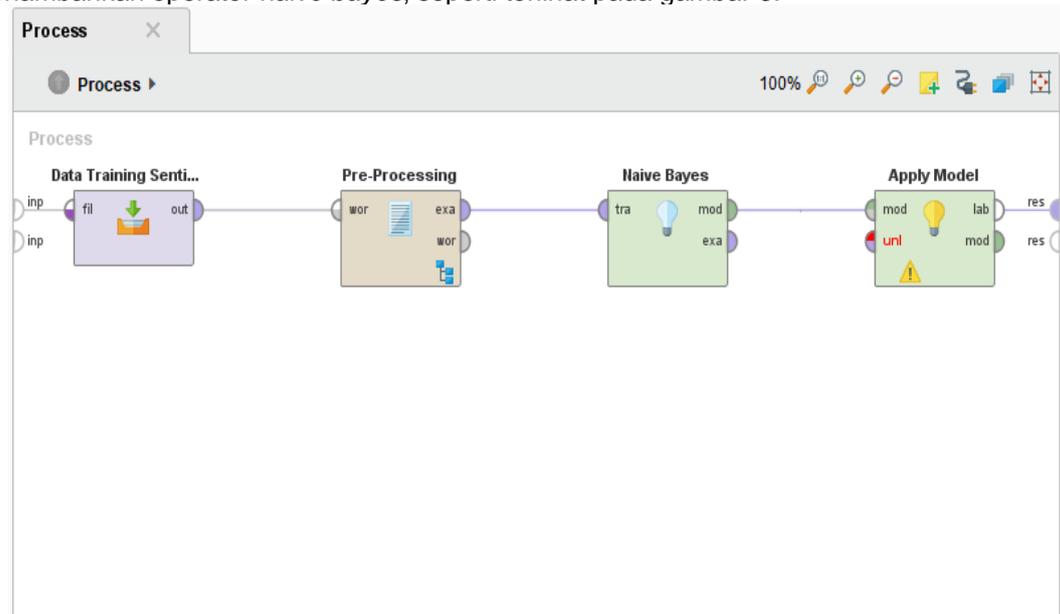
Setelah data training dimasukkan selanjutnya melakukan *Pre-processing* yaitu *tokenizing*, *case folding* dan *removed stopwords*, ditunjukkan pada gambar 2.



Sumber: Hasil Penelitian (2020)

Gambar 1. Tahap Pre-Processing

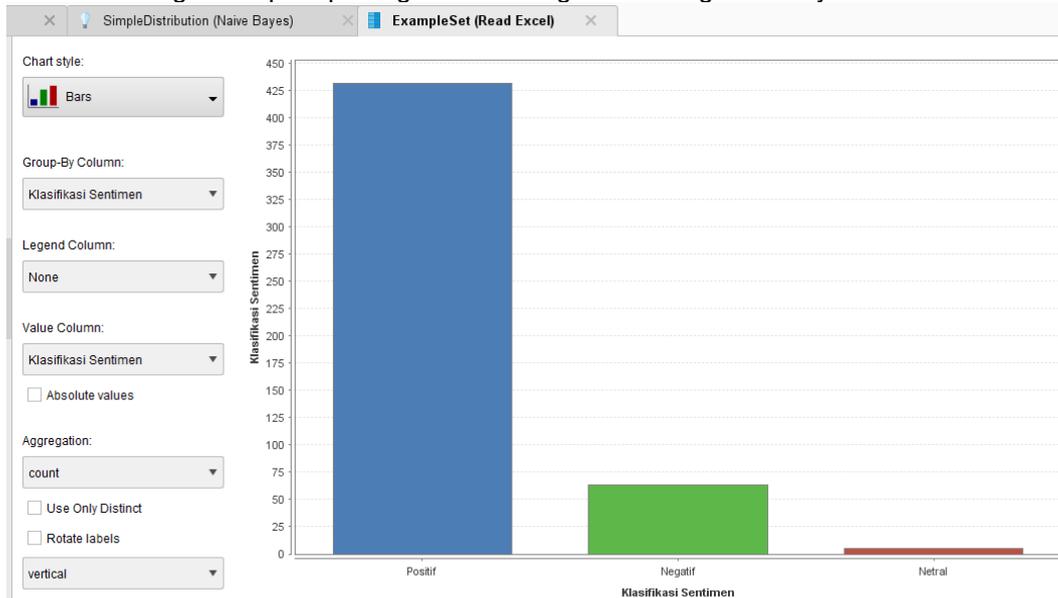
Hasil dari tahap *pre-processing* selesai, langkah selanjutnya yaitu merapkan algoritma *naive bayes* sebagai metode untuk melakukan analisis sentimen/opini terhadap data yang sudah dilakukan *pre-processing*. Pada aplikasi RapidMiner kita akan melakukannya dengan menambahkan operator *naive bayes*, seperti terlihat pada gambar 3.



Sumber: Hasil Penelitian (2020)

Gambar 2. Naïve Bayes pada RapidMiner

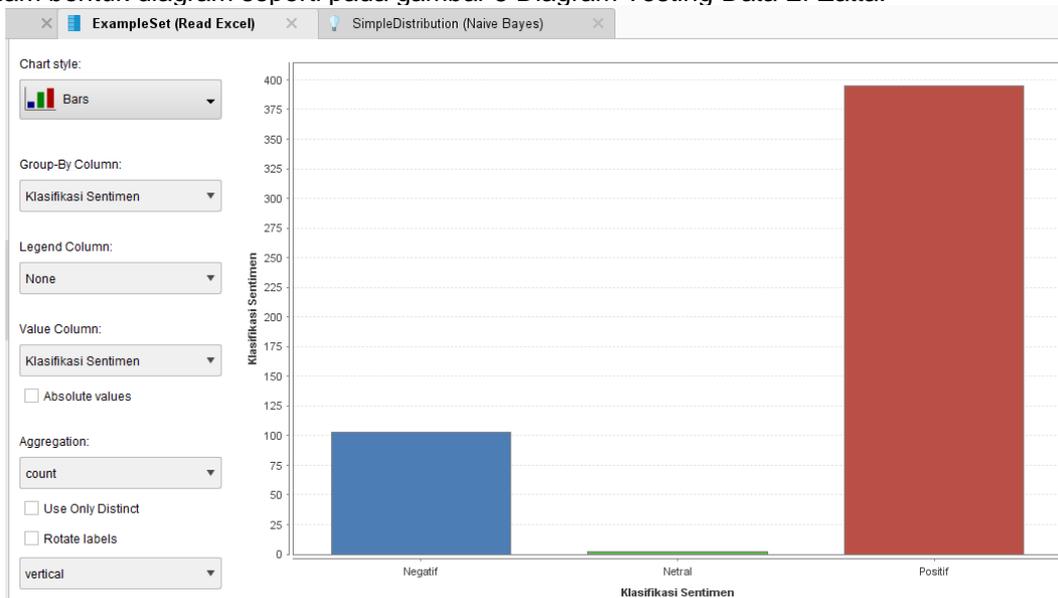
Hasil dari analisis sentimen terhadap 500 data komentar pada *fanspage* Zoya Lovers adalah 432 berupa sentimen positif, 63 berupa sentimen negatif dan 5 komentar netral. Terlihat dalam bentuk diagram seperti pada gambar 4 Diagram Testing Data Zoya.



Sumber: Hasil Penelitian (2020)

Gambar 3. Hasil Testing Zoya

Data komentar satu lagi yang akan kita testing adalah data komentar dari *fanspage* El-Zatta Hijab yaitu berupa 500 komentar dengan menggunakan aplikasi RapidMiner. Terlihat dalam bentuk diagram seperti pada gambar 5 Diagram Testing Data El-Zatta.



Sumber: Hasil Penelitian (2020)

Gambar 4. Hasil Testing El-Zatta

Hasil dari analisis sentimen terhadap 500 data komentar pada *fanspage* El-Zatta Hijab adalah 395 berupa sentimen positif, 103 berupa sentimen negatif dan 2 komentar netral.

4. Kesimpulan

Metode *Naive Bayes Classifier* dapat diterapkan sebagai metode untuk melakukan klasifikasi *sentiment analysis*. Analisis memberikan hasil respon atau pandangan masyarakat terhadap hijab tergolong sangat baik. Untuk pembagiannya sendiri, pada *fanspage* Zoya Lovers

dengan klasifikasi komentar yang masuk kedalam sentimen positif adalah sebanyak 432 komentar, sentimen negatif sebanyak 63 dan kelas netral sebanyak 5. Sedangkan untuk *fanspage* El-zatta Hijab dengan klasifikasi komentar pada sentimen positif sebanyak 395 komentar, sentimen negatif sebanyak 103 dan kelas netral sebanyak 2 komentar. Jadi secara keseluruhan, persepsi masyarakat terhadap dua produk hijab ini yang tertuang pada komentar di *fanspage facebook* memberikan penilaian yang positif.

Referensi

- [1] D. A. Rahmawati and K. Indriani, "Association Rule Mining Dengan Algoritma Apriori Untuk Penjualan Produk Verkol Lubricantes," *Informatics Educ. Prof.*, vol. 2, no. 2, pp. 105–114, 2018.
- [2] G. Vinodhini and R. M. Chandrasekaran, "A comparative performance evaluation of neural network based approach for sentiment classification of online reviews," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 28, no. 1, pp. 2–12, 2016, doi: 10.1016/j.jksuci.2014.03.024.
- [3] A. Of *et al.*, "Analisis Sentimen Terhadap Pemilihan Gubernur Provinsi Sulawesi Selatan Tahun 2018 Menggunakan Data Twitter Dan Metode Sentiment Lexicon Analysis of Sentiment Against the Election of the Governor of the Province of South Sulawesi , the Year 2018 Using the," vol. 6, no. 1, pp. 1928–1934, 2019.
- [4] A. V. Sudiantoro and E. Zuliarso, "Analisis Sentimen Twitter Menggunakan Text Mining Dengan Algoritma NAÏVE BAYES CLASSIFIER," *Pros. SINTAK 2018*, pp. 398–401, 2018.
- [5] F. Handayani and S. Pribadi, "Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110," *J. Tek. Elektro*, vol. 7, no. 1, pp. 19–24, 2015.
- [6] S. Mujilawati, "Pre-Processing Text Mining Pada Data Twitter," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 2089–9815, 2016.
- [7] 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.
- [8] A. Pandhu and W. Diki, "Analisa sentimen dan Klasifikasi Komentar Positif Pada Twitter dengan Naive Bayes Classification Sentiment Analysis and Classification of Positive Comments on Twitter with Naive Bayes Classification," vol. 1, no. 2, 2020.
- [9] H. Pramudita, "Penerapan Algoritma Stemming Nazief & Adriani Dan Similarity Pada Penerimaan Judul Thesis," *Data Manaj. dan Teknol. Inf.*, vol. 15, no. 4, p. 15, 2014.
- [10] A. T. Jaka, "Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining," *Inform. UPGRIS*, vol. 1, pp. 1–9, 2015.
- [11] A. M. Pudjajana and D. Manongga, "Sentimen Analisis Tweet Pornografi Kaum Homoseksual Indonesia Di Twitter Dengan Naive Bayes," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 9, no. 1, pp. 313–318, 2018, doi: 10.24176/simet.v9i1.1922.
- [12] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.