

# ANALISIS SENTIMEN PADA MEDIA SOSIAL *FACEBOOK* TERHADAP *MARKETPLACE ONLINE* DI INDONESIA MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*

Saadillah Razaqi Salam<sup>1</sup>, Agung Nilogiri<sup>2</sup>

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember  
s.razaqi97@gmail.com<sup>1</sup>, agungnilogiri@unmuhjember.ac.id<sup>2</sup>

## ABSTRAK

Keberadaan *marketplace online* di Indonesia membuat berbelanja secara *online* menjadi lebih mudah. Namun terdapat perasaan kepuasan ataupun ketidakpuasan dalam penggunaannya. Media sosial *Facebook* menjadi salah satu media yang digunakan masyarakat untuk menyampaikan perasaan tersebut. Pada penelitian ini dilakukan analisis sentimen yang dapat menyimpulkan penilaian masyarakat terhadap *marketplace online*. Data yang digunakan pada penelitian ini adalah komentar dalam *fans page Facebook* resmi Bukalapak, Shopee dan Tokopedia pada periode waktu bulan Februari sampai dengan September 2018. Metode klasifikasi pada penelitian ini menggunakan klasifikasi *Support Vector Machine* dengan konsep non linear *kernel Radial Basic Function* (RBF). Hasil pengujian dari penelitian ini adalah klasifikasi komentar paling optimal didapatkan nilai akurasi sebesar 80,5% pada Bukalapak dengan persentase jumlah komentar positif 42,5% dan komentar negatif 57,5%, sedangkan pada Shopee didapatkan nilai akurasi sebesar 74% dengan persentase jumlah komentar positif 45% dan komentar negatif 55%, serta pada Tokopedia didapatkan nilai akurasi sebesar 82% dengan persentase jumlah komentar positif 37% dan komentar negatif 63%. Sedangkan kumpulan kata yang mendominasi dari hasil analisis pada Bukalapak sentimen positif antara lain: sukses, mantap, cepat dan sentimen negatif antara lain: tipu, bohong, komplain, sedangkan pada Shopee sentimen positif antara lain: gratis, alhamdulillah, murah dan sentimen negatif antara lain: tipu, susah, kecewa, serta pada Tokopedia sentimen positif antara lain: mantap, hadiah, sukses dan sentimen negatif antara lain: eror, payah, tipu.

**Kata Kunci :** Analisis Sentimen, *Facebook*, *Marketplace Online*, *Support Vector Machine*.

## ABSTRACT

*The existence of an online marketplace in Indonesia makes shopping online easier. But there is a feeling of satisfaction or dissatisfaction in its use. Facebook social media is one of the media used by the community to convey these feelings. In this research a sentiment analysis can be concluded that can conclude the community assessment of the online marketplace. The data used in this study are comments in the official fan page of Bukalapak, Shopee and Tokopedia for the period February to September 2018. The classification method in this study uses the classification of Support Vector Machine with a non linear kernel Radial Basic Function (RBF) concept. The test results from this study are the most optimal classification of comments obtained by the accuracy of 80.5% at Bukalapak with the percentage of the number of positive comments 42.5% and negative comments 57.5%, while in Shopee the accuracy value is 74% with the percentage of comments positive 45% and negative comments 55%, and on Tokopedia obtained an accuracy value of 82% with a percentage of positive comments 37% and negative comments 63%. While the collection of dominating words from the analysis on Bukalapak positive sentiments include: successful, steady, fast and negative sentiments, among others: deception, lying, complaints, while in Shopee positive sentiments include: free, thank God, cheap and negative sentiments include : deception, difficulty, disappointment, and on Tokopedia positive sentiments include: solid, gifts, success and negative sentiments include: error, poor, deceitful.*

**Keywords :** *Sentiment Analysis, Facebook, Online Marketplace, Support Vector Machine*

## PENDAHULUAN

### Latar Belakang

Belanja *online* menjadi pilihan yang paling digemari masyarakat Indonesia untuk berbelanja suatu produk. Menurut gubernur bank Indonesia, sepanjang tahun 2016 jumlah transaksi belanja *online* masyarakat Indonesia mencapai Rp 75 triliun (Nababan, 2017). Hasil survei menyatakan 82,2 juta pengguna internet di Indonesia mengunjungi konten belanja *online* pada tahun 2016 (APJII, 2016). Berdasarkan dari banyaknya jumlah transaksi serta jumlah pengunjung *marketplace online*, tentu terdapat suatu perasaan puas ataupun ketidakpuasan dalam proses bertransaksi. Salah satu media yang digunakan untuk melampiaskan perasaan tersebut adalah melalui media sosial *Facebook*.

Banyak masyarakat yang menyampaikan komentar kepuasan ataupun ketidakpuasan melalui *fans page Facebook marketplace online*. Sehingga apabila diteliti lebih lanjut akan didapatkan sebuah sentimen yang dapat menyimpulkan baik atau buruknya penilaian masyarakat pada *marketplace online* tersebut. Sebuah studi yang dapat dilakukan untuk mendapatkan hasil tersebut adalah dengan melakukan analisis sentimen.

Analisis sentimen merupakan cabang penelitian dari *text mining*, namun lebih terfokus dalam menganalisis opini dari suatu dokumen teks. Salah satu metode yang dapat digunakan dengan baik untuk melakukan analisis sentimen adalah *Support Vector Machine* (SVM).

### Rumusan Masalah

Berdasarkan latar belakang yang dikemukakan di atas, maka rumusan permasalahan yang akan dibahas adalah sebagai berikut: (1) Berapakah tingkat akurasi analisis sentimen pada media sosial *Facebook* terhadap *marketplace online* di Indonesia menggunakan metode *Support Vector Machine* ? (2) Berapakah tingkat persentase komentar positif dan negatif terhadap *marketplace online* di Indonesia ?

(3) Apa kata-kata yang sering muncul pada komentar positif dan negatif terhadap *marketplace online* di Indonesia ?

### Tujuan

Tujuan dari penelitian ini adalah sebagai berikut: (1) Mengetahui tingkat akurasi analisis sentimen pada media sosial *Facebook* terhadap *marketplace online* di Indonesia menggunakan metode *Support Vector Machine*, (2) Mengetahui tingkat persentase komentar positif dan negatif terhadap *marketplace online* di Indonesia, (3) Mengetahui kata-kata yang sering muncul pada komentar positif dan negatif terhadap *marketplace online* di Indonesia.

### Manfaat

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut: (1) Dapat memberikan kontribusi secara keilmuan berupa analisis sentimen pada media sosial *Facebook* dengan menggunakan metode *Support Vector Machine*, (2) Dapat digunakan sebagai referensi untuk penelitian-penelitian selanjutnya dibidang analisis sentimen, (3) Dapat membuktikan respons masyarakat terhadap *marketplace online* di Indonesia.

## METODE PENELITIAN

### Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh data komentar yang akan digunakan untuk objek penelitian. Beberapa proses yang dilakukan antara lain:

1. *Crawling data* pada media sosial *Facebook*, dilakukan untuk mendapatkan data dengan cara mengakses API (*Application Programming Interface*) *Facebook*. Aplikasi *Facepacer* digunakan untuk mengakses API *Facebook* dengan hasil keluaran berupa kumpulan data komentar yang tersimpan ke dalam format *Comma Separated Value* (CSV).
2. Pelabelan data, dilakukan untuk mendapatkan label kelas komentar positif dan komentar negatif pada setiap

data komentar dengan bantuan seorang ahli bahasa Indonesia.

- Perbaikan kata, dilakukan untuk memperbaiki penulisan kata-kata komentar yang digunakan pengguna media sosial *Facebook*. Kata-kata singkatan, kata yang mengandung kesalahan penulisan (*typo*) dilakukan perbaikan untuk menyamakan maksud dari penulisan komentar tersebut.

### Pre-processing

*Pre-processing* dilakukan untuk mempersiapkan data komentar sehingga bersih dari *noise* dan sesuai dengan ketentuan analisis sentimen. Menurut Triawati dkk. (2009), beberapa proses yang dilakukan antara lain:

- Case folding*, yaitu proses mengonversi keseluruhan tulisan dalam data menjadi huruf kecil.
- Tokenizing*, yaitu proses pemotongan tulisan berupa huruf untuk dipisahkan menjadi kata per kata.
- Filtering*, yaitu proses menghapus kata-kata yang kurang penting seperti kata “yang”, “dan”, “di”, “dengan” dari hasil pemotongan tulisan.
- Stemming*, yaitu proses mencari kata dasar dari tiap kata pembentuknya.

### Pembobotan

Metode pembobotan dilakukan untuk mencari representasi nilai dari tiap-tiap kata pada sekumpulan data yang akan dibentuk menjadi sebuah vektor menggunakan algoritma TF-IDF. Penggunaan TF-IDF direpresentasikan pada Tabel 1.

Tabel 1 *Term-document Matrix*

	$t_1$	$t_2$	...	$t_m$
$d_1$	$w_{11}$	$w_{21}$	...	$w_{m1}$
$d_2$	$w_{12}$	$w_{22}$	...	$w_{m2}$
...	...	...	...	...
$d_n$	$w_{1n}$	$w_{2n}$	...	$w_{mn}$

Pembobotan TF-IDF dilakukan dengan cara pemberian bobot hubungan suatu kata atau fitur ( $t$ ) sebanyak  $m$  terhadap data ( $d$ ) sebanyak  $n$ , serta  $w$  merupakan hasil

pembobotan TF-IDF. Beberapa tahapan yang dilakukan pada pembobotan menggunakan TF-IDF antara lain:

- Term frequency* atau TF, merupakan perhitungan frekuensi suatu term atau kata yang terdapat pada setiap data berbentuk dokumen.
- Inverse document frequency* atau IDF, merupakan hubungan ketersediaan suatu term dalam seluruh dokumen. Persamaan yang digunakan pada perhitungan IDF ditunjukkan pada persamaan 1. (Havrlant dan Kreinovich, 2014)

$$idf(t, d) = \ln \frac{N}{df(t)} \quad (1)$$

Pada persamaan 1,  $idf(t, d)$  merupakan perhitungan IDF pada term  $t$  di dalam data komentar  $d$ ,  $N$  merupakan jumlah data komentar, dan  $df(t)$  merupakan jumlah data komentar yang mengandung term  $t$ . Pada persamaan di atas, IDF jika nilai  $N$  sama dengan nilai  $df(t)$  maka akan didapatkan hasil IDF bernilai 0. Untuk meniasasi hal tersebut perlu ditambahkan konstanta bernilai 1 pada perhitungan IDF, sehingga perhitungan IDF menjadi persamaan 2.

$$idf(t, d) = \ln \frac{N}{df(t)} + 1 \quad (2)$$

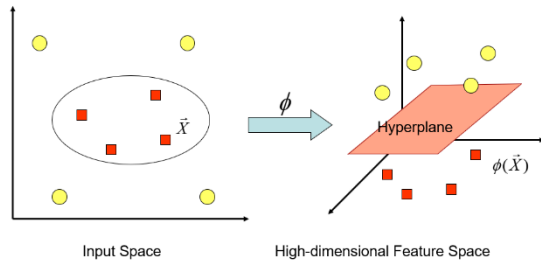
- Term Weighting* TF-IDF, merupakan penggabungan dari rumus TF dengan IDF dengan mengalikan kedua rumus tersebut menjadi sebuah nilai pembobotan. Perhitungan tersebut ditunjukkan pada persamaan 3.

$$w(t, d) = tf(t, d) \times idf(t, d) \quad (3)$$

### Klasifikasi

Metode klasifikasi yang digunakan pada penelitian ini menggunakan metode *Support Vector Machine* (SVM) dengan konsep non linear. *Support Vector Machine* adalah mencari fungsi pemisah (*hyperplane*) terbaik atau paling optimum yang berfungsi sebagai pemisah dua kelas pada ruang vektor *input*. Untuk mengklasifikasikan data pada permasalahan non linear, formula SVM harus dimodifikasi, yaitu dengan

mentransformasikan data yang berada pada ruang vektor *input* (*input space*) ke dalam ruang vektor baru yang berdimensi lebih tinggi (*feature space*) untuk dapat memisahkan data sesuai kelasnya. Berikut ilustrasi pemetaan data dari *input space* ke *feature space*:



Gambar 1 Pemetaan Data dari *Input Space* ke *Feature Space*

Dengan menggunakan fungsi pemetaan  $\vec{x}_i \rightarrow \phi(\vec{x}_i)$  ke dalam *feature space*, maka nilai  $w$  pada persamaan berikut:

$$w = \sum_{i=1}^l \sum_{x_i=SV} \alpha_i \cdot y_i \cdot \phi(\vec{x}_i) \quad (4)$$

Dan persamaan fungsi pemisah adalah sebagai berikut:

$$f(\phi(\vec{x})) = \text{sign}(\sum_{i=1}^l \sum_{x_i=SV} \alpha_i \cdot y_i \cdot K(\vec{x}, \vec{x}_i) + b) \quad (5)$$

Fungsi *kernel* yang digunakan pada permasalahan non linear ini menggunakan persamaan *kernel Radial Basis Function* (RBF) sebagai berikut:

$$K(\vec{x}_i, \vec{x}_j) = \exp(-\gamma \|\vec{x}_i - \vec{x}_j\|^2), \gamma > 0 \quad (6)$$

### Validasi dan Evaluasi

Validasi dilakukan terhadap kumpulan *dataset* komentar setiap *marketplace online* untuk mendapatkan performansi akurasi yang optimal ketika dilakukan proses evaluasi. Metode validasi menggunakan *K-Fold Cross Validation* membagi *dataset* menjadi beberapa bagian. Penentuan nilai *fold K* yang digunakan pada penelitian ini bernilai 2, 4, 5, 8 dan 10.

Evaluasi dilakukan untuk mendapatkan hasil performansi akurasi dari penggunaan metode yang telah dilakukan sebelumnya. Metode evaluasi yang digunakan pada penelitian ini adalah *Confusion Matrix*. Penggunaan *Confusion Matrix* direpresentasikan pada Tabel 2.

Tabel 2 *Confusion Matrix*

		Predicted Class	
		Positif	Negatif
Actual Class	Positif	TP	FN
	Negatif	FP	TN

Berdasarkan tabel di atas, maka perhitungan akurasi ditunjukkan pada persamaan 4.

$$\text{Akurasi} = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (7)$$

Hasil akurasi dari penggunaan *Confusion Matrix* dalam *K-Fold Cross Validation* merupakan hasil rata-rata dari masing-masing *Confusion Matrix* tiap kali iterasi terhadap *dataset*.

## HASIL DAN PEMBAHASAN

### Data

Data yang dikumpulkan berasal dari komentar pada *fans page marketplace online* Bukalapak, Shopee dan Tokopedia. Pengumpulan data komentar pada masing-masing *marketplace online* didapatkan komentar Bukalapak berjumlah 4.138 komentar, Shopee berjumlah 64.806 komentar dan Tokopedia berjumlah 45.216 komentar.

Data yang telah terkumpul selanjutnya diambil 100 data komentar yang dianggap komentar positif dan 100 data komentar yang dianggap komentar negatif pada masing-masing *marketplace online*. Selanjutnya data-data tersebut dilakukan validasi data oleh seorang ahli bahasa Indonesia untuk memastikan label yang telah diberikan, sehingga hasilnya didapat 200 data komentar yang telah tervalidasi pada setiap *marketplace online*.

Selanjutnya data komentar dilakukan perbaikan kata dalam komentar seperti ditunjukkan pada Tabel 3.

Tabel 3 Perbaikan Kata pada Komentar

Komentar	Perbaikan Komentar
BL wokeeeh.. Sukses terus	bukalapak oke.. sukses terus
Bukalapak gak jelas, duit gua kgk dikembalikan	bukalapak tidak jelas, uang saya tidak dikembalikan

### Pre-processing Data

Pre-processing data komentar dilakukan perubahan isi komentar dan label komentar sesuai ketentuan analisis sentimen. Berikut hasil *pre-processing* pada beberapa komentar:

Tabel 4 Pre-processing pada Komentar

Perbaikan Komentar	Label
bukalapak oke.. sukses terus	positif
bukalapak tidak jelas, uang saya tidak dikembalikan	negatif
Hasil Preprocessing	
	ID Label
oke sukses	0
uang kembali	1

### Implementasi Pembobotan

Pengimplementasian pembobotan TF-IDF mendapatkan jumlah fitur atau kata pada semua *marketplace online* Bukalapak, Shopee dan Tokopedia sebanyak 849 kata yang digunakan. Berikut hasil pembobotan pada Bukalapak yang ditunjukkan pada Tabel 5, Shopee yang ditunjukkan pada Tabel 6 dan Tokopedia yang ditunjukkan pada Tabel 7.

Tabel 5 Pembobotan pada Bukalapak

	akibat	belanja	...	suka	usaha
Data 1	0	0	...	3,9	0
Data 2	0	0	...	0	5,6
...	...	...	...	...	...
Data 199	0	2,9	...	0	0
Data 200	6,3	0	...	0	0

Tabel 6 Pembobotan pada Shopee

	barang	belanja	...	suka	uang
Data 1	2,7	0	...	0	0
Data 2	0	0	...	4,2	0
...	...	...	...	...	...
Data 199	0	1,9	...	0	0
Data 200	2,7	0	...	0	2,9

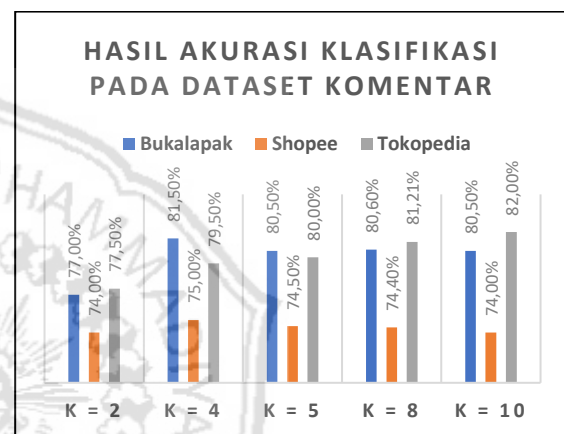
Tabel 7 Pembobotan pada Tokopedia

	barang	jelas	...	terima	uang
Data 1	0	5,6	...	0	3,4
Data 2	0	0	...	0	0
...	...	...	...	...	...
Data 199	2,5	0	...	3,1	0
Data 200	4,9	0	...	0	0

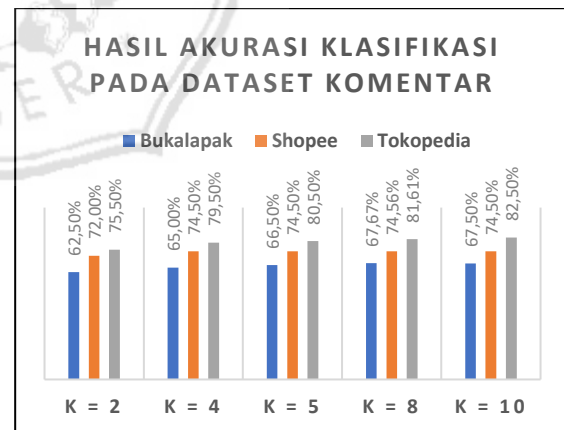
### Hasil Klasifikasi SVM

Pada penelitian ini, klasifikasi dilakukan 2 kali pengujian dalam penggunaan metode SVM dengan konsep non linear. Pengujian pertama dan kedua dilakukan dengan menggunakan parameter

nilai gamma ( $\gamma$ ) yang berbeda, yaitu  $\gamma = 0,01$  dan  $\gamma = \frac{1}{\text{banyak fitur (kata)}}$  untuk mengetahui hasil akurasi paling optimal dalam permasalahan ini. Serta dalam metode validasi dan evaluasi digunakan *K-Fold Cross Validation* dan *Confusion Matrix* dengan nilai *fold* K = 2, 4, 5, 8 dan 10 untuk mengetahui penggunaan *fold* K dengan hasil akurasi yang paling optimal penggunaannya dalam permasalahan ini. Berikut hasil akurasi pada pengujian pertama yang ditunjukkan pada Grafik 1 dan hasil akurasi pada pengujian kedua yang ditunjukkan pada Grafik 2.



Grafik 1 Hasil Akurasi pada Pengujian Pertama

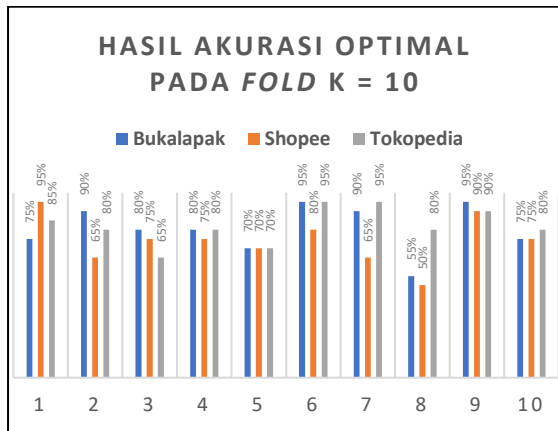


Grafik 2 Hasil Akurasi pada Pengujian Kedua

### Hasil Analisis Komentar

Berdasarkan pengujian klasifikasi yang telah dilakukan, hasil akurasi paling optimal didapatkan pada pengujian pertama dengan parameter nilai gamma  $\gamma = 0,01$  dan nilai *fold* K = 10. Hasil akurasi yang didapatkan pada masing-masing *dataset* komentar *marketplace online* sebesar

80,5% untuk Bukalapak, 74% untuk Shopee dan 82% untuk Tokopedia. Berikut gambaran didapatkan hasil akurasi tersebut pada Grafik 3.



Grafik 3 Hasil Akurasi dengan Parameter Nilai  $\gamma = 0,01$  dan  $K = 10$

Hasil analisis berdasarkan grafik di atas, didapatkan persentase komentar positif dan negatif pada setiap *marketplace online* dengan perolehan hasil *Confusion Matrix* pada masing-masing *marketplace online*. Pada Bukalapak didapatkan *Confusion Matrix* sebagai berikut:

Tabel 8 *Confusion Matrix* pada Bukalapak

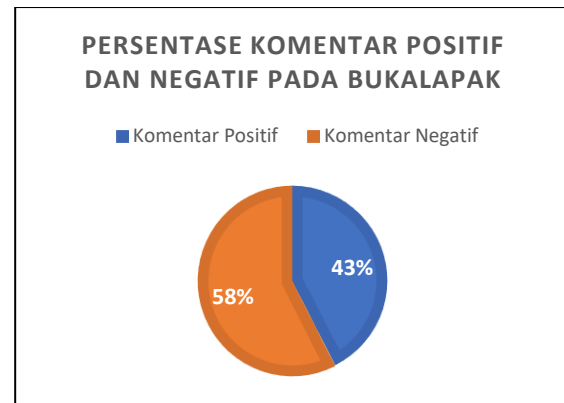
		Predicted Class	
		Positif	Negatif
Actual Class	Positif	73	27
	Negatif	12	88

Sehingga persentase komentar positif dan negatif pada *marketplace online* Bukalapak didapatkan sebagai berikut:

$$\begin{aligned} \text{Komentar Positif} &= \frac{TP + FP}{TP + FP + TN + FN} \times 100\% \\ &= \frac{73 + 12}{73 + 12 + 88 + 27} \times 100\% \\ &= 42,5\% \end{aligned}$$

$$\begin{aligned} \text{Komentar Negatif} &= \frac{TN + FN}{TP + FP + TN + FN} \times 100\% \\ &= \frac{88 + 27}{73 + 12 + 88 + 27} \times 100\% \\ &= 57,5\% \end{aligned}$$

Sehingga pada Bukalapak didapatkan grafik persentase komentar positif dan negatif sebagai berikut:



Grafik 4 Persentase Komentar pada Bukalapak  
Sedangkan pada Shopee didapatkan *Confusion Matrix* sebagai berikut:

Tabel 9 *Confusion Matrix* pada Shopee

		Predicted Class	
		Positif	Negatif
Actual Class	Positif	69	31
	Negatif	21	79

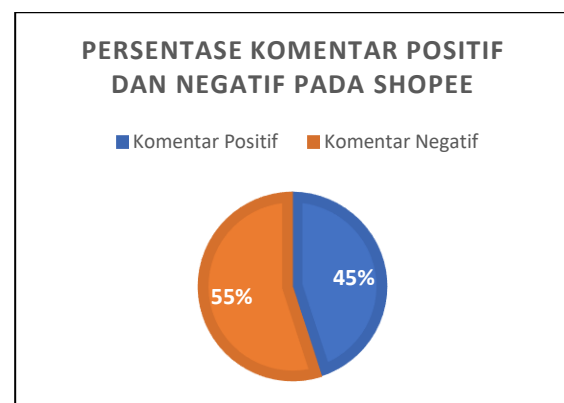
Sehingga persentase komentar positif dan negatif pada *marketplace online* Shopee didapatkan sebagai berikut:

$$\begin{aligned} \text{Komentar Positif} &= \frac{TP + FP}{TP + FP + TN + FN} \times 100\% \\ &= \frac{69 + 21}{69 + 21 + 79 + 31} \times 100\% \end{aligned}$$

$$\text{Komentar Positif} = 45\%$$

$$\begin{aligned} \text{Komentar Negatif} &= \frac{TN + FN}{TP + FP + TN + FN} \times 100\% \\ &= \frac{79 + 31}{69 + 21 + 79 + 31} \times 100\% \\ &= 55\% \end{aligned}$$

Sehingga pada Shopee didapatkan grafik persentase komentar positif dan negatif sebagai berikut:



Grafik 5 Persentase Komentar pada Shopee

Serta pada Tokopedia didapatkan *Confusion Matrix* sebagai berikut:

Tabel 10 *Confusion Matrix* pada Tokopedia

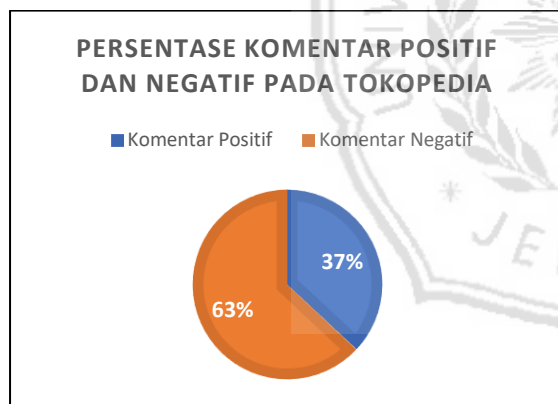
		Predicted Class	
		Positif	Negatif
Actual Class	Positif	69	31
	Negatif	5	95

Sehingga persentase komentar positif dan negatif pada *marketplace online* Tokopedia didapatkan sebagai berikut:

$$\begin{aligned} \text{Komentar Positif} &= \frac{TP + FP}{TP + FP + TN + FN} \times 100\% \\ &= \frac{69 + 5}{69 + 5 + 95 + 31} \times 100\% \\ &= 37\% \end{aligned}$$

$$\begin{aligned} \text{Komentar Negatif} &= \frac{TN + FN}{TP + FP + TN + FN} \times 100\% \\ &= \frac{95 + 31}{69 + 5 + 95 + 31} \times 100\% \\ &= 63\% \end{aligned}$$

Sehingga pada Tokopedia didapatkan grafik persentase komentar positif dan negatif sebagai berikut:



Grafik 6 Persentase Komentar pada Tokopedia

Hasil analisis selanjutnya yaitu kumpulan kata (*word cloud*) dalam komentar positif dan negatif didapatkan berdasarkan frekuensi kemunculan kata (*term frequency*) pada komentar masing-masing *marketplace online* antara lain:

### 1. Bukalapak

Hasil *word cloud* dalam komentar positif pada Bukalapak direpresentasikan pada Gambar 2 dan hasil *word cloud* dalam komentar negatif pada Bukalapak direpresentasikan pada Gambar 3.



Gambar 2 *Word Cloud* dalam Komentar Positif Bukalapak



Gambar 3 *Word Cloud* dalam Komentar Negatif Bukalapak

Berdasarkan Gambar 2 dan Gambar 3, representasi perolehan hasil frekuensi kemunculan kata paling banyak dalam komentar positif dan negatif pada *marketplace online* Bukalapak ditunjukkan pada Tabel 11.

Tabel 11 *Term Frequency* pada Komentar Bukalapak

Label	Kata	Frekuensi
Komentar Positif	sukses	19
	mantap	18
	cepat	18
Komentar Negatif	tipu	23
	bohong	21
	komplain	20

### 2. Shopee

Hasil *word cloud* dalam komentar positif pada Shopee direpresentasikan pada Gambar 4 dan hasil *word cloud* dalam komentar negatif pada Shopee direpresentasikan pada Gambar 5.



Gambar 4 *Word Cloud* dalam Komentar Positif Shopee



Gambar 5 Word Cloud dalam Komentar Negatif Shopee

Berdasarkan Gambar 4 dan Gambar 5, representasi perolehan hasil frekuensi kemunculan kata paling banyak dalam komentar positif dan negatif pada *marketplace online* Shopee ditunjukkan pada Tabel 12.

Tabel 12 Term Frequency pada Komentar Shopee

Label	Kata	Frekuensi
Komentar Positif	gratis	31
	alhamdulillah	27
	murah	25
Komentar Negatif	tipu	39
	susah	37
	kecewa	39

### 3. Tokopedia

Hasil *word cloud* dalam komentar positif pada Tokopedia direpresentasikan pada Gambar 6 dan hasil *word cloud* dalam komentar negatif pada Tokopedia direpresentasikan pada Gambar 7.



Gambar 6 Word Cloud dalam Komentar Positif Tokopedia



Gambar 7 Word Cloud dalam Komentar Negatif Tokopedia

Berdasarkan Gambar 6 dan Gambar 7, representasi perolehan hasil frekuensi kemunculan kata paling banyak dalam komentar positif dan negatif pada *marketplace online* Tokopedia ditunjukkan pada Tabel 13.

Tabel 13 Term Frequency pada Komentar Tokopedia

Label	Kata	Frekuensi
Komentar Positif	mantap	35
	hadiah	33
	sukses	25
Komentar Negatif	eror	28
	payah	27
	tipu	20

## KESIMPULAN

Kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut: (1) Hasil akurasi paling optimal dalam klasifikasi *Support Vector Machine* menggunakan *kernel Radial Basis Function* (RBF) didapatkan nilai sebesar 80,5% untuk komentar Bukalapak, 74% untuk komentar Shopee dan 82% untuk komentar Tokopedia dengan parameter nilai gamma  $\gamma = 0,01$  dan *fold*  $K = 10$ , (2) Tingkat persentase komentar positif dan negatif pada setiap *fans page Facebook marketplace online* didapatkan untuk Bukalapak, komentar positif sebanyak 42,5% dan komentar negatif sebanyak 57,5%. Sedangkan untuk Shopee, komentar positif sebanyak 45% dan komentar negatif sebanyak 55%. Serta untuk Tokopedia, komentar positif sebanyak 37% dan komentar negatif sebanyak 63%, (3) Kata-kata dalam komentar *fans page Facebook marketplace online* yang mendominasi dari hasil analisis yaitu, pada Bukalapak sentimen positif antara lain: sukses, mantap, cepat dan sentimen negatif antara lain: tipu, bohong, komplain, sedangkan pada Shopee sentimen positif antara lain: gratis, alhamdulillah, murah dan sentimen negatif antara lain: tipu, susah, kecewa, serta pada Tokopedia sentimen positif antara lain: mantap, hadiah, sukses dan sentimen negatif antara lain: eror, payah, tipu.



Beberapa saran yang dapat dikembangkan untuk penelitian selanjutnya adalah sebagai berikut: (1) Perlu dikembangkan penelitian yang berfokus pada proses *pre-processing* untuk data komentar yang menggunakan banyak bahasa tidak baku, singkatan, dan lain sebagainya, (2) Perlu dikembangkan menggunakan jumlah data komentar dan *validator* pelabelan data yang lebih banyak untuk lebih mendapatkan banyak kata-kata dan keakuratan dalam pelabelan data sehingga hasil yang didapatkan lebih maksimal, (3) Perlu dikembangkan menggunakan jenis *kernel* SVM non linear lainnya sebagai bahan perbandingan atau kombinasi *kernel* dalam menganalisis.

#### DAFTAR PUSTAKA

- APJII, 2016, Penetrasi & Perilaku Pengguna Internet Indonesia 2016, (Online), (<https://apjii.or.id/download/file/surveipenetrasiinternet2016.pdf>), Diakses tanggal 20 Januari 2018.
- Cahya, P., 2018, 10 *E-Commerce* Terbaik yang Menggeser Eksistensi Toko *Offline*, (Online), (<https://www.idntimes.com/business/economy/putriana-cahya/10-e-commerce-terbaik-yang-menggeser-eksistensi-toko-offline>), Diakses tanggal 18 Februari 2019.
- DeHaff, M., 2010, *Sentiment Analysis, Hard But Worth It!*, (Online), ([http://www.customerthink.com/sentiment\\_analysis\\_hard\\_but\\_worth\\_it](http://www.customerthink.com/sentiment_analysis_hard_but_worth_it)), Diakses tanggal 24 Januari 2018.
- Havrlant, L., dan Kreinovich, V., 2014, *A Simple Probabilistic Explanation of term Frequency-Inverse Document Frequency (TF-IDF) Heuristic (and Variations Motivated by This Explanation)*, University of Texas at El Paso.
- Hidayatullah, A. F., dan SN, A., 2014, Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik pada Twitter, Seminar Nasional Informatika 2014 UPN “Veteran” Yogyakarta.
- Hsu, C. W., Chang, C. C., dan Lin C. J., 2003, *A Practical Guide to Support Vector Classification*, Department of Computer Science, National Taiwan University.
- Jünger, J. dan Keyling, T., 2018, *Facepager. An application for generic data retrieval through APIs*, Source code and releases available at <https://github.com/strohne/Facepager/>.
- Lidya, S. K., 2014, *Sentiment Analysis* pada Teks Bahasa Indonesia Menggunakan *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (K-NN), Tesis, Universitas Sumatera Utara.
- Liu, B., 2012, *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers.
- Mulyana, I., Ramadona, S., Herfina, 2012, Penerapan *Terms Frequency-Inverse Document Frequency* pada Sistem Peringkasan Teks Otomatis Dokumen Tunggal Berbahasa Indonesia, Jurnal.
- Nababan, C. N., 2017, Belanja Online Masyarakat Indonesia Tembus Rp75 Triliun, (Online), (<https://www.cnnindonesia.com/ekonomi/20170809151902-78-233513/belanja-online-masyarakat-indonesia-tembus-rp75-triliun>), Diakses tanggal 12 Januari 2018.
- Nugroho, A. S., 2008, Pengantar *Support Vector Machine*, (Online), ([asnugroho.net/papers/tutorialsvm\\_as\\_nugroho.ppt](http://asnugroho.net/papers/tutorialsvm_as_nugroho.ppt)), Diakses tanggal 20 Januari 2018.
- Pang, B., Lee, L., dan Vithyanathan, S., 2002, *Thumbs Up ? Sentiment Classification Using Machine Learning Techniques. Proceedings of The ACL-02 conference on Empirical methods in natural language processing (pp. 79-86)*, Stroudsburg: Association for Computational Linguistic 4.
- Prakoso, R.W., Novianty, A., dan Setianingsih, C., 2017, Analisis Sentimen Menggunakan *Support Vector Machine* dan *Maximum Entropy*, *E-proceeding of Engineering* : Vol. 4, No. 2 Agustus 2017 halaman 2389.

- Refaeilzadeh, P., Tang, L., Liu, H., 2009, *Cross-Validation*, Arizona State University.
- Sadgotra, W. Y. dan Saputra, E. H., 2013, Perancangan Online Marketplace untuk Usaha Kecil dan Menengah (UKM) di Kabupaten Purworejo, Teknik Informatika STMIK AMIKOM Yogyakarta.
- Suyanto, 2018, *Machine Learning* Tingkat Dasar dan Lanjut, Bandung, Informatika.
- Triawati, C., Bijaksana, M. A., dan Baizal, Z.A, 2009, Metode Pembobotan *Statistical Concept Based* untuk *Clustering* dan Kategorisasi Dokumen Berbahasa Indonesia, Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom.
- Visa, S., Ramsay, B., Ralescu, A., dan Knaap, E. V. D., 2011, *Confusion Martix-based Feature Selection*, *Proceedings of The 22nd Midwest Artificial Intelligence and Cognitive Science Conference 2011*.

