

# Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength

Abdul Aziz<sup>1</sup>, Fauziah<sup>2</sup>

<sup>1,2</sup>Informatika, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional  
abddul.aziz.97@gmail.com, fauziah@civitas.unas.ac.id

## Abstract

*The need to analyze a product or policy becomes an important thing to measure the level of success. Twitter is currently one of the popular applications used by the public to give their impressions and opinions, both positive, negative and neutral opinions. Diverse public opinion on Twitter can be used as a reference material to get the level of community satisfaction on a product, service or policy. In this study, a sentiment analysis system was created using the TF-IDF and SentiStrength Algorithm. The steps in the research are, firstly, crawling Twitter data using the Twitter API, second preprocessing, thirdly doing spell correction, fourth Word weighting (TF-IDF) and lastly SentiStrength classification, where the results of the classification of tweets have positive, negative or neutral sentiments. In the test data taken using the keyword "child vaccines" as many as 1000 tweets, the results obtained were 54% positive sentiment, 20% negative sentiment and 26% neutral sentiment. Comparison with the same negative data analysis using a different algorithm, namely Naïve Bayes, results in positive sentiment of 55%, 16% and neutral 29%. Decision Tree got 61% positive results, 14% negative and 25% neutral.*

**Keywords :** Twitter, Sentiment Analysis, SentiStrength, Opinion, TF-IDF.

## Abstrak

*Kebutuhan dalam melakukan analisis pada sebuah produk atau kebijakan menjadi suatu hal penting untuk mengukur tingkat keberhasilan. Twitter saat ini menjadi salah satu aplikasi populer yang digunakan oleh masyarakat dalam memberikan kesan dan opininya, baik opini yang bersifat positif, negatif maupun netral. Opini masyarakat yang beragam di Twitter dapat dijadikan sebagai bahan acuan untuk mendapatkan tingkat kepuasan masyarakat pada sebuah produk, layanan ataupun kebijakan. Pada penelitian ini sistem analisis sentimen dibuat dengan menggunakan Algoritma TF-IDF dan SentiStrength Langkah-langkah pada penelitian yaitu, pertama crawling data Twitter menggunakan API Twitter, kedua preprocessing, ketiga melakukan spell correction keempat Pembobotan kata (TF-IDF) dan terakhir klasifikasi SentiStrength, dimana hasil klasifikasi tweet memiliki sentimen positif, negatif atau netral. Pada data uji yang diambil dengan menggunakan kata kunci "vaksin anak" sebanyak 1000 tweet, diperoleh hasil sentimen positif 54%, sentimen negatif 20% dan sentimen netral 26%. Perbandingan dengan analisis data yang sama menggunakan algoritma yang berbeda yaitu Naïve Bayes mendapatkan hasil sentimen positif 55%, negatif 16% dan netral 29%. Decision Tree mendapatkan hasil positif 61%, negatif 14% dan netral 25%.*

**Kata kunci:** Twitter, Analisis Sentimen, SentiStrength, Opini, TF-IDF.

## 1. PENDAHULUAN

Jejaring sosial Twitter menyediakan ruang yang dibutuhkan untuk memenuhi kebutuhan manusia akan informasi dan komunikasi pada era modern. Setiap harinya jutaan bahkan milyaran entri baru dibuat oleh



pengguna Twitter di seluruh dunia, mulai dari entri hiburan, berita, iklan promosi dan lain-lain. Dengan banyaknya entri dan pengguna aktif pada Twitter menjadikan data yang ada di dalamnya sangat beragam. Data ini dapat dimanfaatkan untuk mengukur tingkat kepuasan masyarakat pada sebuah produk, kebijakan pemerintah, layanan jasa dan lain-lain.

Analisis sentimen adalah metode yang dipakai dalam menganalisis bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut dapat diklasifikasikan sebagai sentimen positif, negatif dan netral [0]. Analisis sentimen sering digunakan oleh produsen dalam mengukur tingkat kesuksesan suatu produk yang mereka keluarkan. Dalam contoh lain analisis sentiment juga dapat digunakan untuk mengetahui opini masyarakat terhadap seorang public figure ataupun pejabat [0].

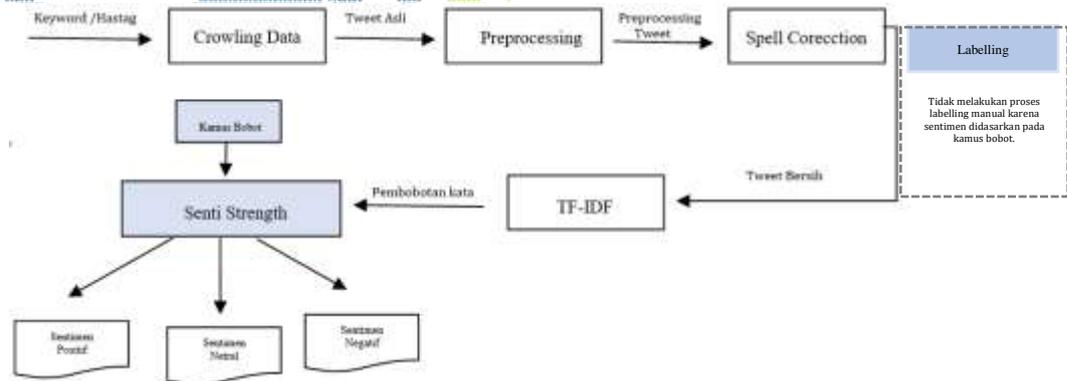
Aplikasi berbasis web untuk analisis sentimen pada Twitter menggunakan Algoritma TF-IDF dan klasifikasi *SentiStrength* dibuat dengan menggunakan Bahasa pemrograman Python. Tampilan web menggunakan bahasa pemrograman PHP dengan framework Laravel 5.4. Metode *SentiStrength* adalah algoritma data mining yang diciptakan oleh *CyberEmotion*. *SentiStrength* merupakan klasifikasi yang memanfaatkan pendekatan berbasis leksikon atau kamus yang berfokus pada analisis kekuatan sentimen.

Pada penelitian Pendeteksian kekuatan sentimen pada teks tweet berbahasa indonesia menggunakan *Sentistrength* dengan hasil akurasi keseluruhan 57.33% [0]. Algoritma TF-IDF untuk meringkas tweet dengan menghasilkan rata-rata recall 0,31, presisi rata-rata 0,34, dan *F1-Measure* gabungan sebesar 0,33 [0]. *SentiStrength* sebagai klasifikasi data pada pengujian sentimen dengan hasil perolehan akurasi rata-rata sebesar 60% [0]. Model aplikasi untuk analisis sentimen pada opini masyarakat dengan algoritma *Naïve Bayes Classifier* [0]. Pendekatan berbasis leksikon untuk mendeteksi pesan terkait bunuh diri di Twitter menggunakan algoritma *Yet Another Two-Stage Idea* atau YATSI memiliki akurasi klasifikasi lebih tinggi (66,67 %) dan *Learning with Local and Global Consistency* atau LLGC memiliki akurasi (59,38%) [0]. Semantik kontekstual untuk analisis sentimen Twitter dengan menggunakan pendekatan berbasis *leksikon* mendapatkan hasil *SentiStrength* dengan akurasi 4-5% [0]. Tujuan penelitian melakukan analisis sentimen tanpa melakukan proses *labelling* manual pada data set. Algoritma *SentiStrength* tidak membutuhkan data training (*unsupervised*) karena klasifikasi didasarkan pada kamus yang telah diberi bobot.

## 2. METODOLOGI PENELITIAN

### 2.1. Alur Kerja Aplikasi

Gambaran alur kerja aplikasi pada penelitian mulai dari sistem berjalan sampai berhenti secara umum dijelaskan pada Gambar 1.



**Gambar 1.** Alur Kerja Aplikasi

Gambar 1 merupakan alur kerja Aplikasi pada penelitian. *Crowling* data merupakan *tweet* yang diambil dengan menggunakan API twitter. Setelah data *tweet* terkumpul dilakukan proses *pre-processing* yang terdiri dari *tokenizer*, *case folding*, *cleansing* dan *stopword*. Setelah tahap *pre-processing* selanjutnya tahap *spell Corection* yang terdiri dari *steaming* dan *lemmatization*. Proses berikutnya pembobotan kata (TF-IDF) dan klasifikasi menggunakan *SentiStrength* dengan hasil akhir sentimen positif, negatif dan netral. Keterbaruan pada penelitian ini yaitu tidak memerlukan tahap labelling manual karena hasil sentimen didasarkan pada kamus bobot yang telah dibuat.

## 2.2. Data Input

Data input dimanfaatkan sebagai data masukan pada aplikasi. Data input diproses dengan tujuan agar mendapatkan nilai *output* yang diharapkan. Data input yang digunakan pada penelitian yaitu:

- a) **Tweet Bahasa Indonesia**, Data Tweet didapat dengan cara *crowling* data menggunakan pasilitas API Twitter. Proses *crowling* pada sistem dengan cara memasukan kata kunci atau hastag yang diinginkan lalu selanjutnya memasukan jumlah data yang ingin diambil. Pada sistem yang dibuat jumlah data yang dapat diambil maksimal 1000 buah dengan periode tweet selama 7 hari.
- b) **Kamus Sentimen**, merupakan kumpulan kata-kata yang dipilih dan diberi bobot dengan nilai sentimen 1 (tidak mempunyai nilai sentimen positif) s.d 6 (mempunyai nilai sentimen positif yang sangat kuat), dan -1 (tidak mempunyai nilai sentimen negatif) s.d -6 (mempunyai nilai sentimen negatif sangat kuat). Kamus sentimen didapatkan dari hasil translasi kamus sentimen berbahasa inggris. Penelitian dalam pembuatan kamus sentimen dilakukan oleh 3 responden ahli, mahasiswa S2 Linguistik Universitas Gadjah Mada 0
- c) **Kamus Emosikon**, berisi simbol yang digunakan untuk mengekspresikan raut wajah yang memiliki emosi. Contoh, simbol “:)” menunjukkan ungkapan emosi senyum. Kamus emosikon juga diberikan bobot untuk menentukan sentimen pada sebuah kalimat.

- d) **Kamus Idiom**, merupakan kata yang mempunyai perbedaan arti dengan kata aslinya. Contoh, kata “Panjang tangan” memiliki makna suka mencuri.
- e) **Boosterword**, merupakan kata yang dapat meningkatkan nilai sentiment pada kata di sebelahnya. Contoh, kata “cantik sekali” memiliki nilai sentimen positif lebih besar dari pada hanya kata “cantik”. Kata Boosterword diberi nilai sentimen 1-2 untuk menambah atau mengurangi nilai kata disampingnya.
- f) **Kata Negasi**, merupakan kata yang dapat mengganti makna pada sebuah kata lain. Contoh, kata “jahat” memiliki nilai sentimen negatif, tapi jika kata “jahat” diawali dengan kata “tidak” maka kata nilai sentimen akan berubah menjadi positif dan susunan kata menjadi “tidak jahat”.
- g) **Kata Tanya**, dapat mengubah sentiment dari suatu opini. Contoh, kalimat “apa anda senang?” meskipun terdapat kata “senang” tapi kalimat tersebut tidak memiliki orientasi positif ataupun negatif.

### 2.3. Data Pre-processing

Preprocessing merupakan teknik awal data mining yang bertujuan mengubah data mentah yang sudah terkumpul menjadi data bersih yang dapat digunakan pada proses berikutnya. Teknik preprocessing yang digunakan pada sistem ini yaitu:

- a) **Cleansing**, tahap untuk membersihkan data dengan menghapus bagian data yang tidak sesuai contohnya seperti symbol, hastag, URL dan lain-lain.
- b) **Tokenizer**, tahap membagi kalimat menjadi perkata dengan acuan spasi sebagai pemisahannya. Contoh kalimat “aku melihatmu mencari buku” menjadi “[‘Aku’, ‘melihatmu’, ‘mencari’, ‘buku’]”.
- c) **Case folding**, tahap untuk mengubah setiap huruf pada kata menjadi lower case.
- d) **Stopword**, tahap menghapus kata penghubung yang tidak memiliki makna. Stopword menggunakan library Sastrawi. Misal : “di”, “yang”, “ke”, dan lain-lain.

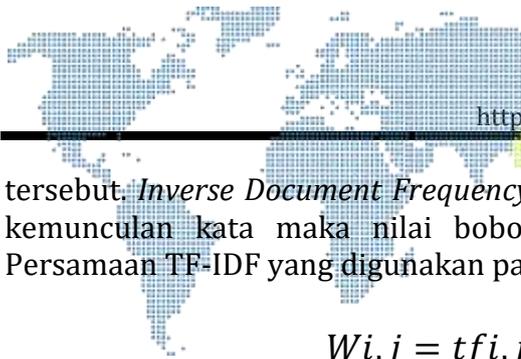
### 2.4. Spell Correction

Pada tahap ini kalimat di cek perkata untuk mendapatkan kata baku dari tiap kalimat. Teknik yang digunakan pada tahap *Spell Correction* yaitu:

- a) **Steaming**, merupakan tahap mengurai suatu kata menjadi bentuk kata dasarnya.
- b) **Lemmatization**, merupakan sebuah proses untuk menemukan bentuk dasar dari sebuah kata.

### 2.5. Term Frequency – Invers Document Frequency (TF-IDF)

TF-IDF berfungsi untuk menghitung frekuensi kemunculan kata pada dokumen lalu selanjutnya mengubah kata tersebut kedalam bentuk numerik 0. *Term Frequency* (TF) adalah semakin besar frekuensi kemunculan kata pada sebuah dokumen maka akan semakin besar nilai bobot untuk kata



tersebut. *Inverse Document Frequency* (IDF) adalah semakin besar frekuensi kemunculan kata maka nilai bobot kata tersebut akan semakin kecil. Persamaan TF-IDF yang digunakan pada penelitian:

$$W_{i,j} = t_{f_{i,j}} \log \left( \frac{N+1}{df_{i+1}} \right) + 1 \quad (1)$$

Penjelasan:

$W_{i,j}$  = banyak data ke-i terhadap kata ke-j

$t_{f_{i,j}}$  = banyaknya kata i yang dicari pada sebuah data j

N = total data

$df_i$  = banyaknya data yang mengandung kata ke-i

## 2.6. *SentiStrength*

*SentiStrength* merupakan algoritma klasifikasi yang memanfaatkan pendekatan leksikon atau kamus. *SentiStrength* memberikan hasil positif dan negatif dengan jangkauan dimulai dari 1 – 6, nilai 1 menunjukkan kalimat netral (tidak mempunyai sentiment), nilai 6 menunjukkan sangat positif atau sangat negatif. Nilai akhir pada kalimat ditentukan dari nilai positif dan nilai negatif tertinggi dari kata-kata pada kalimat tersebut. Contoh kata, “Bela jorok dan kotor tetapi dia penyayang”, maka hasilnya sebagai berikut, “Bela jorok (-4) dan kotor (-3) tetapi penyayang (2). (skor:2, -4)”. Maka dari kalimat tersebut didapatkan nilai akhir positif tertinggi adalah 2 dan nilai akhir negatif tertinggi adalah -4. Selanjutnya nilai sentiment akhir diputuskan berdasarkan aturan **jika** nilai positif > nilai negatif **maka** sentiment positif, **jika** nilai positif < nilai negatif **maka** sentiment negatif, **jika** nilai positif = nilai negatif **maka** sentiment netral 0. Pembobotan term selain menggunakan kamus sentiment juga dapat menggunakan kamus emosikon dan kamus ungkapan yang masing-masing kata dan karakter pada kamus diberi bobot tertentu.

## 2.7. Pengujian Aplikasi

Pengujian Aplikasi dilakukan dengan membandingkan data prediksi (hasil klasifikasi menggunakan aplikasi yang dirancang) dan data actual (hasil klasifikasi dengan labeling manual oleh manusia). Data actual selanjutnya dianalisis menggunakan aplikasi *Rapid Miner* dengan algoritma *Naïve Bayes* dan *Decision Tree* sebagai pembanding algoritma *SentiStrength* yang digunakan pada penelitian. Algoritma *SentiStrength* tidak membutuhkan data training (*unsupervised*) karena klasifikasi didasarkan pada kamus yang telah diberi bobot.

## 3. HASIL DAN PEMBAHASAN

Pada bagian ini akan dijelaskan proses pengembangan prototype aplikasi analisis sentimen dengan menggunakan algoritma TF-IDF dan *SentiStrength* untuk mendapatkan nilai sentimen dari teks pada twitter yang

diambil dengan menggunakan API Twitter berdasarkan *keyword* yang mewakili produk, layanan atau kebijakan yang ingin dianalisis.

### 3.1. Analisis Kebutuhan Aplikasi

Kebutuhan Aplikasi dibagi menjadi dua kategori, yaitu Fungsional dan Non-Fungsional. Penjelasan mengenai Analisis kebutuhan Aplikasi akan dijelaskan pada tabel 1.

**Tabel 1.** Analisis Kebutuhan Aplikasi

No	Kebutuhan Fungsional
1.	Pengguna dapat meng input <i>keyword</i> sesuai kebutuhan dan kemudian <i>keyword</i> tersebut akan digunakan oleh sistem untuk <i>crowling</i> data yang sesuai dengan <i>keyword</i> tersebut
2.	Pengguna dapat menentukan jumlah tweet yang ingin dikumpulkan
3.	Pengguna dapat melihat hasil dari setiap tahapan analisis yang dilakukan sistem dengan hasil akhir berupa sentimen positif, negatif dan netral.
4.	Pengguna dapat menyimpan hasil analisis dalam bentuk file PDF, CSV dan Exel agar dapat digunakan Kembali sebagai bahan analisis di kemudian hari.
5.	Pengguna dapat melihat hasil akhir klasifikasi sentimen dalam bentuk diagram sederhana.
No	Kebutuhan Non-Fungsional
1.	<i>User Interface</i> yang dibuat harus mudah dipahami dan digunakan
2.	Hasil Kahir disajikan dalam Diagram yang jelas dan engan warna yang kontras agar dapat membedakan masing-masing kelas sentimen.
3.	Waktu yang dibutuhkan untuk proses klasifikasi tidak melebihi 5 menit untuk jumlah <i>record</i> antara 300-500 tweet.

### 3.2. Pengembangan *Prototype*

Pada gambar 2. a), b), c) dan d), Model yang ada diubah menjadi unit kerja berbentuk aplikasi web. Tool yang digunakan diantaranya HTML, PHP dengan *framework* *Laravel* 5.4, *Bootstrap* dan *Jquery* sebagai konstruksi user interface, control dan implementasi algoritma *SentiStrength*. Secara keseluruhan sistem terbagi dalam 4 bagian yang meliputi: *Crawling* data, *Pre-Processing* dan *Spell Correction*, *TF-IDF* dan terakhir klasifikasi dengan *SentiStrength*.



a)



b)



c)

d)

**Gambar 2.** a) Pengembangan *Prototype Crawling Data*. b) Pengembangan *Prototype Pre-processing Data*. c) Pengembangan *Prototype TF-IDF*. d) Pengembangan *Prototype* klasifikasi dengan Algoritma *Sentistrength*.

**3.3. Pembahasan**

Untuk uji performa aplikasi maka dilakukan analisis pada salah satu topik yang sedang banyak dibicarakan di masyarakat. Topik yang dianalisis yaitu mengenai vaksin untuk anak, kata kunci/hastag yang digunakan adalah “vaksin anak”, dengan banyak data yang diambil sebanyak 1000 data berbahasa Indonesia. Waktu pengambilan data pada tanggal 4 November 2021. Pada tabel 3 Disajikan contoh tweet hasil preprocessing & spellcorrection.

**Tabel 3.** Contoh hasil *Pre-processing & SpellCorrection*

Tweet Asli	Hasil <i>pre-processing &amp; spellcorrection</i>
@Ria_Bilqis Jadi memang pertimbangan adanya rekomendasi vaksin anak ini udah dipikirkan ya Mba. #VaksinLindungiAnak	Jadi memang pertimbangan ada rekomendasi vaksin anak ini sudah pikir ya
Vaksin aman untuk anak-anak dan dewasa Kita Percaya Jokowi <a href="https://t.co/RIP9ucmL00">https://t.co/RIP9ucmL00</a>	vaksin aman untuk anak anak dan dewasa kita percaya jokowi
@ApuraMeity Sebab Vaksin Sinovac sudah terbukti aman untuk anak usia 6-11 tahun bedasarkan laporan hasil uji klinis fase 2. #VaksinLindungiAnak	sebab vaksin sinovac sudah terbukti aman untuk anak usia tahun dasar lapor hasil uji klinis fase
@republikaonline Untuk sekarang, jgn paksa vaksin anak saya.	untuk sekarang jangan paksa vaksin anak saya
Vaksin anak, genosida terselubung!?! <a href="https://t.co/humI9bDoQD">https://t.co/humI9bDoQD</a>	vaksin anak genosida terselubung

Tabel 3. Menyajikan 5 contoh tweet hasil dari *pre-processing & SpellCorrection*, sedangkan *tweet* keseluruhan yang digunakan pada penelitian sebanyak 1000 *tweet*.

**Tabel 4.** Contoh hasil TF-IDF

kata	TF	IDF	TF-IDF
vaksin	1315	2.09	2745.03
anak	133	2.85	378.69
aman	172	3.27	562.47
bahaya	7	6.91	48.35
bagus	76	3.35	254.78

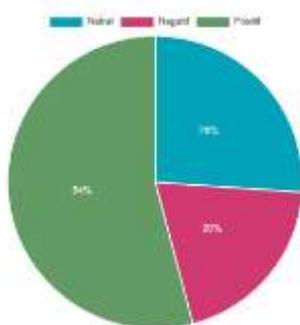
Tabel 4. Menyajikan 5 contoh hasil dari TF-IDF, sedangkan hasil TF-IDF keseluruhan yang didapat pada aplikasi sebanyak 1134 entri. Klasifikasi sentimen yang dihasilkan oleh model *SentiStrength* pada Aplikasi ini dipengaruhi oleh banyaknya data input yaitu, kamus sentiment sebanyak 1732, kamus idiom sebanyak 92, kamus emosicon sebanyak 629, kamus boosterword sebanyak 30, kamus negasi sebanyak 10 dan kamus kata tanya 20, sehingga jumlah kamus yang digunakan sebanyak 2513. sedangkan TF-IDF pada sistem ini digunakan untuk mengimprove model agar hasil yang didapatkan lebih akurat. Untuk perbandingan dilakukan analisis pada data hasil *pre-processing* yang sama dengan menggunakan algoritma *Naïve Bayes* dan *Decision Tree*. Analisis *Naïve Bayes* dan *Decision Tree* dilakukan dengan menggunakan aplikasi *Rapid Miner* setelah data sebelumnya disimpan dalam format excel lalu diberi label positif, negatif dan netral secara manual terlebih dahulu.

### 3.4. Hasil

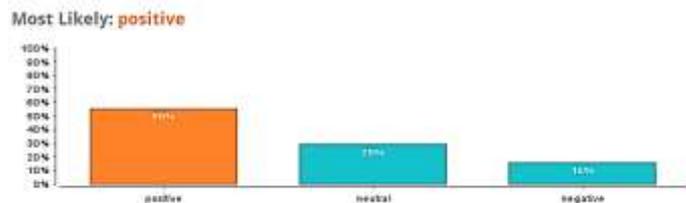
**Tabel 5.** Hasil Analisis Sentimen

Algoritma	Sentiment		
	Positif	Negatif	Netral
<i>SentiStrength</i>	54%	20%	26%
<i>Naïve Bayes</i>	55%	16%	29%
<i>Decision Tree</i>	61%	14%	25%

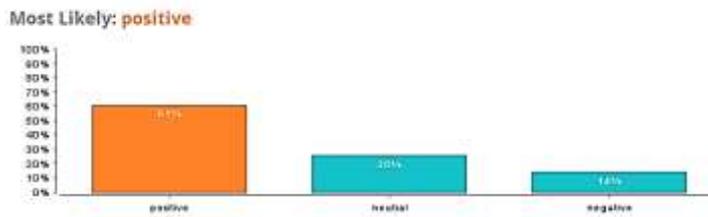
Grafik hasil analisis:



a)



b)



c)

**Gambar 3.** a) grafik hasil analisis dengan sistem yang dibuat menggunakan TF-IDF dan SentiDtrengh. b) grafik hasil analisis menggunakan *RapidMiner* dengan algoritma *Naïve Bayes*. c) grafik hasil analisis menggunakan *RapidMiner* dengan algoritma *Decision Tree*.

Hasil analisis dengan sistem yang dibuat menggunakan algoritma TF-IDF dan *SentiStrength* mendapatkan hasil sentimen positif 54%, Negatif 20% dan netral 26%. Hasil analisis dengan aplikasi Rapid Miner menggunakan algoritma *Naïve Bayes* mendapatkan hasil sentimen positif 55%, negatif 16% dan netral 29%, menggunakan algoritma *Decision Tree* mendapatkan hasil sentimen positif 61%, negatif 14% dan netral 25%.

Dari ketiga analisis dengan data yang sama menggunakan algoritma yang berbeda didapatkan hasil yang sama yaitu, sentiment yang paling dominan adalah sentiment positif, urutan kedua sentiment netral dan yang paling rendah adalah sentiment negatif.

Dengan hasil ini maka sistem yang telah dibuat menggunakan algoritma TF-IDF dan *SentiStrength* sukses menjalankan fungsinya dalam menganalisis sentiment pada twitter meskipun tanpa melakukan tahapan labeling manual seperti yang dilakukan pada percobaan menggunakan algoritma *Naïve Bayes* dan *Decision Tree*.

#### 4. SIMPULAN

Berdasarkan hasil uraian dalam proses pengembangan sistem aplikasi analisis sentimen didapatkan bahwa *prototype* aplikasi dapat berjalan dengan baik sesuai dengan kebutuhan. Aplikasi analisis sentimen dengan menggunakan algoritma TF-IDF dan *SentiStrength* dapat memenuhi analisis kebutuhan Aplikasi baik kebutuhan fungsional maupun non-fungsional. Proses crawling dan preprocessing secara umum bergantung kepada koneksi internet dan spesifikasi PC yang digunakan. pada percobaan yang dilakukan, crawling data yang menghasilkan 1000 tweets membutuhkan waktu sekitar 3-5 menit dan tahapan preprocessing membutuhkan waktu 8-10 menit. Pada percobaan analisis yang dilakukan mengenai topik vaksin untuk anak dengan menggunakan kata kunci “vaksin anak” dengan banyak tweet 1000 didapatkan hasil sentimen Positif sebanyak 54%, sentimen negatif sebanyak 20% dan sentimen netral sebanyak 26%. Hasil ini tidak jauh berbeda dengan analisis perbandingan dengan aplikasi *RapidMiner* menggunakan *Naïve Bayes* yaitu sentimen positif 55%, negatif 16% dan netral 29%, menggunakan

*Decision Tree* sentimen positif 61%, negatif 14% dan netral 25%. Dari ketiga analisis didapatkan hasil yang sama sentiment yang paling dominan adalah positif, urutan kedua netral dan ketiga negatif. Penelitian lanjut dapat dilakukan dengan meningkatkan jumlah dan periode waktu untuk crawling data. Pada sistem ini crawling data yang dapat dilakukan maksimal hanya 1240 tweet dan dalam periode 7 hari sehingga bersifat terbatas. Selanjutnya untuk lebih memaksimalkan performa algoritma *SentiStrength* dapat dilakukan dengan menambahkan lebih banyak kata yang diberi bobot tertentu pada kamus sentimen, kamus idiom, kamus emosikon, kamus *boosterword*, kamus negasi dan kamus kata tanya.

#### DAFTAR PUSTAKA

- [1] Aziz Abdul, Fauziah, & Iskandar Fitri, "Analisis Sentimen Terhadap Kebijakan Pemerintah Tentang Larangan Mudik Hari Raya Idulfitri di Indonesia Tahun 2021 Menggunakan Metode Naïve Bayes," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 5, no. 2, pp. 842-851, September 2021.
- [2] Buntoro, G.A, "Sentiment Analysis Twitter dengan Kombinasi Lexicon Based and Double Propagation" Yogyakarta: Jurusan Teknik Elektro, Fakultas Teknik, Universitas Gajah Mada. 2014.
- [3] Sianipar Raisa, Erwin Budi S., "Pendeteksian Kekuatan Sentimen Pada Teks Tweet Berbahasa Indonesia Menggunakan SentiStrength," e-Proceeding of Engineering: Vol.2, No.3 Desember 2015.
- [4] Sharifi, B.P., Hutton M.A & Kalita J.K., "Experiments in Microblog Summarization," 2010 IEEE Second International Conference on Social Computing, 49-56, 2010.
- [5] Haryalesmana Devid W, Azhari SN, "Peringkasan Sentimen Ekstraksi di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity," *IJCSS*, Vol.10, No.2, pp. 207-218, Juli 2016.
- [6] Affandy, Oktania Nandiyati, "Sentimen Analisis Berbasis Algoritma *Naïve Bayes* Classifier Untuk Identifikasi Persepsi Masyarakat Terhadap Produk/Layanan Perusahaan," *Jurnal Of Information System*, Vol.5, No.1, pp.126-135, Mei 2020.
- [7] Sarsam, S. M., Al-Samarraie, H., Alzahrani, A. I., Alnumay, W., & Smith, A. P., "A lexicon-based approach to detecting suicide-related messages on Twitter." *Biomedical Signal Processing and Control*, 65, 102355, 2021.
- [8] Saif Hassan, He Yulan, Fernandez Miriam & Alani Harith, "Contextual semantics for sentiment analysis of Twitter," *Information Processing & Management*, Januari 2015.
- [9] Nurjannah Musfiroh, Hamdani Hamdani dan Indah Fitri A, "Penerapan Algoritma Term Frequency-Inverse Document Frequency (Tf-Idf) Untuk Text Mining," e-journals, *Jurnal Ilmiah Ilmu Komputer*, Vol 8, No 3 (2013).
- [10] Herwijayanti Bening, Dian Eka Ratnawati & Lailil Muflikhah, "Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine



Similarity,” Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 2, No. 1, hlm. 306-312, Januari 2018.

- [11] Pertiwi Astri, “Analisis Sentimen Dampak Covid-19 Terhadap Perekonomian Indonesia Melalui Media Sosial Menggunakan Metode ANN. Jurnal Mantik”, Vol 4 No. 1, 608-611, 2020