

ANALISIS SENTIMEN PADA ULASAN PRODUK MASKER DI MARKETPLACE SHOPEE MENGGUNAKAN *SUPPORT VECTOR MACHINE* DAN SELEKSI FITUR *CHI SQUARE*

Willy Winata¹, Ati Zaidiah², Nurul Chamidah³

S1 Informatika / Fakultas Ilmu Komputer

Universitas Pembangunan Nasional Veteran Jakarta

Jl. R.S Fatmawati No. 1 Pondok Labu Jakarta Selatan 12450

willyw@upnvj.ac.id¹, atizaidiah@upnvj.ac.id², nurul.chamidah@upnvj.ac.id³

Abstrak. Masker menjadi hal wajib saat ingin keluar rumah yang membuat penjualan masker meningkat drastis. Namun dengan adanya pembatasan kegiatan tersebut membuat aktivitas jual beli secara konvensional menjadi terhambat. Oleh karena itu banyak masyarakat beralih dari berbelanja langsung menjadi berbelanja *online*. *Marketplace* Shopee merupakan salah satu *platform* belanja *online*. Penjualan masker di *marketplace* Shopee mengalami peningkatan yang sangat drastis yang membuat variasi penilaian ulasan produk yang berbeda-beda dari pembeli. Pada penelitian dilakukan analisis sentimen terhadap 2 kelas positif dan negatif berdasarkan *rating* ulasan. Algoritma klasifikasi yang digunakan adalah *Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF) serta menerapkan seleksi fitur *Chi Square*. Tingkat performa yang didapat menggunakan seleksi fitur *Chi Square* dengan nilai taraf nyata (α) 0.3 menghasilkan nilai akurasi sebesar 91,32%, presisi sebesar 93,28%, *recall* sebesar 93,14% dan *f1-score* 93,21%. Dibanding dengan tingkat performa tanpa menggunakan seleksi fitur hanya menghasilkan nilai akurasi sebesar 90,75%, presisi sebesar 92,08%, *recall* sebesar 93,59% dan *f1-score* 92,83%.

Kata Kunci: Masker, Analisis Sentimen, *Marketplace*, Shopee, Ulasan Produk, *Support Vector Machine*, *Chi Square*

1 Pendahuluan

Di era pandemi ini, banyak upaya dari pemerintah dalam menangani pandemi tersebut. Salah satunya dengan Pemberlakuan Pembatasan Kegiatan Masyarakat atau biasa di kenal dengan PPKM. Dengan adanya pembatasan tersebut masyarakat dihibau untuk melakukan aktivitas terbatas diluar rumah guna menekan lonjakan kasus COVID-19. Selain itu pemerintah menerapkan protokol kesehatan untuk mencegah meningkatnya virus COVID-19. Protokol kesehatan yang dimaksud adalah 5M seperti mencuci tangan, menjaga jarak, menjauhi kerumunan, mengurangi mobilitas, dan menggunakan masker.

Masker merupakan hal yang penting pada pandemi saat ini. Penggunaan masker menjadi salah satu kewajiban saat ingin keluar rumah. Oleh karena itu penjualan masker meningkat drastis. Namun dengan adanya PPKM, aktivitas jual beli secara konvensional mengalami hambatan karena pembatasan kegiatan tersebut. Masyarakat sekarang ini lebih banyak melakukan aktivitas jual beli secara *online*. Salah satu tempat untuk berbelanja secara *online* yang banyak digunakan yaitu pada *marketplace* Shopee. Penjualan produk masker pada *marketplace* Shopee mengalami peningkatan yang sangat tinggi. Dengan banyaknya masker yang terjual di *marketplace* Shopee menimbulkan variasi penilaian ulasan produk yang berbeda-beda dari pembeli.

Pemberian ulasan pada produk masker menjadi salah satu indikator calon pembeli dalam memilih produk masker yang sesuai dan mengetahui kekurangan atau kelebihan dari produk tersebut. Ulasan pada produk juga berpengaruh bagi penjual dalam meningkatkan kualitas penjualannya di masa yang akan datang [1]. Namun peningkatan jumlah ulasan suatu produk membuat calon pembeli kesulitan untuk memahami ulasan suatu produk dan akhirnya tidak dapat menarik kesimpulan yang tepat [2]. Oleh sebab itu dibutuhkan analisis sentimen yang dapat membantu calon pembeli dalam mengambil keputusan berdasarkan ulasan atau opini dari pembeli lainnya.

Analisis sentimen merupakan bidang studi yang menganalisis opini, sentimen, evaluasi, sikap dan emosi dari seseorang terhadap suatu entitas seperti produk, jasa, individu, topik dan atribut dari suatu entitas [3]. Analisis

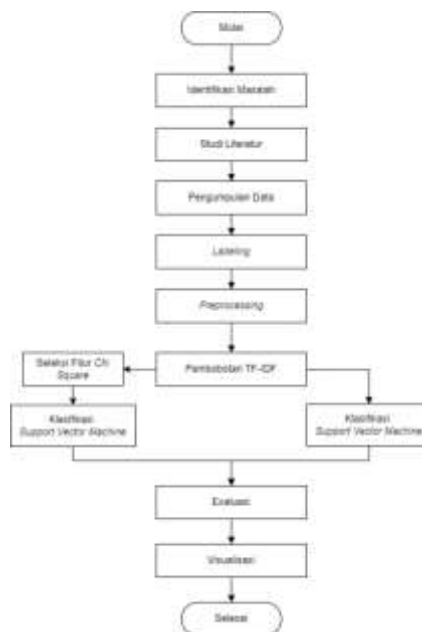
sentimen dilakukan berdasarkan ulasan produk masker di *marketplace* Shopee yang digunakan untuk melihat pendapat atau opini pembeli lain terhadap produk tersebut apakah cenderung beropini negatif atau positif. Ulasan produk masker terdiri dari opini dalam bentuk teks dan penilaian dalam bentuk bintang dari 1 sampai 5.

Berdasarkan uraian diatas, maka peneliti melakukan penelitian tentang analisis sentimen pada ulasan produk masker yang terdapat pada *marketplace* Shopee dengan menggunakan metode *Support Vector Machine* (SVM) dan seleksi fitur *Chi Square* yang dimana dengan menerapkan seleksi fitur diharapkan dapat meningkatkan akurasi klasifikasi nantinya. Tujuan dari penelitian ini adalah menganalisis sentimen terhadap opini pembeli pada produk masker di *marketplace* Shopee.

2 Metodologi Penelitian

2.1 Kerangka Pikir

Kerangka pikir ini menunjukkan gambaran kerangka penelitian yang dilakukan tahap demi tahap dan digunakan sebagai pedoman untuk mempermudah pelaksanaan penelitian. Alur penelitian dilakukan berdasarkan tahapan yang terdapat pada Gambar (1).



Gambar 1. Alur Penelitian

2.2 Pengumpulan Data

Pengumpulan data dilakukan dengan cara *web scraping* dari *website* Shopee. Data diambil dari toko *online* *atmos.official* dengan nama produk “Masker KF94 Korea 4 ply impor *High Quality*”. Alur proses pengambilan data terdapat pada Gambar (2).



Gambar 2. Alur Proses *Web Scraping*

Alur proses *web scraping* dimulai dengan membuka *link* URL dimana akan di *looping* untuk dilakukan proses *scraping* pengambilan data ulasan produk. pada halaman produk dilakukan pengambilan data ulasan produk berdasarkan ulasan dengan komentar.

2.3 Labeling

Proses pemberian label dilakukan berdasarkan jumlah bintang yang diberikan pada produk. Label yang akan digunakan berupa 2 kelas yaitu kelas negatif dan positif untuk setiap ulasan. Menurut [4] proses pelabelan dapat dilihat berdasarkan *rating* ulasan, dimana ulasan yang memiliki *rating* 1 dan 2 termasuk kedalam kelas negatif sedangkan untuk *rating* 4 dan 5 termasuk kedalam kelas positif dan untuk *rating* 3 kebanyakan dipakai untuk kelas netral. Namun untuk kasus ini karena data yang didapat banyak maka ulasan dengan *rating* 3 dihapus.

2.4 Preprocessing

Preprocessing atau Pra proses merupakan tahapan dimana data yang masih mentah akan dijadikan kedalam format data yang konsisten dan mudah diolah nantinya untuk dilakukan klasifikasi sentimen. Adapun tahapan pada pra proses dapat dilihat pada Gambar (3).



Gambar 3. Alur *Preprocessing*

1. Case Folding

Case folding merupakan tahapan yang dilakukan pada data teks dimana huruf yang tidak konsisten besar dan kecilnya kemudian diubah menjadi lowercase, contohnya “Barangnya Bagus Sekali” menjadi “barangnya bagus sekali”.

2. Cleansing

Cleansing memiliki beberapa tahapan diantaranya data teks yang bukan merupakan alfabet akan dilakukan penghapusan seperti *symbol* (@#~!), angka, *unicode*, dll. yang tidak memiliki arti pada proses klasifikasi. Tahapan berikutnya jika terdapat baris baru maka akan diubah menjadi satu baris saja dan tahapan terakhir adalah mengubah kata yang berduplikat berturut-turut menjadi tidak duplikat, contohnya “sukaaaaa” menjadi “suka”.

3. Stopword Removal

Stopword Removal merupakan tahapan dimana kata yang sering muncul namun tidak memiliki makna akan dihapus untuk meningkatkan proses klasifikasi. contohnya kata “ini, dan, ke, di, yang, atau, dll”. Untuk proses ini menggunakan *library* sastrawi.

4. Stemming

Stemming merupakan tahapan dimana kata yang berimbuhan diubah menjadi kata dasar. contohnya kata “terburuk” menjadi “buruk”. Untuk proses ini menggunakan *library* sastrawi.

5. Normalization

Normalization atau normalisasi merupakan tahapan dimana kata yang sebelumnya tidak baku akan diubah menjadi kata baku, contohnya “jaman” menjadi “zaman” dan mengubah kata yang *typo* ataupun kata gaul menjadi kata baku yang baik dan benar, contohnya “cpt” menjadi “cepat”.

6. Tokenization

Tokenization merupakan tahapan dimana suatu teks kalimat dipecah menjadi bagian-bagian kata atau token. Contohnya kalimat “kualitas masker sangat bagus” menghasilkan 4 kata/token yaitu “kualitas”, “masker”, “sangat”, dan “bagus”.

2.5 Pembobotan TF-IDF

Term Frequency – Inverse Document Frequency atau dikenal dengan TF-IDF adalah teknik gabungan dari TF dan IDF untuk memberikan bobot dari setiap term [5]. TF-IDF biasa digunakan pada pemrosesan bahasa alami atau teks yang sangat baik dalam menentukan *term* penting dari suatu teks. *Term Frequency* dimana suatu kata akan dihitung kemunculannya didalam sebuah dokumen. Jika kemunculan suatu kata semakin sering maka kata tersebut memiliki bobot nilai yang besar. Adapun persamaan dari *Term Frequency* dapat dilihat pada rumus persamaan (1).

$$W_{tf_{t,d}} = \begin{cases} 1 + \log_{10}(tf_{t,d}) , & tf_{t,d} > 0 \\ 0 & , \text{ otherwise} \end{cases} \quad (1)$$

Inverse Document Frequency merupakan proses untuk mengukur seberapa penting kata didalam suatu dokumen [6]. Adapun persamaan dari *Inverse Document Frequency* dapat dilihat pada rumus persamaan (2).

$$idf_t = \log_{10} \frac{N}{df_t} \quad (2)$$

Oleh karena itu persamaan dari *Term Frequency – Inverse Document Frequency* dapat dilihat pada rumus persamaan (3).

$$W_{t,d} = W_{tf_{t,d}} \times idf_t \quad (3)$$

Keterangan:

- $W_{tf_{t,d}}$ = bobot kata pada dokumen
- $tf_{t,d}$ = jumlah kemunculan kata dalam dokumen
- df_t = jumlah dokumen yang mengandung kata
- idf_t = bobot inverse dalam nilai df_t
- N = jumlah keseluruhan dokumen
- $W_{t,d}$ = Pembobotan TF-IDF

2.6 Chi Square

Chi Square (χ^2) merupakan salah satu seleksi fitur yang biasa digunakan pada pengklasifikasi teks dimana seleksi fitur ini berguna dalam mereduksi fitur yang tidak relevan. Seleksi fitur dengan *Chi Square* memiliki keunggulan dalam meningkatkan keakuratan seperti dapat mereduksi dengan data yang besar dan tidak mengurangi akurasi sebelumnya. Proses perhitungan yang dilakukan pada seleksi fitur *Chi Square* ini menggunakan teori statistika dalam pengujiannya dan mengevaluasi antar kata apakah saling berkaitan atau tidak. Adapun persamaan dari *Chi Square* dapat dilihat pada rumus persamaan (4).

$$\chi^2(t, c) = \frac{N(AD - CB)^2}{(A + C)(B + D)(A + B)(C + D)} \quad (4)$$

Keterangan:

- t = term atau kata
- c = kategori/kelas pada dokumen
- N = jumlah keseluruhan dokumen
- A = banyaknya jumlah dokumen dalam kategori c yang mengandung kata t
- B = banyaknya jumlah dokumen yang bukan kategori c tetapi mengandung kata t
- C = banyaknya jumlah dokumen dalam kategori c tetapi tidak mengandung kata t
- D = banyaknya jumlah dokumen yang bukan kategori c dan tidak mengandung kata t

Setelah proses penilaian pada tiap kata diketahui, kemudian mengurutkan kata berdasarkan nilai *Chi Square* dari nilai tertinggi hingga terendah dimana semakin besar nilainya maka semakin penting fitur tersebut untuk klasifikasi. Oleh karena itu semakin kecil nilai taraf nyata (α) akan semakin besar nilai kritis. Jika nilai taraf nyata (α) semakin kecil yang diterapkan, maka banyak fitur yang akan terseleksi atau terbuang.

2.7 Support Vector Machine

Support Vector Machine termasuk kedalam kategori *supervised learning* yang cukup baik dalam melakukan pengklasifikasian, regresi, dan prediksi. *Support Vector Machine* bekerja dengan cara mencari *hyperplane* atau garis pembatas yang paling optimal dimana berfungsi untuk memisahkan dua kelas [7]. Metode SVM merupakan hal yang tepat dalam melakukan pengklasifikasian terhadap dua kelas. Untuk memisahkan kedua kelas tersebut dilakukan pengukuran margin antara kedua kelas terhadap *Hyperplane*. Margin adalah jarak antara *Hyperplane* dengan data terdekat pada masing-masing kelas [8]. Untuk mendapatkan margin yang paling optimal dalam memisahkan dua kelas pada garis *Hyperplane* tersebut diperlukan perhitungan untuk menentukan titik maksimal. Adapun persamaan dari *hyperplane* dapat dilihat pada rumus persamaan (5).

$$(w \cdot x_i) + b = 0 \quad (5)$$

Keterangan:

w = parameter hyperplane
 x_i = data input berbentuk angka
b = nilai bias

2.9 Evaluasi

Tahapan berikutnya adalah melakukan evaluasi terhadap hasil yang didapat, dimana hasil tersebut berupa perhitungan dari akurasi, presisi, *recall*, dan *f1-score*. Berikut ini pengertian dari masing-masing pengukuran beserta rumus persamaanya:

Akurasi didapat dengan menghitung rasio dari nilai prediksi benar terhadap data. Artinya nilai akurasi berguna untuk mengukur tingkat akuratnya model klasifikasi dalam memprediksi dengan benar. Adapun persamaan dari akurasi dapat dilihat pada rumus persamaan (6).

$$accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (6)$$

Presisi didapat dengan menghitung rasio dari nilai prediksi benar positif terhadap data yang diprediksi positif. Artinya presisi berguna untuk mengukur kualitas keberhasilan dalam memprediksi data secara benar. Adapun persamaan dari presisi dapat dilihat pada rumus persamaan (7).

$$precision = \frac{TP}{(TP + FP)} \quad (7)$$

Recall didapat dengan menghitung rasio dari nilai prediksi benar positif terhadap data yang sebenarnya positif. Artinya *recall* berguna untuk mengukur kualitas keberhasilan model klasifikasi dalam mengembalikan sebuah informasi secara benar. Adapun persamaan dari *recall* dapat dilihat pada rumus persamaan (8).

$$recall = \frac{TP}{(TP + FN)} \quad (8)$$

F1-score didapat dengan menghitung rata-rata presisi dan *recall* yang dibobokan. Artinya *f1-score* berguna untuk mengukur seberapa baiknya presisi dan *recall* pada model klasifikasi. Adapun persamaan dari *f1-score* dapat dilihat pada rumus persamaan (9).

$$f1 - score = \frac{(2 * recall * precision)}{(recall + precision)} \quad (9)$$

2.10 Visualisasi

Pada tahapan ini diperlukan untuk menampilkan informasi yang berguna bagi calon pembeli dalam mengambil keputusan saat ingin membeli produk tersebut apakah layak untuk dibeli atau tidak. Sedangkan bagi penjual, informasi tersebut dapat digunakan untuk bahan evaluasi dalam peningkatan kualitas produk maupun kualitas pelayanan. Informasi yang akan divisualisasikan berupa frekuensi kata dari masing-masing kelas positif maupun negatif.

3 Hasil Pembahasan

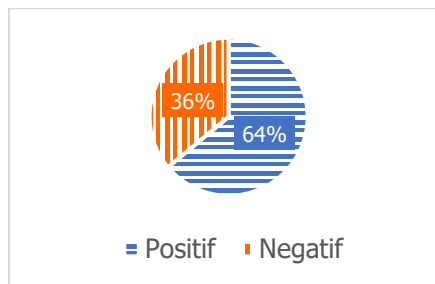
Data yang diperoleh merupakan hasil dari *web scraping* pada salah satu toko *online* yang ada di *marketplace* Shopee dengan nama penjual *atmos.official* dan nama produk “Masker KF94 Korea 4 ply impor *High Quality*”. Proses pengambilan data memakan waktu sekiranya kurang lebih 3 jam dan mendapatkan sebanyak 8.674 data dengan format json. Data yang didapat kemudian dilakukan konversi dari format json menjadi csv yang nantinya dapat digunakan dalam proses klasifikasi sentimen. Berikut merupakan hasil *scraping data* yang diambil pada tanggal 13 April 2022 yang dapat dilihat pada Gambar (4).

1	nama	reviewContent	rating	variant	material	quality	price	tags	like	createdAt	retrievedDate	
2	anjahedyati26	Ada harga ada kualitas ya To harga	5	Peach	tipis	lumayan	murah b		11	06/04/2022	13/04/2022	
3	*****	Harganya murah, dan warna dusty j	5	Dusty Pink	bagus	terbaik	terjangk	Kualitas	2	10/04/2022	13/04/2022	
4	*****		5	Krem / Wa Tebel	Bagus	Murmer	Kualitas		7	07/04/2022	13/04/2022	
5	*****	Harga: terjangkau, masih pas di kant	5	Hitam		bagus dan	oke kok,		1	12/04/2022	13/04/2022	
6	loeva_sakha	Udah langgan, setiap masker habi	5	Dusty Pink	tebal dan b.	bagus	teba	terjangk	Kualitas	0	10/04/2022	13/04/2022
7	*****	Bergiriman cepat dan aman, untuk	5	Peach	Tidak terlalu	lekas baik	sangat ti		22	28/03/2022	13/04/2022	
8	setyaningrumidani	Bergiriman cepat	5	Hitam	lekas,	teba	baik	terjangk	2	07/04/2022	13/04/2022	
9	*****	zsko banget barang datang sesuai p	5	Duckbill Putih	10 pc	baik	murah.	Kualitas	1	05/04/2022	13/04/2022	
10	revisi25	Rekomended sih lucu high warnanya	5	Marcho	Agly	ok	murmer		3	05/04/2022	13/04/2022	
11	*****	Rekomendasi sekali tempat ini amat	5	Hitam	Bagus lang	Terbaik	Bersahai	Kualitas	0	05/04/2022	13/04/2022	
12	p.co_official	Respon cepat penjual baik barang s	5	Dusty Pink	tebal	bagus	terjangk	Kualitas	43	14/03/2022	13/04/2022	
13	suffikayrus	Maskernya bagus banget, tebal, Pol	5	Hitam	tebal	baik	murah		8	15/03/2022	13/04/2022	
14	*****	Worth it! Yuk beli yuk. Karna 10 pcs	5	Biru Tua /	bahannya li	kualitas ba	Harga m	Kualitas	13	23/03/2022	13/04/2022	
15	damaridhan97	Kualitas lumayan bagus tetapi kema	5	Dusty Pink	lumayan	lumayan	sesuai b	Kualitas	19	25/03/2022	13/04/2022	

Gambar 4. Data Hasil Web Scraping

Data yang didapat dari hasil *web scraping* kemudian dilakukan pelebelan kedalam 2 kelas yaitu, kelas positif dan kelas negatif dimana kelas tersebut didapat berdasarkan dari *rating* ulasan yang diberikan oleh pembeli. Ulasan yang memiliki *rating* 1 dan 2 merupakan kelas negatif sedangkan untuk *rating* 4 dan 5 merupakan kelas positif. Namun pada penelitian ini untuk ulasan produk yang memiliki *rating* 3 dihapus dikarenakan ulasan tersebut merupakan ulasan yang *inconsistent rating* atau tidak dapat di lakukan *labeling* berdasarkan *rating*.

Sebelum masuk pada tahap pra proses, data yang akan digunakan terdiri dari gabungan data komentar, penilaian terhadap bahan, penilaian terhadap kualitas, dan penilaian terhadap harga serta label yang telah di dapatkan sebelumnya. Kemudian dilakukan penghapusan pada data yang kosong dan data berkurang menjadi 5.245 data. Data tersebut terdiri dari 1.891 data berlabel negatif dan 3.354 data berlabel positif. Berikut merupakan persentase dari kelas negatif dan kelas positif dalam bentuk diagram lingkaran yang dapat dilihat pada Gambar (5).



Gambar 5. Persentase Kelas Negatif dan Positif

Setelah data yang didapat melalui proses *web scraping* dan sudah dilakukan pelebelan, data tersebut akan dilakukan pembersihan sebelum data diolah kedalam pemodelan. Adapun tahapan yang dilakukan pada pra proses yaitu *case folding*, *cleansing*, *stopword removal*, *stemming*, *normalization*, dan *tokenization*. Berikut ini contoh data ulasan sebelum dan sesudah dilakukan *preprocessing* yang dapat dilihat pada Tabel (1).

Tabel 1. Data Sebelum dan Sesudah *Preprocessing*

Sebelum	Sesudah	Label
'BAGUS MANTAPPP HARGANYAA MURCEE BAHANNYA GA TERLALU TEBEL DAN GA TERLALU TIPIS BAGUS POKONYAA BAKAL REPEAT ORDER SIHH MANTAP👍 lumayan tebal ga terlalu tipis bagus murah'	'bagus mantap harga murah bahan tebal tipis bagus ulang pesan mantap lumayan tebal tipis bagus murah'	Positif
'Tipisssssss, kasar kaakuuuuu. Terburuk dr masker duckbill yg pernah sy beli. Tdk beli lg disiniw. Kapokkk tipissssss kasar, seumur umur barusan dpt putih duckbill kayak gini😞😞😞'	'tipis kasar kaku buruk masker duckbill beli beli kapok tipis kasar umur umur putih duckbil'	Negatif

Pada tahapan pra proses menyebabkan data yang kosong dilakukan penghapusan, dimana data yang awalnya berjumlah 8.674 data menjadi 5.245 data dengan jumlah fitur yang didapat sebanyak 2.187 fitur/kata.

Kata yang sudah dinormalisasi kemudian diberikan bobot dari hubungan suatu kata yang terdapat pada keseluruhan ulasan atau dokumen. Dimana diperlukan pembobotan dengan menggunakan TF-IDF untuk mengubah kata menjadi vektor yang dapat dipahami oleh mesin untuk proses klasifikasi. Berikut ini hasil dari perhitungan TF-IDF yang dapat dilihat pada Tabel (2).

Tabel 2. Perhitungan pembobotan TF-IDF

Term	TF		DF	IDF	TF * IDF	
	D1	D2			D1	D2
bagus	1,477	0	3	0,875	1,292	0
bahan	1	0	1	1,176	1,176	0
beli	0	1,301	2	1	0	1,301
buruk	0	1	1	1,176	0	1,176
duckbil	0	1,301	2	1	0	1,301
harga	1	0	1	1,176	1,176	0
kaku	0	1	1	1,176	0	1,176
kapok	0	1	1	1,176	0	1,176
kasar	0	1,301	2	1	0	1,301
lumayan	1	0	1	1,176	1,176	0
mantap	1,301	0	2	1	1,301	0
masker	0	1	1	1,176	0	1,176
murah	1,301	0	2	1	1,301	0
pesan	1	0	1	1,176	1,176	0
putih	0	1	1	1,176	0	1,176
tebal	1,301	0	2	1	1,301	0
tipis	1,301	1,301	4	0,778	1,012	1,012
ulang	1	0	1	1,176	1,176	0
umur	0	1,301	2	1	0	1,301

Setelah setiap kata yang telah diberi bobot dengan menggunakan TF-IDF. Kemudian dilakukan seleksi terhadap fitur yang tidak relevan terhadap proses klasifikasi nantinya. Berikut contoh ulasan sebelum dan sesudah dilakukan seleksi fitur menggunakan *Chi Square* berdasarkan nilai *p-value* yang didapat dan ditentukan nilai taraf nyata (α) sebesar 0.3 dimana fitur akan terbuang jika nilai *p-value* lebih kecil dari nilai taraf nyata (α).

Tabel 3. Hasil Perhitungan *p-value*

Kata/Fitur	kecewa	banget	beli	parah	...	lumayan	jangkau	bahan	pesan
<i>p-value</i>	0,926	0,716	0,716	0,617	...	0,154	0,073	0,042	0,007

Tabel 4. Contoh Hasil Seleksi Fitur

Sebelum	Sesudah
'bahan bagus kualitas barang bagus harga barang murah jual amanah puas belanja terima kasih'	'bagus kualitas barang bagus harga barang murah terima kasih'
'bahan tipis warna putih sesal beli rekomendasi banget jangkau '	'tipis warna putih sesal beli rekomendasi banget'
'barang bagus kualitas baik harga jangkau seller cepat respon kemas aman terima kasih bagus premium standar '	'barang bagus kualitas harga terima kasih bagus'
'langgan benar kecewa bahan tipis parah kecewa beli bagus tipis pakai banget parah pesan kian kecewa tipis'	'langgan benar kecewa tipis parah kecewa beli bagus tipis pakai banget parah kian kecewa tipis'
'bagus mantap harga murah bahan tebal tipis bagus ulang pesan mantap lumayan tebal tipis bagus murah'	'bagus mantap harga murah tebal tipis bagus ulang mantap tebal tipis bagus murah'

Sebelum dilakukan klasifikasi terhadap data diperlukan pembagian data untuk memisahkan data latih (*training*) dan data uji (*testing*) dengan menggunakan *train_test_split* dari *library* sklearn pada *model_selection*. Pada penelitian ini pembagian dibagi menjadi 20% data *testing* dan 80% data *training* dari 5.245 data. Maka pada data *testing* berjumlah 1.049 data dan untuk data *training* berjumlah 4.196 data. Berikut adalah pembagian data dalam bentuk Tabel (5).

Tabel 5. Pembagian Data

	Positif	Negatif	Total
Data Training (80%)	2.683	1.513	4.196
Data Testing (20%)	671	378	1.049
Total	3.354	1.891	5.245

Setelah mendapatkan data latih dengan jumlah 4.196 data dan 2.187 fitur, berikutnya akan dilakukan beberapa eksperimen klasifikasi menggunakan metode *Support Vector Machine* (SVM) dengan dan tanpa seleksi fitur *Chi Square*. Untuk kernel yang digunakan pada metode *Support Vector Machine* (SVM) adalah kernel *Radial Basis Function* (RBF) dengan parameter *default*, dimana *cost* (C) = 1.0 dan *gamma* (γ) = 0,1. Selanjutnya dilakukan perbandingan evaluasi dari tiap eksperimen dengan menghitung tingkat performa dari hasil klasifikasi, dimana tingkat performa yang dimaksud adalah nilai dari akurasi, presisi, *recall*, dan *f1-score*. Berikut ini hasil yang didapat dari beberapa eksperimen yang dilakukan yang dapat dilihat pada Tabel (6).

Tabel 6. Perbandingan Hasil Evaluasi

Taraf Nyata (α)	Akurasi	Presisi	<i>Recall</i>	<i>f1-score</i>
0 (Tanpa Seleksi Fitur)	90,75%	92,08%	93,59%	92,83%
0.5	90,84%	92,21%	93,59%	92,89%
0.3	91,32%	93,28%	93,14%	93,21%
0.25	91,13%	93,26%	92,84%	93,05%
0.1	90,37%	92,15%	92,84%	92,50%

dapat dilihat bahwa eksperimen yang menggunakan seleksi fitur memiliki tingkat performa yang lebih baik dari pada tanpa seleksi fitur. Untuk penjelasan dari masing-masing hasil evaluasi dapat dilihat sebagai berikut.

Berdasarkan hasil evaluasi terhadap beberapa eksperimen yang telah dilakukan. Peneliti memutuskan untuk memilih hasil evaluasi pada eksperimen dengan nilai taraf nyata (α) 0.3 karena beberapa faktor diantaranya, dilihat dari data FN (*False Negative*) dengan jumlah 46 dan data FP (*False Positive*) dengan jumlah 45 menandakan data tersebut sangat mendekati atau *symmetric* yang artinya nilai akurasi dapat kita gunakan sebagai acuan dalam memilih tingkat performa yang baik.

Faktor lainnya dapat dilihat pada nilai *recall* yang tidak tertinggi dan juga tidak terendah yang menandakan bahwa klasifikasi diharapkan dapat memprediksi ulasan dengan kelas positif tetapi sebenarnya negatif daripada salah memprediksi ulasan dengan kelas negatif padahal sebenarnya ulasan tersebut positif. Namun nilai presisi pada eksperimen ini merupakan tertinggi dari yang lain yang membuat dilema diantara presisi dan *recall* karena perbandingan keduanya terjadi *tradeoff* atau kondisi dimana nilai presisi tinggi, dan nilai *recall* rendah. Oleh karena itu untuk mencari tingkat performa yang terbaik, lebih baik menggunakan nilai *f1-score* sebagai acuannya.

Kemudian tahapan terakhir adalah visualisasi pada ulasan produk masker yang memiliki sentimen positif maupun negatif dengan melihat *term* atau kata yang sering muncul untuk dijadikan informasi yang dapat digunakan untuk calon pembeli dalam mempertimbangkan saat ingin membeli produk masker tersebut. Sedangkan bagi penjual, informasi tersebut dapat digunakan untuk bahan evaluasi dalam peningkatan kualitas produk maupun kualitas pelayanan. Berikut adalah visualisasi terhadap label positif dan negatif yang dapat dilihat pada Gambar (6) terkait visualisasi sentimen positif dan Gambar (7) terkait visualisasi sentimen negatif.



Gambar 6. Wordcloud Sentimen Positif

Pada Gambar 6 dapat dilihat ukuran *font* akan semakin besar jika kata tersebut sering muncul. Adapun kata yang sering muncul pada sentimen positif, yaitu kata ‘bagus’, ‘murah’, ‘lumayan’, ‘cepat’, ‘sesuai’, ‘jangkau’, ‘baik’, ‘tebal’, ‘ok’, ‘mantap’ dan lain-lain. Hasil yang didapat dari sentimen positif tersebut menggambarkan bahwa pembeli lebih banyak memberikan ulasan terkait masker yang dibeli dengan kualitas bagus serta harga yang murah dan terjangkau.



Gambar 7. Wordcloud Sentimen Negatif

Pada Gambar 7 dapat dilihat ukuran *font* akan semakin besar jika kata tersebut sering muncul. Adapun kata yang sering muncul pada sentimen negatif, yaitu kata ‘pesan’, ‘kir-im’, ‘warna’, ‘tipis’, ‘sesuai’, ‘kecewa’, ‘bahan’, ‘kualitas’, ‘jelek’, ‘beda’ dan lain-lain. Hasil yang didapat dari sentimen negatif tersebut menggambarkan bahwa banyak pembeli yang merasa kecewa pada pesanan yang dibeli atau dikirim karena tidak sesuai dengan warna yang di pesan. Sentimen negatif tersebut menggambaran bahwa kualitas pelayanan dari penjual masih kurang dalam kesesuaian pesanan yang dikirim.

4 Kesimpulan dan Saran

4.1 Kesimpulan

Pengaruh seleksi fitur *Chi Square* terhadap algoritma *Support Vector Machine* (SVM) dengan kernel *radial basis function* (RBF) dapat meningkatkan tingkat performa dari pemodelan yang didapat dengan menentukan nilai taraf nyata (α), dimana nilai taraf nyata (α) yang semakin kecil akan mengakibatkan lebih banyak fitur yang terbuang. Analisis hasil evaluasi terhadap beberapa eksperimen yang telah dilakukan menghasilkan model klasifikasi yang terbaik pada eksperimen yang menggunakan seleksi fitur *Chi Square* dengan nilai taraf nyata (α) 0.3. Tingkat performa dari model tersebut menghasilkan nilai akurasi sebesar 91,32%, presisi sebesar 93,28%, *recall* sebesar 93,14% dan *f1-score* sebesar 93,21%. Dibanding dengan tingkat performa tanpa menggunakan seleksi fitur hanya menghasilkan nilai akurasi sebesar 90,75%, presisi sebesar 92,08%, *recall* sebesar 93,59% dan *f1-score* sebesar 92,83%.

4.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, penulis memberikan beberapa saran kepada peneliti selanjutnya sebagai berikut:

1. Pada penelitian selanjutnya diharapkan dapat memperbanyak korpus normalisasi terhadap kata-kata singkat maupun kata typo menjadi kata baku dan formal.
2. Pada penelitian selanjutnya diharapkan dapat melakukan labeling secara manual terhadap ulasan dengan rating 3.
3. Pada penelitian selanjutnya diharapkan dapat mengubah parameter yang ada pada kernel radial basis function (RBF) atau mengubah algoritma klasifikasi lain untuk mencari tingkat performa yang lebih baik.
4. Pada penelitian selanjutnya diharapkan dapat melakukan eksperimen lainnya dengan mengubah nilai taraf nyata (α) pada seleksi fitur Chi Square untuk mendapatkan jumlah fitur yang terbaik.

Referensi

- [1] Farki, A., Baihaqi, I., & Wibawa, M. (2016). *Pengaruh online customer review rating terhadap kepercayaan place di Indonesia*. Jurnal Teknik ITS, 5(2), A614–A619.
- [2] A. D. Saputra, Adiwijaya dan M. S. Mubarak. (2017). *Klasifikasi Sentimen Pada Level Aspek Terhadap Ulasan Produk Berbahasa Inggris Menggunakan Bayesian Network (Case Study : Data Ulasan Produk Amazon)*. e-Proceeding Eng., vol. 4, p. 4882, 2017.
- [3] L. Zhang and B. Liu (2016). *Sentiment Analysis and Opinion Mining*. Encycl. Mach. Learn. Data Min., 2016, doi: 10.1007/978-1-4899-7502-7_907-1.
- [4] Susanti Fransiska, Rianto, & Acep Irham Gufroni. (2020). *Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method*. Scientific Journal of Informatics, Vol. 7, No. 2, November 2020.
- [5] Haq, F. I. N., & Budi, E. (2019). *Implementasi Naive Bayes Classifier untuk Prediksi Kepribadian Big Five pada Twitter Menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) dan Term Frequency-Relevance Frequency (TF-RF)*. e-Proceeding of Engineering, 6(2), 9785–9795.
- [6] Akbari, M.I.H.A.D., Novianty, A. & Setianingsih, C. (2017). *Analisis Sentimen Menggunakan Metode Learning Vector Quantization*. e-Proceeding of Engineering, Vol. 4(2), Agustus 2017, pp. 2283-2292.
- [7] Lidya, S.K., Sitompul, O.S. & Efendi, S. (2015). *Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine (SVM) dan K-Nearest Neighbor (K- NN)*. Seminar Nasional Teknologi Informasi dan Komunikasi.
- [8] M Rushdi Saleh, M.T. Martin–valdivia, A. Montejo-Raez, & L.A. Urena–Lopez. (2011). *Expreiments with SVM to Classiffy Opinion in Different Domain*. Vol. 38. Doi : <https://doi.org/10.1016/j.eswa.2011.05.070>.