

# Klasifikasi Keluhan Masyarakat pada Sosial Media Twitter terhadap Pelayanan Toko Online di Indonesia menggunakan Metode *Cosine TF-IDF*

Iwan Syarif<sup>1</sup>, Rengga Asmara<sup>1,\*</sup>, Nur Ulima Rusmayani<sup>1</sup>

<sup>1</sup> Departemen Teknik Informatika dan Komputer; Politeknik Elektronika Negeri Surabaya; Kampus PENS, Jalan Raya ITS, Keputih, Sukolilo, Surabaya; Telp. +62 31 594 7280, Fax +62 31 594 6114; e-mail: [iwanarif@pens.ac.id](mailto:iwanarif@pens.ac.id), [rengga@pens.ac.id](mailto:rengga@pens.ac.id), [nrmayani2811@gmail.com](mailto:nrmayani2811@gmail.com)

Korespondensi email: [rengga@pens.ac.id](mailto:rengga@pens.ac.id)

Diterima: 23 Maret 2020; Review: 29 Maret 2020; Disetujui: 2 April 2020

Cara sitasi: Syarif, I. Rengga, A. Nur UR. 2020. Klasifikasi Keluhan Masyarakat pada Sosial Media Twitter terhadap Pelayanan Toko Online di Indonesia menggunakan Metode *Cosine TF-IDF*. BINA INSANI ICT Journal. Vol.7 (1): 33-42.

**Abstrak:** Berkembangnya toko online dan transaksi online di Indonesia pada saat ini diiringi dengan berbagai permasalahan seperti keluhan pada pelayanan yang membahas mengenai aplikasi, ketanggapan dan pengiriman. Dengan adanya permasalahan tersebut, perhitungan serta penilaian keluhan yang sering didapatkan oleh masing-masing toko online sangat diperlukan. Dengan memanfaatkan tweet masyarakat yang ditujukan kepada toko online, data tweet tersebut akan diklasifikasikan ke dalam kategori pelayanan yang telah ditentukan. Pengolahan data berupa tweet membutuhkan proses preprocessing yaitu proses untuk mendapatkan keyword dari data tweet yang telah didapatkan, proses preprocessing memiliki tahapan seperti tokenizing, filtering dan stemming. Keyword yang telah didapatkan diolah untuk mendapatkan nilai hasil klasifikasi yang didapatkan. Proses klasifikasi kategori pelayanan pada penelitian ini menggunakan metode *Cosine TF-IDF* dimana metode tersebut membutuhkan bobot dan dokumen pada setiap kategori. Metode yang dikembangkan telah diaplikasikan pada penelitian ini menghasilkan prosentase proses klasifikasi kategori pelayanan menggunakan metode *Cosine TF-IDF* sebesar 63.1%.

**Kata kunci:** analisis sentimen, klasifikasi, *rule based classifier*, *cosine similarity*, *TF-IDF*

**Abstract:** The development of online stores and online transactions in Indonesia at this time is accompanied by various problems such as complaints on services that discuss applications, responsiveness and delivery. With these problems, the calculation and assessment of complaints that are often obtained by each online store is very necessary. By utilizing community tweets aimed at online stores, the tweet data will be classified into predetermined service categories. Data processing in the form of tweets requires a preprocessing process, namely the process of getting keywords from the data tweets that have been obtained, the preprocessing process has stages such as tokenizing, filtering and stemming. The keywords that have been obtained are processed to obtain the classification results obtained. The service category classification process in this study uses the *Cosine TF-IDF* method where the method requires weights and documents in each category. The method developed has been applied in this study to produce a percentage of the service category classification process using the *Cosine TF-IDF* method of 63.1%.

**Keywords:** sentiment analysis, classification, *rule based classifier*, *cosine similarity*, *TF-IDF*

## 1. Pendahuluan

Perkembangan toko online di Indonesia dapat dikatakan sangat pesat beberapa tahun terakhir. Dengan berkembangnya toko online di Indonesia ini diikuti pula dengan berkembangnya transaksi online antara penjual dan pembeli. Semakin berkembangnya interaksi atau transaksi yang dilakukan antara penjual dan pembeli dalam toko online ini juga didasari dengan adanya hubungan saling menguntungkan dari kedua pihak, namun hal ini juga tidak dapat dihindarkan dari permasalahan yang sering muncul jika membeli barang di toko online[1]. Contoh permasalahan tersebut adalah adanya penipuan yang dilakukan oleh oknum toko online nakal yang sengaja menipu calon pembelinya dan adanya pelayanan dari toko online yang tidak semestinya, selain itu masyarakat juga dihadapkan dengan permasalahan beberapa masalah pelayanan yang biasa timbul seperti lambatnya ketanggapan toko online dalam menanggapi komplain maupun pertanyaan yang diajukan oleh masyarakat, kecepatan pengiriman dan permasalahan aplikasi yang sering error dan lain sebagainya[2].

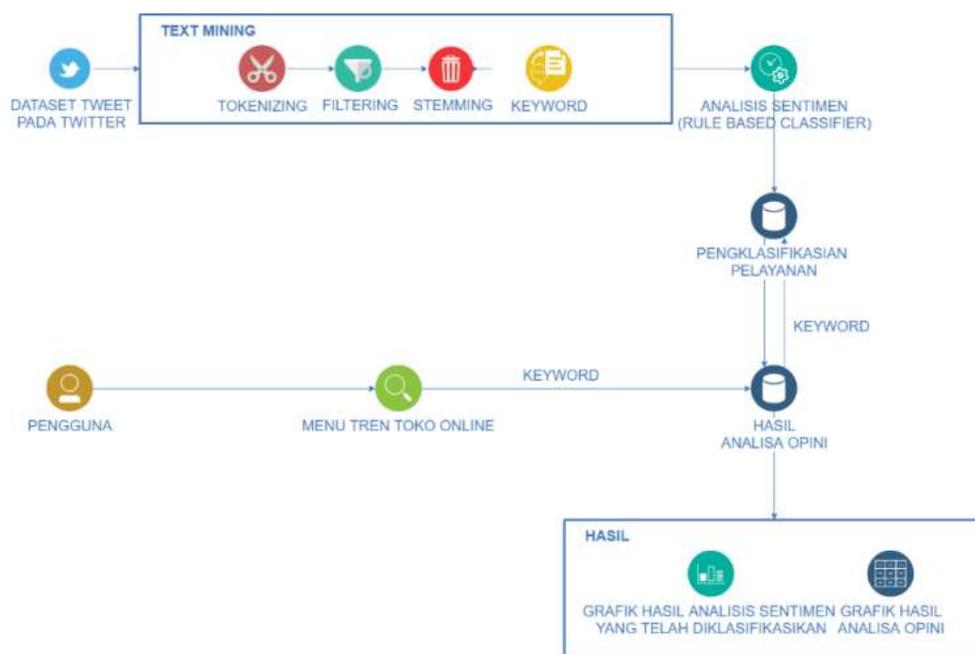
Dengan adanya permasalahan di atas, masyarakat sering kali membuat sebuah ungkapan mengenai permasalahan yang dialami dalam hal berbelanja online salah satunya mereka menuliskan keluhan kesah tersebut pada sosial media Twitter. Sentimen atau opini masyarakat semakin bertambah luas dan bebas diungkapkan di berbagai media. Sentimen dapat menjadi potensi besar bagi perusahaan yang ingin mengetahui umpan balik (feedback) dari masyarakat terhadap merek dagang mereka. Merk telah dianggap sebagai salah satu aset tidak berwujud (intangible asset) yang paling berharga dan manajemen merek adalah prioritas penting bagi pihak manajemen perusahaan atau organisasi[3]. Salah satu wadah masyarakat untuk melihat opini masyarakat mengenai review sebuah toko online dapat dilihat pada sosial media Twitter, seluruh pengguna Twitter dapat melihat berbagai opini atau tweet mengenai dan informasi mengenai pelayanan salah satu toko sesuai dengan keyword yang dituliskan oleh pengguna. Namun, hal tersebut dirasa kurang praktis atau efisien jika harus melihat satu persatu dengan data yang cukup banyak[4].

Untuk mengoptimalkan, membantu dan mengumpulkan informasi mengenai beberapa toko online yang dituju penulis mengajukan suatu pemodelan baru dengan menyediakan hasil informasi yang diperoleh dari pengolahan opini masyarakat terhadap pelayanan toko online di Indonesia yang terdapat pada sosial media Twitter. Pengolahan data tersebut dapat dikumpulkan dan diolah dengan proses analisis sentimen untuk mengetahui apakah sebuah opini tersebut tergolong opini bersentimen positif, negatif atau netral. Analisis sentimen disebut juga opinion mining, adalah bidang ilmu yang menganalisa pendapat, sentimen, evaluasi, penilaian, sikap dan emosi publik terhadap entitas seperti produk, jasa, organisasi, individu, masalah, peristiwa, topik, dan atribut mereka[5]. Pengolahan data opini tersebut bertujuan untuk membantu masyarakat dalam mengetahui informasi mengenai tingkat penilaian pelayanan toko online yang akan dituju sehingga masyarakat dapat membantu masyarakat dalam mengetahui informasi mengenai berbagai macam jenis pelayanan toko online seperti penilaian terhadap pelayanan pengiriman, ketanggapan dan aplikasi pada berbagai macam toko online[6].

## 2. Metode Penelitian

Pada penelitian ini penulis menggunakan pendekatan *Rule Based Classifier* untuk menentukan sentimen positif, negatif dan netral dari sebuah Tweet yang ditujukan kepada empat toko online di Indonesia yaitu Shopee, Tokopedia, Blibli dan Bukalapak. Orisinalitas pada penelitian ini adalah menilai sebuah Tweet yang membahas pelayanan toko online dan mengklasifikasi menjadi tiga jenis pelayanan yaitu ketanggapan, pengiriman dan aplikasi dengan menggunakan metode *Cosine TF-IDF*[7].

Berikut merupakan desain sistem penelitian yang sedang dilakukan. Desain sistem berikut menjelaskan proses pengambilan data hingga proses analisis sentimen dan klasifikasi setiap pelayanan.



Sumber: Hasil Penelitian (2020)

Gambar 1. Diagram Sistem

Seperti pada Gambar 1, pengerjaan sistem diawali dengan pengambilan data Tweet yang berasal dari sosial media Twitter. Dalam pengambilan data Tweet tersebut memanfaatkan API yang telah disediakan oleh Twitter. Setelah mendapatkan data berupa Tweet, setiap Tweet akan masuk ke dalam proses *text mining* atau *text preprocessing* dimana pada proses ini terdapat empat tahapan proses yaitu *tokenizing*, *filtering*, *stemming* dan *keyword*[8]. Di dalam proses *text mining* atau *pre-processing* menghasilkan *keyword* atau kata kunci yang akan digunakan pada proses selanjutnya. Setelah mendapatkan kata kunci pada masing-masing Tweet, proses selanjutnya adalah menganalisa sentimen Tweet tersebut dengan menggunakan *Rule Based Classifier* dimana metode ini membutuhkan beberapa *rule* atau aturan tata tulisan sesuai kaidah bahasa Indonesia sehingga dapat menentukan hasil sentimen positif, negatif atau netral pada tiap Tweet. Setelah mendapatkan hasil dari proses analisis sentimen, Tweet akan masuk pada proses klasifikasi dengan menggunakan metode *Cosine TF-IDF* untuk diklasifikasikan pada tiga jenis pelayanan yaitu penilaian Tweet yang membahas tentang kecepatan pengiriman, aplikasi dan ketanggapan pada toko online tertentu.

### Pengumpulan Data

Penelitian ini diawali dengan melakukan pengambilan data tweet berbahasa Indonesia pada sosial media Twitter berdasarkan kata kunci yang digunakan. Diagram alur pengambilan data pada Twitter dapat dilihat pada Gambar 2.



Sumber: Hasil Penelitian (2020)

Gambar 2. Diagram Alur Pengambilan Data

Pada Gambar 2, user memasukkan kata kunci yang sebelumnya telah ditentukan kemudian program memproses dan mencari data yang mengandung kata kunci tersebut. Proses pengambilan data pada sosial Twitter hanya dapat mengambil 100 data setiap satu kali proses. Pada proses pengambilan data ini, penulis mengambil id string tweet, username pengguna, isi tweet dan tanggal penulisan tweet. Jika data tweet yang mengandung kata kunci sudah didapatkan sebanyak 100 data maka data akan disimpan pada database yang sebelumnya telah dibuat.

Pada penelitian ini menggunakan dataset berupa tweet yang dituliskan oleh masyarakat dalam sosial media Twitter. Dalam pengambilan dan pencarian tweet ini diambil berdasarkan keyword nama toko online dan melalui pencarian keyword username id Twitter dari akun resmi yang dimiliki oleh toko online dengan memanfaatkan API dari Twitter. Tweet yang telah didapatkan dan dijadikan dataset dianggap sebagai tweet yang berisi mengenai pelayanan. Berikut contoh tweet keluhan atau penilaian yang disampaikan kepada akun Twitter toko online dapat dilihat pada Gambar 3.



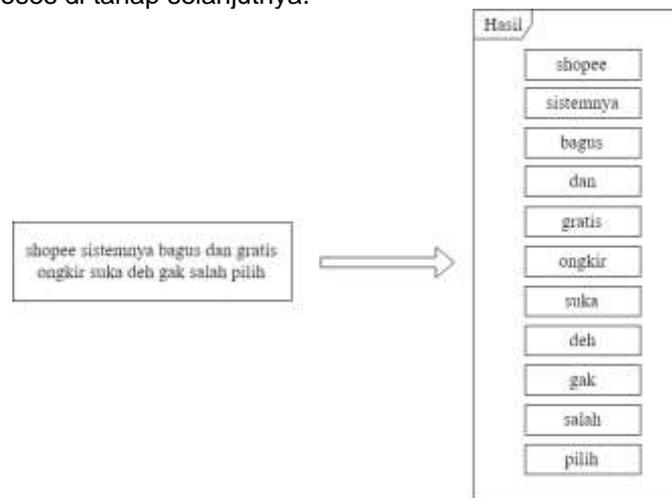
Sumber: Hasil Penelitian (2020)

Gambar 3. Contoh Data Tweet

### Proses Text Mining

*Text Mining* merupakan proses penggalian informasi dari sekumpulan dokumen berupa text yang mengandung informasi tidak terstruktur [9]. Text Mining mempunyai 4 tahapan proses yaitu *Tokenizing*, *Filtering*, *Stemming*, dan *Keyword*.

Tokenizing merupakan tahap dimana adanya pemotongan text atau suatu kalimat menjadi kumpulan kata-kata atau array of words. Kata – kata yang telah dipisahkan atau dipotong akan diproses di tahap selanjutnya.

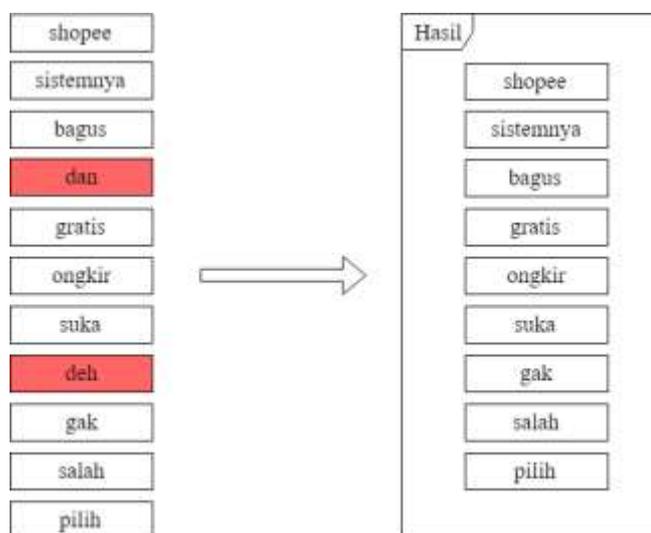


Sumber: Hasil Penelitian (2020)

Gambar 4. Proses Tokenizing

Pada Gambar 4 merupakan gambaran dari proses tokenizing dimana ada sekumpulan kalimat yang pada proses ini akan dipisahkan oleh karakter spasi dan menghasilkan kumpulan kata. Kata-kata inilah yang nantinya akan dijadikan acuan dalam pencarian keyword. Jumlah keyword yang dihasilkan dalam proses Tokenizing tidak selalu sama karena pada proses selanjutnya bisa saja ada beberapa kata yang tidak diperlukan atau dibuang.

Proses filtering adalah proses dimana adanya pembuangan kata yang tidak penting atau bisa disebut dengan stoplist. Kata-kata tersebut antara lain berupa kata penghubung, kata bantu dan sebagainya. Pada proses ini kumpulan kata yang sebelumnya telah didapatkan dari proses Tokenizing akan difilter apakah kata tersebut termasuk kata stoplist atau tidak, jika iya maka kata tersebut akan dibuang[10]. Pada Gambar 5 merupakan gambaran mengenai proses filtering.



Sumber: Hasil Penelitian (2020)

Gambar 5. Gambaran Proses Filtering

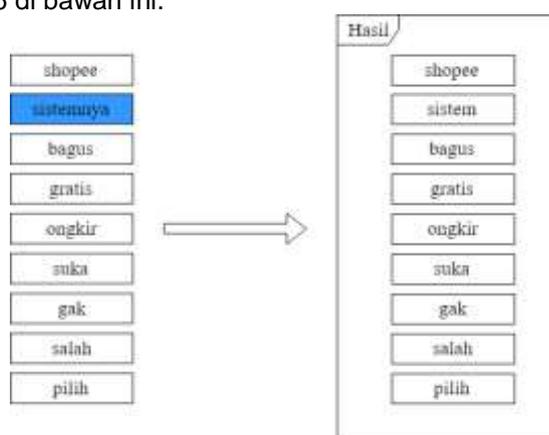
Stemming adalah proses membuang imbuhan pada suatu kata agar konten berisi kata dasar yang dapat mewakili konten tersebut. Proses stemming dimulai dari menghapus akhiran dan dilanjut dengan awalan karena tidak semua kata berawalan bukan merupakan kata dasar[11]. Berikut adalah contoh tabel imbuhan akhiran dan awalan pada Tabel 1.

Tabel 1. Tabel Kata Imbuhan

Imbuhan	Kata Imbuhan
Akhiran	_kan, _pun, i, _nya, _in, _is, _isme, _wan, _ah, _wi
Awalan	Ber_, per_, me_, di_, ter_, ke_, se_, pe_, pem_, peng_

Sumber: Hasil Penelitian (2020)

Pada proses stemming dilakukan perkata, contoh hasil yang didapatkan pada proses ini dapat dilihat pada Gambar 6 di bawah ini.

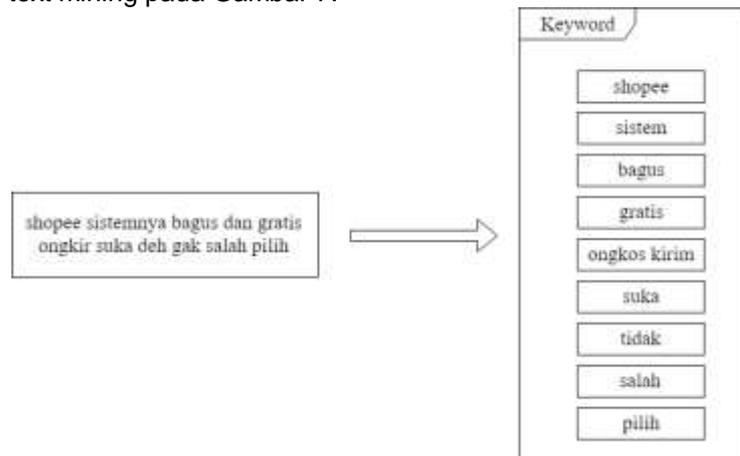


Sumber: Hasil Penelitian (2020)

Gambar 6. Gambaran Proses Stemming

Setelah semua tahap text mining dilalui, kata-kata tadi menjadi keyword untuk teks. Keyword adalah kata-kata penting yang dapat mewakili teks atau dokumen, sehingga pada saat

proses analisis opini, keyword inilah yang akan dijadikan acuan[12]. Berikut adalah keyword yang dihasilkan dari text mining pada Gambar 7.



Sumber: Hasil Penelitian (2020)

Gambar 7. Gambaran Hasil Keyword

**Proses Klasifikasi Keluhan Menggunakan Cosine TF-IDF**

Proses klasifikasi pelayanan menghasilkan tiga macam jenis pelayanan pada toko online yaitu Ketanggapan, Pengiriman, Aplikasi.

Setelah melalui tahap *preprocessing* maka data akan disimpan dalam memori sementara dan nantinya akan diproses lebih lanjut ke dalam tahap *analyzing* menggunakan pembobotan TF-IDF dan klasifikasi dengan algoritma *cosine similarity*.

Metode TF-IDF menggabungkan dua konsep untuk perhitungan bobot, yaitu frekuensi kemunculan sebuah kata di dalam sebuah dokumen tertentu dan *inverse* frekuensi dokumen yang mengandung kata tersebut. Frekuensi kemunculan kata di dalam dokumen yang diberikan menunjukkan seberapa penting kata itu di dalam dokumen tersebut [13]. Frekuensi dokumen yang mengandung kata tersebut menunjukkan seberapa umum kata tersebut. Sehingga bobot hubungan antara sebuah kata dan sebuah dokumen akan tinggi apabila frekuensi kata tersebut tinggi di dalam dokumen dan frekuensi keseluruhan dokumen yang mengandung kata tersebut yang rendah pada kumpulan dokumen. Rumus untuk TF-IDF dapat dilihat pada Rumus (1), (2), (3) :

$$t f = 0,5 + 0,5 \times \frac{t f}{\max(t f)} \dots\dots\dots(1)$$

$$i d f_t = \log \log \left( \frac{D}{d f_t} \right) \dots\dots\dots(2)$$

$$W_{d,t} = t f_{d,t} \times i d f_{d,t} \dots\dots\dots(3)$$

Keterangan:

D = dokumen ke-d

t = term ke-t dari dokumen

W = bobot dokumen ke-d terhadap term ke-t

tf = banyaknya term i pada sebuah dokumen

idf = Inverse Document Frequency

df = banyak dokumen yang mengandung term i

*Cosine Similarity* digunakan untuk melakukan perhitungan kesamaan dari dokumen. Rumus yang digunakan oleh cosine similarity dapat dilihat pada rumus (4) berikut [14]:

$$C o s \alpha = \frac{A \cdot B}{|A| |B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \dots\dots\dots(4)$$

Keterangan :

A = Vektor A, yang akan dibandingkan kemiripannya

B = Vektor B, yang akan dibandingkan kemiripannya  
 $A \cdot B$  = dot product antara vektor A dan vektor B  
 $|A|$  = panjang vektor A  
 $|B|$  = panjang vektor B  
 $|A||B|$  = cross product antara  $|A|$  dan  $|B|$

Metode pengklasifikasian yang digunakan pada sistem ini adalah dengan cara membandingkan kesamaan atau similaritas antara *tweet* masyarakat dengan tiga dokumen kata kunci yaitu ketanggapan, aplikasi dan pengiriman, kemudian cara membandingkan kesamaan atau similaritas antara *tweet* masyarakat dengan kata kunci berikutnya, begitu seterusnya. Kemudian dicari jumlah similaritas yang tertinggi antara dokumen, hasil similaritas tertinggi adalah hasil klasifikasi akhir pada sebuah *tweet*. Apabila total similaritas yang didapatkan pada keseluruhan dokumen adalah nol (0) maka *tweet* diklasifikasikan menjadi “*Uncategorized*”.

### 3. Hasil dan Pembahasan

Hasil dari penelitian ini mulai dari tahapan pengumpulan data hingga pengujian dan evaluasi dijabarkan sebagai berikut.

#### Pengumpulan Data

Data yang didapatkan berasal dari opini masyarakat terkait dengan pelayanan toko online di Indonesia berupa sebuah *tweet* yang terdapat pada sosial media twitter. Contoh data *tweet* yang didapatkan disajikan pada Tabel 2.

Tabel 2. Tabel Data Contoh Data Tweet

No.	Tweet
1	Halo min, knp ya saya udah beberapa kali ini mau belanja lewat group hemat tapi gak gratis ongkir terus? @ShopeeCare
2	@ShopeeCare Halo shope. Saya ada pesan barang di toko. Hingga masa garansi shope habis belum dikirim saya bayar tgl... <a href="https://t.co/hqvnF7211M">https://t.co/hqvnF7211M</a>
3	@TokopediaCare tolong dong gimana cara nya membatalkan pesanan, saya sudah email tapi belum. Ada balasan
4	Min @TokopediaCare asuransi pengiriman itu misal barang rusak saat perjalanan bisa nge klaim senilai barang tersebut... <a href="https://t.co/70k3xpVHvM">https://t.co/70k3xpVHvM</a>
5	@BibliCare @JNECare Iya kemarin CS Bibli sudah ada yg kontak langsung via telp dan sedang tunggu respon dari... <a href="https://t.co/S3tZfX32YI">https://t.co/S3tZfX32YI</a>
6	pagi kak ini tracking nya kenapa belum berubah ya dari tgl 17 skrg udah tgl 20. tolong segera di follow up karena ba... <a href="https://t.co/0VPisZGu0e">https://t.co/0VPisZGu0e</a>
7	@BukaBantuan dear admin. Mohon DM saya ditanggapi. Trims
8	@BukaBantuan Lalu apa hubungannya komplain tentang penggantian nomor telepon dengan tiba2 ada tambahan permintaan,... <a href="https://t.co/556UYgAXwr">https://t.co/556UYgAXwr</a>
9	@BibliCare Pesanan nomor 12027631944 atas nama Teguh Yuana kok belum sampai juga ya? Udah meleset dari estimasi. P... <a href="https://t.co/3njKI5Qe0D">https://t.co/3njKI5Qe0D</a>
10	@ShopeeCare gimana kelanjutan pesanan saya ditgl 3 /april 2019 brg saya gk terima tetapi kok bisa barang ditls udh... <a href="https://t.co/Enq6fkHvJV">https://t.co/Enq6fkHvJV</a>

Sumber: Hasil Penelitian (2020)

#### Hasil Proses Klasifikasi Menggunakan Metode Cosine TF-IDF

Pada pengujian pengklasifikasian yang digunakan pada sistem ini adalah dengan cara membandingkan kesamaan atau similaritas antara *tweet* masyarakat dengan tiga dokumen kata kunci yaitu ketanggapan, aplikasi dan pengiriman, kemudian cara membandingkan kesamaan atau similaritas antara *tweet* masyarakat dengan kata kunci berikutnya, begitu seterusnya. Kemudian dicari jumlah similaritas yang tertinggi antara dokumen, hasil similaritas tertinggi adalah hasil klasifikasi akhir pada sebuah *tweet*. Apabila total similaritas yang didapatkan pada keseluruhan dokumen adalah nol (0) maka *tweet* diklasifikasikan menjadi “*Uncategorized*”. Hasil pengujian klasifikasi pelayanan dapat dilihat pada Tabel 3.

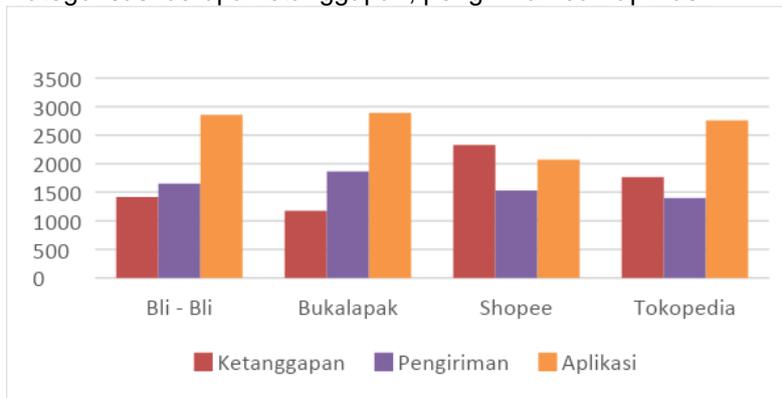
Tabel 3. Tabel Hasil Klasifikasi Pelayanan

No.	Tweet	Hasil Klasifikasi Pelayanan	
		Program	Manual
1	@ShopeeCare ini belum dikirim ya.. gmn cara komplain	Pengiriman	Pengiriman
2	Coba contoh @ShopeeCare . Kan enak pembayarannya mudah, opsinya bisa di switch2 dgn mudah tanpa harus batalkan transaksi dulu????	Aplikasi	Aplikasi
3	@TokopediaCare Saya mau lihat detail invoice IVR/20190121/XIX/I/111223198 , ga bisa, mau bayar lewat V.A	Aplikasi	Aplikasi

No.	Tweet	Hasil Klasifikasi Pelayanan	
		Program	Manual
	takutnya ga masuk, makasih		
4	@TokopediaCare Oke itu pesanan telah selesai memang karena sebelumnya saya mengajukan komplain karena saya pikir ke... <a href="https://t.co/KSXnF8UhJR">https://t.co/KSXnF8UhJR</a>	Ketanggapan	Pengiriman
5	@FatrialN @TokopediaCare Barang gak jadi dikirim? Ya harusnya ongkir jg ikut balik. Kecuali barang hilang setelah d... <a href="https://t.co/XzlxSsQlhx">https://t.co/XzlxSsQlhx</a>	Pengiriman	Pengiriman
6	@bukalapak Males akh karena sudah mengecewakan kami #uninstal Mending @LazadaID pesan mlm besok siang udah nyampe rumah pesanan nya.	Ketanggapan	Pengiriman
7	Halo min, knp ya saya udah beberapa kali ini mau belanja lewat group hemat tapi gak gratis ongkir terus? @ShopeeCare	Aplikasi	Aplikasi
8	@BlibliCare Baik..makasih info dan sarannya.	Ketanggapan	Ketanggapan
9	@bukalapak min mnta bantuannya saya salah ngirim isi paket ap bsa d tarik buat ganti isinya?	Pengiriman	Pengiriman
10	@BlibliCare Oke kak sudah aku balas, mohon cek dm lagi?????????	Ketanggapan	Ketanggapan

Sumber: Hasil Penelitian (2020)

Hasil yang didapatkan dari proses klasifikasi pelayanan pada program dapat dilihat pada Gambar 8 Pada proses klasifikasi pelayanan dihasilkan data tiap toko online dengan klasifikasi atau kategorisasi berupa ketanggapan, pengiriman dan aplikasi.



Sumber: Hasil Penelitian (2020)

Gambar 8. Grafik Hasil Klasifikasi Pelayanan

Detail hasil yang didapatkan dari keseluruhan proses klasifikasi pelayanan pada tiap toko online dapat dilihat pada Tabel 4.

Tabel 4. Tabel Jumlah Keseluruhan Klasifikasi

	Ketanggapan	Pengiriman	Aplikasi
Bli - Bli	1416	1654	2857
Bukalapak	1173	1862	2892
Shopee	2328	1531	2068
Tokopedia	1766	1400	2761

Sumber: Hasil Penelitian (2020)

### Pengujian Proses Klasifikasi Menggunakan Metode Cosine TF-IDF

Perhitungan nilai akurasi terhadap proses klasifikasi pelayanan menggunakan metode *Cosine* TF-IDF menggunakan pengujian persentase rata-rata. Pada perhitungan akurasi pada penelitian ini menggunakan 1000 data dengan komposisi seperti Tabel 5 di bawah ini.

Tabel 5. Tabel Uji Klasifikasi Pelayanan

Percobaan	Jumlah Tweet	Jumlah Klasifikasi Benar	Prosentase Akurasi (%)
1	100	70	70 %
2	100	76	76 %
3	100	36	36 %
4	100	65	65 %
5	100	52	52 %
6	100	58	58 %
7	100	70	70 %

Percobaan	Jumlah <i>Tweet</i>	Jumlah Klasifikasi Benar	Prosentase Akurasi (%)
8	100	67	67 %
9	100	67	67 %
10	100	70	70 %
Persentase Akurasi Rata – Rata			63.1 %

Sumber: Hasil Penelitian (2020)

Pada persamaan persentase akurasi rata-rata didapatkan 63.1 % untuk pengujian klasifikasi pelayanan menggunakan metode *cosine* TF-IDF. Hasil yang didapatkan kurang maksimal karena adanya kemungkinan keyword yang ada pada tiap dokumen kategori belum bertambah sesuai kata yang sering muncul pada tweet mengenai keluhan atau penilaian mengenai pelayanan toko online di Indonesia.

#### 4. Kesimpulan

Proses analisis sentimen pada sebuah komentar atau tanggapan adalah hal yang sangat penting, contoh pada penelitian ini digunakan untuk menilai suatu kinerja dari sebuah pelayanan berbagai macam toko online. Dengan adanya proses analisis sentiment dapat memudahkan dalam memproses banyak data yang telah didapatkan. Kedepannya, sentimen analisis ini dapat digunakan untuk kepentingan artificial intelligence. Analisis sentimen juga berguna untuk pemahaman atau penilaian sebuah tanggapan yang disampaikan secara tekstual dengan menampilkan hasil yang didapatkan yaitu apakah pendapat itu termasuk pendapat positif atau negatif.

Selain proses analisis sentiment yang dapat menghasilkan dan mengetahui mana review atau komentar yang positif, negatif dan netral, pada penelitian ini juga menunjukkan bahwa proses klasifikasi juga membantu untuk mengklasifikasikan jenis komentar atau *tweet* ke dalam kategorisasi yang telah ditentukan sebelumnya.

Pada percobaan yang telah dilakukan, penulis menyimpulkan bahwa untuk menilai kinerja toko online tidak cocok jika menggunakan data yang didapatkan dari Twitter karena data yang telah didapatkan cenderung lebih mengarah pada keluhan bukan untuk menilai kinerja suatu toko online.

Proses klasifikasi pelayanan dengan menggunakan metode *Cosine* TF-IDF mendapatkan akurasi sebesar 63.1 % dengan menggunakan uji coba dari 1000 data yang telah didapatkan.

#### Referensi

- [1] R. Habibi, D. B. Setyohadi, and E. Wati, "Analisis Sentimen Pada Twitter Mahasiswa Menggunakan Metode Backpropagation," *J. Inform.*, vol. 12, no. 1, pp. 103–109, 2016.
- [2] D. N. Pratiwi, "Kepercayaan Mahasiswa Jurusan Ilmu Komunikasi Universitas Muhammadiyah Yogyakarta terhadap Situs Belanja Online Bukalapak.com, Elevenia.co.id dan Tokopedia.com," *Skripsi UMY*, vol. 3, no. 1, p. 56, 2016.
- [3] Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayesite," *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, p. 113, 2018.
- [4] A. P. Syah, Adiwijaya, and S. Al Faraby, "Analisis Sentimen Pada Data Ulasan Produk Toko Online Dengan Metode Maximum Entropy," in *e-Proceeding of Engineering*, 2017, pp. 4632–4640.
- [5] S. Gusriani, K. D. K. Wardhani, and M. I. Zul, "Analisis Sentimen Terhadap Toko Online di Sosial Media Menggunakan Metode Klasifikasi Naïve Bayes (Studi Kasus: Facebook Page BerryBenka)," in *4th Applied Business and Engineering Conference*, 2016, pp. 1–7.
- [6] D. R. Lazuardi, "Analisis Sentimen Untuk Mengetahui Persepsi Kualitas Merek Menggunakan Text Mining Dan Social Network Analysis Pada Konten Percakapan Di Media Sosial Twitter," *e-Proceeding Eng.*, pp. 1–9, 2014.
- [7] V. Amrizal, "Penerapan Metode Term Frequency Inverse Document Frequency (Tf-Idf) Dan Cosine Similarity Pada Sistem Temu Kembali Informasi Untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Hadits Shahih Bukhari-Muslim)," *J. Tek. Inform.*, vol. 11, no. 2, pp. 149–164, 2018.

- [8] H. A. Vidyastuti, "Analisis Perbandingan mengenai e-Commerce berdasarkan persepsi Online User dan Online Shopper," *STIE Ekuitas*, pp. 1–14, 2014.
- [9] M. H. Adnan, W. Maharani, and A. Herdiani, "Analisis Klasifikasi Sentimen dan Peringkasan pada Review Produk Menggunakan Metode Unsupervised," *e-Proceeding Eng.*, pp. 35–44, 2016.
- [10] D. A. Prasyda, "Analisis Sentimen berdasarkan Aspek Pada Review Restoran Menggunakan Bayesian Networks Untuk Dokumen Berbahasa Inggris," in *Indonesia Symposium on Computing (IndoSC)*, 2016, pp. 307–318.
- [11] M. N. Ardiansyah, R. Umar, and Sunardi, "Analisis Sentimen pada Twitter menggunakan Metode Support Vector Machine," in *Prosiding Semnastek*, 2019, pp. 739–742.
- [12] M. Rizal, "Analisis Sentimen Pengguna Twitter terhadap Objek Pariwisata di Indonesia Menggunakan Algoritma Pengolahan Deep Natural Language dari IBM Insights untuk Twitter," *Skripsi Univ. Islam Negeri Alauddin Makassar*, 2017.
- [13] A. B. T. I. Wahyudi, W. Maharani, "Analisis Sentimen terhadap Produk di Twitter menggunakan Metode Support Vektor Machine," *e-Proceeding Eng.*, pp. 1–6, 2012.
- [14] I. F. Rozi, "Implementasi Rule-Based Document Subjectivity Pada Sistem Opinion Mining," *J. ELTEK*, vol. 11, no. 1, pp. 29–41, 2013.