

**ANALISIS SENTIMEN PADA ULASAN APLIKASI CHAT  
GENERATIVE PRE-TRAINED TRANSFORMER GPT MENGGUNAKAN  
METODE KLASIFIKASI K-NEAREST NEIGHBOR(KNN)  
*Sistematic Literature Review***

**Muhammad Nanda Fahriza, Noviana Riza**

Program Studi Teknik Informatika, Universitas Logistik dan Bisnis Internasional  
Jl. Sariosih No. 54 Bandung 40151 Telp. 022-2009570, fax. 022-2009568  
nandafahriza06@gmail.com

**ABSTRAK**

Klasifikasi K-Nearest Neighbor (KNN) akan dipakai dalam penelitian ini untuk melakukan analisis sentimen terhadap ulasan pengguna terhadap aplikasi Chat Generative Pre-Trained Transformer GPT. Model bahasa reguler GPT telah dipersiapkan secara luas dan dapat menghasilkan teks yang dapat dimengerti dan signifikan. Bagaimanapun, untuk memahami reaksi klien terhadap aplikasi, diperlukan pemeriksaan opini atas audit. Dalam penelitian ini, hasil analisis sentimen akan memberikan pemahaman yang lebih mendalam tentang respon pengguna terhadap aplikasi Chat Generative Pre-Trained Transformer GPT. Informasi ini dapat berguna bagi pengembang dan pemilik aplikasi untuk memahami kelebihan dan kekurangan aplikasi serta meningkatkan pengalaman pengguna secara keseluruhan. Penelitian ini ingin mengidentifikasi penelitian terkait sebelumnya, membandingkan pendekatan dan teknik yang digunakan, menganalisis efektivitas metode KNN, mengidentifikasi kelebihan dan kekurangan metode tersebut, serta memberikan rekomendasi untuk pengembangan lebih lanjut. Diharapkan hasil penelitian ini dapat memberikan pemahaman yang mendalam tentang respon pengguna terhadap aplikasi tersebut dan memberikan sumbangan bagi pengembangan dan peningkatan pengalaman pengguna.

**Kata Kunci:** Analisis sentimen, ulasan pengguna, aplikasi Chat Generative Pre-Trained Transformer GPT, K-Nearest Neighbor (KNN).

**1. PENDAHULUAN**

Teknologi chatbot semakin populer dan aplikasi yang menggunakan model generatif seperti GPT semakin banyak digunakan. GPT merupakan model bahasa yang dilatih dengan data besar untuk menghasilkan teks yang mirip dengan teks manusia. Aplikasi chat generatif seperti ini dapat digunakan dalam berbagai konteks, seperti layanan pelanggan, asisten virtual, atau interaksi manusia-mesin lainnya. Namun, penting untuk memahami sentimen atau pendapat pengguna terhadap aplikasi chat generatif ini. Analisis sentimen bertujuan untuk mengekstrak dan memahami opini, sikap, atau perasaan yang terkandung dalam teks ulasan. Dengan menganalisis sentimen, pengembang aplikasi dapat memperoleh wawasan yang berharga untuk meningkatkan kualitas layanan atau memperbaiki masalah yang ada. Dalam konteks ini, penelitian ini menggunakan metode klasifikasi K-Nearest Neighbor (KNN).

KNN adalah sebuah metode klasifikasi yang umum digunakan dalam pemrosesan bahasa alami. Metode ini mengklasifikasikan suatu teks berdasarkan kesamaannya dengan teks-teks lain dalam dataset. Dalam penelitian ini, KNN digunakan untuk mengklasifikasikan ulasan-ulasan aplikasi chat GPT menjadi kategori sentimen tertentu, seperti positif, negatif, atau netral. Untuk mendukung analisis sentimen pada ulasan aplikasi chat GPT, penelitian ini melakukan tinjauan literatur sistematis.

Tinjauan literatur sistematis merupakan suatu metode penelitian yang sistematis dan komprehensif

untuk mengumpulkan, mengevaluasi, dan mensintesis literatur yang relevan dalam suatu domain penelitian tertentu. Dengan melakukan tinjauan literatur sistematis, penelitian ini bertujuan untuk menyediakan gambaran yang komprehensif tentang analisis sentimen pada ulasan aplikasi chat GPT menggunakan metode klasifikasi KNN berdasarkan penelitian-penelitian terkait yang telah dilakukan sebelumnya.

Sistematic literature review adalah proses pengumpulan, evaluasi, dan sintesis dari studi sebelumnya yang relevan dengan topik penelitian yang sedang dipelajari. Dengan menggunakan metode ini, kami dapat mengidentifikasi, menganalisis, dan membandingkan studi sebelumnya yang telah dilakukan oleh peneliti lain dalam bidang analisis perbandingan algoritma klasifikasi lainnya.

Penelitian ini juga memiliki tujuan untuk melakukan tinjauan literatur sistematis. Tinjauan literatur sistematis dilakukan untuk mengumpulkan, mengevaluasi, dan mensintesis penelitian-penelitian terkait yang telah dilakukan sebelumnya tentang analisis sentimen pada ulasan aplikasi chat GPT menggunakan metode klasifikasi KNN. Tujuan dari tinjauan literatur ini adalah untuk menyediakan gambaran yang komprehensif tentang penelitian-penelitian terkait dan memperoleh pemahaman yang lebih baik tentang topik penelitian ini.

Dalam penelitian ini, peneliti akan mengumpulkan artikel-artikel ilmiah terkait yang telah dipublikasikan dalam jurnal-jurnal dan konferensi terkemuka. Selanjutnya akan melakukan evaluasi

terhadap studi-studi tersebut berdasarkan kriteria-kriteria yang telah ditentukan, seperti metode yang digunakan, dataset yang digunakan, dan metrik evaluasi yang digunakan.

Selanjutnya, menganalisis dan membandingkan hasil-hasil yang ditemukan dalam studi-studi tersebut. Mengeksplorasi kelebihan dan kelemahan K-Nearest Neighbor dibandingkan dengan algoritma lainnya dalam hal akurasi prediksi, kompleksitas model, dan kemampuan mengatasi masalah dengan atribut yang hilang atau kontinu. Algoritma Machine learning bersifat konstruktif dalam membuat prediksi yang akurat dan memberikan analisis yang akurat.

Kesimpulan dari analisis sentimen ini akan memberikan bukti bahwa metode K-Nearest Neighbor efektif dalam mengklasifikasikan sentimen pada ulasan aplikasi chat generatif yang menggunakan model GPT. Metode KNN dapat mengelompokkan ulasan-ulasan ke dalam kategori sentimen yang sesuai, seperti positif, negatif, atau netral, berdasarkan kesamaan dengan teks-teks lain dalam dataset. Penggunaan metode KNN dalam analisis sentimen memberikan hasil yang dapat diandalkan dan akurat.

## 2. TINJAUAN PUSTAKA

### 2.1. Analisis Sentimen

Pemeriksaan perasaan, atau disebut juga dengan *assessment mining*, adalah bidang tinjauan komputasi yang berencana untuk menemukan, memahami, dan menawarkan sudut pandang, opini, penilaian, mentalitas, perasaan, subjektivitas, keputusan, atau perspektif yang terkandung dalam teks.

Tujuannya adalah untuk mendapatkan tingkat opini positif dan perasaan negatif yang terkait dengan individu, organisasi, yayasan, barang, atau kondisi tertentu. Hasil akhir dari investigasi perasaan dapat disusun ke dalam tiga klasifikasi utama, khususnya opini yang baik, opini pesimis, dan opini nonpartisan, atau dapat juga dipisahkan untuk mengenali sumber perasaan yang baik atau perasaan pesimis dari pertemuan atau pertemuan tertentu. Dalam penerapan analisis sentimen pada data teks, terdapat beberapa langkah yang perlu diikuti:

- Langkah pertama adalah mengumpulkan dataset yang berisi opini dan penilaian dari masyarakat terhadap restoran, produk, dan hal lainnya.
- Pre-processing: Tahap ini melibatkan proses pengolahan awal data teks, termasuk tokenisasi (mengubah teks menjadi token atau kata-kata), penghapusan stopwords (kata-kata umum yang tidak memberikan makna penting), dan stemming (mengubah kata-kata menjadi kata dasar).
- Transformation: Melibatkan pembobotan data teks, di mana bobot diberikan kepada kata-kata berdasarkan frekuensi kemunculannya atau metode lainnya.
- Feature Selection: Tahap ini mengurangi jumlah fitur atau yang tidak relevan atau tidak diperlukan pada analisis sentimen.

- Classification: Tahap klasifikasi merupakan proses yang melibatkan pengelompokan teks memakai berbagai metode seperti Naive Bayes, KNN, SVM, dan lainnya. Di antara metode yang sering dipakai untuk klasifikasi dalam analisis sentimen adalah metode K-Nearest Neighbor.
- Analisis dan Interpretasi: Akurasi dan nilai Area Under the Curve (AUC) dari model klasifikasi yang dibangun diukur selama fase evaluasi

### 2.2. K-NN

Dalam buku "Algoritma Data Mining" yang ditulis oleh Kusri, terdapat penjelasan mengenai Algoritma (K-NN) dalam konteks analisis sentimen. Data dapat direpresentasikan dalam ruang multi-dimensi dalam analisis sentimen, dengan setiap dimensi mencerminkan fitur data. Pentingnya memilih nilai  $k$  yang tepat dalam perhitungan K-NN terkait dengan dampaknya terhadap hasil karakterisasi. Oleh karena itu, pemilihan nilai  $k$  harus mempertimbangkan karakteristik data yang digunakan. Tujuan utama dari algoritma K-NN dalam analisis sentimen adalah melakukan klasifikasi objek berdasarkan atribut-atribut dan sampel pelatihan yang ada. Algoritma K-NN menggunakan pendekatan kedekatan titik untuk memperkirakan nilai dari instance query yang baru. Dalam metode K-NN, jarak kedekatan data dapat dihitung menggunakan metrik jarak seperti Euclidean atau Manhattan. Dengan mengikuti langkah-langkah dalam metode K-NN seperti yang dijelaskan di atas, analisis sentimen dapat dilakukan pada objek berdasarkan atribut-atributnya dan dapat memprediksi hasil klasifikasi dari instance query yang baru.

### 2.3. Text Mining

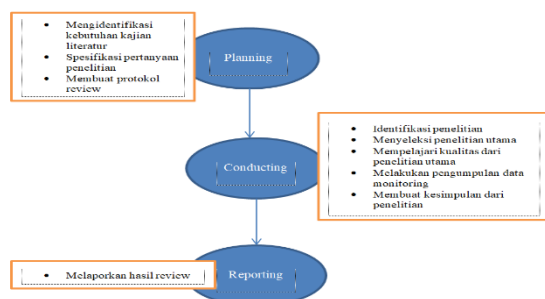
Text Mining adalah proses memilah-milah teks yang tidak terstruktur dan menemukan informasi penting. Tujuan dari text mining adalah untuk mengidentifikasi pola, tren, dan wawasan yang terkandung dalam teks yang sangat besar dan kompleks. Metode text mining melibatkan penggunaan teknik-teknik analisis linguistik, statistik, dan machine learning untuk memproses, mengorganisasi, dan menganalisis teks. Proses text mining meliputi beberapa tahapan, antara lain:

- Pre-processing: Tahap ini termasuk membersihkan dan menormalkan teks, misalnya, menghilangkan aksentuasi, menghilangkan kata-kata yang tidak perlu (stopwords), dan melakukan stemming atau lemmatization untuk mengubah kata-kata menjadi bentuk dasar.
- Tokenization: Memecah teks menjadi unit-unit terkecil yang disebut token, seperti kata-kata atau frasa.
- Pembentukan representasi teks: Mengubah teks menjadi representasi numerik yang dapat diproses oleh algoritma. Metode yang umum digunakan termasuk pembobotan term (seperti TF-IDF), pemodelan tema (seperti Latent Dirichlet

- Allocation), atau pemodelan jaringan saraf (seperti word embeddings).
- Analisis dan pemodelan: Menerapkan metode analisis seperti analisis frekuensi kata, analisis sentimen, klasifikasi teks, atau analisis topik untuk mendapatkan wawasan yang berguna dari teks.
  - Evaluasi dan interpretasi: Mengevaluasi hasil analisis dan menginterpretasikan temuan untuk mengambil keputusan atau membuat kesimpulan yang informatif. Text mining dapat digunakan dalam berbagai bidang, seperti analisis sentimen, analisis opini, penambangan data, pengenalan entitas, pemrosesan bahasa alami, dan banyak lagi.

### 3. METODE PENELITIAN

Metode Penelitian ini menggunakan metode tinjauan literatur sistematis. Tinjauan literatur sistematis atau yang dikenal luas sebagai Systematic Literature Review adalah tinjauan literatur yang mengikuti serangkaian aturan baku digunakan untuk mengidentifikasi dan mensintesis semua studi yang relevan serta memberikan penilaian terhadap apa yang diketahui dari topik metodologi penelitian yang digunakan pada penelitian ini menggunakan metode studi literatur yang didapat melalui informasi dari jurnal, buku, atau sumber lain yang telah dibuat sebelumnya. Adapun alur dari penelitian yang dilakukan dalam penelitian ini digambarkan pada gambar 1



Gambar 1. Alur metodologi Penelitian

#### 3.1. Planning

Dalam tahap ini penting untuk membuat Research Question. Tahapan ini sangat penting karena didasari dari ketertarikan atau pusat tujuan dari penulisan Systematic Literature Review. Planning merupakan tahap awal dalam melakukan tinjauan literatur secara sistematis. SLR adalah pendekatan sistematis untuk mengumpulkan, mengevaluasi, dan menyintesis studi literatur yang relevan dengan topik penelitian tertentu. Tahap perencanaan dalam SLR penting untuk mengarahkan langkah-langkah selanjutnya dan memastikan bahwa tinjauan literatur dilakukan secara terorganisir dan komprehensif. Perencanaan yang cermat dan terorganisir pada tahap awal SLR sangat penting untuk memastikan validitas dan kualitas tinjauan literatur yang dilakukan. Hal ini juga membantu menghindari bias dan meminimalkan

kesalahan dalam proses pencarian dan evaluasi literatur yang relevan.

#### 3.2. Research Question

Pertanyaan penelitian dibuat untuk menentukan metode yang akan digunakan dalam penelitian yang akan dilakukan oleh penulis nantinya.

Berikut beberapa pertanyaan penelitiannya:

Tabel 1. Research Question

| ID  | Pertanyaan Penelitian                                                              | Motivasi                                                                                              |
|-----|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
| RQ1 | Akurasi yang dihasilkan dari penelitian sebelumnya?                                | Mengidentifikasi pada akurasi yang dihasilkan dari penelitian sebelumnya                              |
| RQ2 | Metode apa saja yang sering digunakan untuk melakukan penerapan analisis sentimen? | Mengidentifikasi metode yang sering digunakan pada analisis sentimen                                  |
| RQ3 | Jumlah data yang dilibatkan dalam analisis sentimen                                | Mengidentifikasi sumber data yang digunakan untuk analisis sentimen dalam ulasan Aplikasi chatbot GPT |

#### 3.3. Conducting

Conducting sebagai Tahapan pelaksanaan SLR terdiri dari tahapan-tahapan yang sudah ditetapkan. Menetapkan “search string” merupakan hal utama serta mengacu kepada PICOC (Population, Intervention, Comparison, Outcomes, Context). Pada tahap pelaksanaan atau penerapan metode yang telah direncanakan sebelumnya untuk melakukan tinjauan literatur sistematis. Tahap ini melibatkan proses praktis dalam mengumpulkan, mengevaluasi, dan menganalisis studi literatur yang relevan.

Tabel 2. Conducting

|                 |                                                                                      |
|-----------------|--------------------------------------------------------------------------------------|
| Population(P)   | Apa populasi target yang akan digunakan dalam penelitian ini?                        |
| Intervention(I) | K-Nearest Neighbor yang digunakan untuk mengklasifikasikan apa dalam penelitian ini? |
| Comparison(C)   | Bagaimana membandingkan ulasan positif dan negatif?                                  |
| Outcomes(O)     | Apa hasil yang akan diukur atau dinilai dalam penelitian ini?                        |
| Context(C)      | Apakah ada konteks khusus yang perlu diperhatikan dalam penelitian ini?              |

#### 3.4. Reporting

Hasil dari tinjauan literatur secara sistematis ini berupa tulisan, baik untuk dipublikasikan dalam bentuk paper ke jurnal ilmiah atau untuk menyusun Literature Review dari skripsi/tesis/disertasi. Struktur penulisan dari tinjauan literatur secara sistematis biasanya terdiri dari 3 bagian besar, yaitu: Pendahuluan (Introduction), Utama

(Main Body) dan Kesimpulan (Conclusion). Penyusunan laporan akhir atau makalah ilmiah yang menggambarkan secara terperinci proses, temuan, dan kesimpulan dari tinjauan literatur yang telah dilakukan.

Tahap ini penting untuk mengkomunikasikan secara jelas metodologi yang digunakan, hasil tinjauan literatur, serta implikasi atau kontribusi penelitian terhadap pengetahuan yang ada. Tahap "reporting" pada SLR memastikan bahwa temuan dan kesimpulan dari tinjauan literatur disampaikan secara terperinci dan transparan kepada pembaca atau pihak yang berkepentingan.

Hal ini memungkinkan pengulangan dan peneliti lainnya untuk memahami dan memanfaatkan pengetahuan yang diperoleh dari tinjauan literatur yang telah dilakukan.

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

- Menyeleksi paper sesuai dengan kriteria yang telah ditentukan
- Melakukan pencarian paper pada berbagai basis data seperti, Google Scholar, Crossref
- Melakukan analisis pada bukti empiris dari paper yang terpilih. Dalam proses pencarian sumber penelitian didapatkan sebanyak 10 paper yang sesuai dan dengan kriteria.

#### 4. HASIL DAN PEMBAHASAN

Tabel 4. Penelitian Sebelumnya

| No | Judul                                                                                                        | Tahun | Hasil Penelitian                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Metode             | QA1 | QA2 | QA3 |
|----|--------------------------------------------------------------------------------------------------------------|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|-----|-----|-----|
| 1. | "SENTIMENT ANALYSIS AGAINST THE DANA E-WALLET ON GOOGLE PLAY REVIEWS USING THE K-NEAREST NEIGHBOR ALGORITHM" | 2021  | "Algoritma k-nearest neighbor mampu secara efektif mengklasifikasikan ulasan aplikasi DANA e-Wallet berdasarkan sentimen positif dan negatif, karena Kelas pertama mencakup semua penilaian bintang (1-5), kelas kedua mencakup penilaian bintang 1 dan 5, dan kelas ketiga mencakup penilaian bintang 1 dan 2 sebagai label negatif, bintang 3 sebagai label netral, serta bintang 4 dan 5 sebagai label positif. pengklasifikasian ulasan aplikasi DANA e-Wallet dengan kelas kedua (penilaian bintang 1 dan 5) dan menggunakan nilai $k=1$ menghasilkan akurasi sebesar 86,64%". [1] | K-Nearest Neighbor | Y   | Y   | Y   |
| 2. | "ANALISIS KLASIFIKASI SENTIMEN ULASAN WORD CLOUD DENGAN METODE NAIVE BAYES DAN K-NEAREST NEIGHBOR"           | 2022  | "Penelitian dilakukan untuk menganalisis platform Shopee menggunakan metode 500 ulasan dari Shopee sebagai dataset. Metode word cloud juga digunakan untuk mengidentifikasi hal yang sering dibahas oleh user. Sementara itu, metode KNN memiliki akurasi sebesar 0,928, presisi 0,929, recall 0,928, dan F1 score 0,926 Dalam analisis word cloud." [2]                                                                                                                                                                                                                                | K-Nearest Neighbor | Y   | Y   | Y   |
| 3. | "Aplikasi Algoritma K-Nearest Neighbor pada Analisis Sentimen Omicron Covid-19"                              | 2022  | "Nilai persentase prediksi euclidean distance menggunakan $k = 15$ terhadap analisis sentimen data tweet dengan hashtag #OmicronVariant dan #Covid19 yaitu diperoleh persentase kategori sentimen positif menggunakan cosine similarity distance sebesar 65.04% artinya tweet tersebut tersebut mempercayai adanya omicron, sedangkan persentase kategori sentimen negatif menggunakan cosine similarity distance sebesar 17.67% artinya                                                                                                                                                | K-Nearest Neighbor | Y   | Y   | Y   |

#### 3.5. Seleksi Studi

Setelah tahapan pencarian sumber penelitian dengan strategi yang telah ditentukan pada penelitian, tahap selanjutnya sejumlah 10 paper diseleksi kembali. Tahapan seleksi studi dilakukan dengan membaca dan menganalisis bagian abstrak dari setiap sumber yang berkaitan dengan topik 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.

#### 3.6. Quality Assessment

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

Tabel 3. Quality Assessment

|     |                                                                                                                   |
|-----|-------------------------------------------------------------------------------------------------------------------|
| 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? |

| No | Judul                                                                                                                       | Tahun | Hasil Penelitian                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Metode                | QA1 | QA2 | QA3 |
|----|-----------------------------------------------------------------------------------------------------------------------------|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|-----|-----|-----|
|    |                                                                                                                             |       | tweet tersebut kontra atau tidak mempercayai;adanya omicron dan persentase kategori sentimen netral menggunakan cosine similarity distance sebesar 17.29% artinya tweet tersebut hanya sebuah berita atau tidak pro dan kontra terhadap omicron.”[3]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                       |     |     |     |
| 4. | “PENERAPAN ANALISIS SENTIMEN PADA PENGGUNA TWITTER MENGGUNAKAN METODE K-NEAREST NEIGHBOR”                                   | 2018  | “Akurasi tertinggi yang dicapai; yaitu sebesar 67,2% saat menggunakan parameter k=5. Hal ini menunjukkan bahwa algoritma K-NearestNeighbor tingkat keberhasilan sistem dalam mengklasifikasikan sentimen tweet dengan benar. bahwa algoritma K-Nearest Neighbor”[4]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | K-Nearest Neighbor    | Y   | Y   | Y   |
| 5. | “ANALISIS SENTIMEN APLIKASI KAI ACCESS METODE SUPPORT VECTOR MACHINE”                                                       | 2023  | “Hasil klasifikasi menggunakan SVM menunjukkan tingkat akurasi, presisi, recall, dan F-measure yang baik untuk setiap aspek yang diteliti. Dengan menggunakan teknik SMOTE, penelitian ini dapat mengatasi masalah imbalance class dalam dataset. penerapan model dengan menggunakan algoritma Support Vector Mahine (SVM) terhadap dataset yang sudah didapatkan sebanyak 1.261 ulasan dari Oktober 2021 hingga April 2022. Penggunaan Support Vector Mahine (SVM) dapat dijadikan opsi dalam melakukan analisis sentimen level aspek untuk penilaian kepuasan pengguna aplikasi KAI Access. Terbukti berdasarkan pengujian dan nilai dari akurasi yang dihasilkan dari aspek Learnability sebesar 94,73%, aspek Efficiency sebesar 94,38%, aspek Errors sebesar 85,13%, dan aspek Satisfaction sebesar 87,26%. Secara umum sentimen positif pada aplikasi KAI Access berkaitan dengan bisa membeli tiket tanpa perlu pergi ke stasiun. Sentimen negatif terhadap aplikasi KAI Access terkait adanya error pada aplikasi setelah update dan kesulitan mendapatkan tiket kereta saat Lebaran.”[5] | Support Vector Mahine | Y   | Y   | Y   |
| 6. | “Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP” | 2022  | “Dari hasil penelitian, 77,67% data uji berhasil diklasifikasikan dengan benar ke dalam kelas ulasan positif. Hasil precision (ketepatan) dan recall (ketepatan pemulihan) untuk kelas ulasan positif adalah 82,67% dan 86,92%. b. Tingkat akurasi klasifikasi sebesar 76,68%. c. Ulasan pengguna terhadap aplikasi Flip menunjukkan hasil positif, sejalan dengan rating yang diterima di Google Play Store. d. Algoritma K-Nearest Neighbor dengan pembobotan TF-IDF berhasil digunakan dalam proses analisis sentimen pengguna terhadap aplikasi Flip.”[6]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | K-Nearest Neighbor    | Y   | Y   | Y   |
| 7. | “Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor”                       | 2023  | “Hasil dari penelitian ini menunjukkan bahwa tingkat akurasi model pelatihan sebesar 69,5%, recall sebesar 69,5%, dan presisi sebesar 68,7%. Hal ini mengindikasikan bahwa model dapat mengklasifikasikan sentimen pengguna Twitter terhadap perpanjangan PPKM dengan tingkat keberhasilan yang moderat. Dalam analisis sentimen, ditemukan bahwa sebagian besar tanggapan masyarakat terhadap perpanjangan PPKM cenderung negatif, namun juga terdapat sebagian yang mendukung kebijakan tersebut dengan alasan-alasan masing-masing. Hasil ini dapat memberikan perspektif yang berharga dalam memahami respons masyarakat terhadap kebijakan PPKM dan dapat menjadi landasan dalam                                                                                                                                                                                                                                                                                                                                                                                                             | K-Nearest Neighbor    | Y   | Y   | Y   |

| No  | Judul                                                                                                                                        | Tahun | Hasil Penelitian                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Metode                             | QA1 | QA2 | QA3 |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|-----|-----|-----|
|     |                                                                                                                                              |       | pengambilan keputusan kebijakan yang lebih baik di masa depan.”[7]                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                    |     |     |     |
| 8.  | “Analisis Sentimen Tokopedia Pada Ulasan di Google Playstore Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor on Twitter” | 2022  | “Hasil penelitian menunjukkan bahwa akurasi algoritma Naïve Bayes adalah 75,30%, sedangkan akurasi algoritma K-Nearest Neighbor adalah 86,09%. Hal ini menunjukkan bahwa algoritma K-Nearest Neighbor memberikan hasil analisis sentimen yang lebih akurat dibandingkan dengan algoritma Naïve Bayes. Selain itu, hasil penelitian juga memberikan insight tentang bagaimana pengguna menilai performa aplikasi Tokopedia di Google Playstore.”[8]                                                   | K-Nearest Neighbor dan Naïve Bayes | Y   | Y   | Y   |
| 9.  | “Aplikasi Myim3 Sentimen Ulasan Pada Situs Google Play Menggunakan SVM”                                                                      | 2022  | “Hasil penelitian menunjukkan bahwa pada skenario (70:30) untuk kernel linear dengan akurasi 87% dan skenario (90:10) dengan akurasi 87% di kernel RBF adalah yang terbaik. Proses klasifikasi juga menghasilkan kata-kata yang paling sering muncul pada masing-masing kelas sentimen, yang divisualisasikan dengan word cloud. Kata "bagus" merupakan kata yang paling dominan pada data ulasan positif, sedangkan kata "jaringan" yang paling dominan pada data ulasan negatif aplikasi MyIM3.[9] | K-Nearest Neighbor                 | Y   | Y   | Y   |
| 10. | “Sentiment Analysis SEA Games in Philippines 2019 Based on Opinion of Internet User of Social Media Twitter with KNN and SVM Twitter”        | 2023  | “Berdasarkan penelitian ini, hasil yang diperoleh menunjukkan bahwa klasifikasi support vector machine dengan kernel polinomial memiliki akurasi terbaik sebesar 92,96%. Hasil ini menunjukkan bahwa prediksi dari klasifikasi Support Vector Machine cenderung negatif terhadap SEA Games 2019 di Filipina.”[10]                                                                                                                                                                                    | K-Nearest Neighbor                 | Y   | Y   | Y   |

#### 4.1. Pembahasan Hasil

Tahap pembahasan hasil menjawab pertanyaan yang telah ditentukan pada Research Question.

##### RQ1 Akurasi yang dihasilkan dari penelitian sebelumnya?

Dari tabel 4. Review jurnal yang diperoleh dari beberapa metode yang dipakai mengenai Analisis Sentimen dari hasil riset 10 jurnal, terdapat beberapa metode dengan akurasi berbeda.

Tabel 5. Hasil akurasi

| Jurnal ke- | Algoritma 1                        | Akurasi           |
|------------|------------------------------------|-------------------|
| 1          | K-Nearest Neighbor                 | 86,64%            |
| 2          | K-Nearest Neighbor                 | 91,40%            |
| 3          | K-Nearest Neighbor                 | -                 |
| 4          | K-Nearest Neighbor                 | 67,2%             |
| 5          | Support Vector Mahine              | 94,73%            |
| 6          | K-Nearest Neighbor                 | 69,5%,            |
| 7          | K-Nearest Neighbor                 | 96,67%            |
| 8          | K-Nearest Neighbor dan Naïve Bayes | 86,09% dan 75,30% |
| 9          | K-Nearest Neighbor                 | 87%               |
| 10         | K-Nearest Neighbor                 | 92,96%            |

##### RQ 2 Metode apa aja yang sering digunakan untuk melakukan penerapan analisis sentimen?

Dari tabel 1. Review jurnal yang saya dapatkan mengusulkan beberapa metode yang dipakai mengenai Analisis Sentimen dari hasil riset 10 jurnal, terdapat beberapa metode yang diusulkan oleh peneliti sebelumnya diantara lain :

|                       |          |
|-----------------------|----------|
| K-Nearest Neighbor    | 9 Jurnal |
| Support Vector Mahine | 1 Jurnal |
| Naïve Bayes           | 1 Jurnal |

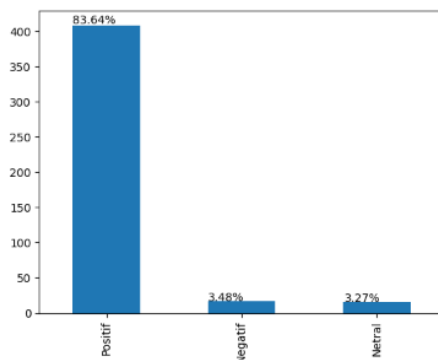
##### RQ 3 Jumlah data yang dilibatkan dalam analisis sentimen

Pada jurnal penelitian sebelumnya memiliki data untuk melakukan analisis sentimen yaitu ada pada tabel dibawah:

Tabel 6. Hasil jumlah data

| Jurnal    | Jumlah Data |
|-----------|-------------|
| Jurnal 1  | 500         |
| Jurnal 2  | 500         |
| Jurnal 3  | 12.951      |
| Jurnal 4  | 4.799       |
| Jurnal 5  | 2.152       |
| Jurnal 6  | 11.002      |
| Jurnal 7  | 2000        |
| Jurnal 8  | 992         |
| Jurnal 9  | 3721        |
| Jurnal 10 | 300         |

#### 4.2. Hasil analisis sentimen aplikasi Chatbot GPT



Gambar 2. hasil sentiment



Gambar 3. hasil evaluasi

Nilai K terbaik: 9 Akurasi: 0.9662921348314607 Presisi: 0.48314606741573035 Recall: 0.5 F1 Score: 0.4914285714285714 Dalam hasil yang Anda berikan, akurasi model yang dihasilkan adalah sekitar 0.966 atau sekitar 96.6%. Ini menunjukkan bahwa model k-NN yang Anda gunakan cukup baik dalam memprediksi sentimen pada data testing. Perlu diperhatikan bahwa akurasi bisa berbeda-beda tergantung pada dataset yang digunakan dan pengaturan spesifik lainnya. Namun, dengan akurasi sebesar 96.6%, dapat dikatakan bahwa model k-NN yang Anda terapkan memberikan hasil yang cukup baik dalam memprediksi sentimen.

## 5. KESIMPULAN

Berdasarkan penelitian terdahulu hasil penelitian metode yang sering digunakan untuk melakukan Analisis sentiment berdasarkan tabel 2 yaitu menggunakan K-Nearest Neighbor, Support Vector Mahine, Naïve Bayes, pada metode KNN memiliki akurasi yang tinggi dalam analisis sentimen.

Analisis sentimen ulasan pada aplikasi Chat Generative Pre-Trained Transformer GPT menggunakan metode klasifikasi K-Nearest Neighbor (KNN), diperoleh hasil sebagai berikut: Nilai K terbaik Dalam penelitian ini, setelah melakukan evaluasi, ditemukan bahwa nilai K terbaik adalah 9. Hal ini berarti dalam proses klasifikasi sentimen, menggunakan 9 tetangga terdekat memberikan hasil

yang optimal. Akurasi: Model yang dihasilkan mencapai akurasi sekitar 96.6%, yang menunjukkan bahwa metode klasifikasi KNN berhasil dalam memprediksi sentimen pada data ulasan aplikasi.

## DAFTAR PUSTAKA

- [1] S. Masturoh, ); Achmad, B. Pohan, and C. Author, "SENTIMENT ANALYSIS AGAINST THE DANA E-WALLET ON GOOGLE PLAY REVIEWS USING THE K-NEAREST NEIGHBOR ALGORITHM." [Online]. Available: [www.bsi.ac.id](http://www.bsi.ac.id).
- [2] S. Masturoh, ); Achmad, B. Pohan, and C. Author, "SENTIMENT ANALYSIS AGAINST THE DANA E-WALLET ON GOOGLE PLAY REVIEWS USING THE K-NEAREST NEIGHBOR ALGORITHM." [Online]. Available: [www.bsi.ac.id](http://www.bsi.ac.id)
- [3] J. A. Josen Limbong, I. Sembiring, K. Dwi Hartomo, U. Kristen Satya Wacana, and P. Korespondensi, "ANALISIS KLASIFIKASI SENTIMEN ULASAN PADA E-COMMERCE SHOPEE BERBASIS WORD CLOUD DENGAN METODE NAIVE BAYES DAN K-NEAREST NEIGHBOR ANALYSIS OF REVIEW SENTIMENT CLASSIFICATION ON E-COMMERCE SHOPEE WORD CLOUD BASED WITH NAIVE BAYES AND K-NEAREST NEIGHBOR METHODS", doi: 10.25126/jtiik.202294960.
- [4] A. Firdaus, "Aplikasi Algoritma K-Nearest Neighbor pada Analisis Sentimen Omicron Covid-19," *Jurnal Riset Statistika*, pp. 85–92, Dec. 2022, doi: 10.29313/jrs.v2i2.1148.
- [5] A. Deviyanto, M. R. Didik Wahyudi, and T. Informatika UIN Sunan Kalijaga Yogyakarta Jl Marsda Adi Sucipto No, "PENERAPAN ANALISIS SENTIMEN PADA PENGGUNA TWITTER MENGGUNAKAN METODE K-NEAREST NEIGHBOR," *Jurnal Informatika Sunan Kalijaga*, vol. 3, no. 1, pp. 1–13, 2018, [Online]. Available: <https://twitter.com/search?l=id&q=AHY%20sin ce%3A2017-01-01%20until%3A2017-01->
- [6] G. Radiena and A. Nugroho, "ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI KAI ACCESS MENGGUNAKAN METODE SUPPORT VECTOR MACHINE," 2023.
- [7] S. Rahayu, Y. MZ, J. E. Bororing, and R. Hadiyat, "Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP," *Edumatic: Jurnal Pendidikan Informatika*, vol. 6, no. 1, pp. 98–106, Jun. 2022, doi: 10.29408/edumatic.v6i1.5433.
- [8] A. A. P. T. Mara, E. Sediyo, and H. Purnomo, "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Metode Pembelajaran

- Dalam Jaringan (DARING) Di Universitas Kristen Wira Wacana Sumba.”
- [9] A. Fitriyani and A. Triayudi, “SENTIMENT ANALYSIS OF REKSADANA ON BIBIT APPLICATIONS USING THE NAÏVE BAYES METHOD AND K-NEAREST NEIGHBOR (KNN)id 2) (\*) Corresponding Author.”
- [10] A. P. Pandian, R. Palanisamy, and K. Ntalianis, “Advances in Intelligent Systems and Computing 1272 Proceedings of International Conference on Intelligent Computing, Information and Control Systems.” [Online]. Available: <http://www.springer.com/series/11156>
- [11] M. Riefky and W. Pramesti, “Sentiment Analysis of Southeast Asian Games (SEA Games) in Philippines 2019 Based on Opinion of Internet User of Social Media Twitter with K-Nearest Neighbor and Support Vector Machine Analisis Sentimen terhadap Southeast Asian Games (SEA Games) di Filipina Tahun 2019 berdasarkan Opini Netizen dari Media Sosial Twitter dengan Metode K-nearest Neighbor dan Support Vector Machine,” vol. 17, no. 1, pp. 26–41, 2020, doi: 10.20956/jmsk.v%vi%i.9947.