

Analisis Sentimen Pengguna Terhadap Smart Ring Berbasis Data Twitter Menggunakan Algoritma Naive Bayes

Faiq Fauzi¹, Nizar Ali Rifai², Anisa Az Zahra³, M Yoka Fathoni⁴

^{1,2,3,4}Sistem Informasi, Fakultas Rekayasa Industri, Telkom University; Jl. DI Panjaitan No.128, Karangreja, Purwokerto Kidul, Kec. Purwokerto Sel., Kabupaten Banyumas, Jawa Tengah 53147

*e-mail: 2311103011@ittelkom-pwt.sch.id

Abstrak

Smart ring merupakan perangkat wearable berbentuk cincin yang dilengkapi teknologi canggih untuk memantau aktivitas pengguna. Teknologi ini semakin populer namun masih menghadapi berbagai tantangan dalam penerimaan masyarakat. Penelitian ini bertujuan untuk menganalisis opini pengguna di media sosial, khususnya Twitter, serta bagaimana sentimen tersebut memengaruhi pandangan dan niat beli konsumen. Data dikumpulkan menggunakan teknik crawling dengan Google Colab dan alat tweet-harvest, kemudian dianalisis menggunakan algoritma Naive Bayes untuk mengklasifikasikan sentimen menjadi positif dan negatif. Hasil penelitian menunjukkan bahwa dari 580 tweet yang dikumpulkan, terdapat 147 sentimen positif dan 63 sentimen negatif setelah penghapusan duplikat, dengan tingkat akurasi sebesar 76,19%. Penelitian ini memberikan kontribusi penting dalam memahami sentimen pengguna terhadap teknologi smart ring serta mengusulkan strategi berbasis data yang dapat membantu produsen dalam mengembangkan produk yang lebih sesuai dengan kebutuhan pengguna dan meningkatkan strategi pemasaran agar lebih efektif dalam meningkatkan penerimaan masyarakat.

Kata kunci—Analisis Sentimen, Crawling Data, Media Sosial, Naive Bayes, Smart Ring

Abstract

A smart ring is a ring-shaped wearable device equipped with advanced technology to monitor user activity. This technology is growing in popularity but still faces challenges in public acceptance. This study aims to analyse user opinions on social media, particularly Twitter, and how these sentiments influence consumer views and purchase intentions. Data was collected using a crawling technique with Google Colab and a tweet-harvest tool, then analysed using the Naive Bayes algorithm to classify sentiment into positive and negative. The results showed that of the 580 tweets collected, there were 147 positive sentiments and 63 negative sentiments after duplicate removal, with an accuracy rate of 76.19%. This research makes an important contribution in understanding user sentiment towards smart ring technology and proposes data-driven strategies that can assist manufacturers in developing products that better suit user needs and improve marketing strategies to be more effective in increasing public acceptance.

Keywords—Sentiment Analysis, Data Crawling, Social Media, Naive Bayes, Smart Ring

1. PENDAHULUAN

Teknologi Informasi merupakan sebuah teknologi yang digunakan dalam pengolahan data, yakni memproses, mendapatkan, menyusun, menyimpan, dan memanipulasi data dengan menggunakan berbagai macam cara guna menghasilkan informasi yang relevan, akurat, dan berkualitas sehingga nantinya bisa dimanfaatkan dalam kebutuhan individu, bisnis, ataupun pemerintahan. Informasi ini juga memiliki nilai strategis yang penting dalam pengambilan keputusan yang bijak. Teknologi ini memanfaatkan perangkat komputer untuk mengolah data, kemudian ada juga sistem jaringan yang digunakan untuk menghubungkan komputer satu dengan komputer yang lainnya, serta teknologi telekomunikasi guna memastikan data dapat disebarluaskan dan di akses secara global [1].

Kemajuan Teknologi Informasi mendorong hal baru dalam kehidupan kita. Teknologi Informasi memacu suatu cara baru dalam kehidupan. Saat ini dunia teknologi sedang di suguhkan dengan inovasi hangat yaitu hadirnya perangkat *wearable*, perangkat ini memungkinkan pengguna untuk mengakses informasi, memantau kesehatan, dan juga berinteraksi dengan dunia digital hanya melalui perangkat yang di pasang pada tubuh. Secara keseluruhan, perangkat *wearable* ini menggunakan teknologi yang memiliki biosensor canggih dan komunikasi data nirkabel yang memungkinkan pemakainya dapat mengakses dan juga mengirimkan informasi di berbagai bidang kehidupan manusia. Dengan dibekali biosensor miniature yang mampu melakukan komunikasi nirkabel, perangkat ini telah dikembangkan menjadi teknologi pemantauan yang inovatif dan non-invasif [2].

Dari berbagai macam jenis *wearable* yang bisa dikenakan, *smartwatch* merupakan salah satu *wearable* yang paling banyak digunakan di seluruh dunia. Namun, saat ini, *smartwatch* memiliki pesaing baru yaitu *smart ring*, *smart ring* hadir dengan dibekali fitur yang serupa dan memiliki bentuk dan ukuran yang lebih kompak dan nyaman. Cara penggunaan *smart ring* ini juga cukup mudah, cukup dikenakan di jari, *smart ring* akan melacak berbagai macam metrik kesehatan, seperti saturasi oksigen, detak jantung, tingkat aktifitas fisik, stress, dan kualitas tidur [3]. Dengan menghubungkan *smart ring* ini ke *smartphone* atau perangkat lain, pengguna sudah bisa melihat dan juga memantau metrik kesehatan mereka secara *real time* [4]. Menurut data yang dilansir dari web Grand View Research pada tahun 2024, pasar perangkat *wearable* global diperkirakan mencapai USD 348,56 juta pada tahun 2024 dengan tingkat pertumbuhan tahunan sebesar 21,1% dari tahun 2025 hingga 2030. Tren ini menunjukkan adanya potensi besar dalam adopsi *smart ring* di berbagai segmen masyarakat.

Meskipun inovasi *smart ring* ini terus berkembang, ada banyak faktor juga yang mempengaruhi penerimaan dan penggunaan *smart ring* ini di kalangan masyarakat. Sosial media merupakan platform yang mendominasi para pengguna untuk berbagi pengalaman, pendapat, dan juga opini mereka terhadap penggunaan berbagai jenis teknologi, salah satunya adalah *smart ring* ini. Maka dari itu, kami akan menganalisis sentimen yang terkait dengan *smart ring* dengan tujuan untuk memberikan wawasan yang lebih baik kepada produsen dan juga konsumen.

Berbagai penelitian sebelumnya telah dilakukan terkait analisis sentimen pada perangkat *wearable technology*. Namun, penelitian terkait *smart ring* masih tergolong terbatas, khususnya dalam konteks pasar Indonesia. Kami akan melakukan penelitian ini dengan menggunakan data yang dikumpulkan dari Twitter atau X melalui metode *crawling*. Kemudian algoritma yang digunakan adalah *Naïve Bayes*, algoritma ini akan digunakan untuk mengklasifikasikan sentimen Twitter atau X menjadi positif, atau negatif. Penelitian ini berusaha untuk mengisi kesenjangan (*research gap*) tersebut dengan menganalisis opini pengguna Twitter di Indonesia terhadap *smart ring*. Dengan adanya penelitian ini, diharapkan dapat memperoleh wawasan mengenai prospektif masyarakat terhadap adanya *smart ring* ini, sehingga para produsen bisa menciptakan sebuah produk yang sesuai dengan kebutuhan pengguna.

Dalam beberapa tahun terakhir, penelitian terkait teknik *crawling* untuk data yang terstruktur mengalami perkembangan yang pesat, terutama dengan meningkatnya kebutuhan dalam

mengelola data dalam jumlah besar untuk mendukung pengambilan keputusan strategis [5]. Penelitian yang dilakukan oleh Kontogiannis pada tahun 2021 mengungkapkan bahwa dalam konteks crawling terfokus yang memanfaatkan teknik pembelajaran mesin. Salah satu pendekatan yang menonjol adalah penggunaan Reinforcement Learning (RL) untuk meningkatkan efisiensi dalam mengidentifikasi dan mengakses halaman web yang relevan [6]. Penelitian oleh Liu pada tahun 2023 mengintegrasikan skor TF-IDF dengan kesamaan semantik untuk membangun vektor topik dan teks, yang memungkinkan pemilihan URL yang lebih tepat berdasarkan relevansi konten. Pendekatan ini menunjukkan bahwa pemahaman mendalam tentang konteks kata kunci dapat meningkatkan akurasi dalam penentuan halaman yang akan di-crawl [7].

Terdapat beberapa tantangan dalam *crawling*, salah satunya adalah keseimbangan antara eksploitasi dan eksplorasi. Pada tahun 2024, Singh merepresentasikan masalah ini menjadi persoalan bandit, di mana halaman-halaman baru yang ditemukan dikelompokkan berdasarkan host-nya. Dari pendekatan ini, pemilihan halaman untuk di jelajahi dilakukan berdasarkan nilai potensial yang di harapkan. Hasilnya menunjukkan bahwa dengan adanya kombinasi pembelajaran online dan strategi pemilihan berbasis bandit, *crawler* mampu memperoleh lebih dari 10% halaman relevan tambahan dibanding dengan metode classifier statis, tanpa melebihi batasan anggaran yang ada [8].

Murugudu juga melakukan penelitian di tahun 2023, mereka menemukan bahwa seiring dengan meningkatnya layanan web dalam deep web, berbagai teknik efisien untuk mengeksplorasi lokasi antarmuka deep web juga telah dikembangkan, yang pada gilirannya meningkatkan pengalaman pengguna dalam eksplorasi informasi terkait dengan kueri dan klik pengguna pada data pencarian web, penggunaan algoritma crawling yang lebih cerdas juga dapat meningkatkan relevansi hasil pencarian dengan mengorganisir informasi pencarian pengguna berdasarkan preferensi mereka, sehingga membentuk kueri dinamis yang lebih efektif [9].

Penelitian menunjukkan bahwa teknologi *wearable*, seperti *smart ring*, menunjukkan bahwa teknologi teknologi ini semakin populer, khususnya dalam konteks kesehatan dan juga kebugaran. Jan Boge pada tahun 2021 menyoroti bahwa faktor psikologis, seperti sikap dan norma subjektif, dapat mempengaruhi keinginan konsumen untuk membeli perangkat ini. Hal ini sesuai dengan model *Theory of Planned Behavior* (TPB), yang menjelaskan bahwa niat seseorang dipengaruhi oleh sikap, norma sosial, dan juga persepsi kontrol atas perilakunya. Selain itu, menurut *Technology Acceptance Model* (TAM), menunjukkan bahwa kemudahan penggunaan dan manfaat yang dirasakan menjadi faktor utama yang mendorong adopsi teknologi baru. Pada kasus *smart rings*, konsumen lebih tertarik pada perangkat yang mudah digunakan dan memberikan manfaat kesehatan yang nyata, seperti contoh memantau tidur atau aktifitas fisik lainnya [10].

Seiring dengan perkembangan teknologi, analisis sentiment juga salah satu yang terpenting juga, karena meningkatnya penggunaan media sosial dan platform digital sebagai sumber informasi. Pendekatan seperti *machine learning*, analisis leksikom dan juga hibrida telah digunakan untuk meningkatkan akurasi dan efektivitas analisis sentimen [11]. Seperti penelitian yang dilakukan oleh Mashuri pada tahun 2022 menunjukkan bahwa pendekatan sistematis, seperti wawancara kinerja yang terstruktur, dapat membantu memahami opini pengguna secara lebih efektif [12].

Penelitian oleh Nguyen pada tahun 2022 menunjukkan bahwa metode seperti *Support Vector Machine* (SVM) dan *Artificial Neural Network* (ANN) dapat digunakan secara efektif untuk menilai kerentanan pada jaringan pipa gas perkotaan. Hal ini mencerminkan bagaimana metode tersebut dapat diterapkan dalam konteks yang lebih luas [13]. Penemuan yang dilakukan oleh Li pada tahun 2020 mengungkapkan bahwa manipulasi ulasan online dapat mempengaruhi cara konsumen melihat suatu produk, sehingga diperlukan pengembangan model yang dapat mendeteksi dan mengatasi bias dalam data [14]. Penelitian lainnya yang dilakukan oleh He pada tahun 2022 menekankan bahwa analisis sentimen tidak hanya berfokus pada pengolahan teks, tetapi juga harus paham akan konteks dan emosi yang ada pada ulasan tersebut. Hal ini menunjukkan pentingnya pendekatan yang lebih menyeluruh dalam analisis sentiment, yang memperhitungkan faktor-faktor luar yang dapat mempengaruhi pandangan pengguna [15].

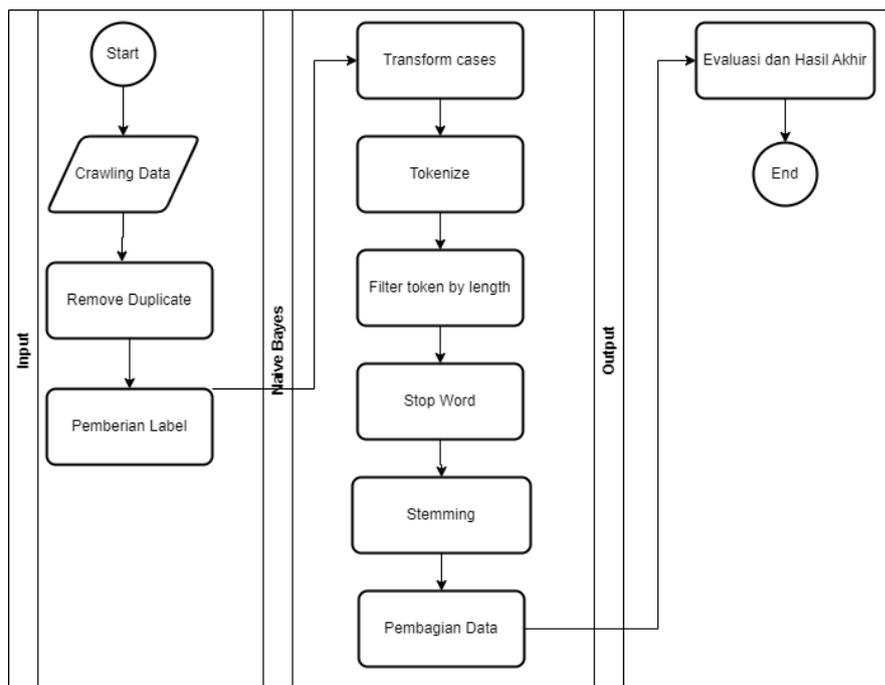
Algoritma seperti Naive Bayes dapat menjadi solusi yang efektif untuk masalah serupa. Walaupun Naive Bayes menggunakan asumsi bahwa atributnya independen, yaitu atribut yang tidak selalu benar, algoritma ini masih sering digunakan karena kemampuan mereka untuk memberikan hasil klasifikasi yang baik dengan perhitungan yang efisien. Dari keunggulan inilah naive bayes menjadi pilihan yang baik untuk berbagai aplikasi, seperti analisis sentimen. Namun, dalam konteks yang lebih kompleks, seperti yang dijelaskan oleh penelitian sebelumnya, penggunaan metode yang lebih canggih seperti SVM dan ANN mungkin lebih diperlukan untuk mendapatkan hasil yang lebih akurat dan menyeluruh [16].

2. METODE PENELITIAN

Penelitian ini merupakan penelitian kuantitatif, yang berfokus pada analisis sentimen terhadap data ulasan pengguna terkait teknologi smart ring. Tujuannya adalah merancang dan mengimplementasikan sebuah sistem klasifikasi sentimen menggunakan algoritma Naive Bayes. Sistem ini dirancang untuk mengelompokkan sentimen ke dalam kategori positif atau negative berdasarkan data yang telah diproses sebelumnya.

2.1 Alur Penelitian

Dalam penelitian ini menggunakan algoritma Naive Bayes, dimana dilakukan beberapa langkah seperti yang dijelaskan pada gambar di bawah ini:



Gambar 1. Alur Penelitian

Penelitian ini dimulai dengan pengumpulan dataset, tahap selanjutnya adalah algoritma naive bayes berupa Transform cases, Tokenize, Filter token by length, Stopword dan Stemming. Setelah algoritma naive bayes, tahap terakhir adalah evaluasi dan hasil akhir.

2.1.1 Crawling Data

Proses pengumpulan data dilakukan melalui proses crawling menggunakan platform Twitter sebagai sumber data utama dengan alat tweet-harvest di Google Colab. Kata kunci pencarian "smart ring" digunakan dengan filter bahasa Indonesia untuk mendapatkan data yang sesuai.

2.1.2 *Remove Duplicate*

Pada tahap *remove duplicate*, proses dilakukan dengan menggunakan aplikasi Microsoft Excel untuk memastikan bahwa data yang digunakan tidak memiliki duplikasi yang dapat memengaruhi hasil analisis. Data yang telah dikumpulkan dimasukkan ke dalam lembar kerja Excel, kemudian fitur bawaan "Remove Duplicates" digunakan untuk mengidentifikasi dan menghapus entri yang sama berdasarkan kolom tertentu, seperti teks atau kriteria lainnya.

2.1.3 Pemberian Label

Selanjutnya, dilakukan tahap pemberian label secara manual untuk menentukan sentimen dari setiap tweet. Pada tahap ini, setiap tweet dianalisis satu per satu dan diberikan label sesuai dengan konteksnya, yaitu sentimen positif atau negatif. Penentuan sentimen dilakukan berdasarkan interpretasi terhadap isi tweet, seperti penggunaan kata-kata yang menunjukkan kepuasan atau keluhan, sehingga data siap digunakan untuk proses analisis lebih lanjut.

2.1.4 *Transform Cases*

Tahap *transform cases* dilakukan untuk menyamakan format teks pada data tweet, yaitu dengan mengubah seluruh huruf menjadi huruf kecil. Proses ini bertujuan untuk menghindari perbedaan interpretasi akibat perbedaan huruf besar dan kecil, misalnya antara kata "Smart" dan "smart," yang seharusnya dianggap sama. Transformasi ini dilakukan menggunakan perangkat lunak Visual Studio Code, di mana setiap teks dalam data tweet diproses secara otomatis untuk memastikan konsistensi format.

2.1.5 *Tokenize*

Tahap *tokenize* merupakan proses memecah setiap teks tweet menjadi unit-unit kata yang lebih kecil, yang disebut *token*. Proses ini dilakukan untuk mempermudah analisis pada tahap selanjutnya, seperti penghilangan kata yang tidak relevan atau penghitungan frekuensi kata.

2.1.6 *Filter Token by Length*

Tahap *filter token by length* adalah proses menyaring token berdasarkan panjang karakter. Pada tahap ini, token yang memiliki panjang kurang dari atau sama dengan dua karakter dihapus karena dianggap tidak memberikan makna signifikan dalam analisis. Hal ini bertujuan untuk mengurangi noise pada data dan meningkatkan efisiensi pemrosesan.

2.1.7 *Stopword*

Tahap *stopword removal* adalah proses menghapus kata-kata umum yang sering muncul tetapi tidak memberikan kontribusi signifikan terhadap analisis sentimen. Kata-kata ini disebut *stopwords*, seperti "dan," "di," "yang," "dari," dan lainnya dalam bahasa Indonesia. Dalam penelitian ini, *stopword removal* dilakukan dengan menggunakan pustaka NLTK di Visual Studio Code. Pustaka ini menyediakan daftar kata-kata umum (*stopwords*) yang telah didefinisikan sebelumnya untuk berbagai bahasa, termasuk bahasa Indonesia. Proses ini membantu mengurangi dimensi data dan memastikan bahwa analisis hanya berfokus pada kata-kata yang lebih bermakna.

2.1.8 *Stemming*

Tahap *stemming* merupakan proses mengubah kata-kata dalam data menjadi bentuk dasar atau akar katanya. Tujuan dari *stemming* adalah untuk menyederhanakan variasi kata yang memiliki makna serupa sehingga model dapat

lebih mudah mengenali pola. Pada penelitian ini, proses *stemming* dilakukan menggunakan pustaka NLTK di Visual Studio Code. Pustaka ini dilengkapi dengan *stemmer* yang mendukung bahasa Indonesia, yaitu *Sastrawi Stemmer*. Melalui pustaka ini, setiap token yang telah melalui tahap *stopword removal* diproses untuk diubah menjadi bentuk akar katanya.

2.1.9 Pembagian Data

Setelah melalui tahap *stemming*, data dibagi menjadi dua bagian, yaitu data training dan data testing, yang digunakan untuk melatih dan menguji model klasifikasi sentimen. Pembagian ini bertujuan untuk mengevaluasi performa model dalam memprediksi sentimen berdasarkan data yang telah diproses sebelumnya. Data dibagi menggunakan perangkat lunak Microsoft Excel dengan memanfaatkan rumus RAND() untuk menghasilkan angka acak pada setiap baris data. Selanjutnya, data diurutkan berdasarkan nilai acak yang dihasilkan, dan pembagian dilakukan dengan perbandingan 80:20, di mana 80% dari total data digunakan sebagai data training dan 20% sebagai data testing.

2.1.10 Evaluasi dan Hasil Akhir

Pada tahap ini, model yang telah dibuat dievaluasi untuk menentukan tingkat akurasi dan performanya dalam mengklasifikasikan data sentimen. Proses evaluasi dilakukan menggunakan metode *Performance Classification* yang disediakan oleh RapidMiner. Data testing digunakan untuk mengukur performa model, sedangkan metrik evaluasi yang digunakan adalah akurasi, presisi, recall, dan F1-score. Selain itu, hasil akhir dari proses evaluasi disajikan dalam bentuk matriks konfusi yang menunjukkan jumlah prediksi benar dan salah pada masing-masing kategori sentimen.

2.2 Algoritma Naïve Bayes

Tahap ini merupakan inti dari penelitian, yaitu penerapan algoritma *Naive Bayes* untuk mengklasifikasikan data sentimen. Algoritma *Naive Bayes* digunakan karena memiliki keunggulan dalam menangani data berbasis teks, khususnya dalam analisis sentimen. Naïve Bayes adalah salah satu bentuk Bayesian Network Classifier yang didasarkan pada aturan Bayes

$$P(y | x) = P(y)P(x | y)/P(x) \quad (1)$$

Keterangan:

$P(y | x)$ = Probabilitas bersyarat y yang diberikan oleh x

$P(y)$ = Probabilitas kejadian y

$P(x | y)$ = Probabilitas bersyarat x yang diberikan oleh y

$P(x)$ = Probabilitas kejadian x

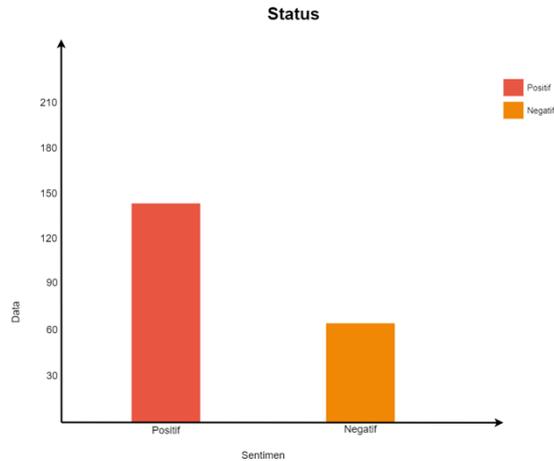
3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian yang diperoleh dari setiap tahap dalam proses pengolahan data dan penerapan algoritma *Naive Bayes*. Bab ini juga membahas performa model yang diukur berdasarkan akurasi dan evaluasi hasil klasifikasi.

3.1 Pengolahan Data

Sebanyak 580 tweet berhasil dikumpulkan dalam format CSV, yang mencakup informasi seperti isi tweet, nama pengguna, dan waktu pengunggahan. Namun setelah dilakukan

penghapusan duplikat, data yang didapatkan menjadi 210 data tweet. Lalu pelabelan data dilakukan secara manual untuk memastikan ketepatan sentimen pada data yang ada dan data sentimen dikelompokkan menjadi 2 bagian yaitu data positif dan negatif. Pengelompokkan data ini menghasilkan 147 data sentimen positif dan 63 data sentimen negatif.



Gambar 2. Statistik Hasil Pengumpulan Data

Setelah data dikelompokkan, tahap berikutnya adalah pengolahan data yang melibatkan serangkaian langkah preprocessing untuk mempersiapkan data agar siap digunakan dalam analisis sentimen. Berikut adalah tahapan preprocessing yang dilakukan:

3.1.1 Transform Case

Semua teks dalam data diubah menjadi huruf kecil untuk memastikan konsistensi dan mengurangi duplikasi akibat perbedaan huruf kapital.

Tabel 1. Hasil Transform Case

Sebelum	Sesudah
Ternyata smart ring ini cocok buat monitoring aktivitas sehari-hari walau baterainya cepat habis	ternyata smart ring ini cocok buat monitoring aktivitas sehari-hari walau baterainya cepat habis

3.1.2 Tokenize

Setiap teks dipecah menjadi unit-unit kata atau token. Proses ini memungkinkan analisis kata-kata individual dalam sebuah kalimat atau paragraf.

Tabel 2. Hasil Tokenize

Sebelum	Sesudah
smart ring sangat berguna	["smart", "ring", "sangat", "berguna"]

3.1.3 Filter Token by Length

Token yang terlalu pendek atau kurang dari 2 karakter akan dihapus untuk menghindari ketidaksesuaian data. Hal ini dilakukan agar token yang diproses lebih relevan dan bermakna.

3.1.4 Stopword Removal

Token-token yang termasuk dalam daftar stopwords bahasa Indonesia, seperti "dan", "yang", dan "di", dihapus karena tidak memiliki makna signifikan dalam analisis sentimen.

Tabel 3. Hasil Stopword Removal

Sebelum	Sesudah
Smart Ring ini kok sepertinya menarik ya	Smart Ring menarik

3.1.5 *Stemming*

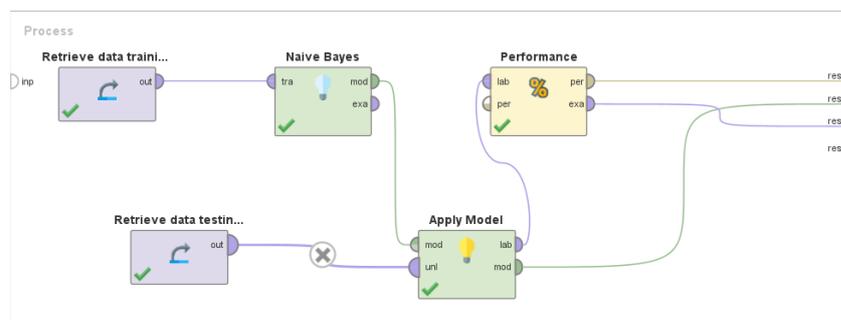
Setiap token dikembalikan ke bentuk dasarnya menggunakan proses stemming. Misalnya, kata "membantu", "dibantu", dan "bantuan" akan diubah menjadi "bantu" untuk mengurangi redundansi data.

Tabel 4. Hasil Stemming

Sebelum	Sesudah
Beliin smart ring dong	Beliin smart ring dong

3.2 *Proses Validasi Algoritma*

Algoritma Naive Bayes diterapkan dengan menggunakan perangkat lunak RapidMiner. Langkah pertama adalah menghubungkan data training yang telah melalui tahap preprocessing ke operator Naive Bayes. Pada tahap ini, model Naive Bayes dilatih untuk mengenali pola dan hubungan antara data input dengan label klasifikasinya. Setelah model terbentuk, output dari operator Naive Bayes berupa model klasifikasi dihubungkan ke operator Apply Model. Data testing yang telah disiapkan juga dihubungkan ke operator Apply Model untuk menghasilkan prediksi berdasarkan model yang telah dilatih. Selanjutnya, hasil prediksi dari operator Apply Model dihubungkan ke operator Performance Classification untuk mengevaluasi performa model. Evaluasi ini mencakup penghitungan metrik seperti akurasi, presisi, dan recall dengan membandingkan hasil prediksi dengan data aktual dari data testing. Output dari operator Performance Classification memiliki dua bagian utama, yaitu Per yang berisi metrik evaluasi, dan Exa yang berisi contoh data hasil evaluasi. Kedua output ini, bersama dengan output prediksi (mod) dari operator Apply Model, dihubungkan ke result viewer untuk ditampilkan dan dianalisis lebih lanjut.



Gambar 3. Model Pengujian Naive Bayes dengan RapidMiner

3.3 *Evaluasi dan hasil akhir*

Dari tahapan - tahapan pengujian yang sudah dilakukan dengan menggunakan dataset sebanyak 210 respon tweet pengguna twitter mengenai smart ring, dataset dibagi menjadi 2 yaitu data training berisi 168 data tweet dan data test yang berisi 42 data tweet. Maka hasil Akurasi Algoritma Naive Bayes yaitu sebesar 76,19 %. Berikut ini adalah Tabel hasil Algoritma Naive Bayes.

Tabel 5. Hasil Algoritma Naive Bayes

	True Positive	True Negative	
Pred Positive	32	10	76,19 %
Pred Negative	0	0	0 %
	100 %	0%	

Berdasarkan hasil dan pembahasan, sentimen negatif terhadap smart ring yang muncul dari analisis data tweet mencakup beberapa faktor utama. Pertama, banyak pengguna mengeluhkan harga yang masih dianggap mahal, sehingga mereka ragu untuk membeli. Kedua, keterbatasan kompatibilitas smart ring dengan perangkat lain menjadi sorotan, karena pengguna menginginkan integrasi yang lebih baik dengan aplikasi yang mereka gunakan. Ketiga, daya tahan baterai yang dianggap tidak memadai juga menjadi keluhan, mengganggu pengalaman penggunaan.

Untuk mengatasi sentimen negatif ini, produsen perlu meninjau harga agar lebih kompetitif, meningkatkan kompatibilitas dengan berbagai perangkat, dan memperbaiki daya tahan baterai. Dengan langkah-langkah ini, produsen dapat meningkatkan kepuasan pengguna dan memperbaiki citra produk di pasar, sehingga meningkatkan penerimaan masyarakat terhadap smart ring.

4. KESIMPULAN

Penelitian ini berhasil mengungkap persepsi masyarakat terhadap smart ring melalui analisis sentimen yang dilakukan terhadap data tweet yang dikumpulkan dari Twitter. Dengan menggunakan metode crawling dan algoritma Naive Bayes, penelitian ini menemukan bahwa dari total 210 tweet yang dikumpulkan, terdapat 147 sentimen positif dan 63 sentimen negatif. Hasil ini menunjukkan bahwa mayoritas pengguna memiliki pandangan positif terhadap smart ring, meskipun masih ada sejumlah sentimen negatif yang perlu diperhatikan oleh produsen. Tujuan penelitian untuk memahami opini masyarakat dan faktor-faktor yang mempengaruhi penerimaan smart ring telah tercapai, memberikan wawasan berharga bagi pengembang produk dalam merespons kebutuhan dan harapan konsumen.

Selain itu, penelitian ini juga membuka potensi pengembangan metodologi analisis sentimen di masa depan. Dengan kemajuan teknologi dan algoritma yang terus berkembang, pendekatan yang lebih canggih seperti penggunaan deep learning atau analisis konteks yang lebih mendalam dapat diterapkan untuk meningkatkan akurasi dan efektivitas analisis sentimen. Penelitian lebih lanjut dapat mengeksplorasi integrasi berbagai sumber data, termasuk ulasan produk dan interaksi pengguna di platform lain, untuk memberikan gambaran yang lebih komprehensif tentang sentimen pengguna. Dengan demikian, hasil penelitian ini tidak hanya memberikan gambaran tentang sentimen pengguna, tetapi juga berkontribusi pada pengembangan metodologi analisis sentimen yang lebih efektif dalam era digital.

5. SARAN

Untuk penelitian lebih lanjut, kami menyarankan untuk bisa memperluas jumlah data yang dikumpulkan, baik dari tweet maupun platform sosial media lainnya, guna untuk meningkatkan representativitas hasil yang lebih komprehensif. Disisi lain, walaupun algoritma Naive Bayes telah digunakan, penelitian lebih lanjut bisa juga menggunakan algoritma lain seperti SVM atau Random Forest guna memperoleh hasil yang lebih akurat dalam klasifikasi sentimen. Untuk penelitian selanjutnya juga bisa lebih fokus menganalisis sentimen negatif, menggali lebih dalam faktor-faktor yang menyebabkan ketidakpuasan pengguna sehingga para produsen bisa membenahi hal tersebut. Terakhir, penting untuk meneliti lebih dalam mengenai hubungan antara

sentimen yang diungkapkan di media sosial dan niat beli konsumen. Untuk memahami seberapa besar pengaruh sentimen terhadap keputusan pembelian.

UCAPAN TERIMA KASIH

Penelitian ini tidak akan bisa terwujud jika tanpa adanya dukungan dan bantuan dari berbagai pihak, terutama para dosen kami. Kami mengucapkan terimakasih banyak kepada dosen pembimbing yang telah meluangkan waktu, memberikan arahan, dan saran yang sangat berharga sepanjang proses penelitian. Terimakasih juga kami sampaikan kepada pengguna Twitter, yang mana tweet-tweetnya adalah bagian penting dalam pengumpulan data penelitian ini. Yang terakhir, kami juga ingin mengapresiasi pihak-pihak yang telah memberikan dukungan teknis, khususnya dalam penggunaan alat dan platform untuk crawling data. Kami harap, hasil penelitian ini bisa memberikan kontribusi yang bermanfaat dalam pengembangan produk dan strategi pemasaran tentang smart ring ke depannya.

DAFTAR PUSTAKA

- [1] C. A. Cholik, "PERKEMBANGAN TEKNOLOGI INFORMASI KOMUNIKASI / ICT DALAM BERBAGAI BIDANG," *Jurnal Fakultas Teknik*, vol. 2, no. 2, 2021.
- [2] A. Sharma, M. Badea, S. Tiwari, and J. L. Marty, "Wearable biosensors: An alternative and practical approach in healthcare and disease monitoring," *Molecules*, vol. 26, no. 3, Feb. 2021, doi: 10.3390/molecules26030748.
- [3] S. M. Reitz, M. E. Scaffa, and J. Dorsey, "Occupational Therapy in the Promotion of Health and Well-Being.," *American Journal of Occupational Therapy*, vol. 74, no. 3, pp. 1–14, May 2020, doi: 10.5014/ajot.2020.743003.
- [4] M. Fiore *et al.*, "The Use of Smart Rings in Health Monitoring—A Meta-Analysis," *Applied Sciences (Switzerland)*, vol. 14, no. 23, Dec. 2024, doi: 10.3390/app142310778.
- [5] I. Bifulco, S. Cirillo, C. Esposito, R. Guadagni, and G. Polese, "An intelligent system for focused crawling from Big Data sources," *Expert Syst Appl*, vol. 184, Dec. 2021, doi: 10.1016/j.eswa.2021.115560.
- [6] A. Kontogiannis, D. Kelesis, V. Pollatos, G. Paliouras, and G. Giannakopoulos, "Tree-based Focused Web Crawling with Reinforcement Learning," Dec. 2021, [Online]. Available: <http://arxiv.org/abs/2112.07620>
- [7] W. Liu *et al.*, "A semantic and intelligent focused crawler based on semantic vector space model and membrane computing optimization algorithm," *Applied Intelligence*, vol. 53, no. 7, pp. 7390–7407, 2023, doi: 10.1007/s10489-022-03180-5.
- [8] R. Singh, F. Liu, Y. Sun, and N. Shroff, "Multi-armed bandits with dependent arms," *Mach Learn*, vol. 113, no. 1, pp. 45–71, 2024, doi: 10.1007/s10994-023-06457-z.
- [9] M. R. Murugudu and L. S. S. Reddy, "Efficiently harvesting deep web interfaces based on adaptive learning using two-phase data crawler framework.," *Soft Computing-A Fusion of Foundations, Methodologies & Applications*, vol. 27, no. 1, 2023.
- [10] J. Boge, "Smart rings-German consumers' perception and factors influencing the purchase intention," pp. 8–9, 2021.
- [11] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.
- [12] S. Mashuri, M. Sarib, A. Rasak, and F. Alhabsyi, "Semi-structured Interview: A Methodological Reflection on the Development of a Qualitative Research Instrument in Educational Studies Ruslin," vol. 12, no. 1, pp. 22–29, 2022, doi: 10.9790/7388-1201052229.

- [13] L. Van Nguyen, D. T. Bui, and R. Seidu, "Comparison of Machine Learning Techniques for Condition Assessment of Sewer Network," *IEEE Access*, vol. 10, pp. 124238–124258, 2022, doi: 10.1109/ACCESS.2022.3222823.
- [14] L. Li, T. T. Goh, and D. Jin, "How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis," *Neural Comput Appl*, vol. 32, no. 9, pp. 4387–4415, May 2020, doi: 10.1007/s00521-018-3865-7.
- [15] S. He, B. Hollenbeck, G. Overgoor, D. Proserpio, and A. Tosyali, "Detecting fake-review buyers using network structure: Direct evidence from Amazon," *Proceedings of the National Academy of Sciences*, vol. 119, no. 47, p. e2211932119, 2022, doi: 10.1073/pnas.2211932119.
- [16] I. Wickramasinghe and H. Kalutarage, "Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation," *Soft comput*, vol. 25, no. 3, pp. 2277–2293, Feb. 2021, doi: 10.1007/s00500-020-05297-6.