

# SENTIMEN ANALISIS TERHADAP OPINI MASYARAKAT MENGENAI DRAMA KOREA PADA TWITTER MENGGUNAKAN METODE NAÏVE BAYES

## Literature Review

Nur Raisa, Noviana Riza\*

Program Studi D4 Teknik Informatika, Universitas Logistik dan Bisnis Internasional  
novianariza@ulbi.ac.id\*

## ABSTRAK

Data yang menyebabkan perluasan informasi, khususnya informasi teks dapat digunakan sebagai hotspot yang sangat diharapkan untuk pengungkapan tambahan. Data yang tidak terstruktur, terutama data tekstual, mencapai 80% dari jumlah data yang terus bertambah. Banyak sekali informasi yang tidak terstruktur yang dihasilkan dari hiburan virtual, misalnya, Twitter. Kata Korea untuk latar belakang sejarah adalah sageuk, dan drama Korea dapat berlatar belakang masa kini atau masa lalu. Kedua jenis drama ini termasuk dalam berbagai genre, termasuk komedi romantis, serial aksi, dan drama fiksi ilmiah. Naïve Bayes adalah sebuah keluarga algoritma pembelajaran mesin berbasis kemandirian statistik, bukan hanya satu algoritma. Terlepas dari namanya, pengklasifikasi Naïve Bayes yang menjadi lebih akurat ketika lebih banyak data ditambahkan, seperti ketika pengguna menandai pesan email di kotak masuk mereka sebagai spam. Secara keseluruhan terdapat berbagai jurnal melalui search process. setelah di review dan diseleksi berdasarkan inclusion and exclusion criteria dengan kata kunci “analisis sentiment” terdapat 15 paper jurnal yang kemudian diberi kualitas penilaian. Kemudian menggunakan scrapping data dan dimasukkan ke dalam “Google Colab” agar di simpan dan diolah kedalam bentuk excel.csv. kemudian akan dilanjutkan dengan mencari atribut yang mana saja penting dan yang akan diolah kemudian Sentimen Analisis hasil akhir akan menjadi suatu data yang akan digunakan sebagai data latih pada metode algoritma yang akan digunakan.

**Kata kunci:** Naïve Bayes, Twitter, Drama Korea, Analisis Sentimen

## 1. PENDAHULUAN

Peningkatan cepat dari dunia data dan inovasi korespondensi tidak dapat dipisahkan dari organisasi spesialis web yang memberikan data yang berbeda. Data yang menyebabkan perluasan informasi, khususnya informasi teks dapat digunakan sebagai hotspot yang sangat diharapkan untuk pengungkapan tambahan. Data yang tidak terstruktur, terutama data tekstual, mencapai 80% dari jumlah data yang terus bertambah. Banyak sekali informasi yang tidak terstruktur yang dihasilkan dari hiburan virtual, misalnya, Twitter.

Twitter merupakan salah satu jenis media sosial yang sangat diminati dan banyak digunakan di seluruh dunia, termasuk di Indonesia. Hiburan berbasis web ini diciptakan oleh Jack Dorsey dengan tujuan utama untuk mengirimkan pesan antar klien melalui kicauan atau tweet. Pada tahun 2020, Twitter akan memiliki 19,5 juta pengguna dan menduduki peringkat kelima sebagai platform media sosial yang paling banyak digunakan. Menurut data tersebut, Twitter merupakan platform media sosial yang memiliki banyak potensi pengguna, terutama di Indonesia. Orang-orang dapat dengan mudah berpendapat atau mengkritik sesuatu secara *real time* di Twitter, salah satu jejaring sosial. Salah satunya adalah mengenai drama Korea yang sering dikonsumsi oleh masyarakat Indonesia. Hal ini menimbulkan saran dan keluhan tentang manfaat dan kekurangannya. Ketidakpastian muncul sebagai akibat dari banyaknya opini publik yang positif dan negatif tentang budaya drama Korea. Oleh karena itu, metode

naive bayes dapat digunakan untuk menganalisis opini masyarakat mengenai budaya drama Korea dan membaginya ke dalam kategori positif, netral, dan negatif untuk mencari solusinya.

Naive Bayes adalah teknik untuk model klasifikasi probabilitas sederhana yang akan digunakan dalam perhitungan kumpulan probabilitas dengan menjumlahkan semua frekuensi dan campuran nilai yang menjadi bagian dari kumpulan data. Ketika diterapkan pada database yang berisi data dalam jumlah besar, Naive Bayes terbukti sangat akurat dan cepat.

Salah satu pendekatan yang paling banyak digunakan untuk klasifikasi teks adalah metode Naive Bayes, yang berkinerja baik. Tujuan dari penelitian ini adalah menggunakan Naive Bayes untuk memeriksa sentimen publik dengan merayapi, memproses teks, dan mengklasifikasikan data. Untuk karakterisasi tambahan memunculkan jenis positif, pesimis dan tidak memihak yang akan digunakan sebagai tolak ukur dalam memperkirakan seberapa beruntung atau tidak beruntungnya penilaian bangsa Indonesia terhadap dramatisasi Korea.

## 2. TINJAUAN PUSTAKA

### 2.1. Analisis Sentimen

Sebuah laporan komputasi untuk melihat dan menawarkan sudut pandang, perasaan, penilaian, mentalitas, perasaan, subjektivitas, keputusan, atau perspektif yang terkandung dalam sebuah pesan

dikenal sebagai pemeriksaan opini atau penggalan opini [1].

Instrumen penggalan opini adalah pengelolaan sekumpulan daftar yang diindeks dari suatu hal tertentu untuk menyusun daftar manfaat dari hal tersebut (seperti kualitas, sorotan, dan sebagainya), serta menentukan jumlah penilaian dari setiap kualitas tersebut (rendah, sedang, dan tinggi) [2]. Kamus Besar Bahasa Indonesia (KBBI) mendefinisikan sentimen sebagai:

1. Penilaian atau pandangan berdasarkan kecenderungan hati yang keliru terhadap sesuatu (berlawanan dengan akal sehat). Contoh: pilihan berikutnya akan menjadi tidak masuk akal bila disertai dengan perasaan yang bersifat individual.
2. Kecenderungan yang ekstrem. Contoh: Jika kita jauh dari bangsa ini, perasaan sebagai orang Indonesia akan semakin kuat.
3. Iri hati; ketidaksenangan; kebencian.
4. Tanggapan negatif. Contoh: Menurut KBBI, opini adalah pendapat, pikiran, atau pendirian. Sentimen pasar saja yang menjadi penyebab turunnya harga saham [3].

## 2.2. Drama Korea

Pertunjukan Korea (Hangul: 한국드라마; Hanja: 韓國드라마; RR: han-guk deurama, yang lebih sering disebut sebagai K-show, adalah serial TV berbahasa Korea Selatan. Drama Korea populer di seluruh dunia sebagian karena "Korean Wave" dan ketersediaan budaya populer Korea yang meluas melalui layanan streaming yang sering menawarkan subtitle dalam berbagai bahasa. Banyaknya adaptasi internasional dari drama Korea telah memberikan dampak yang signifikan di berbagai negara. Stasiun televisi di berbagai negara telah digunakan untuk menyiarkan acara-acara yang paling terkenal. Dae Jang Geum, misalnya, telah terjual ke 150 negara. Mode, gaya, dan budaya K-drama telah menarik perhatian seluruh dunia. Meningkatnya ketenaran acara-acara Korea telah mendorong dorongan yang sangat besar untuk garis gaya [4].

Drama Korea, juga dikenal sebagai K-drama korea, adalah jenis drama televisi Korea yang diproduksi di Korea Selatan. Sebagian besar drama berdurasi miniseri dan memiliki karakteristik yang berbeda yang membedakannya dari serial televisi atau opera sabun Barat biasa. Kata Korea untuk latar belakang sejarah adalah sageuk, dan drama Korea dapat berlatar belakang masa kini atau masa lalu. Kedua jenis drama ini termasuk dalam berbagai genre, termasuk komedi romantis, serial aksi, dan drama fiksi ilmiah. Korea Selatan memulai serial TV telekomunikasi pada tahun 1960-an, serial deurama yang berlangsung dengan 12-24 episode yang lebih kecil dari yang diharapkan dimulai pada tahun 1990-an, mengalihkan serial konvensional yang dapat diverifikasi sepenuhnya ke organisasi ini, juga,

membuat kemungkinan "kombinasi sageuk". Drama Korea biasanya direkam pada jadwal yang sangat dekat, sering kali beberapa jam sebelum penayangan yang sebenarnya, situasi yang mudah beradaptasi dan dapat berubah kapan saja selama pembuatan, tergantung pada kritik pengamat, membuat para penghibur berada dalam posisi yang sulit. Organisasi kreasi sering mengalami masalah keuangan. Layanan streaming sekarang menawarkan terjemahan berbagai bahasa dari drama Korea, yang berkontribusi pada popularitasnya di seluruh dunia [5].

## 2.3. Naïve Bayes

Naïve Bayes adalah sebuah keluarga algoritma pembelajaran mesin berbasis kemandirian statistik, bukan hanya satu algoritma. Mereka lebih efisien daripada algoritma Naïve Bayes yang lebih rumit dan relatif mudah ditulis. Aplikasi yang paling terkenal adalah saluran spam. Saluran spam melihat pesan email untuk semboyan tertentu dan menemukannya di amplop spam dengan asumsi bahwa pesan tersebut cocok. Terlepas dari namanya, pengklasifikasi Naïve Bayes yang menjadi lebih akurat ketika lebih banyak data ditambahkan, seperti ketika pengguna menandai pesan email di kotak masuk mereka sebagai spam [6]. Rumus untuk teori dari Naïve Bayes diberikan sebagai:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

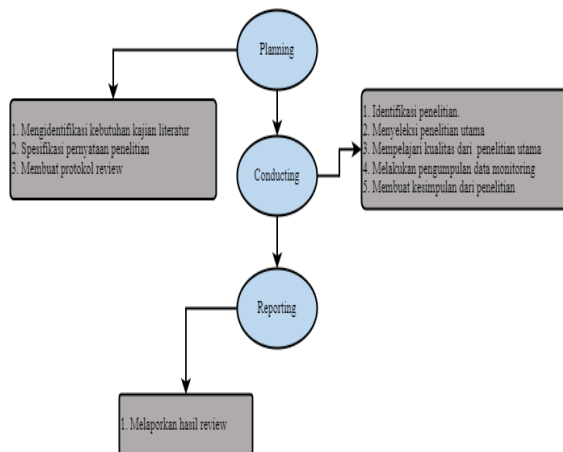
Gambar 1 rumus Naïve Bayes

Keterangan:

- $P(A|B)$  adalah Probabilitas Posterior: Probabilitas hipotesis A pada peristiwa yang diamati B.
- $P(B|A)$  adalah Probabilitas Kemungkinan: Probabilitas bukti yang diberikan bahwa probabilitas suatu hipotesis benar.
- $P(A)$  adalah Prior Probability: Probabilitas hipotesis sebelum mengamati bukti.
- $P(B)$  adalah Probabilitas Marginal: Probabilitas Bukti [7].

## 3. METODE PENELITIAN

Dalam penelitian ini terdapat diagram alur yang menggambarkan proses alur penelitian yang akan dilakukan berupa tahap-tahapan sehingga mudah dimengerti:



Gambar 2. Diagram SLR

### 3.1. Planning

Tujuan dari perencanaan adalah untuk mengantisipasi Research Question (RQ) atau pertanyaan penelitian. Tahap ini sangat penting karena akan memastikan bahwa tinjauan dilakukan secara terstruktur dan metodis serta meningkatkan kemungkinan untuk mendapatkan hasil yang relevan dan akurat. Kriteria inklusi dan eksklusi yang akan digunakan untuk memilih artikel untuk tinjauan harus ditentukan oleh peneliti selama tahap perencanaan, serta tujuan dan pertanyaan penelitian yang spesifik dan jelas. Selain itu, analisis juga harus membedakan sumber data yang penting dan sesuai dengan titik pemeriksaan, memutuskan sistem pencarian yang layak, dan menyusun metode untuk penentuan artikel, evaluasi kualitas, ekstraksi informasi, dan investigasi informasi yang akan diselesaikan. kumpulan data, teknik, model, dan metode.

### 3.2. Conducting

Tahap yang mencakup eksekusi aktual SLR adalah tahap conducting. Tahap ini dimulai dengan memilih slogan pencarian penulisan berbasis PICOC (string pencarian). Pemahaman Anda tentang sinonim dan alternatif pengganti akan menentukan ketepatan pencarian literatur. Identifikasi perpustakaan digital sebagai sumber pencarian literatur merupakan langkah selanjutnya. Karena jumlah literatur yang akan dikumpulkan - mungkin ratusan atau ribuan makalah - akan sangat banyak, maka disarankan agar Anda menggunakan alat bantu perangkat lunak untuk mempermudah pengelolaan literatur dengan menggunakan Mendeley.

### 3.3. Reporting

Pelaporan adalah tahap di mana hasil SLR dituliskan, baik untuk publikasi dalam jurnal ilmiah atau sebagai bab dalam tesis atau disertasi tentang tinjauan pustaka. Biasanya, struktur penulisan SLR terdiri dari tiga bagian utama: Bagian utama, pendahuluan, dan kesimpulan. Latar belakang dan dasar mengapa SLR pada suatu topik penting dan

harus dilakukan akan diberikan di bagian Pendahuluan. Protokol SLR, hasil analisis dan sintesis, serta diskusi mengenai implikasi dari hasil temuan tersebut akan dimasukkan ke dalam Bagian Utama. Sesuai dengan RQ yang telah ditetapkan di awal, ringkasan hasil yang diperoleh akan dimasukkan ke dalam bagian Kesimpulan.

### 3.4. Research Question

Pernyataan penelitian dibuat untuk menentukan metode yang akan digunakan dalam penelitian yang akan dilakukan oleh penulis nantinya. Berikut beberapa pernyataan penelitiannya:

Tabel 1. Conducting

Population(P)	Apa populasi target yang akan digunakan dalam penelitian ini ?
Intervention(I)	Apa metode naïve bayes yang akan digunakan sebagai intervensi dalam penelitian ini?
Comparison©	Dengan apa metode naïve bayes akan dibandingkan ?
Outcomes(O)	Apa hasil yang akan diukur atau dinilai dalam penelitian ini ?
Context©	Apakah ada konteks khusus yang perlu diperhatikan dalam penelitian ini?

Tabel 2. Research Question

ID	Pertanyaan Penelitian	Motivasi
RQ1	Metode penelitian manakah yang telah diusulkan peneliti Sebelumnya mengenai sentimen analisis?	Mengidentifikasi metode yang pernah diusulkan peneliti sebelumnya mengenai penerapan sentimen analisis pada pembelajaran.
RQ2	Sumber data dan jumlah data yang digunakan dalam analisis sentiment?	Mengidentifikasi sumber data dan jumlah data yang digunakan peneliti sebelumnya pada jurnal sebelumnya.
RQ3	Bahasa pemrograman yang sering digunakan dalam analisis sentiment?	Mengidentifikasi Bahasa pemrograman yang sering digunakan dalam penerapan analisis sentiment.

### 3.5. Strategi Pencarian

Dalam penelitian ini, strategi pencarian sumber yang digunakan adalah sebagai berikut:

1. Data diperoleh melalui *scraping* data di twitter.
2. Data di *scraping* menggunakan Google Colabration <https://colab.research.google.com/>
3. Menyeleksi Paper sesuai dengan kriteria yang telah ditentukan.
4. Melakukan pencarian paper pada berbagai basis data seperti Publish or perish, Open Knowledge Maps, Google Scholar.
5. Dalam proses pencarian sumber penelitian didapatkan sebanyak 15 paper yang sesuai dan dengan kriteria

### 3.6. Seleksi Studi

Setelah tahapan sumber penelitian dengan strategi yang telah ditentukan, tahap selanjutnya sejumlah 15 paper diseleksi 1315embali. Tahapan seleksi studi dilakukan dengan membaca dan menganalisis bagian abstrak dari setiap sumber yang berkaitan dengan topik analisis sentiment pada tahapan seleksi studi didapatkan paper yang relevan dan sesuai dengan kriteria yang telah ditentukan. Kemudian dilakukan review secara mendetail dengan membaca seluruh bagian dari paper yang sesuai dengan pertanyaan penelitian dengan topik analisis sentiment.

### 3.7. Quality Asessment

Pada tahap selanjutnya yaitu melakukan pengecekan dan evaluasi pada data sumber paper dengan beberapa pertanyaan sebagai berikut:

Tabel 3. *Quality Asessment*

QA1	Apakah literature tersebut diunggah dalam waktu 5 tahun terakhir ?
QA2	Apakah literature tersebut menggunakan metode yang valid dan dapat dipertanggung jawabkan ?
QA3	Apakah proses penafsiran hasil yang dilakukan telah mencakup semua aspek relevan, efisien, dan bersifat konsisten?

## 4. HASIL DAN PEMBAHASAN

No	Judul	Tahun	Hasil
1.	Sentimen Analisis Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier.	2022	Evaluasi opini berdasarkan data Twitter dilakukan dengan menggunakan perhitungan Unsophisticated Bayes Classifier (NBC). Untuk membuat asumsi tentang akun Twitter @Sandiuno, survei ini mengandalkan 450 tweet. Tweet-tweet tersebut diformat secara efektif dengan cara yang akurat, kritis, dan adil. Emosi dibagi menjadi urusan sosial yang adil, urusan sosial yang positif, dan urusan sosial yang negatif dengan menggunakan ekspresi yang muncul setelah data dibersihkan dari bahasa yang tidak biasa dan kata-kata yang bersifat mencegah. Sebuah dataset berisi 450 opini berbahasa Indonesia dari akun Twitter @sandiuno digunakan untuk mengevaluasi sistem sebelum diimplementasikan di Jawa dengan alat WEKA. Menurut temuan, ekstraksi kata kunci selama pembelajaran diperlukan untuk analisis sentimen yang akurat. Tingkat akurasi tertinggi dicapai, dengan ketepatan umum 65,18 persen dan tingkat informasi pengujian persiapan 70:30. Terlepas dari itu, akurasi jujur adalah 48,88%, akurasi negatif 68,88%, dan akurasi positif 77,77%, tergantung pada permintaan. Hasil tinjauan ini menunjukkan bahwa struktur stemming berpengaruh pada ketepatan rencana karena mengubah pentingnya evaluasi penting secara teratur.
2.	Implementasi Algoritma Naïve Bayes Terhadap Analisis Sentimen Opini Masyarakat Indonesia Mengenai Drama Korea Pada Twitter.	2022	Implementasi kondisi Bayes langsung untuk menganalisis evaluasi Twitter Indonesia yang terkenal terhadap acara televisi Korea. Data yang dikumpulkan melalui penggarukan dianalisis menggunakan perangkat lunak oranye. Dengan menggunakan perhitungan Innocent Bayes, tinjauan ini memiliki opsi untuk mengkarakterisasi perasaan sebagai baik, tidak bias, atau pesimis dengan tingkat presisi 69%. Artikel ini juga membahas pentingnya penilaian perasaan dalam memisahkan evaluasi yang akrab dengan contoh-contoh yang tidak akrab. Proses penilaian kecenderungan meliputi penggarukan data, pra-pengolahan, permintaan Naive Bayes, pengujian, penilaian, dan analisis hasil dengan matriks kekacauan dan awan kata. Menurut penelitian tersebut, algoritme tersebut memiliki skor F1 sebesar 69 persen, presisi 73 persen, recall 69 persen, dan akurasi 69 persen. Artikel tersebut mengasumsikan bahwa penelitian perasaan dapat memberikan evaluasi umum terhadap berbagai subjek, termasuk dramatisasi Korea. Web scratching, pemeriksaan perasaan, mesin vektor pendukung, dan Naive Bayes adalah beberapa teknik pemeriksaan teks yang dirujuk dalam artikel tersebut untuk digunakan dalam pengelompokan dan pemeriksaan informasi hiburan online.
3.	Analisis Sentimen Cyberbullying Kpop di Media Sosial Twitter Menggunakan Metode Naïve Bayes.	2022	Analisis sentimen Naive Bayes dari cyberbullying terhadap penggemar K-Pop di Twitter. Tujuan dari ulasan ini adalah untuk mengklasifikasikan tweet negatif, positif, dan netral untuk memahami arti penting penggemar K-Pop muda. Selain itu, artikel ini juga meneliti kebutuhan untuk mengakhiri cyberbullying dan pengaruhnya terhadap kesejahteraan pribadi seseorang. Dalam penelitian ini, strategi berbasis referensi kata yang disebut SentiStrength digunakan untuk mengidentifikasi kecenderungan, dan dasbor online disiapkan untuk evaluasi cepat. Artikel ini membahas metode Naive Bayes untuk mengevaluasi cyberbullying terkait KPOP di Twitter. Metode ini mengkonsolidasikan pengaturan informasi, pembersihan, pengujian, dan arahan pra-independen. Hasilnya memiliki tingkat presisi 67,27 persen, tetapi ada masalah dengan informasi yang tidak merata dan penandaan manual. Artikel ini menganalisis penilaian Naive Bayes untuk perasaan waspada terhadap informasi Twitter. Temuan dari tinjauan ini dipengaruhi oleh distribusi yang tidak merata dari informasi yang ditandai dengan kuat dan negatif. Untuk hasil yang lebih tepat, para penulis merekomendasikan untuk menggunakan beberapa algoritme klasifikasi dan meningkatkan jumlah data pelatihan. Mereka juga merekomendasikan untuk memperluas penelitian tentang analisis sentimen

No	Judul	Tahun	Hasil
			cyberbullying dan memasukkan konversi gambar ke teks ke dalam sistem. Dalam artikel ini, evaluasi terkait ditinjau secara singkat.
4.	Analisis Sentimen Terhadap Penggemar K-pop di Media Sosial Twitter Menggunakan Naïve bayes (Studi Kasus Penggemar Grup BTS)	2020	Meneliti sentimen yang diungkapkan oleh penggemar BTS dan K-pop di Twitter. Ulasan ini menggunakan perhitungan Innocent Bayes dan bahasa pemrograman Python untuk mengelompokkan 1.000 tweet sebagai yakin, negatif, atau tidak bias. Menurut temuannya, 34,2% dari tweet adalah negatif, 58,5 persen netral, dan 7,3% positif. Penelitian ini sampai pada kesimpulan bahwa analisis sentimen dapat membantu memahami dan mengatasi isu-isu terkait ujaran kebencian dan pelecehan online. Tiga sumber yang dibahas berfokus pada penelitian perasaan yang memanfaatkan kecerdasan buatan (AI) untuk mengatur penilaian umum tentang berbagai subjek melalui platform media sosial seperti Instagram dan Twitter. Pengujian difokuskan pada pemisahan tweet dan komentar untuk memilih opini populer tentang transgender, profesional pemerintah, dan organisasi. Dengan menggunakan pembobotan berdasarkan jumlah retweet, Naive Bayes Classification dan K-Nearest Neighbor digunakan.
5.	Analisis Sentimen Persepsi penggunaan Pedulilindungi menggunakan Algoritma Naïve Bayes Classier (NBC) dan Support Vector Machine (SVM)	2022	Ulasan aplikasi PeduliLindungi dari versi 4.3 hingga 4.4.3.1 diposting di situs web Google Play antara tanggal 30 Maret 2022 dan 14 Mei 2022. Sebanyak 5.077 survei telah dikumpulkan, menurut laporan. Setelah melakukan prapemrosesan dan menyelidiki perasaan, sebanyak 2.105 survei digunakan, dengan 747 survei (35 persen) memiliki opini positif dan 1.358 survei (65 persen) memiliki opini negatif. Istilah-istilah audit aplikasi, bantuan, antibodi, akan, cinta, mengakui, baru, hebat, masuk, dan manfaat sering dibahas dalam kelas opini positif. Kelas opini negatif sering membahas tentang aplikasi, inokulasi, pernyataan, merepotkan, login, bantuan, buka, informasi, pembaruan, tanggal, dan kelahiran, serta muncul sebagai survei. Seharusnya tidak mengejutkan siapa pun bahwa beberapa klien aplikasi dapat menggunakannya, sementara yang lain melacaknya dengan susah payah dalam variasi ini. Sebagai hasil langsung dari hal ini, adalah mungkin untuk sampai pada kesimpulan bahwa, pada dasarnya, subjek yang biasanya diperiksa dalam survei kelas perasaan positif dan negatif sangat sebanding. Kesimpulan ini dapat dicapai dengan melihat data. Pada setiap pengujian pembagian datum, eksekusi akurasi menunjukkan bahwa penilaian Support Vector Machine (SVM) berkinerja lebih baik dibandingkan dengan perhitungan Naive Bayes Classifier (NBC). Hal ini mempertimbangkan informasi yang sudah jadi dari para pemimpin dengan mengisolasi informasi pengujian dari informasi yang mengatur dalam proporsi 90:10, 80:20, dan 70:30. Untuk setiap hubungan, data dapat diuji dua kali atau beberapa kali. Berdasarkan temuan dari evaluasi garis besar fishbone, survei kelas kecenderungan mencurigakan dari klien aplikasi PeduliLindungi mencakup delapan pertanyaan yang disusun dalam kelas 3P (Individu, Hal, dan Siklus). Hasil eksekusi akurasi juga menunjukkan bahwa aplikasi PeduliLindungi digunakan tanpa peringatan, sering mengalami crash, terdapat masalah penyaringan, administrasi klien yang tidak memadai dan membingungkan, sulit untuk menyimpan catatan vaksinasi dari aplikasi, pembaruan terbaru menyebabkan masalah, pengguna tidak dapat masuk, dan tidak dapat menyimpan catatan vaksinasi.
6.	Sentimen Analisis Stay Home Menggunakan Metode Klasifikasi NaïveBayes, Support Vector Machine, dan K- Nearest Neighbor.	2022	sebuah investigasi tentang kesan klien Twitter terhadap prosedur "tetap di rumah" selama pandemi Covid. Untuk mengantisipasi penilaian perasaan yang dikomunikasikan dalam tweet, survei ini menggunakan tiga teknik administratif: Naive Bayes, k-Closest Neighbor, dan Support Vector Machine Perhitungan Algoritma Vector Machine + Destroyed dianggap paling dapat diandalkan dengan ketepatan 80,05 persen. Makna pemeriksaan opini dalam pemeriksaan kesimpulan, penilaian, mentalitas, dan sentimen sehubungan dengan berbagai elemen, termasuk barang, administrasi, asosiasi, orang, masalah, kejadian, dan poin, serta kualitas zat-zat ini, juga merupakan subjek artikel ini.
7.	Klasifikasi Sentimen Pada VideoChannel Youtube Raditya Dika Menggunakan Metode Naïve Bayes Classifier.	2021	Video di kanal YouTube Raditya Dika dapat dikategorikan sentimennya secara tepat dengan bantuan pengklasifikasi naive bayes. Pengumpulan kecenderungan di saluran YouTube Raditya Dika berhasil diselesaikan menggunakan pengklasifikasi naif bayes dengan kecepatan presisi 72,33 persen dalam evaluasi 90:10..
8.	Sentimen Analisis Pelanggan	2021	Penilaian kecenderungan studi klien pada Shopee, sebuah komunitas bisnis elektronik yang mencolok di Indonesia, menggunakan estimasi Straightforward Bayes. Dengan menggunakan jaringan yang berantakan, penelitian ini bertujuan untuk menentukan

No	Judul	Tahun	Hasil
	Shopeedi Twitter Menggunakan Algoritma Naïve Bayes.		keakuratan komentar positif dan negatif dari grup Twitter Shopee. Hasil penelitian menunjukkan bahwa terdapat 150 komentar positif dan 150 komentar negatif, dengan tingkat akurasi sebesar 97%. Selain itu, garis besar penelitian sebelumnya tentang pemeriksaan perasaan menggunakan Naive Bayes dan perhitungan yang berbeda diberikan dalam artikel ini. Pengklasifikasi yang digunakan dalam kemungkinan dan pengukuran adalah perhitungan Naive Bayes. Instrumen untuk memilih ketepatan adalah kerangka harapan yang digagalkan. Dengan akurasi 97%, penelitian ini menyelidiki sentimen dalam tweet pelanggan Shopee menggunakan Naive Bayes. Penelitian ini menyarankan untuk membandingkan berbagai pendekatan untuk penelitian di masa depan. R Studio adalah aplikasi sumber terbuka yang digunakan untuk membuat program dalam bahasa R.
9.	Analisis Sentimen Review Halodoc Menggunakan Naïve Bayes Classifier.	2021	Investigasi perasaan dengan perhitungan Innocent Bayes Classifier untuk audit Halodoc di Google Play Store. Artikel ini menggambarkan proses preprocessing, tokenisasi, perubahan ke huruf kecil, saluran stopwords, dan saluran token (berdasarkan panjangnya) informasi. Validasi silang dan nilai validasi k-fold sebesar 10 digunakan untuk mengklasifikasikan data ke dalam dua kategori positif dan negatif. Klasifikasi dilakukan dengan cukup baik, dengan AUC 0,756 dan akurasi 81,68 persen. Selain itu, artikel tersebut membahas tentang pertumbuhan jumlah pengguna internet di Indonesia dan pentingnya layanan kesehatan digital. Artikel tersebut sampai pada kesimpulan bahwa Naive Bayes Classifier adalah metode yang mudah dan efektif untuk analisis sentimen; namun, metode ini tidak memiliki kemampuan untuk mengukur akurasi prediksi dan pemilihan atribut secara akurat. Penggunaan particle swarm optimization dalam studi optimasi pengklasifikasi Naive Bayes untuk klasifikasi teks dan analisis sentimen juga menjadi topik dalam artikel tersebut. Artikel ini juga secara singkat menjelaskan tentang ide Information Mining dan pengaruh Industri 4.0 terhadap layanan medis dan bioteknologi.
10.	Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naïve Bayes Classification.	2021	Menurut temuan eksperimen, 72,72% responden setuju bahwa virus Corona lazim digunakan dalam hal rencana, terutama PSBB, persyaratan untuk kesimpulan, dan waktu check-in. Tiga kata yang paling sering digunakan dalam konteks positif adalah "kerja", "jakarta", dan "demo". Di sisi lain, istilah "individu", "Jakarta", dan "demo" paling banyak digunakan dalam konteks negatif. Penggunaan kata "demo" dan "jakarta" dalam dua konteks luar biasa ini menunjukkan bahwa tweet dari sistem Covid the board tidak dapat dilepaskan dari data-data yang terkait dengan peristiwa, terutama selama pelaksanaan Peraturan Cipta Karya dan di saat-saat paling akhir PSBB Menuju Masyarakat Jakarta yang Kokoh, Terjamin, dan Indah. Selain itu, penerapan prosedur Covid the board di Jakarta, khususnya PSBB, kewajiban berkerudung, dan pembatasan waktu, belum terlalu berhasil, sebagaimana dibuktikan dengan masih adanya publikasi berbagai pemberitaan Jakarta yang vulgar. Strategi ini cukup untuk menyampaikan emosi optimis dan pesimis, sebagaimana dibuktikan oleh skor ajakan situasi sebesar 87,34 persen ketepatan, 93,44 persen kepedulian, dan 71,74 persen kekhususan. Hal ini menunjukkan bahwa data dan sumber informasi pilihan seperti Twitter harus menjadi fokus dari tanggapan masyarakat terhadap masalah terkini. Sangat mungkin untuk menentukan apa yang dikomentari orang dalam kalimat-kalimat tweet dengan mempertimbangkan pengulangan kata-kata dengan frekuensi tinggi. Salah satu tanda bahwa rencana permainan saat ini tidak berhasil adalah bahwa orang-orang di Jakarta tidak senang dengan bagaimana teknik reaksi Covid digunakan. Pemerintah daerah dapat melibatkan hal ini sebagai penilaian untuk memperbaiki kerangka kerja yang sedang berjalan. Hal ini dapat memeriksa berbagai masalah yang sedang berlangsung dan, pada akhirnya, memberikan informasi baru untuk pemeriksaan berikutnya dengan memanfaatkan berbagai metode pengamanan yang dipantau dan sumber informasi dari internet atau platform permainan berbasis web lainnya.
11.	Analisis Sentimen Terhadap Pembelajaran Daring dengan Algoritma Naïve Bayes Classifier.	2022	Perhitungan Blameless Bayes Classifier (NBC) dapat memperkirakan dan mempertimbangkan penilaian dari sudut pandang wali murid terhadap pembelian online dari rumah dengan hasil yang fenomenal mengingat fakta bahwa tanggapan tambahan diyakini sebagai tanggapan yang baik dan bukannya meragukan, di mana derajatnya adalah 62,5% untuk tanggapan yang baik dan 37,5% untuk tanggapan yang biasa saja. Ini berasal dari 40 informasi ulasan yang masuk ke dalam model yang telah dibuat dan menyelesaikan tahap penilaian. Pengujian ini memiliki tingkat keakuratan sebesar 65 persen. Kepercayaan terhadap keakuratan data ini didasarkan pada 200 dataset berita, 100 di antaranya berisi berita positif dan 100 di antaranya berisi berita negatif. Hasilnya, tingkat ketepatan model investigasi opini dipengaruhi oleh seberapa banyak dataset yang dikumpulkan digunakan untuk mengklasifikasikan berita sebagai positif atau negatif. Dalam banyak kasus, model ini tidak membedakan antara data dari jajak pendapat positif dan negatif atau sebaliknya. Seberapa pasti dan benar-benar berita dapat dipartisi ke dalam kelas positif dan negatif mempengaruhi hal ini.

No	Judul	Tahun	Hasil
12.	Analisa Sentimen Mengenai Kenaikan Harga Bbm Menggunakan Metode Naïve Bayes Dan Support Vector Machine.	2023	Meskipun ada beragam pendapat, kenaikan harga bahan bakar ini seharusnya mempengaruhi perekonomian Indonesia secara keseluruhan. Dengan menggunakan perhitungan dengan Naive Bayes dan Support Vector Machine, artikel ini melihat opini di Twitter. Perhitungan Support Vector Machine ditemukan sebagai yang paling akurat, dengan tingkat akurasi 90%. Artikel ini juga membahas bagaimana kenaikan biaya bahan bakar mempengaruhi ekspansi dan ekonomi secara keseluruhan. Pada bulan Juni 2013, kenaikan harga bahan bakar menghasilkan tingkat pertumbuhan tahunan sebesar 8,38 persen untuk ekonomi Indonesia. Perspektif pengguna Twitter terhadap kenaikan harga bahan bakar dikumpulkan melalui penggunaan Naive Bayes dan Support Vector Machine dalam penelitian ini. Algoritma Support Vector Machine dan Naive Bayes akan digunakan dalam penelitian ini untuk mengetahui bagaimana perasaan masyarakat terhadap subsidi BBM.
13.	Analisis Sentimen Produk Kecantikan Jenis Serum Menggunakan Algoritma Naïve Bayes Classifier.	2022	18.587 tweet, atau 68,3 persen dari semua ulasan negatif, menunjukkan bahwa produk kecantikan jenis serum tidak boleh digunakan. Hasil ulasan tersebut membuat beberapa orang percaya bahwa produk kecantikan jenis serum tidak boleh digunakan. Terdapat 8.644 tweet dengan sentimen positif, atau 35%, dan 356 tweet, atau 1,2%, dianggap netral terhadap penggunaan produk unggulan jenis serum. Berdasarkan hasil eksperimen yang dilakukan pada grid yang berantakan, model Naive Bayes Classifier dianggap sangat baik dalam melakukan pengorganisasian dengan nilai akurasi sebesar 80,4%. Perasaan negatif merupakan faktor pameran terbaik model ini dengan nilai 91 persen. Sementara model tersebut menemukan bahwa enam puluh persen dari survei yang dilakukan adalah positif.
14.	Analisis Sentimen Pada Komentar Sosial Media Instagram Layanan Kesehatan BPJS Menggunakan Naïve Bayes Classifier.	2021	Algoritma Nave Bayes dapat digunakan untuk mengklasifikasikan komentar pada sebuah postingan Instagram BPJS Kesehatan berdasarkan hasil nilai probabilitas Nave Bayes yang termurah. Penulis dapat mempermudah masyarakat dalam mengidentifikasi sebuah opini atau pendapat mengenai akun Instagram BPJS Kesehatan berdasarkan hasil penelitian pada bagian akhir laporan. Sistem yang dikembangkan bebas dari kesalahan dan beroperasi sesuai dengan masukannya, yang ditunjukkan dengan hasil temuan pengujian blackbox. Penelitian ini menggunakan 50 kometer data testing selain 250 kometer data training. Sebanyak 31 komentar terbesar dari 50 data pengujian menghasilkan klasifikasi yang benar, sementara 19 komentar menghasilkan klasifikasi yang salah. Dengan tingkat akurasi 62%, hasil pengujian klasifikasi menunjukkan performa yang cukup baik.
15.	Sentimen Analisis Tentang Kebijakan Pemerintah terhadap kasus carona menggunakan metode Naïve Bayes.	2021	Diketahui dari perhitungan sentimen penerapan sistem New Normal, sebanyak 56,39 persen masyarakat berpendapat positif dan 43,61 persen berpendapat negatif terhadap kebijakan pemerintah terkait penerapan sistem New Normal. Informasi ini dikumpulkan dari 1823 cuitan yang dituliskan di akun Twitter pada periode 6 Juli 2020 hingga 25 Juli 2020. Hasil eksperimen menggunakan <i>Disarray Network</i> mendapatkan nilai ketepatan yang paling kecil dengan karakterisasi Naïve Bayes menggunakan TF-IDF meliputi ekstraksi, yang menambahkan hingga 81%, dengan nilai Precision 78%, Review 91%, dan f1-Score 84%. Menggunakan parameter algoritma Naive Bayes dengan N-Gram tipe Trigram menghasilkan nilai akurasi tertinggi 84 % tetapi ada juga nilai untuk <i>precision</i> 84%, <i>recall</i> 86%, dan <i>f1-Score</i> 85%. Hal ini menunjukkan seberapa baik algoritma Naive Bayes dengan ekstraksi fitur <i>TF-IDf</i> dan <i>N-Gram</i> dapat digunakan untuk mengklasifikasikan data tweet publik terkait dengan kebijakan pemerintah mengenai penerapan sistem <i>New Normal</i> . Peningkatan presisi yang didapat setelah menggunakan <i>N-Gram</i> adalah karena banyak Prarasa bahasa Indonesia yang memiliki 2 hingga 3 kosakata.

### Pembahasan Hasil

Bagian ini akan menjelaskan *Research Question (RQ)*

direview berdasarkan metode penelitian yang sering digunakan untuk menjawab research question.

#### RQ1. Metode Penelitian mana yang sudah diusulkan dalam penelitian sebelumnya mengenai analisis sentiment?

Secara keseluruhan terdapat berbagai jurnal melalui search process. setelah di review dan diseleksi berdasarkan inclusion and exclusion criteria dengan kata kunci (keyword) "analisis sentiment" terdapat 15 paper jurnal yang kemudian diberi kualitas penilaian (quality assessment). Dari hasil Quality Assesment (QA) terdapat 15 jurnal yang kemudian dikelompokkan berdasarkan usulan dari paper yang

Tabel 4. Metode Jurnal

Metode Penelitian	Jumlah Jurnal
Naïve Bayes Classifier	7 Jurnal
Naïve Bayes	5 Jurnal
Naïve Bayes Classifier dan Support Vector Machine	2 Jurnal
Naïve Bayes Classifier, Support Vector Machine dan K-Nearest Neighbor	1 Jurnal

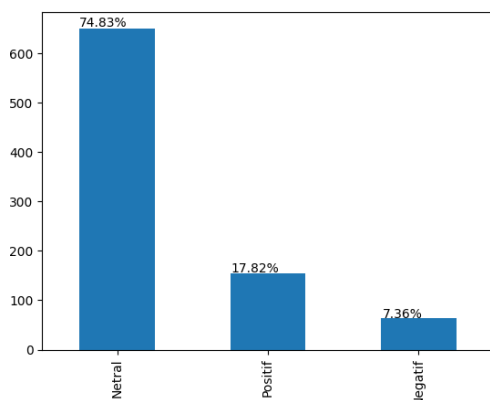
## RQ2. Sumber Data dan Jumlah Data yang digunakan dalam analisis sentimen?

Dari paper yang sudah direview sebelumnya diambil dari sumber data twitter melakukan proses pengumpulan data yaitu masuk ke dalam web halaman "Twitter" tentang data drama korea dan mengambil sebanyak 7 judul drama korea dan masing judul mengambil 200 data pada twitter.

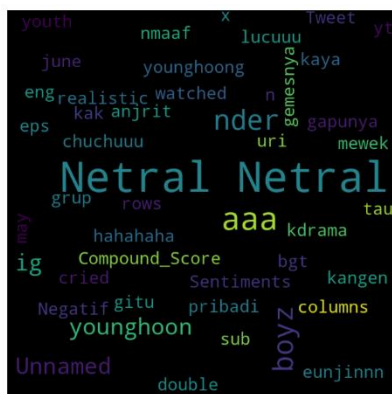
Tabel 5. Pengumpulan Data

Vincenzo	Action	201 Data
Happiness	Fantasi	210 Data
Bad Prosecutor	Law	210 Data
The Glory	Thriller	208 Data
Welcome to Waikiki	Comedy	200 Data
Start Up	Romance	202 Data
The Penthouse	Melodrama	202 Data

Kemudian menggunakan scrapping data dan dimasukkan ke dalam "Google Colab" agar di simpan dan diolah kedalam bentuk excel.csv. kemudian akan dilanjut dengan mencari atribut yang mana saja penting dan yang akan diolah kemudian Sentimen Analisis hasil akhir akan menjadi suatu data yang akan digunakan sebagai data latih pada metode algoritma yang akan digunakan. Dari 1400 data setelah melakukan *preprocessing* menjadi 600 data dari data tersebut dilakukan *sentimen analisis* terdapat Positif, Negatif, dan Netral.



Gambar 3. Grafik Sentimen



Gambar 4. Hasil Sentimen

## RQ3. Bahasa Pemrograman yang sering digunakan dalam analisis sentimen?

Bahasa pemrograman yang paling mudah dan sering digunakan yaitu Python bahasan pemrograman yang paling populer untuk analisis sentiment. Dan dalam python memiliki alat bantu seperti sudah mempunyai *library*, NLTK (Natural Language Toolkit), TextBlob, dan scikit-learn, pemrosesan NLP yang dapat memberikan ekstraksi kata, penerjemah kata.

## 5. KESIMPULAN DAN SARAN

Dari Penelitian yang telah dilakukan Naive Bayes adalah metode yang efektif untuk analisis sentimen pada data teks, terutama pada data yang diambil dari media sosial seperti Twitter, Instagram, dan YouTube. Namun, terdapat beberapa keterbatasan dalam penggunaan metode ini, seperti akurasi prediksi dan seleksi atribut. Oleh karena itu, beberapa penelitian merekomendasikan penggunaan beberapa algoritma klasifikasi dan peningkatan jumlah data latih untuk hasil yang lebih akurat. pentingnya analisis sentimen dalam memahami opini publik dan isu-isu yang berkaitan dengan ujaran kebencian. Hasil dari sentimen analisis yang dilakukan terdapat 74,83% sentiment netral 17,82% positif, 7,36% negatif yang ada pada tweet pada 7 drama korea tersebut. Drama terbaik yaitu Welcome to Waikiki dengan 121 data, 43 positif, 25 Negatif dan 53 Netral dan Drama Terfavorit yaitu Vincenzo dengan 121 data, 80 Positif, 8 Negatif, 33 Netral.

Saran untuk peneliti selanjut untuk dapat menambahkan atau membandingkan dengan metode yang bisa membuat perbandingan atau menambah keakuratan hasil pada penelitian.

## DAFTAR PUSTAKA

- [1] Muhammad Imam Syafi. 2022. Sentimen Analisis Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier.
- [2] Resti Amelia, Darmansyah, Nanda Sesty Pratiwi, dan Muhammad Eka Purbaya. 2022. Implementasi Algoritma Naïve Bayes Terhadap Analisis Sentimen Opini Masyarakat Indonesia Mengenai Drama Korea Pada Twitter.
- [3] Anisa Putri, Ari Muzakir. 2022. Analisis Sentimen Cyberbullying Kpop di Media Sosial Twitter Menggunakan Metode Naïve Bayes.
- [4] Puji Nurmawati, Endang Supriyati, dan Tri Listyorini. 2020. Analisis Sentimen Terhadap Penggemar K-pop di Media Sosial Twitter Menggunakan Naïve Bayes (Studi Kasus Penggemar Grup BTS).
- [5] Rizki Rahmattullah. 2022. Analisis Sentimen Persepsi penggunaan Pedulilindungi menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM).
- [6] Ikhwanul Hakim, Arifin Nugroho, Sulaeman Sukmana, Windu Gata. 2022. Sentimen Analisis Stay Home Menggunakan Metode Klasifikasi Naïve Bayes, Support Vector Machine, dan K-



- Nearest Neighbor.
- [7] Muhammad Danil Saputra. 2021. Klasifikasi Sentimen Pada Video Channel Youtube Raditya Dika Menggunakan Metode Naïve Bayes Classifier.
- [8] Aditya Hastami Ruger, M Suyanto, dan Mei P Kurniawan. 2021. Sentimen Analisis Pelanggan Shopee di Twitter Menggunakan Algoritma Naïve Bayes.
- [9] Asep Hendra, Fitriyanti. 2021. Analisis Sentimen Review Halodoc Menggunakan Naïve Bayes Classifier.
- [10] Ni Putu Gita Naraswati, Delvira Cindy Rosmilda, Dinda Desinta, Fadhilatul Khairi, Riska Damaiyanti, Rani Nooraeni. 2021. Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naïve Bayes Classification.
- [11] Luthfia Oktasari, Yulison Herry Chrisnanto, Rezki Yuniarti. 2016. Text Mining Dalam Analisis Sentimen Asuransi Menggunakan Metode Naïve Bayes Classifier.
- [12] Fajar Siddik, Ibnu Suhada, Azhar Haikal Anwar, Firman Noor Hasan. 2022. Analisis Sentimen Terhadap Pembelajaran Daring dengan Algoritma Naïve Bayes Classifier.
- [13] Ferdi Rejeki, Vina Ayumi. 2023. Analisa Sentimen Mengenai Kenaikan Harga Bbm Menggunakan Metode Naïve Bayes Dan Support Vector Machine.
- [14] Muhammad Hamka, Naila Alfatari, Dhani Ratna Sari. 2022. Analisis Sentimen Produk Kecantikan Jenis Serum Menggunakan Algoritma Naïve Bayes Classifier.
- [15] Ahmad Karim. 2021. Analisis Sentimen Pada Komentar Sosial Media Instagram Layanan Kesehatan BPJS Menggunakan Naïve Bayes Classifier.