

Review Produk Iphone dengan Analisis Sentimen menggunakan Algoritma Text Mining TF-IDF

Devanta Abraham Tarigan

Jurusan Teknik Komputer dan Informatika, Program Studi Teknologi Rekayasa Perangkat Lunak, Politeknik Negeri Medan, Medan, Indonesia

Email: devanta.at@polmed.ac.id

(*:coresponding author)

Abstrak—Iphone merupakan salah satu produk yang telah menjadi perhatian utama di lingkungan masyarakat dan telah menjadi salah satu kebutuhan utama di kehidupan sehari-hari. Namun terkadang iphone sering menghadapi beberapa permasalahan yang perlu diperhatikan. Salah satu masalah yang sering kali menjadi sorotan utama ialah harga yang cukup tinggi. Oleh sebab itu dibutuhkan suatu sistem yang dapat menentukan pandangan masyarakat terhadap produk iphone tersebut. Penelitian ini menggunakan text mining dan TF-IDF proses penentuan pandangan masyarakat terhadap produk iphone. Text mining dapat diartikan sebagai penemuan informasi yang baru yang sebelumnya tidak diketahui dan ekstraksi informasi yang berharga dari teks secara otomatis dari sumber yang berbeda. Sedangkan TF-IDF digunakan untuk menentukan nilai frekuensi kata-kata dalam dokumen. Dalam penelitian ini, sentimen merujuk pada pandangan masyarakat terhadap produk iphone, baik itu positif atau negatif. Hasil akhir dari analisa sentimen ini terdapat nilai sentimen positif sebesar 68,65% sedangkan nilai sentimen negatif sebesar 31,35%. Hal ini diharapkan dapat memberikan informasi tentang sejauh mana produk iphone diterima oleh masyarakat. Dengan memahami sentimen masyarakat, perusahaan Apple dapat mengambil tindakan yang diperlukan untuk meningkatkan kualitas produk dan kepuasan pengguna. Selain itu, penelitian ini juga memperkenalkan konsep Text Mining dan algoritma TF-IDF sebagai alat yang kuat untuk menganalisis data teks dalam konteks analisa sentimen.

Kata Kunci: Iphone; Analisa Sentimen; Text Mining; TF-IDF

Abstract—The iPhone is a product that has become a major concern in society and has become one of the main needs in everyday life. However, sometimes the iPhone often faces several problems that need attention. One problem that is often the main focus is the fairly high price. Therefore, we need a system that can determine the public's view of the iPhone product. This research uses text mining and TF-IDF to determine people's views on iPhone products. Text mining can be defined as the discovery of new, previously unknown information and the automatic extraction of valuable information from text from different sources. Meanwhile, TF-IDF is used to determine the frequency value of words in a document. In this research, sentiment refers to people's views on iPhone products, whether positive or negative. The final result of this sentiment analysis is that the positive sentiment value is 68.65% while the negative sentiment value is 31.35%. This is expected to provide information about the extent to which iPhone products are accepted by the public. By understanding people's sentiments, Apple company can take necessary actions to improve product quality and user satisfaction. Apart from that, this research also introduces the concept of Text Mining and the TF-IDF algorithm as a powerful tool for analyzing text data in the context of sentiment analysis.

Keywords: Iphone; Sentiment Analysis; Text Mining; TF-IDF

1. PENDAHULUAN

Algoritma text mining melewati beberapa tahapan Preprocessing yang diantaranya case folding, tokenizing, filtering dan stemming. Text mining telah menjadi alat yang kuat untuk mengolah dan menganalisis jumlah besar teks yang ada, membantu organisasi dan individu dalam menggali informasi berharga dari teks yang sebelumnya tidak terstruktur[1].

Signifikansi produk teknologi, khususnya iphone, dalam era digital. Meskipun memiliki popularitas tinggi dan penggemar setia, iphone menghadapi beberapa masalah seperti harga tinggi, bug perangkat lunak, masalah baterai, dan keterbatasan penyimpanan. Keluhan-keluhan ini sering diungkapkan melalui media sosial, mempengaruhi sentimen masyarakat terhadap produk. Penelitian ini menggunakan algoritma text mining dan TF-IDF untuk menganalisis sentimen masyarakat terhadap iphone berdasarkan data komentar di media sosial. Masalah yang ditemui mencakup harga, daya tahan baterai, keterbatasan penyimpanan, dan kritik terhadap kebijakan harga dan privasi Apple. Text mining digunakan untuk mengumpulkan dan menganalisis keluhan-keluhan ini, sedangkan TF-IDF digunakan untuk menghitung akurasi sentimen positif dan negatif[2].

Banyaknya masalah yang ada pada produk iphone disebut sebagai faktor penyebab kurangnya daya tarik masyarakat untuk menggunakan smartphone ini. Dengan adanya media sosial saat ini memudahkan masyarakat mengungkapkan keluhannya lewat sosial media. Dengan banyaknya keluhan yang terjadi sehingga kita perlu menganalisa keluhan tersebut dengan cara mengumpulkan semua keluhan yang diungkapkan masyarakat lewat sosial media tersebut. Sentimen merupakan pandangan positif atau negatif, serta berbagai perasaan atau sikap yang terkait dengan penggunaan dan pengalaman produk iphone. Maka sentimen yang dimaksud dalam penelitian merujuk pada pandangan masyarakat atau opini yang dimiliki masyarakat terhadap produk iphone dari apple. Penelitian ini melihat aplikasi yang dibangun dengan menggunakan bahasa pemrograman. Hasil pengujian aplikasi ini dapat digunakan sebagai sarana untuk mendiskusikan keamanan informasi atau pesan untuk memastikan keamanan tanpa diketahui oleh orang yang tidak berwenang.

Text mining Term Frequency-Inverse Document Frequency (TF-IDF) dapat menghitung skor kemiripan. Dengan kata-kata dalam sumber informasi sesuai genre. Maka penelitian ini dilakukan untuk mengetahui cara menganalisis sentimen produk iphone berdasarkan data komentar yang dikumpulkan pada media sosial[3]. Data komentar yang

dikumpulkan dan diproses menggunakan algoritma text mining untuk mengetahui tujuan komentar masyarakat yang akan dibagi menjadi dua kategori yaitu positif dan negatif[4]. Selanjutnya menerapkan algoritma Term Frequency-Inverse Document Frequency (TF-IDF) untuk menghasilkan jumlah akurasi sentimen positif dan negatif terkait analisa sentimen produk iphone.

Beberapa penelitian terdahulu seperti yang dilakukan oleh Muhammad Hafizh Mahendra, dkk pada tahun 2023 dimana hasil dari penelitian yang didapatkan bahwasannya penggunaan TF-IDF sebagai feature extraction dengan max features 10000, pada rasio data splitting 70:30, menghasilkan akurasi tertinggi yaitu sebesar 74,4%[5]. Penelitian lainnya yang dilakukan oleh Vincentius Westley Dimitrius Thomas dan Fitrah Rumaisa pada tahun 2022 dimana hari dari penelitian menggunakan TF-IDF dengan melakukan tuning terhadap parameter dan memilih parameter terbaik yaitu parameter dengan min_df sebesar 3, max_df sebesar 0.25, dan max_features sebesar 5000 Pada eksperimen yang dilakukan penelitian ini berdasarkan parameter tuning diperoleh hasil F1-Score terbaik sebesar 83% dan pada eksperimen apabila tanpa implementasi slang handling diperoleh nilai F1-Score sebesar 80% sedangkan eksperimen apabila tanpa stopword diperoleh hasil F1-Score sebesar 82%. Kemudian pada evaluasi kinerja model dengan K-fold, nilai akurasi tertinggi berada pada Fold-7 sebesar 87%[6].

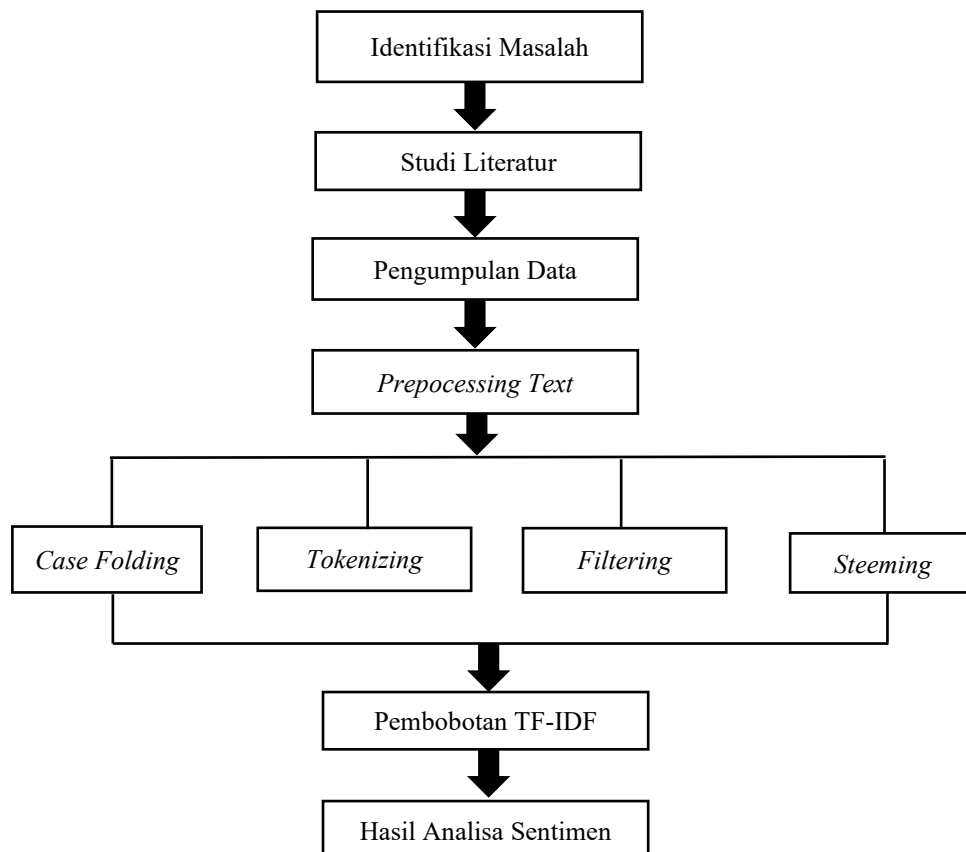
Selain itu, juga dilakukan penelitian oleh Okta Ihza Gifari, dkk pada tahun 2022 dimana dari proses pengujian dapat digunakan untuk kasus review film dengan nilai Accuracy 85%, nilai Precision 100%, nilai Recall 70%, dan nilai F1-Score sebesar 82%[7]. Dan penelitian terakhir yang dilakukan oleh Jessica Emarapenta Br Sinulingga dan Hizkya Cesar Kayika Sitorus pada tahun 2024 dimana proses pada penelitian menghasilkan Hasil yang diperoleh dari analisis ini adalah sebesar 82.51% untuk akurasi, 5.28% untuk presisi, 7.26% untuk recall, dan 6.12% untuk F1 Score[8].

Maka penelitian ini dilakukan untuk mengetahui cara menganalisis sentimen produk iphone berdasarkan data komentar yang dikumpulkan pada media sosial. Data komentar yang dikumpulkan dan diproses menggunakan algoritma text mining untuk mengetahui tujuan komentar masyarakat yang akan dibagi menjadi dua kategori yaitu positif dan negatif dan diharapkan dapat memberikan wawasan tentang pandangan masyarakat terhadap iPhone dan kontribusi algoritma text mining serta TF-IDF dalam menganalisis sentimen produk.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian adalah level atau tingkatan bisa disebut juga jenjang dalam sebuah aktivitas penelitian. Dimana tahapan tersebut terdapat memiliki proses yang dilakukan secara terstruktur, baku, logis dan sistematis. Adapun urutan langkah-langkah pada penelitian ini dapat dilihat pada struktur berikut:



Gambar 1. Kerangka Kerja Penelitian

2.2 Natural Language Processing (NLP)

Natural Language Processing (NLP) salah satu ilmu Artificial Intelligence pada ilmu komputer yang mempelajari interaksi antara komputer dan manusia dimana komputer dibuat untuk dapat mengerti dan memanipulasi bahasa alami yang berupa kalimat atau sebuah ungkapan. Model komputasi ini dapat memudahkan manusia untuk mendapatkan informasi pada komputer. Manipulasi sebuah teks sudah dikenal dan termasuk bidang penelitian yang penting dalam Natural Language Processing (NLP)[9], [10].

2.3 Analisis Sentimen

Analisa sentimen adalah proses menganalisis komentar masyarakat mengenai suatu peristiwa yang sedang terjadi. Biasanya, analisa sentimen digunakan untuk mengumpulkan dan memahami komentar masyarakat yang terdapat dalam postingan di berbagai media sosial seperti Twitter, Facebook, Blog dan lain sebagainya. Tujuan dari analisa sentimen adalah untuk mengetahui tujuan dari komentar yang masyarakat sampaikan terhadap sebuah objek. Opini-opini tersebut dapat beragam, baik berupa positif dan negatif itu tergantung pada pandangan yang disampaikan masyarakat terhadap peristiwa yang telah terjadi[11], [12].

2.4 Text Mining

Text mining salah satu teknik yang dapat diterapkan untuk membuat klasifikasi, penentuan maupun clustering yang dimana text mining adalah hasil penambangan data dari banyaknya data teks yang berjumlah besar. Dalam memproses suatu data yang besar pastinya akan membutuhkan sumber daya yang banyak kaitannya dengan proses pengolahan tersebut. Text mining juga dikatakan sebagai konsep yang diterapkan untuk mencari pola dalam teks dengan tujuan tertentu. Agar dapat melakukan penarikan informasi untuk menentukan kualitas layanan rumah sakit maka ada beberapa tahapan yang harus dilakukan agar sumber data yang akan diolah menjadi baik dan terstruktur. Terdapat beberapa tahapan dalam text mining sebagai berikut[13], [14]:

1. Case folding tahapan mengkonversi semua huruf kapital menjadi huruf kecil. Peran case folding sangat dibutuhkan untuk dapat mengubah semua huruf ke dalam bentuk standar agar dapat memudahkan untuk lanjut ke tahap selanjutnya.
2. Tokenizing tahapan yang dilakukan untuk membagi kalimat sesuai susunannya menjadi sebuah kata tunggal untuk mempermudah proses perhitungannya.
3. Filtering tahapan yang dilakukan untuk menghilangkan kata yang tidak penting dari hasil proses tokenizing. Tahapan filtering memanfaatkan library stopwords menggunakan tala untuk mempermudah dalam menghapus kata-kata yang tidak penting.
4. Stemming tahapan yang dilakukan untuk mengembalikan kata-kata yang telah di filtering menjadi kata-kata dasar dengan menghilangkan imbuhan awal atau akhir pada kata.

2.5 Term Frequency-Inverse Document Frequency (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) metode yang digunakan untuk mencari bobot hubungan kata (term) pada dokumen. Metode ini gabungan antara dua konsep untuk menghitung bobot, yaitu Term Frequency (TF) untuk mencari bobot kemunculan kata dalam sebuah dokumen dan Inverse Document Frequency (IDF) untuk mencari frekuensi kemunculan kata. Mencari frekuensi kemunculan kata di dalam dokumen dilakukan agar dapat mengetahui seberapa penting kata tersebut di dalam dokumen sehingga bobot antara kata dan dokumen akan tinggi apabila frekuensi kemunculan kata dalam dokumen tersebut tinggi[15], [16].

2.5.1 Term Frequency (TF)

Penelitian ini menggunakan jenis perhitungan Term Frequency Raw (TF Murni) dimana perhitungan ini jenis perhitungan yang sering digunakan. Perhitungan bobot pada jenis perhitungan Term Frequency Raw (TF Murni) dapat menggunakan rumus, sebagai berikut[17]:

$$tf = tf_{d,t} \quad (1)$$

Keterangan:

tf = Bobot Term Frequency (TF).

$tf_{d,t}$ = Jumlah kemunculan kata kunci (t) pada setiap dokumen (d).

2.5.2 Inverse Document Frequency (IDF)

Inverse Document Frequency (IDF) untuk mencari bobot kemunculan dokumen yang mengandung kemunculan kata tersebut. Perhitungan bobot pada Inverse Document Frequency (IDF) dapat menggunakan rumus sebagai berikut[18]:

$$idf_t = \log \left(\frac{N}{df_t} \right) + 1 \quad (2)$$

Keterangan:

idf_t = inverse document-frequency.

N = Jumlah seluruh dokumen.

df_t = Jumlah dokumen yang mengandung kata (t)

Selanjutnya tahap perhitungan bobot Term Frequency-Inverse Document Frequency (TF-IDF) dengan menggabungkan perhitungan bobot dari persamaan 1 dan 2, sebagai berikut[19]:

$$W_{d,t} = tf_{d,t} \times idf_t = tf_{d,t} \times \log\left(\frac{N}{df_t}\right) + 1 \tag{3}$$

Keterangan:

$W_{d,t}$ = Bobot Term Frequency-Inverse Document Frequency (TF-IDF).

$tf_{d,t}$ = Jumlah kemunculan kata kunci (t) pada setiap dokumen (d).

N = Jumlah seluruh dokumen.

df_t = Jumlah dokumen yang mengandung kata kunci (t).

Selanjutnya tahap menghitung total keseluruhan bobot Term Frequency-Inverse Document Frequency (TF-IDF) untuk mendapatkan hasil keseluruhan sentimen dengan menggunakan rumus persamaan 4 dan persamaan 5, sebagai berikut:

$$\text{Sentimen Positif} = \frac{100}{\text{Konotasi Positif} + \text{Konotasi Negatif}} \times \text{Konotasi Positif} \tag{4}$$

$$\text{Sentimen Negatif} = \frac{100}{\text{Konotasi Positif} + \text{Konotasi Negatif}} \times \text{Konotasi Negatif} \tag{5}$$

3. HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan analisis sentimen masyarakat yang berupa tanggapan atau opini terhadap informasi tentang iphone yang di dalamnya terdapat keluhan masyarakat terhadap produk iphone melalui media sosial twitter. Tahap ini merupakan tahapan paling awal yang dilakukan untuk memecahkan permasalahan yang ada. Peneliti menganalisa komentar masyarakat terhadap kualitas pelayanan rumah sakit dengan data yang didapat dari proses penarikan data dan dibuat menjadi dua kategori yaitu positif dan negatif.

Data diambil melalui aplikasi twitter dengan kata kunci “pelayanan rumah sakit”. Data yang telah didapat akan dilakukan tahap pengolahan data dengan memilih data yang akan digunakan sebagai data penerapan algoritma. Tahap pengolahan dilakukan untuk mengambil data yang hanya berisikan komentar yang disampaikan oleh masyarakat, karena data yang dicrawling biasanya berisikan data campur antara data postingan berita dan data komentar yang disampaikan oleh masyarakat pada sebuah postingan berita. Setelah melewati tahap pengolahan data maka selanjutnya dilakukan proses text mining sebagai preprocessing text dengan tahapan case folding, tokenizing, filtering, dan stemming. Text mining dilakukan untuk mengidentifikasi dan mengklasifikasikan opini-opini yang disampaikan masyarakat dan membuat data komentar menjadi lebih terstruktur. Dengan menggunakan text mining peneliti dapat mengolah data yang belum diketahui bentuk dan maksud bentuk opininya. Setelah tahapan dari text mining selesai maka selanjutnya diterapkan algoritma Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini gabungan dari dua konsep untuk perhitungan bobot, yaitu Term Frequency (TF) yang digunakan untuk mencari frekuensi kemunculan kata di dalam dokumen dan Inverse Document Frequency (-IDF) yang digunakan untuk menghitung frekuensi dokumen yang mengandung term atau kata tersebut. Frekuensi kemunculan kata pada sebuah dokumen akan menunjukkan seberapa penting term atau kata dalam dokumen tersebut. Proses ini akan membantu peneliti memahami sejauh mana suatu term atau kata dapat berkontribusi terhadap sentiment positif, negatif dalam data yang dianalisis.

Dari penjelasan tersebut dapat disimpulkan bahwa hasil yang didapat dari analisis sentimen berdasarkan penilaian, yaitu kata yang mengandung konotasi positif dan negatif yang bertujuan untuk mendapatkan jumlah hasil sentimen positif dan negatif dari masyarakat terhadap kualitas produk iphone .

3.1 Text Mining

Setelah data diolah pada algoritma text mining, maka diperoleh hasil data dari penerapan preprocessing text. Hasil data yang telah melewati penerapan algoritma text mining dapat dilihat pada tabel berikut:

Tabel 1. Hasil Preprocessing Text

No	Komentar	Case Folding	Tokenizing	Filtering	Stemming
1	Berkat fitur Pesan	berkat fitur pesan	[berkat] [fitur]	[berkat] [fitur]	[berkat] [fitur]
	SMS ke Ponsel	sms ke ponsel	[pesan] [sms] [ke]	[pesan] [sms]	[pesan] [sms]
	Satelit di	satelit di iphone	[ponsel] [satelit] [di]	[ponsel] [satelit]	[ponsel] [satelit]
	iPhone 14	14 berhasil	[iphone] [14]	[iphone] [14]	[iphone] [14]
	berhasil	menyelamat	[berhasil]	[berhasil]	[berhasil]
	menyelamat	remaja yang	[menyelamat]	[menyelamat]	[selamat] [remaja]
	remaja yang	terdampar di	[remaja] [yang]	[remaja] [terdampar]	[dampar][ngarai]
	terdampar di	ngarai santa paula	[terdampar]	[ngarai] [santa]	[santa] [paula]
	Ngarai Santa		[di][ngarai] [santa]	[paula]	
	Paula		[paula]		

Setelah dilakukan hasil penerapan algoritma text mining pada analisa layanan rumah sakit, langkah selanjutnya adalah melakukan pembobotan menggunakan algoritma Term Frequency-Inverse Document Frequency (TF-IDF).

3.2 Algoritma Term Frequency-Inverse Document Frequency (TF-IDF)

Setelah dilakukan pengolahan data dengan menggunakan text mining maka tahapan selanjutnya adalah mengambil beberapa kata kunci yang berkonotasi positif dan negatif dari hasil penerapan algoritma text mining. Cara kerjanya dengan memahami makna kalimat berdasarkan konteks yang dibicarakan dari setiap data. Kata kunci yang berkonotasi positif dan negatif dapat dilihat pada tabel 2 dan 3 berikut.

Tabel 2. Konotasi Positif

No	Konotasi Positif
1	Iphone
2	Fitur
3	Enak
4	Keren
5	Jernih
6	Flagship
7	Manteb
8	Bagus
9	Unggul
10	Target

Tabel 3. Konotasi Negatif

No	Konotasi Negatif
1	Beda
2	Benci
3	Nyesel
4	Ragu
5	Ancur
6	Mahal
7	Jelek
8	Rusak
9	Miskin
10	Jauh

Berdasarkan dari hasil kata kunci konotasi positif dan negatif yang telah ditentukan, maka dapat dilakukan pembobotan Term Frequency (TF). Pembobotan Term Frequency (TF) bertujuan untuk mencari frekuensi kemunculan kata kunci dalam sebuah dokumen. Pembobotan Term Frequency (TF) dapat dilakukan dengan menggunakan rumus persamaan 1 yang sudah ada pada metode penelitian dan hasil yang diperoleh dari pembobotan Term Frequency (TF) disajikan pada tabel 4 dan tabel 5 berikut.

Tabel 4. Hasil Konotasi Postitif Pembobotan Term Frequency (TF)

No	Kata Kunci	D1	D2	D3	D4	D5	D16	D17	...	D19	D20
1	Iphone	1	1	1	1	1	1	1	...	1	1
2	Fitur	1	0	1	1	0	1	1	...	1	1
3	Enak	0	0	0	0	0	0	0	...	0	0
4	Keren	0	0	0	0	0	0	0	...	0	0
5	Jernih	0	0	0	0	0	0	0	...	0	0
6	Flagship	0	0	0	0	0	0	0	...	0	0
7	Manteb	0	0	1	0	0	0	0	...	0	0
8	Bagus	0	0	0	0	0	0	0	...	0	0
9	Unggul	0	0	0	0	0	0	0	...	0	0
10	Target	0	0	0	0	0	0	0	...	0	0

Tabel 5. Hasil Konotasi Negatif Pembobotan Term Frequency (TF)

No	Kata Kunci	D1	D2	D3	D4	D5	D16	D17	...	D19	D20
1	Beda	0	1	1	0	0	0	0	...	0	1
2	Benci	0	0	0	1	0	0	0	...	0	0
3	Nyesel	0	0	0	0	0	0	0	...	1	0
4	Ragu	0	0	0	0	0	0	0	...	0	0
5	Ancur	0	0	0	0	0	0	0	...	0	0
6	Mahal	0	0	0	0	1	0	0	...	0	0
7	Jelek	0	0	0	0	0	0	0	...	0	0
8	Rusak	0	0	0	0	0	0	0	...	0	0

No	Kata Kunci	D1	D2	D3	D4	D5	D16	D17	...	D19	D20
9	Miskin	0	0	0	0	0	0	0	...	0	0
10	Jauh	0	0	0	0	0	0	0	...	0	0

Setelah diperoleh hasil dari pembobotan Term Frequency (TF), maka langkah selanjutnya adalah melakukan pembobotan Inverse Document Frequency (IDF). Pembobotan Inverse Document Frequency (IDF) bertujuan untuk menghitung frekuensi seluruh dokumen yang mengandung kata kunci. Pembobotan Inverse Document Frequency (IDF) dapat dilakukan dengan menggunakan rumus persamaan 2 dan satu contoh kata kunci “bagus” yang akan digunakan sebagai perhitungan pada pembobotan Inverse Document Frequency (IDF). Proses perhitungan bobot selanjutnya dilakukan dengan cara yang sama dan hasil keseluruhan dari pembobotan Inverse Document Frequency (IDF) disajikan pada tabel 6 dan 7 berikut.

$$idf_t = \log\left(\frac{N}{df_t}\right) + 1$$

$$= \log\left(\frac{20}{22}\right) + 1$$

$$= 0,9586$$

Tabel 6. Hasil Konotasi Positif Pembobotan Inverse Document Frequency (IDF)

Konotasi Positif				
No	Kata Kunci	DF	N	$\log\left(\frac{N}{df_t}\right) + 1$
1	Iphone	22	20	0,958
2	Fitur	12	20	1,221
3	Enak	3	20	1,823
4	Keren	1	20	2,301
5	Jernih	1	20	2,301
6	Flagship	1	20	2,301
7	Manteb	1	20	2,301
8	Bagus	1	20	2,301
9	Unggul	1	20	2,301
10	Target	1	20	2,301

Tabel 7. Hasil Konotasi Negatif Pembobotan Inverse Document Frequency (IDF)

Konotasi Negatif				
No	Kata Kunci	DF	N	$\log\left(\frac{N}{df_t}\right) + 1$
1	Beda	3	20	1,823
2	Benci	1	20	2,301
3	Nyesel	1	20	2,301
4	Ragu	1	20	2,301
5	Ancur	1	20	2,301
6	Mahal	1	20	2,301
7	Jelek	1	20	2,301
8	Rusak	1	20	2,301
9	Miskin	1	20	2,301
10	Jauh	1	20	2,301

Setelah diperoleh hasil dari pembobotan Inverse Document Frequency (IDF), maka langkah selanjutnya adalah melakukan pembobotan Term Frequency-Inverse Document Frequency (TF-IDF). Pembobotan Term Frequency-Inverse Document Frequency (TF-IDF) bertujuan untuk menghitung total frekuensi setiap kata kunci konotasi positif dan negatif. Pembobotan Term Frequency-Inverse Document Frequency (TF-IDF) dapat dilakukan dengan menggunakan rumus persamaan 2.3 dan satu contoh yang akan digunakan dalam perhitungan pada pembobotan Term Frequency-Inverse Document Frequency (TF-IDF). Perhitungan yang digunakan sebagai contoh pembobotan yaitu menggunakan hasil dari kata kunci “bagus” pada Inverse Document Frequency (IDF) pada tabel 6 dan dokumen nomor satu (D1) dari hasil Term Frequency (TF) pada tabel 4. Proses perhitungan bobot selanjutnya dilakukan dengan cara yang sama dan hasil keseluruhan dari pembobotan pada Term Frequency-Inverse Document Frequency (TF-IDF) disajikan pada tabel 8 dan 9 berikut :

$$W_{d,t} = tf_{d,t} \times \log\left(\frac{N}{df_t}\right) + 1$$

$$= 22 \times 0,9586$$

$$= 21,089$$

Tabel 8. Hasil Konotasi Positif Pembobotan Term Frequency-Document Frequency (TF-IDF)

No	Kata Kunci	D1	D2	D3	D4	D5	D17	D18	...	D19	D20
1	Iphone	0,9586	0,9586	0,9586	0,9586	0,9586	0,9586	0,9586	...	0,9586	0,9586
2	Fitur	1,2218	0	1,2218	1,2218	0	0	0	...	1,2218	1,2218
3	Enak	0	0	0	0	0	1,8239	0	...	3,6478	0
4	Keren	0	0	0	0	0	0	0	...	0	0
5	Jernih	0	0	0	0	0	2,3010	0	...	0	0
6	Flagship	0	0	0	0	0	0	0	...	0	0
7	Manteb	0	0	2,3010	0	0	0	0	...	0	0
8	Bagus	0	0	0	0	0	0	0	...	0	0
9	Unggul	0	0	0	0	0	0	0	...	0	0
10	Target	0	0	0	0	0	0	0	...	0	0

Tabel 9. Hasil Konotasi Negatif Pembobotan Term Frequency-Document Frequency (TF-IDF)

No	Kata Kunci	D1	D2	D3	D4	D5	D17	D18	...	D19	D20
1	Beda	0	1,8239	1,8239	0	0	0	0	...	0	1,8239
2	Benci	0	0	0	2,3010	0	0	0	...	0	0
3	Nyesel	0	0	0	0	0	0	0	...	2,3010	0
4	Ragu	0	0	0	0	0	0	0	...	0	0
5	Ancur	0	0	0	0	0	0	0	...	0	0
6	Mahal	0	0	0	0	2,3010	0	0	...	0	0
7	Jelek	0	0	0	0	0	0	0	...	0	0
8	Rusak	0	0	0	0	0	0	0	...	0	0
9	Miskin	0	0	0	0	0	0	0	...	0	0
10	Jauh	0	0	0	0	0	0	0	...	0	0

Berdasarkan hasil pembobotan Term Frequency-Inverse Document Frequency (TF-IDF) pada tabel 8 dan 9 dapat disimpulkan hasil pembobotan Term Frequency-Inverse Document Frequency (TF-IDF) dari setiap kata kunci konotasi positif dan negatif. Menyimpulkan perhitungan dilakukan dengan cara menjumlahkan keseluruhan nilai bobot dokumen dari setiap kata kunci, misalnya kata kunci “bagus” pada tabel 8 dilakukan penjumlahan nilai bobot dari seluruh dokumen mulai dokumen satu (D1) sampai dokumen 20 (D20) sehingga menghasilkan jumlah nilai bobot dari keseluruhan dokumen sebesar 21,0892. Proses perhitungan nilai bobot dokumen dari setiap kata kunci selanjutnya dilakukan dengan cara yang sama dan hasil keseluruhan dari pembobotan pada Term Frequency-Inverse Document Frequency (TF-IDF) disajikan pada tabel 10 dan 11 berikut.

Tabel 10. Hasil Pembobotan Konotasi Positif

No	Konotasi Positif	Bobot Konotasi Positif
1	Iphone	21,0892
2	Fitur	14,6616
3	Enak	5,4717
4	Keren	2,3010
5	Jernih	2,3010
6	Flagship	2,3010
7	Manteb	2,3010
8	Bagus	2,3010
9	Unggul	2,3010
10	Target	2,3010
Total		57,329

Tabel 11. Hasil Pembobotan Konotasi Negatif

No	Konotasi Negatif	Bobot Konotasi Negatif
1	Beda	5,4717
2	Benci	2,3010
3	Nyesel	2,3010
4	Ragu	2,3010
5	Ancur	2,3010
6	Mahal	2,3010
7	Jelek	2,3010

No	Konotasi Negatif	Bobot Konotasi Negatif
8	Rusak	2,3010
9	Miskin	2,3010
10	Jauh	2,3010
Total		26,180

Berdasarkan hasil dari kesimpulan total nilai bobot dari konotasi positif dan negatif pada tabel 10 dan 11, maka selanjutnya melakukan perhitungan nilai sentimen dari keseluruhan total nilai konotasi positif dan negatif. Perhitungan sentimen ini bertujuan untuk mengetahui total nilai sentimen positif dan negatif mengenai kualitas layanan rumah sakit di Indonesia. Menghitung total nilai sentimen positif dan negatif dapat dilakukan dengan menggunakan rumus persamaan 4 dan 5 dan hasil dari perhitungan sentimen positif dan negatif disajikan pada tabel 12 berikut:

$$\text{Sentimen Positif} = \frac{100}{\text{Konotasi Positif} + \text{Konotasi Negatif}} \times \text{Konotasi Positif}$$

$$= \frac{100}{57,329 + 26,180} \times 57,329 = 68,65 \%$$

$$\text{Sentimen Negatif} = \frac{100}{\text{Konotasi Positif} + \text{Konotasi Negatif}} \times \text{Konotasi Negatif}$$

$$= \frac{100}{26,180 + 57,329} \times 26,180 = 31,35 \%$$

Tabel 12. Hasil Sentimen

Sentimen	Hasil Sentimen
Positif	68,65%
Negatif	31,35%
Total	100%

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan maka dapat dibuat kesimpulan sebagai berikut Penelitian ini menggunakan data tweet dimana mayoritas pengguna twitter memberikan tanggapan negatif terhadap kualitas layanan rumah sakit di Indonesia. Kata yang mengandung konotasi positif dan negatif sangat berpengaruh untuk mendapatkan jumlah hasil sentimen positif dan negatif dari masyarakat terhadap kualitas layanan rumah sakit di Indonesia. Algoritma text mining dan TF-IDF mampu menganalisa produk iphone berdasarkan sentimen pada komentar tweet.

REFERENCES

- [1] A. Hermawan, I. Jowensen, J. Junaedi, and Edy, "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine," *JST (Jurnal Sains dan Teknol.*, vol. 12, no. 1, pp. 129–137, 2023, doi: 10.23887/jstundiksha.v12i1.52358.
- [2] E. Harieby, H. Hoiriyah, and M. Walid, "Twitter Text Mining Mengenai Isu Vaksinasi Covid-19 Menggunakan Metode Term Frequency, Inverse Document Frequency (Tf-Idf)," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 532–537, 2022, doi: 10.36040/jati.v6i2.5129.
- [3] Ida Bagus Ketut Surya Arnawa, "Analisis Sentimen pada Media Sosial Terhadap Perkuliahan Hybrid Menggunakan Algoritma TF IDF dan K Nearest Neighbor," *J. Sist. dan Inform.*, vol. 17, no. 1, pp. 40–46, 2023, doi: 10.30864/jsi.v17i1.495.
- [4] R. Wati, S. Ernawati, and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH," *J. Manaj. Inform.*, vol. 13, no. 1, pp. 84–93, 2023, doi: 10.34010/jamika.v13i1.9424.
- [5] M. H. Mahendra, D. T. Murdiansyah, and K. M. Lhaksana, "Analisis Sentimen Tweet COVID-19 menggunakan K-Nearest Neighbors dengan TF-IDF dan Ekstraksi Fitur CountVectorizer," *DIKE J. Ilmu Multidisiplin*, vol. 1, no. 2, pp. 37–43, 2023, doi: 10.69688/dike.v1i2.35.
- [6] V. W. D. Thomas and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," *J. Media Inform. Budidarma*, vol. 6, no. 3, p. 1767, 2022, doi: 10.30865/mib.v6i3.4218.
- [7] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, "Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine," *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022.
- [8] J. E. Br Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF," *J. Manaj. Inform.*, vol. 14, no. 1, pp. 42–53, 2024, doi: 10.34010/jamika.v14i1.11946.
- [9] R. C. Rivaldi and T. D. Wismarini, "Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP) (Studi Kasus Zalika Store 88 Shopee)," *J. Ilm. Elektron. DAN Komput.*, vol. 17, no. 1, pp. 120–128, 2024.
- [10] N. Nurwanda, N. Suarna, and W. Prihartono, "Penerapan Nlp (Natural Language Processing) Dalam Analisis Sentimen Pengguna Telegram Di Playstore," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 1841–1846, 2024, doi: 10.36040/jati.v8i2.8469.
- [11] I. H. Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 302–307, 2023, doi: 10.30591/jpit.v8i3.5734.
- [12] W. Sejati, A. Singh Bist, and A. Tambunan, "Pengembangan Analisis Sentimen Dalam Rekayasa Software Engineering Menggunakan Tinjauan Literatur Sistematis," *J. MENTARI Manaj. Pendidik. dan Teknol. Inf.*, vol. 2, no. 1, pp. 95–103, 2023, [Online]. Available: <https://journal.pandawan.id/mentari/article/view/377>
- [13] K. C. Astuti, A. Firmansyah, and A. Riyadi, "Implementasi Text Mining Untuk Analisis Sentimen Masyarakat Terhadap Ulasan

- Aplikasi Digital Korlantas Polri pada Google Play Store,” *REMIK Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 8, no. 1, pp. 383–394, 2024.
- [14] A. W. V. Hutabarat, N. L. S. S. Adnyani, and K. Suryadi, “Analisis Sentimen Data Ulasan Pengguna MyPertamina di Twitter dengan Metode Text Mining,” *J. Rekayasa Sist. Ind.*, vol. 13, no. 1, pp. 145–154, 2024, doi: 10.26593/jrsi.v13i1.6958.145-154.
- [15] A. Aziz and Fauziah, “Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, p. 115, 2022, [Online]. Available: <https://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/download/430/407>
- [16] M. Musfiroh, A. Tholib, and Z. Arifin, “Analisis Sentimen Terhadap Ulasan Aplikasi Shopee di Google Play Store Menggunakan Metode TF-IDF dan Long Short-Term Memory),” *J. Electr. Eng. Comput.*, vol. 6, no. 2, pp. 371–381, 2024, doi: 10.33650/jeeecom.v6i2.8713.
- [17] R. Rahmadani, A. Rahim, and R. Rudiman, “Analisis Sentimen Ulasan ‘Ojol the Game’ Di Google Play Store Menggunakan Algoritma Naive Bayes Dan Model Ekstraksi Fitur Tf-Idf Untuk Meningkatkan Kualitas Game,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4988.
- [18] S. A. Helmayanti, F. Hamami, and R. Y. Fa’rifah, “Penerapan Algoritma Tf-Idf Dan Naïve Bayes Untuk Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Flip Pada Google Play Store,” *J. Indones. Manaj. Inform. dan Komun.*, vol. 4, no. 3, pp. 1822–1834, 2023, doi: 10.35870/jimik.v4i3.415.
- [19] N. Hidayah and Dodiman, “Implementasi Algoritma Multinomial Naïve Bayes, TF-IDF dan Confusion Matrix dalam Pengklasifikasian Saran Monitoring dan Evaluasi Mahasiswa Terhadap Dosen Teknik Informatika Universitas Dayanu Ikhsanuddin,” *J. Akad. Pendidik. Mat.*, vol. 10, no. 1, pp. 8–15, 2024, [Online]. Available: <https://doi.org/10.55340/japm.v10i1.1491>