

TREN *MARKETPLACE* BERDASARKAN KLASIFIKASI ULASAN PELANGGAN MENGGUNAKAN PERBANDINGAN KERNEL *SUPPORT VECTOR MACHINE*

Dwi Latifah Rianti¹, Yuyun Umaidah², Apriade Voutama³

Program Studi Teknik Informatika, Universitas Singaperbangsa Karawang^{1,2,3}
dwi.latifah17001@student.unsika.ac.id¹, yuyun.umaidah@staff.unsika.ac.id²,
apriade.voutama@staff.unsika.ac.id³

Submitted June 24, 2021; Revised June 24, 2021; Accepted August 1, 2021

Abstrak

Saat ini banyak masyarakat Indonesia yang gemar melakukan transaksi perdagangan secara *online*. Akan tetapi sejumlah pelaku bisnis kesulitan untuk memilih *marketplace* untuk memasarkan produknya. Salah satu penyebabnya karena mereka jarang memperhatikan tren *marketplace* yang sedang diperbincangkan konsumen. Oleh karena itu, menganalisis tren di media sosial seperti Twitter, menjadi sangat penting bagi para pebisnis untuk memahami pola kecenderungan konsumen terhadap jasa atau produknya. Maka tujuan penelitian ini yaitu membuat model yang dapat menganalisis tren *marketplace* berdasarkan klasifikasi ulasan pelanggan di Twitter menggunakan algoritme SVM. Kernel yang digunakan yaitu *linear*, RBF, *sigmoid*, dan *polynomial* dengan optimasi parameter menggunakan *grid search*. Metodologi yang digunakan yaitu KDD. Hasil evaluasi model klasifikasi terbaik yaitu kernel *sigmoid* dengan nilai *accuracy* 92%, *precision* 92%, *recall* 92%, dan *F1 score* 92% serta parameter $C=100$, $\gamma=0.01$, dan $r=1$. Hasil tren *marketplace* berdasarkan persentase ulasan positif tertinggi adalah Tokopedia, Shopee, dan yang terakhir Bukalapak.

Kata Kunci : Kernel SVM, Klasifikasi, *Marketplace*, *Support Vector Machine*

Abstract

Currently, many Indonesian people like to conduct online trading transactions. However, a number of business people find it difficult to choose a marketplace to market their products. One of the reasons is because they rarely pay attention to the marketplace trends that consumers are discussing. Therefore, analyzing trends on social media such as Twitter, it becomes very important for business people to understand the pattern of consumer tendencies towards their services or products. So the purpose of this study is to create a model that can analyze marketplace trends based on the classification of customer reviews on Twitter using the SVM algorithm. The kernels used are linear, RBF, sigmoid, and polynomial with parameter optimization using grid search. The methodology used is KDD. The results of the evaluation of the best classification model are the sigmoid kernel with 92% accuracy, 92% precision, 92% recall, and 92% F1 score and parameters $C=100$, $\gamma=0.01$, and $r=1$. Market trend results based on the highest percentage of positive reviews are Tokopedia, Shopee, and lastly Bukalapak.

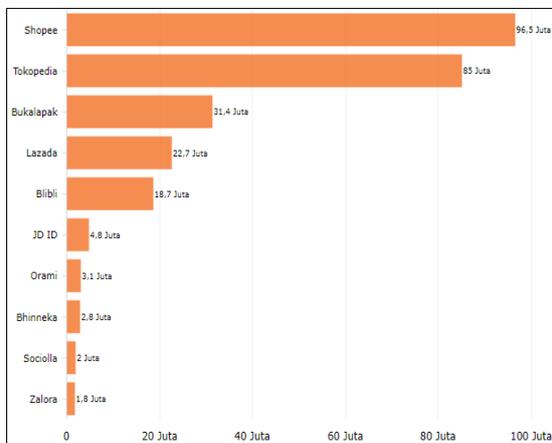
Key Words : Classification, *Marketplace*, *Support Vector Machine*, SVM Kernel

1. PENDAHULUAN

Saat ini banyak pelaku bisnis dan masyarakat Indonesia yang gemar melakukan transaksi perdagangan secara *online* melalui *e-commerce*. Hal ini dikarenakan berbagai kelebihan yang, salah satunya yaitu wilayah dan waktu yang tidak terbatas [1]. Fenomena ini

didukung dengan data dari *We Are Social* pada situs *teknioia.com*, yang menunjukkan bahwa persentase pengguna *e-commerce* di Indonesia menduduki peringkat pertama dibandingkan beberapa negara di dunia dengan persentase mencapai 88% pada tahun 2020 [2].

Gambar 1 merupakan grafik data dari iPrice mengenai rata-rata pengguna yang mengunjungi beberapa website *e-commerce* Indonesia setiap bulannya di kuartal III tahun 2020 [3].



Sumber: databoks.katadata.co.id, 2020

Gambar 1. Rata-rata Kunjungan Web E-commerce (Kuartal III-2020)

Dapat dilihat pada Gambar 1 bahwa tiga urutan kunjungan pengguna terbanyak diraih oleh jenis *e-commerce marketplace*, yaitu Shopee dengan 96,5 juta, Tokopedia 85 juta, dan Bukalapak urutan ketiga dengan 31,4 juta pengunjung.

Berdasarkan hasil pengamatan, sejumlah pelaku bisnis kesulitan untuk memilih *marketplace* yang akan mereka pilih untuk memasarkan produknya. Salah satu penyebabnya karena mereka jarang memperhatikan tren *marketplace* yang sedang diperbincangkan konsumen. Akibatnya, produk hanya berhasil terjual sedikit. Banyak pengguna membahas tren *marketplace* di media sosial, satu diantaranya yaitu Twitter. Twitter saat ini merupakan media sosial yang populer dan banyak digunakan dikalangan masyarakat. Hal ini berdasarkan laporan keuangan Twitter yang mencatat peningkatan pengguna harian sebesar 17% pada kuartal ketiga tahun 2019 menjadi 145 juta [4].

Menganalisis tren di Twitter sangat penting bagi para pebisnis agar dapat memahami pola kecenderungan konsumen pada jasa

atau produknya. Analisis tren juga dibutuhkan untuk membantu dalam peningkatan efisiensi pemasaran dan juga *forecasting* [4]. Hal ini dapat dilakukan dengan menggunakan analisis sentimen pada ulasan pelanggan *marketplace* di media sosial Twitter. Analisis sentimen merupakan bidang ilmu yang dapat menemukan tingkat kepuasan pelanggan mengenai suatu produk atau jasa. Hal ini dilakukan dengan mengklasifikasikan polaritas ulasan ke dalam bentuk opini positif atau negatif [5].

Maka berdasarkan permasalahan di atas, diperlukan suatu model yang dapat menganalisis tren *marketplace* berdasarkan klasifikasi ulasan pelanggan pada *platform* Twitter. Berbagai penelitian mengenai klasifikasi ulasan pelanggan yang telah dilakukan oleh [6]–[8], algoritme *Support Vector Machine* (SVM) digunakan untuk pemodelannya. Hal tersebut dikarenakan SVM merupakan algoritme yang akurat karena mampu membuat model dengan berbagai dimensi fitur (*linear* dan *nonlinear*). SVM juga jauh lebih tidak rentan terhadap *over-fitting* data dibandingkan algoritme lainnya [9]. Keakuratan SVM dapat dilihat dari penelitian Singla et al. mengenai ulasan produk ponsel web *Amazon.com* dengan jumlah data 400.000 untuk 4500 jenis ponsel. Penelitian ini menghasilkan algoritme SVM merupakan algoritme terbaik dengan akurasi 81.75%, dibandingkan *Decision Tree* sebesar 74.76% dan *Naive Bayes* 66.95% [10]. Penelitian lainnya juga dilakukan oleh Lutfi et al. yang membandingkan algoritme SVM dengan *Naive Bayes* untuk menganalisis ulasan pada *marketplace* BukaLapak. Hasil akurasi tertinggi yaitu algoritme SVM dengan nilai 93.65%, sedangkan algoritme *Naive Bayes* hanya 89.21% [6].

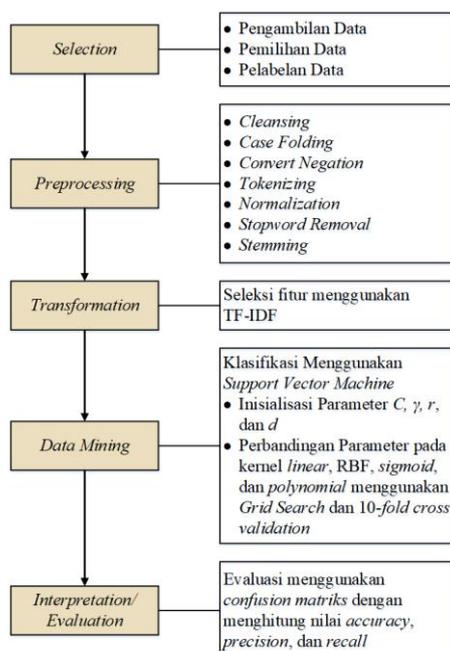
Meskipun SVM memiliki beberapa kelebihan, kinerja algoritme ini sangat

bergantung pada kernel yang dipilihnya. Sehingga, pemilihan kernel pada SVM sangat penting dilakukan karena dapat menentukan *feature space* yaitu tempat untuk mencari fungsi *classifier*. Teknik yang dapat digunakan untuk menentukan kernel dalam SVM yaitu *kernel trick*. Adapun beberapa jenis kernel yang terdapat pada SVM diantaranya yaitu kernel *linear*, *Radial Basis Function (RBF)*, *polynomial* dan *sigmoid* [8].

Sehingga penelitian ini memiliki tujuan untuk melakukan perbandingan kernel SVM pada model klasifikasi ulasan pelanggan *marketplace* di media sosial Twitter. Manfaat dari hasil penelitian ini diharapkan mampu membantu para pebisnis untuk menentukan *marketplace* yang sesuai dan menjadi evaluasi bagi pengelola *marketplace* agar dapat meningkatkan kepercayaan pelanggannya.

2. METODE PENELITIAN

Metode yang digunakan pada penelitian ini yaitu *Knowledge Discovery in Database (KDD)* dengan rancangan penelitian yang ditunjukkan pada Gambar 2.



Gambar 2. Rancangan Penelitian

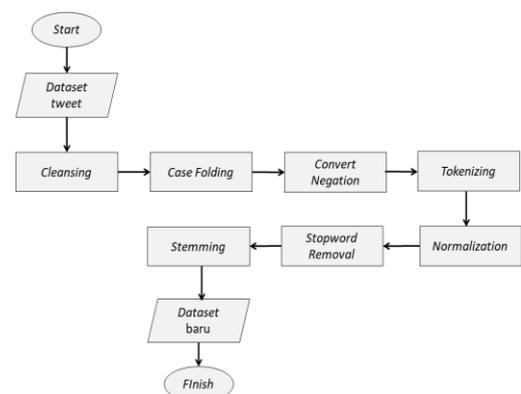
1. Data Selection

Tahap *data selection* merupakan tahap untuk pengambilan, pemilihan, dan pelabelan data. Pengambilan data dilakukan dengan proses *crawling* dari media sosial Twitter menggunakan API Twitter.

Pengambilan data dilakukan dengan menggunakan *tools* Jupyter Notebook dan bahasa pemrograman Python. Data yang diambil yaitu ulasan pelanggan dalam Bahasa Indonesia mengenai *marketplace* Shopee, Tokopedia, dan BukaLapak dari media sosial Twitter. Data diambil dengan *keyword* @ShopeeID, @ShopeeCare, @ShopeePay_ID, @Tokopedia, @TokopediaCare, @BukaLapak, dan @BukaBantuan. Data hasil *crawling* kemudian dicek dan dipilih untuk diproses ke tahap berikutnya. Kemudian data tersebut dikategorikan menjadi kelas sentimen positif atau negatif, selanjutnya data akan divalidasi oleh ahli Bahasa Indonesia.

2. Preprocessing

Tahap *preprocessing* merupakan proses untuk menyesuaikan data dengan kebutuhan penelitian. Gambar 3 merupakan alur dari tahap *preprocessing* yang dilakukan.



Gambar 3. Tahap Preprocessing

Penjelasan dari Gambar 3 yaitu sebagai berikut:

- Cleansing*: membersihkan atribut dan karakter yang tidak berpengaruh terhadap penelitian..

- b. *Case Folding*: perubahan huruf kapital menjadi huruf kecil (*lower case*) pada dokumen..
- c. *Convert Negation*: mengkonversi kata-kata negasi dengan menggabungkannya ke dalam kata berikutnya.
- d. *Tokenizing*: pemotongan kalimat menjadi potongan kata tunggal untuk keperluan pembobotan kata.
- e. *Normalization*: memperbaiki kata-kata yang tidak sesuai pada dokumen.
- f. *Stopword Removal*: menghilangkan kata yang tidak berpengaruh terhadap penelitian.
- g. *Stemming*: mengubah kata berimbuhan pada dokumen menjadi kata dasar.

3. Transformation

Tahap ini akan mengubah data yang berupa nominal menjadi numerik dengan menerapkan teknik seleksi fitur TF-IDF. TF-IDF bekerja dengan melakukan pembobotan terhadap kata yang muncul pada dokumen. Pemilihan metode TF-IDF dikarenakan hasil evaluasi yang lebih baik dan juga waktu eksekusi yang lebih cepat dibandingkan seleksi fitur lain [11].

4. Data mining

Tahap *data mining* yaitu pencarian pola menggunakan teknik klasifikasi. Algoritme yang digunakan yaitu *Support Vector Machine* dengan kernel *linear*, *polynomial*, *RBF*, dan *sigmoid*. Penelitian ini juga akan menerapkan *Grid Search* untuk menemukan nilai parameter yang terbaik pada masing-masing kernel dengan rasio pembagian data *training* dan data *testing* yaitu 80:20 serta menerapkan *10-fold cross validation*. Adapun parameter yang akan digunakan yaitu C (*cost*), γ (*gamma*), r (*coefficient*), dan d (*degree*). Penentuan parameter optimasi pada masing-masing kernel ditunjukkan pada Tabel 1.

Tabel 1. Parameter Optimasi

No.	Kernel	Parameter Optimasi
1	Linear	C
2	RBF	C dan γ
3	Sigmoid	C, γ , dan r
4	Polynomial	C, γ, r , dan d

Sedangkan untuk nilai parameter yang akan digunakan pada penelitian ini akan disajikan pada Tabel 2.

Tabel 2. Parameter Optimasi

Parameter	Nilai
C	0.001, 0.01, 0.1, 1, 10, 100
γ	0.001, 0.01, 0.1, 1, 10, 100
r	0, 1, 5, 8, 10
d	0, 1, 2, 3

Penentuan nilai parameter tersebut didasarkan pada beberapa penelitian sebelumnya yaitu penelitian Tuttle et al. [12] menggunakan parameter $C = \{0.001, 0.01, 0.1, 1, 10, 100\}$, $\gamma = \{0.001, 0.01, 0.1, 1, 10, 100\}$, $r = \{0, 1, 5, 8, 10\}$, dan $d = \{2, 3, 4, 5, 6, 7, 8\}$. Kemudian penelitian Rudd [13] menggunakan parameter $C = \{0.001, 0.01, 0.1, 1, 10\}$, $\gamma = \{0.001, 0.01, 0.1, 1\}$, dan $d = \{2, 3\}$. Sedangkan penelitian Nanda et al. [14] parameter yang digunakan adalah $C = \{2^{-15}, 2^{-14}, \dots, 2^2\}$, $r = \{2^{-10}, 2^{-9}, \dots, 2^2\}$, $\gamma = \{2^{-10}, 2^{-9}, \dots, 2^2\}$, dan $d = \{0, 1, 2, 3\}$. Maka untuk penelitian ini telah dilakukan beberapa penyesuaian dan memilih nilai yang memungkinkan untuk digunakan.

5. Interpretation/Evaluation

Tahap ini akan menampilkan hasil dari penelitian yang telah dilakukan berupa urutan tingkat kepopuleran *marketplace* berdasarkan persentase ulasan pelanggan pada media sosial Twitter. Selain itu akan dilakukan juga evaluasi model yang telah dibuat menggunakan *confusion matrix* dengan perhitungan nilai *accuracy*, *precision*, *recall*, dan *f1 score*.

3. HASIL DAN PEMBAHASAN

1. Data Selection

Data yang diambil merupakan ulasan pelanggan terhadap *marketplace* Shopee, Tokopedia, dan BukaLapak pada media sosial Twitter dalam bahasa Indonesia. Proses pengambilan data dilakukan mulai tanggal 10 Februari 2021 sampai dengan 8 April 2020. Pengambilan data dilakukan setiap tujuh hari sekali sebanyak kurang

lebih 3000 *tweet/marketplace*. Data yang berhasil diperoleh dari hasil *crawling* berjumlah 91.033 *tweet*.

Data tersebut kemudian diseleksi dengan menghapus data duplikat dan mengambil *tweet* berupa ulasan pelanggan. Dari total 91.033 data hasil *crawling*, diperoleh data mengenai ulasan pelanggan sebanyak 4.699 *tweet*. Selanjutnya *tweet* tersebut diberikan label berupa ulasan positif atau negatif. Proses pelabelan dilakukan secara manual kemudian divalidasi oleh Ahli Bahasa Indonesia dengan rincian data disajikan pada Tabel 3.

Tabel 3. Rincian Data Hasil Validasi

Objek	Jumlah	Positif	Negatif
Shopee	1.543	590	953
Tokopedia	1.609	464	1.145
BukaLapak	1.547	769	778
TOTAL	4.699	1.823	2.876

Adapun contoh data yang telah diseleksi dan divalidasi oleh Ahli Bahasa Indonesia dapat dilihat pada Tabel 4.

Tabel 4. Contoh Data Hasil Validasi

Tweet	Object	Sentiment
b@jntexpressid gw kan salah kasih alamat waktu beli di @bukalapak terus udah telpon ke jnt bisa diganti kealamat yg sesuai malah ga nyampe paket gw cek dm min'	Bukalapak	Negatif
b@ShopeeCare halo min mau tanya ini tuh kenapa ya? Padahal limit masih banyak dan tagihan bulan ini juga sudah payment. Terimakasih https://t.co/mrpqVaF86p'	Shopee	Negatif
b@tokopedia Makasih min, udh ngundang 5 bujang tiang listrik moa \xf0\x9f\x98\x87'	Tokopedia	Positif

2. Preprocessing

Data yang telah diseleksi kemudian akan melalui tahap *preprocessing* untuk membersihkan data dan mengurangi atribut yang kurang berpengaruh terhadap proses klasifikasi. Tabel 5 adalah contoh dari *text preprocessing* yang telah dilakukan.

Tabel 5. Contoh Text Preprocessing

Tahap Preprocessing	Hasil
Dataset Awal	b'@bukalapak Betul, pihak bikalapak meminta link produk yg di blokir. Sedangkan produk yg di blokir tdk bisa di klik, lalu ambil link produknya bagaimana ?'
Cleansing	Betul pihak bikalapak meminta link produk yg di blokir Sedangkan produk yg di blokir tdk bisa di klik lalu ambil link produknya bagaimana
Case Folding	<u>betul</u> pihak bikalapak meminta link produk yg di blokir <u>sedangkan</u> produk yg di blokir tdk bisa di klik lalu ambil link produknya bagaimana
Convert Negation	betul pihak bikalapak meminta link produk yg di blokir sedangkan produk yg di blokir <u>tdkbisa</u> di klik lalu ambil link produknya bagaimana
Tokenizing	['betul', 'pihak', 'bikalapak', 'meminta', 'link', 'produk', 'yg', 'di', 'blokir', 'sedangkan', 'produk', 'yg', 'di', 'blokir', 'tdkbisa', 'di', 'klik', 'lalu', 'ambil', 'link', 'produknya', 'bagaimana']
Normalization	['betul', 'pihak', ' <u>bukalapak</u> ', 'meminta', ' <u>tautan</u> ', 'produk', ' <u>yang</u> ', 'di', 'blokir', 'sedangkan', 'produk', ' <u>yang</u> ', 'di', 'blokir', ' <u>tidakbisa</u> ', 'di', 'klik', 'lalu', 'ambil', ' <u>tautan</u> ', 'produknya', 'bagaimana']
Stopword Removal	['bukalapak', 'tautan', 'produk', 'blokir', 'produk', 'blokir', 'tidakbisa', 'klik', 'ambil', 'tautan', 'produknya']
Stemming	['bukalapak', ' <u>taut</u> ', 'produk', 'blokir', 'produk', 'blokir', 'tidakbisa', 'klik', 'ambil', ' <u>taut</u> ', ' <u>produk</u> ']

3. Transformation

Tahap *transformation* dilakukan dengan pencarian nilai TF-IDF. Gambar 4 menunjukkan hasil dari TF-IDF.

	Nilai TF	Nilai IDF
(0, 191)	0.07142857142857142	8.76217061 8.76217061
(0, 244)	0.07142857142857142	6.51087881 8.76217061
(0, 368)	0.07142857142857142	8.06902343 5.09860896
(0, 373)	0.07142857142857142	7.25809321 8.3567055
(0, 496)	0.07142857142857142	8.76217061 8.76217061
(0, 658)	0.07142857142857142	8.76217061 6.31982357
(0, 1360)	0.07142857142857142	8.76217061 8.3567055
(0, 1942)	0.14285714285714285	8.76217061 8.3567055
Nilai TFIDF		
(0, 191)	0.26269039031967806	
(0, 244)	0.21329053363588407	
(0, 368)	0.5268483032870224	
(0, 373)	0.26940985859710487	
(0, 496)	0.32936647865996843	
(0, 658)	0.3146758414606152	
(0, 1360)	0.2672560701563415	
(0, 1942)	0.7890703078534034	

Gambar 4. Hasil Nilai TF dan IDF

4. Data mining

Tahap *data mining* merupakan tahap pemodelan menggunakan algoritme SVM. Pembagian *data training* dan *testing* dilakukan dengan skala 80:20 dan menerapkan *10-fold cross validation* untuk pembagiannya. Tabel 6 menunjukkan jumlah *dataset* penelitian.

Tabel 6. Pembagian Dataset Penelitian

Data Training (80%)	Data Testing (20%)
3759	940

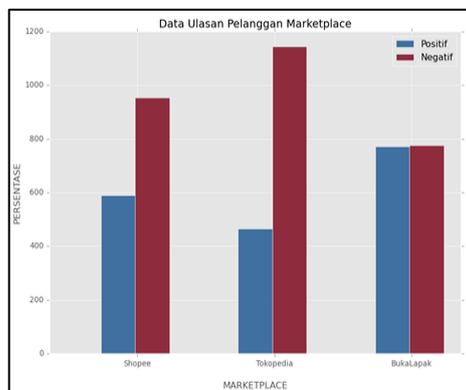
5. Interpretation/Evaluation

Gambar 5 menunjukkan hasil evaluasi masing-masing kernel pada model klasifikasi ulasan pelanggan *marketplace* yang telah dibuat.

Kernel SVM	Hasil Evaluasi (%)				Best Parameter
	Accuracy	Precision	Recall	F1 Score	
Linear	92	92	92	91	$C = 0,1$
RBF	92	92	92	91	$C = 10;$ $\gamma = 0,01$
Sigmoid	92	92	92	92	$C = 100;$ $\gamma = 0,01;$ $r = 1$
Polynomial	91	91	91	91	$C = 0,001;$ $\gamma = 1;$ $r = 10;$ $d = 3$

Gambar 5. Akumulasi Hasil Evaluasi

Berdasarkan Gambar 5 diperoleh bahwa model terbaik yaitu menggunakan kernel *sigmoid* dengan parameter $C=100$, $\gamma=0.01$, dan $r=1$. Hasil evaluasi dari model tersebut memiliki nilai *accuracy* 92%, *precision* 92%, *recall* 92%, dan *F1 score* 92%. Selanjutnya pada Gambar 6 disajikan grafik mengenai data ulasan pelanggan pada masing-masing *marketplace*.



Gambar 6. Perbandingan Data Masing-Masing Marketplace

Berdasarkan Gambar 6, dapat disimpulkan bahwa ulasan pelanggan positif terbanyak diraih oleh *marketplace* BukaLapak dengan total sebanyak 769 *tweet*. Sedangkan untuk jumlah ulasan pelanggan negatif tertinggi dimiliki oleh *marketplace* Tokopedia dengan jumlah 1.145 *tweet*. Persentase ulasan pelanggan masing-masing *marketplace* dapat dilihat pada Tabel 7.

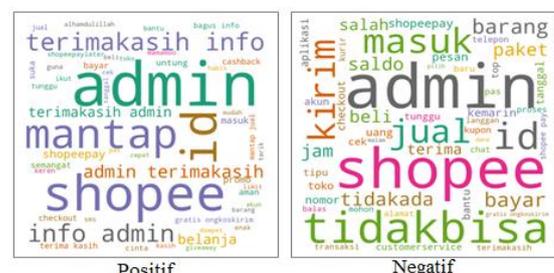
Tabel 7. Persentase Ulasan Pelanggan Marketplace

Objek	Positif	Negatif
Shopee	38,24%	61,76%
Tokopedia	28,84%	71,16%
BukaLapak	49,71%	50,29%

Tabel 8 menunjukkan bahwa tren *marketplace* berdasarkan persentase ulasan positif tertinggi yaitu BukaLapak dengan nilai 49,71%, pada urutan kedua yaitu *marketplace* Shopee dengan nilai 38,24%, sedangkan Tokopedia berada pada urutan ketiga dengan persentase sebanyak 28,84%.

Ulasan pelanggan pada masing-masing *marketplace* dapat divisualisasikan dalam bentuk *wordcloud*. Visualisasi ini dapat digunakan untuk mengetahui gambaran atau informasi umum mengenai data ulasan pelanggan *marketplace* pada media sosial Twitter. Ukuran suatu kata pada visualisasi *wordcloud* menunjukkan besar frekuensi kemunculan kata pada suatu dokumen.

Kata yang sering muncul pada ulasan Shopee akan disajikan pada Gambar 7.



Gambar 7. Visualisasi Wordcloud Ulasan Pelanggan Shopee

Berdasarkan Gambar 7 dapat dilihat bahwa ulasan positif Shopee didominasi

pelanggan yang mengucapkan terimakasih kepada admin karena sering memberikan informasi yang bermanfaat. Selain itu terdapat juga beberapa kata yang menunjukkan berbagai strategi pemasaran Shopee yang mendapat respons positif dari para pelanggannya seperti *giveaway* dan gratis ongkos kirim. Sedangkan ulasan negatif didominasi oleh pelanggan yang menyampaikan kendala dalam pengiriman suatu produk atau *top up* ShopeePay yang tidak bisa masuk.

Sedangkan ulasan positif dan negatif pada Tokopedia dapat dilihat pada Gambar 8.



Gambar 8. Visualisasi Wordcloud Ulasan Pelanggan Tokopedia

Berdasarkan Gambar 8 dapat dilihat bahwa ulasan positif didominasi dengan respons positif pelanggan terhadap program Waktu Indonesia Belanja (WIB) yang diadakan oleh Tokopedia. Program WIB pada bulan Maret banyak dinantikan oleh para pelanggannya karena dihadiri oleh *boygroup* Korea Selatan yaitu TXT. Sedangkan ulasan negatif menunjukkan bahwa pelanggan Tokopedia banyak mengalami kendala dalam hal pembayaran, penjualan, dan barang yang dibeli/dijual.

Wordcloud ulasan positif dan negatif pada Bukalapak akan disajikan pada Gambar 9.



Gambar 9. Visualisasi Wordcloud Ulasan Pelanggan Bukalapak

Berdasarkan Gambar 9 dapat dilihat ulasan positif *marketplace* Bukalapak berkaitan dengan respons positif pelanggan terhadap program Rabu Lokal yang diadakan oleh Bukalapak. Hal ini dikarenakan program Rabu Lokal merupakan upaya Bukalapak untuk mendukung produk lokal dan UMKM Indonesia. Sedangkan untuk ulasan negatif menunjukkan bahwa pelanggan bukalapak banyak yang mengalami kendala dalam transaksi, pengiriman, pembelian, dan barang yang dijual/dibeli.

4. SIMPULAN

Berdasarkan penelitian yang telah dilakukan, model dengan hasil evaluasi terendah yaitu menggunakan kernel *polynomial* yang memiliki *accuracy* 91%, *precision* 91%, *recall* 91%, dan *f1 score* 91%. Urutan selanjutnya yaitu kernel *linear* dan RBF dengan nilai yang sama yaitu *accuracy* 92%, *precision* 92%, *recall* 92%, dan *f1 score* 91%. Model dengan hasil terbaik yaitu kernel *sigmoid* dengan parameter $C=100$, $\gamma=0.01$, dan $r=1$. Hasil evaluasi model *sigmoid* yaitu *accuracy* 92%, *precision* 92%, *recall* 92%, dan *F1 score* 92%. Sehingga dapat ditarik kesimpulan bahwa pemilihan kernel pada algoritme SVM dapat mempengaruhi hasil akhir klasifikasi.

Hasil analisis tren *marketplace* pada media sosial Twitter berdasarkan persentase ulasan positif tertinggi diraih oleh BukaLapak dengan nilai 49,71%, urutan kedua yaitu Shopee dengan nilai 38,24%, sedangkan Tokopedia berada pada urutan ketiga dengan persentase 28,84%.

DAFTAR PUSTAKA

[1] Muljono, D. P. Artanti, A. Syukur, A. Prihandono, and D. R. I. M. Setiadi, "Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes," *Konf. Nas. Sist. Inf.*

- 2018, pp. 165–170, 2018.
- [2] B. Ramadhan, “Data Internet di Indonesia dan Perilakunya Tahun 2020,” *teknioa.com*, 2020. <https://teknioa.com/data-internet-di-indonesia-dan-perilakunya-880c7bc7cd19> (accessed Nov. 24, 2020).
- [3] D. H. Jayani, “Shopee, E-Commerce dengan Pengunjung Situs Tertinggi Kuartal III 2020,” *databooks*, 2020. <https://databoks.katadata.co.id/datapublish/2020/11/20/shopee-e-commerce-dengan-pengunjung-situs-tertinggi-kuartal-iii-2020> (accessed Nov. 24, 2020).
- [4] U. A. Nasron and M. Habibi, “Analysis of Marketplace Conversation Trends on Twitter Platform Using K-Means,” *Compiler*, vol. 9, no. 1, pp. 51–61, 2020, doi: 10.28989/compiler.v9i1.579.
- [5] D. Pajri, Y. Umaidah, and T. N. Padilah, “K-Nearest Neighbor Berbasis Particle Swarm Optimization untuk Analisis Sentimen Terhadap Tokopedia,” *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 242–253, 2020, doi: <http://dx.doi.org/10.28932/jutisi.v6i2.2658>.
- [6] A. A. Lutfi, A. E. Permanasari, and S. Fauziati, “Corrigendum: Sentiment Analysis in the Sales Review of Indonesian Marketplace by Utilizing Support Vector Machine,” *J. Inf. Syst. Eng. Bus. Intell.*, vol. 4, no. 1, pp. 57–64, 2018, doi: <http://dx.doi.org/10.20473/jisebi.4.1.57-64>.
- [7] J. H. Jaman and R. Abdulrohman, “Sentiment Analysis of Customers on Utilizing Online Motorcycle Taxi Service at Twitter with the Support Vector Machine,” *Int. Conf. Electr. Eng. Comput. Sci.* 2019, pp. 231–234, 2019.
- [8] W. D. Wicaksono, “Klasifikasi Analisis Sentimen Movie Review dengan Metode Support Vector Machine Menggunakan Kernel Radial Basis Function dan Information Gain,” Thesis. Universitas Negeri Semarang, Semarang, 2020.
- [9] J. Han, M. Kamber, and J. Pei, *Data mining: Data mining concepts and techniques*, Third Edit. Waltham: Morgan Kaufmann, 2012.
- [10] Z. Singla, S. Randhawa, and S. Jain, “Sentiment Analysis on Product Reviews using Machine Learning Techniques,” *2017 Int. Conf. Intell. Comput. Control*, pp. 1–5, 2017, doi: 10.1109/I2C2.2017.8321910.
- [11] M. N. Akbar, “Klasifikasi Bibliografi Otomatis Menggunakan C4.5 dan Information Gain,” *J. Inform. Sains dan Teknol.*, vol. 6, no. 1, pp. 46–55, 2021, doi: <https://doi.org/10.24252/instek.v6i1.18636>.
- [12] J. F. Tuttle, L. D. Blackburn, and K. M. Powell, “On-line Classification of Coal Combustion Quality using Nonlinear SVM for Improved Neural Network NOx Emission Rate Prediction,” *Comput. Chem. Eng.*, vol. 141, pp. 1–11, 2020, doi: 10.1016/j.compchemeng.2020.106990.
- [13] J. M. Rudd, “Application of Support Vector Machine Modeling and Graph Theory Metrics for Disease Classification,” *Model Assist. Stat. Appl.*, vol. 13, pp. 341–349, 2018, doi: 10.3233/MAS-180444.
- [14] M. A. Nanda, K. B. Seminar, D. Nandika, and A. Maddu, “A comparison study of kernel functions in the support vector machine and its application for termite detection,” *Information*, vol. 9, no. 5, pp. 1–14, 2018, doi: 10.3390/info9010005.