

Analisis Sentimen Terhadap *Smartphone Flagship* Pada Media Sosial Berbasis *Python TextBlob* Menggunakan Algoritma Naive Bayes

Alif Ilman Nafian^{*1)}, Muhammad Farhan Fauzaan²⁾, dan Riyandi Firman Pratama³⁾

Program Studi Rekayasa Perangkat Lunak Kampus Daerah Cibiru

Universitas Pendidikan Indonesia

Bandung, Indonesia

alif.ilmannafian@upi.edu, faruzaan@upi.edu, riyandi.firman08@upi.edu

Abstract

Today, social media plays an important role in the advancement of organizations, industries and businesses. This is an important fundamental aspect that must be known by both business people and individuals. One way or another, everyone today is associated with social media. Each individual can even interact and exchange knowledge because of a mixture of technology and social relations. Sentiment analysis is an analytical technique that makes it possible to get information from users who express emotions, perspectives and opinions on the internet. Today there are many websites where users can express themselves briefly and easily. One of the strategic sectors in implementing sentiment analysis is the technology sector, especially the smartphone sector. The number of smartphone variants currently circulating is a problem for individuals to find the best smartphone product. The flagship smartphone sentiment analysis conducted in this article aims to find the best solution for 2 flagship smartphones from well-known smartphone manufacturers, namely the Samsung S22 Ultra and Xiaomi 12 Pro. The data taken comes from various social media, such as Twitter, YouTube, and GSMArena. The data set was then analyzed using Python Textblob and the results of the analysis were negative, positive, and neutral sentiments displayed using various visualizations. The final result is a Net Brand Reputation assessment to assess a brand's reputation on a number of social media.

Keywords: *Smartphone Flagship, Social Media, Python, TextBlob, Naive Bayes*

Abstrak

Dewasa ini, media sosial memainkan peran yang penting dalam kemajuan organisasi, industri, dan bisnis. Ini merupakan suatu aspek fundamental penting yang harus diketahui baik oleh pelaku bisnis maupun individu. Dengan satu atau lain cara, setiap orang saat ini dikaitkan dengan media sosial. Masing-masing individu bahkan sudah dapat berinteraksi dan bertukar ilmu karena campuran teknologi dan hubungan sosial. Analisis sentimen merupakan suatu teknik analisa yang memungkinkan untuk mendapatkan informasi dari pengguna yang mengekspresikan emosi, perspektif, dan pendapat di internet. Saat ini banyak situs web di mana pengguna dapat mengekspresikan diri mereka secara singkat dan mudah. Salah satu sektor yang strategis dalam implementasi sentimen analisis adalah sektor teknologi terkhusus sektor *smartphone*. Banyaknya varian *smartphone* yang beredar saat ini menjadi suatu problematika bagi individu untuk menemukan produk *smartphone* terbaik. Analisis sentimen *smartphone flagship* yang dilakukan pada artikel ini bertujuan untuk mencari solusi terbaik dari 2 *smartphone flagship* dari pabrikan ternama *smartphone*, yakni Samsung S22 Ultra dan Xiaomi 12 Pro. Data yang diambil bersumber dari berbagai media sosial, seperti Twitter, YouTube, dan GSMArena. Kumpulan data kemudian dianalisis menggunakan Python Textblob dan hasil analisis berupa sentimen negatif, positif, dan netral yang ditampilkan menggunakan berbagai visualisasi. Hasil akhir berupa penilaian *Net Brand Reputation* untuk menilai reputasi suatu merek di sejumlah media sosial.

Kata kunci: *Smartphone Flagship, Media Sosial, Python, TextBlob, Naive Bayes*

1. Pendahuluan

Pada masa kini, banyak pelaku bisnis melakukan berbagai cara untuk meningkatkan produk yang dihasilkan. Menjangkau para konsumen adalah metode yang sering digunakan oleh pelaku bisnis untuk belajar tentang perasaan pelanggan mereka. Penilaian kepuasan pelanggan, survey ulasan serta pelacakan kegiatan pelanggan adalah metode umum yang digunakan untuk mendapatkan tanggapan dan insight dari pelanggan. Tanggapan tersebut dapat menjadi sebuah informasi berharga yang dapat digunakan oleh organisasi bisnis untuk meningkatkan kualitas produk yang dihasilkan.

Beberapa tahun belakangan, bisnis *smartphone* mengalami peningkatan yang signifikan, tidak hanya dalam penjualan konvensional tetapi juga penjualan secara online. Namun, tidak semua *smartphone* memiliki kualitas yang tinggi untuk memenuhi kebutuhan dari pelanggan dan ini merupakan sesuatu yang harus diketahui oleh pelanggan. Saat membeli *smartphone*, pelanggan seharusnya dapat menyadari dan memahami karakteristik dan fungsi suatu *smartphone*, yang dapat diperoleh dari sebuah pendapat maupun testimoni pengguna terkait *smartphone* yang mana pada saat ini dapat dengan mudah dijumpai di internet.

Hal ini kemudian menjadi lumrah di internet dimana para individu yang berperan sebagai pelanggan dapat berbagi pemikiran dan pengalaman mereka pribadi terkait suatu produk smartphone. Namun, diperlukan waktu yang cukup lama bagi pelanggan untuk membaca review tersebut secara keseluruhan. Klasifikasi sentimen hadir untuk menyelesaikan permasalahan tersebut dengan cara mengkategorikan review pengguna menjadi kategori positif atau negatif.

Analisis sentimen adalah sebuah teknik untuk mendeteksi opini terhadap suatu subyek (misalnya individu, organisasi ataupun produk) dalam sebuah kumpulan data [1]. Maraknya pemanfaatan internet memanjakan pengguna untuk dapat berdiskusi mengenai suatu produk melalui postingan di blog/web, debat online, situs survei produk, dan media sosial dalam internet. Media sosial merupakan suatu media yang dapat digunakan untuk berekspresi oleh penggunanya [2]. Pengguna dapat memanfaatkan media sosial seperti Twitter, YouTube, Facebook, dan Google untuk mengkomunikasikan pendapat dan perasaan mereka. Keluaran dari kegiatan tersebut menghasilkan sejumlah besar data yang kaya akan konsep dan sangat strategis untuk keperluan bisnis.

Artikel ini menggambarkan suatu sistem yang akan membantu pelanggan dalam memahami emosi global mengenai produk smartphone flagship berdasarkan pada ulasan pengguna dan akan membantu dalam proses pengambilan keputusan apakah akan membeli produk terkait atau tidak. Sebuah preprocessing data akan dilakukan pada korpus data yang terkumpul untuk kedua smartphone dari 2 brand ternama yaitu Samsung S22 Ultra dan Xiaomi 12 Pro, dimana keduanya merupakan smartphone flagship dari masing-masing brand. Smartphone flagship merupakan ponsel terbaik yang dimiliki masing-masing brand yang memiliki kualitas serta spesifikasi yang mumpuni.

Analisis sentimen yang diimplementasikan pada artikel ini menggunakan Python dengan package TextBlob. TextBlob merupakan suatu library yang mudah dan biasa digunakan untuk melakukan pemrosesan Natural Language Processing (NLP). Dimana nantinya akan dilakukan pelabelan sentimen secara otomatis menggunakan TextBlob sebelum akhirnya akan dilakukan analisa dan pengujian teks akurasi, presisi, recall, dan f_1 score dengan algoritma Naive Bayes.

Beberapa penelitian sebelum nya metode Naive Bayes memiliki tingkat akurasi lebih tinggi dari K-NN. Akurasi Naive Bayes sebesar 87.48% dan k-NN 85.40% dalam analisis sentimen tweet pornografi kaum homoseksual indonesia di twitter [3]. Dalam suatu penelitian diperlukan dukungan hasil-hasil penelitian yang telah ada sebelumnya yang berkaitan dengan penelitian tersebut. Sebagai contoh yang pertama ada penelitian [4] yang melakukan Analisis Sentimen Pada Review Film Dengan Menggunakan Algoritma Klasifikasi Naive Bayes Berdasarkan Term Objects Keywords. Memberikan hasil akhir hasil pengujian metode Naive Bayes classifier sebesar 28% dimana sampel yang digunakan adalah 100 data sentimen positif dan 100 data sentimen negatif. Penelitian dari [5] melakukan Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. Sehingga memperoleh hasil akhir dari 2000 data tweet berbahasa indonesia didapatkan akurasi optimal pada nilai $k=23$ sejumlah 79.99%. Penelitian dari [6] melakukan Analisis Sentimen Dokumen Twitter mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier. Didapatkan hasil akhir dengan metode Naive Bayes mendapatkan akurasi 67% dan error rate sebesar 33%. Kemudian ada Penelitian dari [7] yang melakukan Analisis Sentimen Pemindahan Ibu Kota Indonesia Pada Twitter. Memberikan hasil akhir dari klasifikasi Bernoulli Naive Bayes mendapatkan akurasi sebesar 68,10% dengan menggunakan skenario 11. Perbedaan penelitian yang penulis lakukan dengan penelitian sebelumnya dari keyword, pelabelan, dan source data. Dimana penulis menggunakan keyword smartphone flagship, pelabelan sentimen menggunakan Python TextBlob, dan source data dengan melakukan crawling data dari Twitter, YouTube, dan situs GSMArena.

2. Metode Penelitian

Metode dalam penelitian diperlukan agar penelitian lebih terstruktur, sehingga hasil yang akan diperoleh sesuai dengan tujuan pada penelitian [13]. Adapun tahapan metode penelitian seperti pada gambar 1.



Gambar 1 Tahapan Penelitian

A. Crawling Data

Crawling adalah aktivitas dalam membuat salinan dengan porsi yang relevan yang diambil dari sebuah World Wide Web [9]. Penelitian ini melakukan crawling di 3 sumber yang berbeda, yakni Twitter, YouTube, dan Website GSMArena. Data yang di crawl adalah Samsung S22 Ultra dan Xiaomi 12 Pro dan crawling data dilakukan pada tanggal 19 Desember 2022. Setelah itu dataset yang masih mentah akan disimpan dalam bentuk CSV (Comma Separated Values) yang nantinya akan diproses lebih lanjut. Hasil data crawling untuk kata

kunci Samsung S22 Ultra di Twitter, YouTube, dan GSMArena berturut-turut adalah 381, 220, dan 2605. Sedangkan hasil data crawling untuk kata kunci Xiaomi 12 Pro di Twitter, YouTube, dan GSMArena berturut-turut adalah 199, 226, dan 725.

B. Preprocessing Data

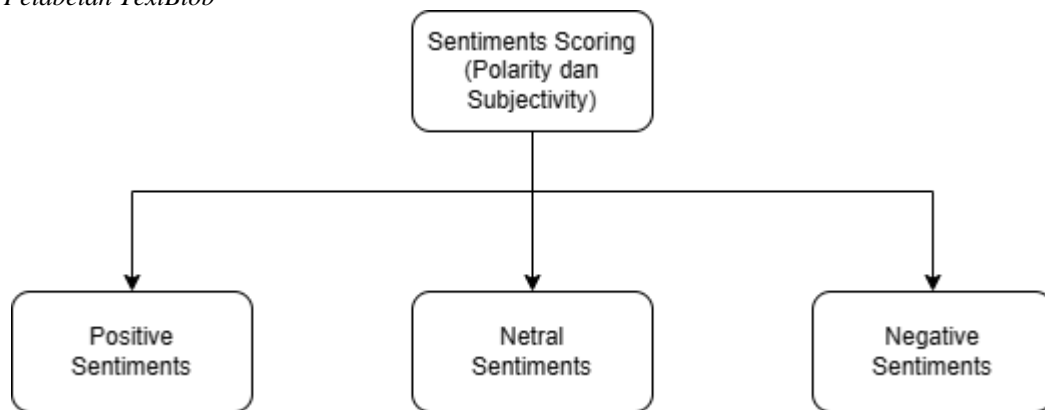
Setelah proses crawling data dan disimpan dalam bentuk CSV maka langkah selanjutnya adalah preprocessing data, dikarenakan dataset tersebut belum terstruktur. Tugas utama preprocessing data adalah untuk menghilangkan serta mengatasi noise data agar hasil perhitungan optimal. Adapun tahapan dalam preprocessing data diantara lain, cleaning data, case folding, tokenizing dan filtering. Berikut ini adalah tahapan preprocessing data yang ditunjukkan pada gambar 2.



Gambar 2 Tahapan Preprocessing Data

Tahapan preprocessing data yang pertama adalah cleaning data, ini dilakukan karena masih terdapat noise dalam data. Proses cleaning menghilangkan karakter, tanda baca, hashtag, alamat URL, dan angka pada data. Selanjutnya adalah case folding yang merupakan proses penyeragaman huruf menjadi huruf kecil. Setelah itu akan dilakukan tahap tokenizing, dimana kalimat dipecah menjadi sebuah token atau bagian-bagian yang lebih kecil. Lalu tahap terakhir dari preprocessing data adalah filtering, yang merupakan tahapan untuk menghilangkan kata yang tak bermakna (stopword).

C. Pelabelan TextBlob



Gambar 3 Tahapan Pelabelan TextBlob

Pada tahap ini dilakukan pelabelan sentimen secara otomatis menggunakan library TextBlob. Library TextBlob hanya dapat mengenali bahasa inggris. Dimana sebuah sentimen itu dapat dikatakan negatif, positif, atau netral karena telah diklasifikasikan dalam tahap ini. Nilai sentimen "+1" untuk sentimen positif, nilai "-1" untuk sentimen negatif, dan nilai "0" untuk sentimen netral.

D. Pengujian Algoritma Naive Bayes

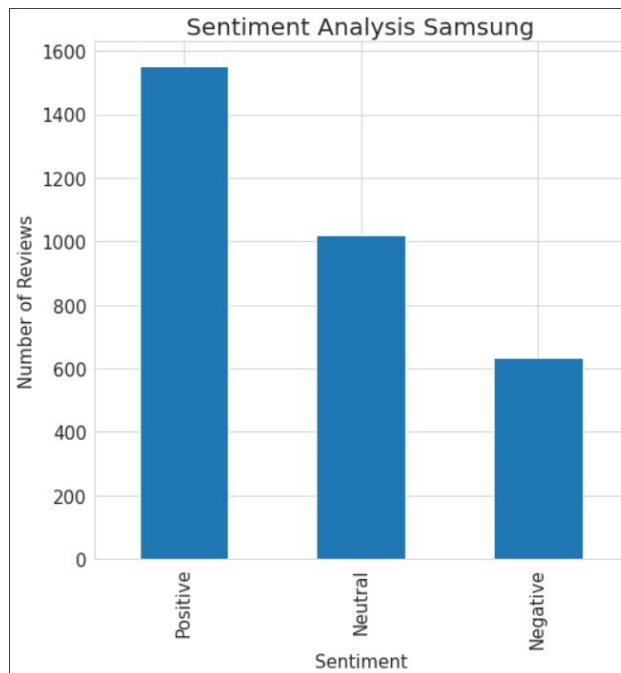
Metode Naive Bayes adalah metode yang memanfaatkan statistika sederhana dengan dasar teorema bayes yang mengasumsikan ada atau tidaknya suatu kelas dengan fitur yang lain. Metode Naive Bayes digunakan untuk proses klasifikasi untuk menentukan f1_score accuracy, recall, precision [10]. Pada penelitian ini, penulis menggunakan library Sklearn untuk implementasi rumus Naive Bayes.

3. Hasil dan Pembahasan

Hasil pemrosesan sentimen untuk keyword Samsung S22 Ultra menunjukkan hasil 1552 sentimen positif, 1019 sentimen netral, dan 635 sentimen negatif, seperti yang tertera pada tabel 1

Tabel 1 Hasil Sentimen Samsung S22 Ultra

Positif	Netral	Negatif	Total Sentimen
1552	1019	635	3206

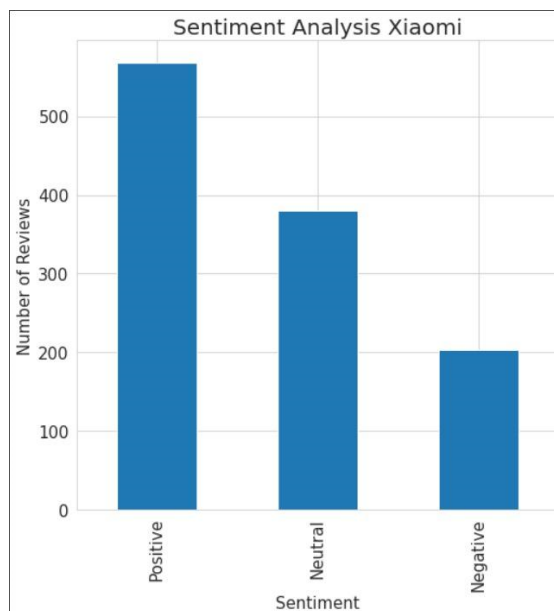


Gambar 4 Diagram Batang Sentimen Samsung S22 Ultra

Hasil pemrosesan sentimen untuk keyword Xiaomi 12 Pro menunjukkan hasil 568 sentimen positif, 380 sentimen netral, dan 203 sentimen negatif, seperti yang tertera pada tabel 2.

Tabel 2 Hasil Sentimen Xiaomi 12 Pro

Positif	Netral	Negatif	Total Sentimen
568	380	203	1151



Gambar 5 Diagram Batang Sentimen Xiaomi 12 Pro

Setelah dilakukan crawling, preprocessing, pelabelan otomatis dengan TextBlob, dan pengklasifikasian dengan Naïve Bayes maka selanjutnya adalah mengukur accuracy, precision, recall dan F1_Score menggunakan library Sklearn di dalam bahasa pemrograman Python dengan data testing sebesar 0.2 dan data training sebesar 0.8. Nilai pengujian Naive Bayes pun didapat untuk setiap merk (Samsung, Xiaomi) yang telah dikalkulasi dan tertera pada tabel 3.

Tabel 3 Hasil Pengujian Naive Bayes

	Accuracy	Precision	Recall	F1 Score
Samsung S22 Ultra	66	0.66	0.66	0.63
Xiaomi 12 Pro	64	0.62	0.64	0.61

A. Net Brand Reputation

Net Brand Reputation merupakan metode untuk menilai reputasi suatu merek di sejumlah media sosial. Penerapan NBR sederhananya mengukur tingkat loyalitas pelanggan. Rumus dari NBR adalah :

$$\text{NBR} = (\text{Sentimen Positif} - \text{Sentimen Negatif}) / (\text{Sentimen Positif} + \text{Sentimen Negatif}) \times 100 \quad (1)$$

Menggunakan rumus NBR di atas, menghasilkan brand reputation dimana hasilnya adalah Xiaomi muncul sebagai brand yang paling disukai ketimbang Samsung. Dengan detail nilai NBR sebagai berikut

Tabel 4 Nilai NBR

Samsung	Xiaomi
41.9	47.3

4. Kesimpulan

Dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan bahwa Analisis Sentimen dapat digunakan untuk mengetahui sentimen masyarakat dunia terhadap smartphone flagship dari 2 brand ternama yakni Samsung S22 Ultra dan Xiaomi 12 Pro. Tujuannya untuk mengetahui seberapa positif, netral atau negatif terhadap kedua keyword smartphone flagship tersebut. Setelah dilakukan pelabelan otomatis dengan Python TextBlob dengan total keseluruhan 4357 korpus data maka didapat jumlah sentimen positif, negatif, dan netral untuk Samsung S22 Ultra berturut-turut adalah 48.4%, 19.8%, dan 31.7%. Sedangkan untuk jumlah sentimen positif, negatif, dan netral untuk Xiaomi 12 Pro berturut-turut adalah 49.3%, 17.6%, dan 33%.. Dari hasil proses pengujian menggunakan algoritma Naive Bayes dengan masing-masing data testing dan data training adalah 0.2 dan 0.8 sehingga menghasilkan nilai akurasi sebesar 66%, precision 66%, recall 66%, dan F1 Score 63% untuk keyword Samsung S22 Ultra. Sedangkan hasil pengujian algoritma Naive Bayes untuk keyword Xiaomi 12 Pro menghasilkan nilai akurasi sebesar 64%, precision 62%, recall 64%, dan F1 Score 61%. Selain itu, dilakukan kalkulasi Net Brand Reputation dengan hasil 41.9% untuk brand Samsung dan 47.3% untuk brand Xiaomi. Dari perhitungan dan temuan di atas dapat diambil kesimpulan bahwa masyarakat lebih tertarik dengan brand Xiaomi dibanding Samsung dalam hal smartphone flagship, karena brand Xiaomi memiliki nilai reputasi yang lebih baik dibandingkan dengan brand Samsung.

5. Daftar Rujukan

- [1] Nasukawa, Tetsuya, and Jeonghee Yi. "Sentiment analysis: Capturing favorability using natural language processing." *Proceedings of the 2nd international conference on Knowledge capture*. 2003.
- [2] Sari, Bety Wulan, and Fadholi Fat Haranto. "Implementasi Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Terhadap Pelayanan Telkom Dan Biznet." *Jurnal Pilar Nusa Mandiri* 15.2 (2019): 171-176.
- [3] Pudjajana, Andre Maureen, and Danny Manongga. "Sentimen Analisis Tweet Pornografi Kaum Homoseksual Indonesia Di Twitter Dengan Naive Bayes." *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer* 9.1 (2018): 313-318.
- [4] Nirawana, I. Wayan Sugianta, and I. Gede Indrawan. "Analisis Sentimen Pada Review Film Dengan Menggunakan Algoritma Klasifikasi Naive Bayes Berdasarkan Term Objects Keywords." *Semin. Nas. Pendidik. Tek. Inform.(SENAPATI 2016)*, no. Senapati (2016): 172-174.
- [5] Septian, J. A., T. M. Fahrudin, and A. Nugroho. "Journal of Intelligent Systems and Computation 43." (2019): 43-49.
- [6] Astari, Ni Made Ayu Juli, Dewa Gede Hendra Divayana, and Gede Indrawan. "Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier." *Jurnal Sistem Dan Informatika (JSI)* 15.1 (2020): 27-29.

- [7] Mas'udah, Erica, Eka Dyar Wahyuni, and Amalia Anjani. "Analisis sentimen: Pemindahan ibu kota Indonesia pada twitter." *Jurnal Informatika dan Sistem Informasi (JIFoSI)* 1.2 (2020): 397-401.
- [8] Rosdiana, Rosdiana, et al. "Analisis Sentimen pada Twitter terhadap Pelayanan Pemerintah Kota Makassar." *Seminar Nasional Teknik Elektro dan Informatika. Jurusan Teknik Elektro Politeknik Negeri Ujung Pandang*, 2019.
- [9] Husada, Ivan Nathaniel, et al. "Ekstraksi dan Analisis Produk di Marketplace Secara Otomatis dengan Memanfaatkan Teknologi Web Crawling." *Jurnal Teknik Informatika dan Sistem Informasi* 5.3 (2019).
- [10] Anjasmosros, Mochamad Tri, Istiadi Istiadi, and Fitri Marisa. "ANALISIS SENTIMEN APLIKASI GO-JEK MENGGUNAKAN METODE SVM DAN NBC (STUDI KASUS: KOMENTAR PADA PLAY STORE)." *Conference on Innovation and Application of Science and Technology (CIASTECH)*. Vol. 3. No. 1. 2020.