



Klasifikasi Sentimen Terhadap Dinamika Metavers Menggunakan Algoritma Naive Bayes

Cahaya Novita Sari¹, Dea Nur Azizah², Dira Trisnadia³, Ilham Ma'ruf⁴, Muhammad
Ilham Herwandani⁵, Fandi Kurniawan⁶

¹ Sistem Teknologi Informasi, Universitas Muhammadiyah Kotabumi, kotabumi, Indonesia
Email: ¹cahya.2059201020, ²Deanu.2059201093@umko.ac.id,
³dira.2059201040@umko.ac.id, ⁴ilham.2059201003@umko.ac.id,
⁵ilham.2059201022@umko.ac.id ⁶fandi.kurniawan@umko.ac.id

Abstrak

Metaverse, sebuah inovasi teknologi yang menakjubkan dengan pengaruh global, merupakan kumpulan realitas digital yang menyatukan elemen media sosial, permainan daring, realitas tambahan (AR), realitas maya (VR), dan mata uang kripto. Pendekatan manual dalam menentukan sifat positif atau negatif suatu opini sering kali memakan waktu dan tenaga, terutama dengan melimpahnya jumlah opini publik. Untuk mengatasi hal ini, diusulkan penggunaan teknik pembelajaran mesin guna efisien mengklasifikasikan beragam sumber data opini. Studi ini mengedepankan penerapan Penambangan Teks sebagai metode untuk klasifikasi dokumen. Metode Klasifikasi Naive Bayes (NBC) dipilih oleh peneliti karena efektivitasnya dalam menggunakan perhitungan probabilitas. Hasil penelitian mengindikasikan bahwa sistem atau model yang dikembangkan berhasil mencapai presisi sempurna (100%) dalam mengidentifikasi kelas opini positif, artinya semua prediksi positif yang dihasilkan tepat sasaran. Namun, terdapat margin kecil dalam recall untuk kelas positif, yakni 93,37%, mengindikasikan adanya beberapa kasus positif yang luput dari deteksi. Sementara itu, presisi untuk kelas negatif berada pada angka 43,46%, yang menyiratkan kemungkinan adanya kesalahan dalam sebagian kecil prediksi negatif. Namun, recall untuk kelas negatif mencapai 100%, menandakan semua kasus negatif berhasil diidentifikasi dengan benar.

Kata kunci: Metaverse, Naive Bayes, Penambangan Teks

1. PENDAHULUAN

Perkembangan teknologi yang pesat telah memotivasi pengembang dan pengguna untuk mengeksplorasi cara-cara baru dalam berinteraksi dan bekerja. Metaverse, sebagai jawaban inovatif, menciptakan dunia virtual yang merangkul aspek-aspek media sosial, permainan online, augmented reality (AR), virtual reality (VR), dan mata uang kripto. Terinspirasi dari novel fiksi ilmiah "Snow Crash" oleh Neal Stephenson pada tahun 1992, metaverse menyajikan lingkungan 3D di mana manusia berinteraksi melalui avatar.



Potensi metaverse dalam mengubah cara kita bekerja, belajar, dan bermain sangat signifikan. Di metaverse, berbagai aktivitas seperti bertemu teman dan keluarga, menghadiri acara, hingga berbelanja, menjadi mungkin. Fenomena ini menawarkan sebuah pengalaman kreatif yang berbeda dari aplikasi komputer konvensional, memungkinkan kita untuk 'berteleportasi' ke dalam hologram, membuka dimensi baru dalam interaksi sosial [1].

Dengan dampak globalnya yang luas, metaverse belum sepenuhnya dieksplorasi dalam konteks teologi Kristen. Penelitian ini mengadopsi pendekatan teologi kontekstual yang diperkenalkan oleh Pierre Teilhard de Chardin untuk mengkaji fenomena metaverse, yang diartikan sebagai ruang transparan tiga dimensi yang meliputi seluruh dunia [2]. Dalam konteks ini, metaverse menggabungkan dunia virtual dan nyata, menciptakan lingkungan unik untuk berbagai aktivitas manusia.

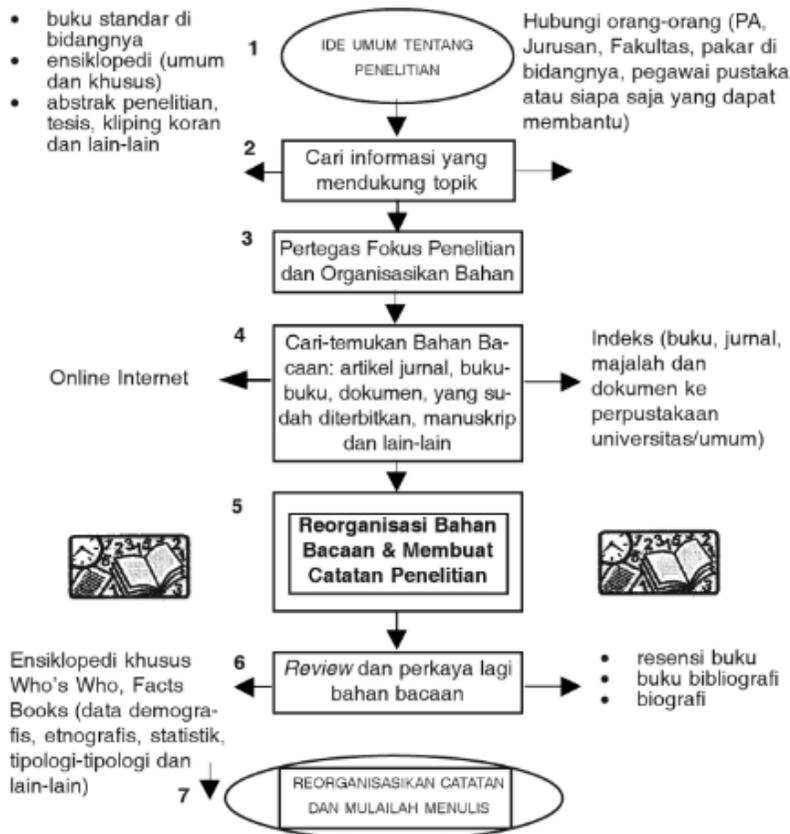
Media sosial, khususnya di Indonesia, telah menjadi alat komunikasi penting, termasuk platform seperti YouTube, yang memungkinkan pengguna untuk mengunggah, menonton, dan berbagi video, serta mengemukakan berbagai opini [3]. Komentar di YouTube memberikan wawasan tentang opini publik, yang sering kali mencakup pandangan positif dan negatif. Namun, menilai opini ini secara manual bukanlah pendekatan yang efisien, terutama mengingat volume besar data yang tersedia. Oleh karena itu, penelitian ini mengusulkan penggunaan teknik pembelajaran mesin seperti text mining untuk mengklasifikasikan opini publik tersebut.

Menggunakan metode klasifikasi Naive Bayes, yang efektif dalam menganalisis teks dan klasifikasi data berbasis probabilitas, penelitian ini bertujuan untuk mengkategorikan opini publik dengan lebih efisien dan akurat. Metode ini dianggap ideal karena kemudahan implementasinya dan kemampuannya untuk menghasilkan hasil yang memuaskan dengan data pelatihan yang relatif sedikit [4].

Penelitian ini bertujuan untuk mengaplikasikan teknik Naive Bayes dalam analisis opini publik di YouTube terkait metaverse. Fokus utama adalah untuk mengembangkan sebuah model yang dapat secara efektif mengklasifikasikan opini menjadi kategori positif atau negatif, dengan memanfaatkan data yang tersedia secara luas di platform ini. Dengan demikian, penelitian ini berupaya memberikan wawasan yang lebih mendalam tentang persepsi publik terhadap metaverse, sekaligus menunjukkan potensi pembelajaran mesin dalam memahami dan menginterpretasikan data opini besar secara otomatis.

2. METODE

Kajian penelitian klasifikasi sentimen terhadap dinamika metavers menggunakan algoritma Naive Bayes dilakukan dengan tahapan seperti Gambar 1.



Gambar 1. Pendalaman

Berdasarkan tahapan pendalaman literatur seperti pada Gambar 1 [5], berikut adalah penjelasan masing-masing langkah dalam melakukan penelitian kepustakaan:

1. Menentukan topik penelitian: Memilih topik yang jelas dan sempit adalah kunci dalam penelitian kepustakaan, Membantu menentukan ruang lingkup. Topik yang terlalu luas dapat mengakibatkan kurangnya fokus dan kedalaman.

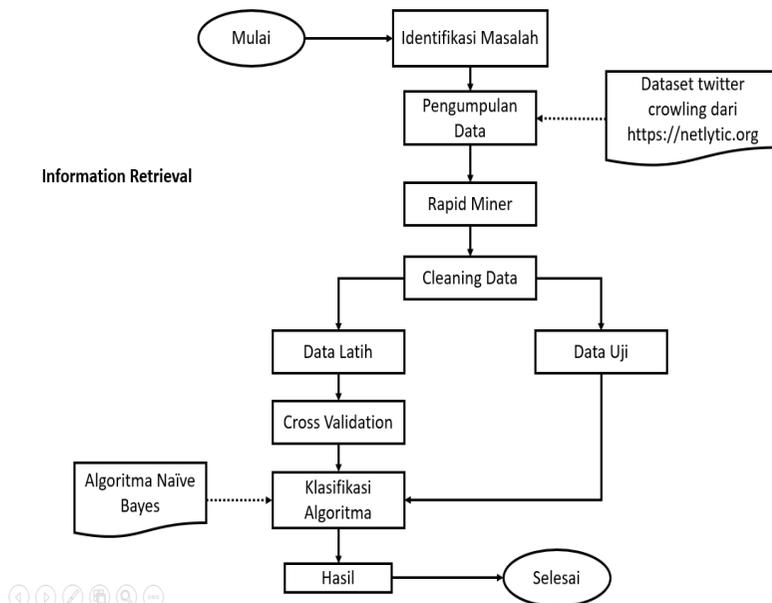
2. Identifikasi Kata Kunci: Kata kunci adalah istilah atau frasa penting yang berkaitan dengan topik penelitian Anda. Pengenalan kata kunci membantu Anda melakukan riset literatur secara efektif dan spesifik [6].
3. Pencarian Sumber Literatur: Pencarian literatur dapat dilakukan melalui berbagai sumber seperti perpustakaan, database online, dan repositori. kata kunci untuk menemukan buku, jurnal, artikel, dan dokumen lain yang berkaitan dengan topik penelitian [5].
4. Membaca dan memahami isi literatur: Setelah mengumpulkan literatur, membaca dan memahami isi setiap sumber. Perhatikan konsep, metode penelitian, temuan, dan argumen yang muncul dalam literatur.
5. Analisis dan Sintesis Data: Menganalisis informasi yang ditemukan dalam literatur. Identifikasi pola, persamaan, perbedaan, dan tren yang muncul dari berbagai sumber. Menggabungkan wawasan utama dan mensintesis data untuk menciptakan pemahaman komprehensif tentang topik penelitian [6].
6. Menulis laporan penelitian: Sajikan hasil penelitian kepustakaan Anda dalam bentuk laporan atau tinjauan pustaka. Struktur laporan dapat mencakup pendahuluan, kerangka teori, metodologi penelitian (jika ada), hasil, pembahasan, dan kesimpulan. Menulis dengan bahasa yang jelas, logis, dan tepat.

Tambahkan referensi dengan format yang mengikuti pedoman penulisan yang digunakan. Melalui langkah-langkah tersebut, penelitian kepustakaan dapat memberikan kontribusi yang berharga dalam memahami perkembangan pengetahuan yang ada, mengidentifikasi kesenjangan pengetahuan, dan merumuskan landasan teori untuk penelitian selanjutnya [6].

Selanjutnya pengumpulan data dilakukan berupa data opini dari YouTube dilakukan dengan bantuan website <https://netlytic.org/> yang menyediakan tools untuk melakukan proses crawling dataset dengan cepat. Dengan website netlytic Anda dapat mengambil data Twitter untuk jangka waktu tertentu, kata kunci tertentu, atau mengeksportnya ke format tertentu. Penelitian ini menggunakan 2500 opini yang bernilai positif dan negatif dan diberi label secara manual [7].

3. HASIL DAN PEMBAHASAN

Alur dalam membuat program analisis dilakukan dalam beberapa tahap seperti terlihat pada Gambar 2.



Gambar 2. Alur Project

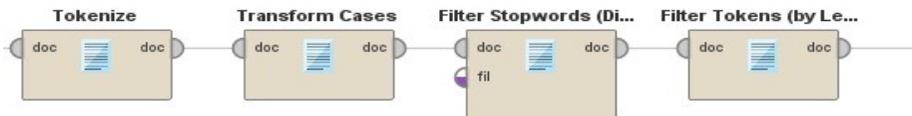
Sesuai Gambar 2 pengumpulan data opini dari YouTube dilakukan dengan bantuan website <https://netlytic.org/> yang menyediakan tools untuk melakukan proses crawling dataset dengan cepat. Dengan website netlytic Anda dapat mengambil data Twitter untuk jangka waktu tertentu, kata kunci tertentu, atau mengekspornya ke format tertentu. Penelitian ini menggunakan 2500 opini yang bernilai positif dan negatif dan diberi label secara manual[8]. Untuk contoh twit yang telah dilabeli dapat dilihat pada tabel dibawah ini.

Tabel 2. Opini Youtube

Text	Sentimen
jawaban om ded setuju bgt soal gejed aja udah bikin orang malas palgi kalo d metabvres	Negatif
buset kesukaan aku tuh roblox wkwk	Positif

Label positif artinya yang berkomentar setuju atau mendukung selebriti yang dibicarakan, begitu pula sebaliknya, jika label negatif maka komunitas Saya tidak setuju atau mendukung orang tersebut dan saya tidak mendukungnya. Subproses penelitian ini selanjutnya adalah menghilangkan emoticon dan mengubahnya menjadi kata-kata yang muncul di setiap komentar guna memudahkan analisis terhadap opini yang diungkapkan masyarakat melalui

media sosial YouTube. Setelah subproses selesai, data yang telah disiapkan diproses terlebih dahulu pada tahap berikutnya, dimana opini yang ada menjalani pemrosesan tokenisasi, stopwords, dan rootword. Pada aplikasi Rapid Miner, process preprocessing dilakukan oleh operator pengolah dokumen[9]. Preprocessing data dilakukan seperti pada Gambar 3.



Gambar 3. Preprocessing data

Untuk memudahkan pemrosesan, operator Transform Case mengubah seluruh teks menjadi huruf kecil. Selanjutnya, operator Tokenize memecah setiap kalimat menjadi kata-kata. Operator Filter Stopwords kemudian menghilangkan kata-kata yang tidak relevan, seperti kata sambung, kata depan, dan kata keterangan. Daftar kata-kata yang tidak relevan tersebut dapat ditemukan dalam sebuah file yang harus dibuka menggunakan operator Open File Stopword. TF-IDF (Term Frequency-Inverse Document Frequency) adalah metode untuk mengukur tingkat kepentingan suatu kata dalam suatu dokumen. Metode ini mempertimbangkan konteks koleksi dokumen yang lebih besar. Teknik ini umumnya digunakan dalam pemrosesan bahasa alami, pengambilan informasi, dan penambangan teks. [10]. Gambar 4 dan 5 merupakan Hasil dari TF-IDF yang telah dilakukan.

Row No.	word	in documents	total
1	metaverse	601	766
2	dunia	397	595
3	orang	165	228
4	nyata	182	222
5	online	189	205
6	manusia	135	203
7	film	162	200
8	hidup	127	155
9	game	104	146
10	sword	142	146
11	teknologi	103	137
12	kalo	106	127
13	anime	97	120
14	player	103	105
15	udah	94	105

ExampleSet (20 examples, 0 special attributes, 3 regular attributes)

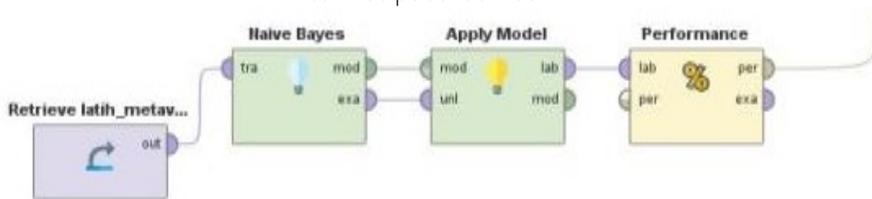
Gambar 5 Hasil Total TF-IDF

Word Cloud merupakan representasi visual dari kumpulan kata dalam teks yang ditampilkan dalam bentuk awan kata, di mana kata-kata yang lebih sering muncul dalam teks memiliki ukuran yang lebih besar dan lebih menonjol dalam tampilan visual, visualisasi data yang di tampilkan untuk mempermudah seseorang untuk melihat dan mengartikan maksud dan hasil yang didapatkan[11]. Gambar 6 merupakan lanjutan dari proses TF-IDF yang mana pada gambar besar yang memiliki tulisan PDIP adalah jumlah text yang sering muncul dibandingkan dengan text/ kata yang lain.



Gambar 6 Word Cloud Hasil Analisa

Penelitian ini bertujuan untuk mengukur seberapa akurat algoritma Naive Bayes Classifier dalam menentukan sentimen dari teks. [12]. Dengan bantuan aplikasi Rapidminer, operator Naive Bayes Classifier digunakan untuk percobaan seperti terlihat pada Gambar 7.



Gambar 7. Naive Bayes Classifier Operator

Pada algoritma Naive Bayes Classifier dataset yang digunakan berjumlah 588 total keseluruhan yang sudah dibersihkan atau difilter pada proses sebelumnya, selanjutnya data dibagi menjadi 70% data latihan dan 30% data uji dengan pembagian data latihan yang digunakan 410 yang sudah diberi sentimen dan data uji 178 yang tidak diberi sentiment. Hasil nilai akurasi berada pada angka 93.70% seperti pada Gambar 8.

accuracy: 93.70%

	true positif	true negatif	class precision
pred. positif	1522	0	100.00%
pred. negatif	108	83	43.46%
class recall	93.37%	100.00%	

Gambar 8. Akurasi Naive Bayes Classifier

4. KESIMPULAN

Hasil ini menunjukkan bahwa sistem atau model memiliki presisi yang sempurna (100%) dalam mengklasifikasikan kelas positif, yang berarti semua prediksi positif yang diberikan adalah benar. Namun, recall untuk kelas positif adalah 93.37%, yang berarti ada beberapa kasus positif yang tidak terdeteksi oleh sistem atau model. Untuk kelas negatif, presisi adalah 43.46%, yang berarti sebagian kecil dari prediksi negatif mungkin salah. Namun, recall untuk kelas negatif adalah 100%, yang diberikan adalah benar. Namun, recall untuk kelas positif adalah 93.37%, yang berarti ada beberapa kasus positif yang tidak terdeteksi oleh sistem atau model. Untuk kelas negatif, presisi adalah 43.46%, yang berarti sebagian kecil dari prediksi negatif mungkin salah. Namun, recall untuk kelas negatif adalah 100%, yang berarti semua kasus negatif teridentifikasi dengan benar.

REFERENSI

- [1] A. Deolika, K. Kusriani, and E. T. Luthfi, "Analisis Pembobotan Kata Pada Klasifikasi Text Mining," *J. Teknol. Inf.*, vol. 3, no. 2, p. 179, 2019, doi: 10.36294/jurti.v3i2.1077.
- [2] F. Kurniawan and Q. Al Qorni, "Exploring Sentimen Analysis Using Machine Learning: A Case Study on Partai Demokrasi Indonesia Perjuangan (PDIP) in the 2024 General Election," vol. 2, no. 4, pp. 911–920, 2024.
- [3] S. Sahar, "Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Classifier Pada Dataset Penyakit Jantung," *Indones. J. Data Sci.*, vol. 1, no. 3, pp. 79–86, 2020, doi: 10.33096/ijodas.v1i3.20.
- [4] V. B. Kobayashi, S. T. Mol, J. Vrolijk, and G. Kismihók, "Text mining in career studies: Generating insights from unstructured textual data1," *Handb. Res. Methods Careers*, pp. 139–163, 2021, doi: 10.4337/9781788976725.00015.
- [5] M. Sari and A. Asmendri, "Penelitian Kepustakaan (Library Research) dalam Penelitian Pendidikan IPA," *Nat. Sci.*, vol. 6, no. 1, pp. 41–53, 2020, doi: 10.15548/nsc.v6i1.1555.
- [6] R. Fatha Pringgar and B. Sujatmiko, "Penelitian Kepustakaan (Library Research) Modul Pembelajaran Berbasis Augmented Reality pada Pembelajaran Siswa," *J. IT-EDU*, vol. 05, no. 01, pp. 317–329, 2020.
- [7] W. Darmalaksana, "Metode Penelitian Kualitatif Studi Pustaka dan Studi Lapangan," *Pre-print Digit. Libr. UIN Sunan Gunung Djati Bandung*, pp. 1–6, 2020.
- [8] S. Tedmori and A. Awajan, "Sentiment analysis main tasks and

- applications: A survey," *J. Inf. Process. Syst.*, vol. 15, no. 3, pp. 500–519, 2019, doi: 10.3745/JIPS.04.0120.
- [9] M. Christianto, J. Andjarwirawan, and A. Tjondrowiguno, "Aplikasi analisa sentimen pada komentar berbahasa Indonesia dalam objek video di website YouTube menggunakan metode Naïve Bayes classifier," *J. Infra*, vol. 8.1, pp. 255–259, 2020.
- [10] D. F. Zhafira, B. Rahayudi, and I. Indriati, "Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube," *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 2, no. 1, pp. 55–63, 2021, doi: 10.25126/justsi.v2i1.24.
- [11] S. Alkhuly, "Wordcloud Based R Programming as a Supplementary Tool for Critical Discourse Analysis," *Scientific in Arts*, vol. 24, no. 1, pp. 62–93, 2023, doi: 10.21608/jssa.2023.168941.1437.
- [12] J. W. Iskandar and Y. Nataliani, "Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 6, pp. 1120–1126, 2021, doi: 10.29207/resti.v5i6.3588.