**UNIVERSITAS GUNADARMA**

**FAKULTAS TEKNOLOGI INDUSTRI**



**ANALISIS SENTIMEN PADA ULASAN VISION+ DI GOOGLE PLAY STORE MENGGUNAKAN METODE ALGORITMA BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT)**

**Disusun Oleh:**

|  |  |  |
| --- | --- | --- |
| Nama | : | Didin Nur Yahya |
| NPM | : | 51419788 |
| Jurusan | : | Informatika |
| Pembimbing | : | Dr. Dea Adlina, ST, MMSI. MSc. |

**Diajukan Guna Melengkapi Sebagian Syarat**

**Dalam Mencapai Gelar Sarjana Strata Satu (S1)**

**JAKARTA**

**2023**

**PERNYATAAN ORIGINALITAS DAN PUBLIKASI**

Saya yang bertanda tangan dibawah ini,

|  |  |  |
| --- | --- | --- |
| Nama | : | Didin Nu Yahya |
| NPM | : | 51419788 |
| Judul Skripsi | : Analisis Sentimen Pada Ulasan Vision+ di | |
|  |  | Google Play store Menggunakan Metode |
|  |  | Algoritma Bidirectional Encoder Representations from  Transformers (BERT) |
| Tanggal Sidang | : |  |
| Lulus Sidang | : |  |

Menyatakan bahwa tulisan ini adalah merupakan hasil karya saya sendiri dan dapat dipublikasikan sepenuhnya oleh Universitas Gunadarma. Segala kutipan dalam bentuk apapun telah mengikuti kaidah, etika yang berlaku. Mengenai isi dan tulisan adalah merupakan tanggung jawab Penulis, bukan Universitas Gunadarma.

Demikian pernyataan ini dibuat dengan sebenarnya dan penuh kesadaran

Depok, 30 Agustus 2023

(Didin Nur Yahya)

**LEMBAR PENGESAHAN**

**KOMISI PEMBIMBING**

|  |  |  |
| --- | --- | --- |
| **NO** | **NAMA** | **KEDUDUKAN** |
| 1. | Dr. Dea Adlina.ST, MMSI, Msc. | Ketua |
| 2. |  |  |
| 3. |  |  |

Tanggal Sidang:

**PANITIA UJIAN**

|  |  |  |
| --- | --- | --- |
| **NO** | **NAMA** | **KEDUDUKAN** |
| 1. | Dr. Ravi Ahmad Salim | Ketua |
| 2. | Prof. Dr. Wahyudi Priyono |  |
| 3. |  |  |

Tanggal Lulus:

Jakarta, ………… 2023

Mengetahui,

            Pembimbing     Kepala Bagian Sidang Ujian

(Dr. Dea Adlina.ST, MMSI, Msc.)                           (Didin Nur Yahya)

**ABSTRAK**

Didin Nur Yahya. 51419788.

**ANALISIS SENTIMENT PADA ULASAN VISION+ DI GOOGLE PLAY STORE MENGGUNAKAN METODE ALGORITMA BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT)**

Skripsi. Informatika. Fakultas Teknologi Industri.

Universitas Gunadarma. 2022.

Kata Kunci: Analisis Sentimen, BERT,Vision+*.*

(xii + 60 + Lampiran)

Penulisan ini bertujuan untuk analisis sentiment bahasa Indonesia terhadap ulasan pengguna aplikasi Vision+ di Google Play Store. Ulasan aplikasi yang bervariasi dari ulasan positif, negatif maupun netral dapat digunakan untuk studi sosial sebagai alat bantu dalam pengambilan keputusan. Hasil ulasan *rating* biasanya digunakan oleh product manager untuk mengontrol apakah fitur yang di aplikasi berjalan atau tidak bisa dilihat dari komentar dari ulasan Vision+ yang ada di Google Play Store.Namun pada implementasinya *review* aplikasi terkadang ulasan tidak sesuai atau relevan dengan rating yang diberikan, maka dibutuhkan suatu otomatisasi dalam mengklasifikasikan ulasan berdasarkan polaritas sentimennya. Analisis sentimen dengan penggunaan algoritma machine learning tradisional seperti Naive Bayes, SVM, dan lain-lain tidak dapat memahami konteks dari komentar secara mendalam tentang semantik kata yang ada karena hanya mempelajari pola-pola yang diberikan seperti frekuensi kemunculan kata.Untuk itu dibutuhkan sebuah pendekatan deep learning seperti BERT (*Bidirectional Encoder Representations from Transformers*) yang menghasilkan sebuah model bahasa dua arah (*bidirectional*). Dataset yang digunakan melalui tahap pre-processing yang terdiri dari *case folding, data cleaning*, *tokenisasi*, dan normalisasi dengan library NLTK sebelum dilakukan analisis sentimen. Dalam penelitian ini *hyperparameters* yang digunakan adalah 7 epoch, *learning rate* 3e-6, dan batch size 32. Pengujian analisis sentimen menggunakan *pre-trained* mode/ IndoBERT dengan akurasi sebesar 82%.

Daftar Pustaka (1997 - 2022)

**ABSTRACT**

Didin Nur Yahya. 51419788.

**SENTIMENT ANALYSIS ON VISION+ REVIEWS ON GOOGLE PLAY STORE USING THE BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT) ALGORITHM**

Undergraduate Thesis. Informatics. Faculty of Industrial Technology. Gunadarma University. 2022. Keywords: BERT, Vision+, Sentiment Analysis.

(xii + 48 + Appendix)

This study aims to conduct sentiment analysis on user reviews of the Vision+ application in the Indonesian language on the Google Play Store. The reviews vary from positive, negative, to neutral, making them suitable for social studies and decision-making support tools. Product managers often utilize review ratings to assess the performance of application features based on user comments. However, in practice, reviews may not always align with the given ratings, necessitating an automated approach to classify reviews based on their sentiment polarity. Traditional machine learning algorithms like Naive Bayes, SVM, and others struggle to comprehend the contextual nuances and semantic meanings of words within comments, as they only learn patterns based on word frequencies.For this purpose, a deep learning approach using BERT (Bidirectional Encoder Representations from Transformers) is employed. BERT is a bidirectional language model that can capture intricate contextual relationships within text data. The dataset is pre-processed through case folding, data cleaning, tokenization, and normalization using the NLTK library before performing sentiment analysis. The study utilizes hyperparameters including 7 epochs, a learning rate of 3e-6, and a batch size of 32. The sentiment analysis is conducted using a pre-trained IndoBERT model, achieving an accuracy of 82%.

Bibliography (1997-2022)

**KATA PENGANTAR**

Puji syukur Penulis panjatkan kepada Allah SWT karena atas rahmat dan karunia-Nya Penulis dapat menyelesaikan Tugas Akhir tentang “ANALISIS SENTIMENT PADA ULASAN VISION+ DI GOOGLE PLAY STORE MENGGUNAKAN METODE ALGORITMA BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT)”.

Tugas Akhir ini disusun guna melengkapi syarat dalam mencapai gelar sarjana strata satu pada jurusan Informatika, Fakultas Teknologi Industri, Universitas Gunadarma. Dalam penyusunan Tugas Akhir ini perkenankanlah Penulis untuk mengucapkan terima kasih yang sebesar-besarnya atas bimbingan serta motivasi yang telah diterima sehingga dapat menyelesaikan Tugas Akhir ini, kepada:

* 1. Prof. Dr. E.S. Margianti, SE, MM, selaku Rektor Universitas Gunadarma.
  2. Prof. Dr.-Ing. Adang Suhendra S.Si., S.Kom., M.Sc, selaku Dekan Fakultas Teknologi Industri Universitas Gunadarma.
  3. Dr. Lintang Yuniar B., S.Kom., M.Sc, selaku Ketua Jurusan Teknik Informatika
  4. Dr.Edi Sukirman, SSi., MM, M.Ikom., selaku Kepala Bagian Sidang Ujian.
  5. Dr. Dea Adlina, S.T., MMSI, Msc, selaku Dosen Pembimbing Penulisan Skripsi yang telah memberikan pengarahan serta saran, bantuan dan bimbinganya sehingga Penulisan Ilmiah ini dapat terselesaikan dengan baik
  6. Ayah Trio Purnomo serta Ibu Mariyati, yang selalu memberikan dorongan semangat serta do’a dari awal tulisan hingga akhir penyusunan tulisan ilmiah ini.
  7. Seluruh teman, dan segenap pihak-pihak yang tidak dapat penulis sebutkan satu persatu tanpa mengurangi rasa hormat, karena telah memberikan rasa semangat, bantuan tenaga maupun pikiran serta dukungan dalam penyelesaian penulisan ini.
  8. Seluruh teman-teman kelas 4IA07 angkatan 2019 yang tidak mungkin disebutkan satu persatu, yang telah banyak memberikan dorongan, semangat dan bantuan.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kata sempurna. Penulis mengharapkan kritik dan saran yang membangun untuk dijadikan perbaikan sehingga dapat bermanfaat bagi dunia pendidikan dan teknologi informasi.

Depok,

Didin Nur Yahya

**Daftar Isi**

Halaman

[TULISAN ILMIAH i](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682583)

[PERNYATAAN ORIGINALITAS DAN PUBLIKASI ii](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682584)

[LEMBAR PENGESAHAN iii](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682585)

[ABSTRAK iv](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682586)

[ABSTRACT iv](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682586)

[KATA PENGANTAR v](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682587)

[DAFTAR ISI vii](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682588)

[DAFTAR GAMBAR x](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682589)i

[DAFTAR TABEL xiv](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682590)

[DAFTAR LAMPIRAN xv](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682591)

[1. PENDAHULUAN 1](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682592)

[1.1. Latar Belakang 1](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[1.2. Rumusan Masalah 3](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682594)

[1.3. Tujuan Penulisan 3](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682595)

[1.4. Manfaat Penelitian 3](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682596)

[1.5. Batasan Penelitian 4](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682597)

[1.6. Metodologi Penelitian 4](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682597)

[1.7. Sistematika Penelitian Ilmiah 4](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682597)

[2. TINJAUAN PUSTAKA 6](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682598)

2.1. Crawling Data 6

2.2. Aplikasi Streaming 6

2.3. Text Mining 6

2.4. Vision+ 6

2.5. Analisis Sentimen

2.6. Neural Network 6

2.7. Artificial Neural Network 6

2.8. Recurrent Neural Network 6

2.9. Machine Learning 6

2.10. Deep Learning 6

2.11. Bidirectional Encoder Representations from Transformers (BERT) 6

2.12. IndoNLU 6

2.13. IndoBERT-BASE 6

2.14. *Confusion Matrix* 6

2.15. Dataset 6

2.16. *Web Scrapping* 6

2.17. *Case Folding* 6

2.18. Data *Cleaning* 6

2.19. Normalisasi Teks 6

2.20. *Dataset Splitting* 6

2.21. Python 6

2.22. Google Colaboratory 6

2.23. Penelitian Terkait 6

3.Metodologi Penelitian 24

[3.1. Tahapan Penelitian 24](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[3.2. Data *Labelling* 24](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[3.3. Scrapping 24](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[3.4. Preprocessing 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[3.4.1. Case Folding 25](#_Toc109842189)

[3.4.2. Data Cleaning 27](#_Toc109842189)

[3.4.3. Tokenisasi 27](#_Toc109842189)

[3.4.4. Stemming 27](#_Toc109842189)

[3.4.5. Stopwords Removal 27](#_Toc109842189)

[3.4.6. Normalisasi Kata Tidak Baku 27](#_Toc109842189)

[3.5. Database Splitting 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[3.6. Implementasi BERT 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

3.7. Memuat Model Pre-trained 25

[3.8. Training Data(Fine Tuning) 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[3.9. Evaluasi 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

4.HASIL DAN PEMBAHASAN 24

[4.1. Evaluasi Hasil 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[4.2. Uji Akurasi 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[4.3. Visualisasi Frekuensi Data 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[4.4. Visualisasi *Word Cloud* 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

[4.5. Ketidakaukuratan Hasil *Testing* 25](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682593)

5. Penutup 156

[5.1 Kesimpulan 157](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682661)

[5.2. Saran 147](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682662)

DAFTAR PUSTAKA 148

[LAMPIRAN L-1](file:///C:\Users\didin\Downloads\Abdan%20Ahkam%20-10117008%20(1)%20-%20Copy%20-%20Copy%20-%20Copy%20(1).docx#_Toc47682664)

**Daftar Gambar**

[Gambar 2.1 *Recurrent Neural Network* 9](file:///D:\PI\PI%20Arya%20Javas%20Fatih.docx#_Toc109851285)

[Gambar 2.2 Arsitektur Model BERT 15](file:///D:\PI\PI%20Arya%20Javas%20Fatih.docx#_Toc109851286)

[Gambar 3.1 Tahapan Penelitian 26](file:///D:\PI\PI%20Arya%20Javas%20Fatih.docx#_Toc109851292)

[Gambar 3.2 Proses *Pre-processing* 27](file:///D:\PI\PI%20Arya%20Javas%20Fatih.docx#_Toc109851293)

[Gambar 3.3 Ilustrasi *Layer* untuk Analisis Sentimen 28](file:///C:\Users\didin\Downloads\PI%20Arya%20Javas%20Fatih%20(1).docx#_Toc109851294)

[Gambar 4.1 Proses Uji Akurasi 30](file:///C:\Users\didin\Downloads\PI%20Arya%20Javas%20Fatih%20(1).docx#_Toc109851295)

[Gambar 4.2 *Learning Curve* 31](file:///C:\Users\didin\Downloads\PI%20Arya%20Javas%20Fatih%20(1).docx#_Toc109851296)

[Gambar 4.3 *Confusion Matrix* 32](file:///C:\Users\didin\Downloads\PI%20Arya%20Javas%20Fatih%20(1).docx#_Toc109851297)

[Gambar 4.4 Hasil Evaluasi 54](file:///C:\Users\didin\Downloads\PI%20Arya%20Javas%20Fatih%20(1).docx#_Toc109851298)

[Gambar 4.5 Visualisasi Frekuensi Data 55](file:///C:\Users\didin\Downloads\PI%20Arya%20Javas%20Fatih%20(1).docx#_Toc109851299)

[Gambar 4.6 *Word Cloud* 56](file:///C:\Users\didin\Downloads\PI%20Arya%20Javas%20Fatih%20(1).docx#_Toc109851300)

**Daftar Tabel**

[Tabel 2.1 *Confusion Matrix* 33](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842295)

[Tabel 2.2 Penelitian Terkait 34](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842296)

[Tabel 3.1 Contoh Hasil *case folding* 34](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842297)

[Tabel 3.2 Contoh Hasil *Data Cleaning* 35](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842298)

[Tabel 3.3 Contoh Hasil *Tokenisasi* 36](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842299)

[Tabel 3.4 Daftar *Stopwords* 37](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842300)

[Tabel 3.5 Contoh Hasil Normalisasi 38](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842300)

[Tabel 3.6 Contoh Hasil Tahapan *Tokenizer* 39](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842297)

[Tabel 3.7 Tahap Evaluasi 40](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842298)

[Tabel 4.1 *True Positive* 40](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842299)

[Tabel 4.2 *True Neutral* 41](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842300)

[Tabel 4.3 *True Negative* 43](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842300)

[Tabel 4.4 *False Positive True* 44](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842297)

[Tabel 4.5 *False Positive False* 45](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842298)

[Tabel 4.6 False Neutral True 47](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842299)

[Tabel 4.7 False Neutral False 49](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842300)

[Tabel 4.8 False Negative True 50](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842300)

[Tabel 4.9 False Negative False 52](file:///C:\Users\didin\AppData\Local\Temp\Temp2_Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai%20(2).zip\Penelitian%20Ilmiah%20Final-Revisi-Dospeng-Selesai\Bab_1_Bab_2_Bab3_Lampiran.docx#_Toc109842300)

**Daftar Lampiran**

Evaluasi Hasil Uji Training Data…………..………………………………………………………..L-1

**BAB 1: PENDAHULUAN**

* 1. **Latar Belakang**

(Perbaikan latar belakang) (dan tulisan bhs inggris tanda miring)

(seaborn harus dijelaskan di bab2 beserta mathlolib)

Dalam era digital yang semakin berkembang pesat, penggunaan teknologi untuk memenuhi kebutuhan hiburan semakin meningkat. Salah satu bentuk hiburan yang sedang populer saat ini adalah platform streaming, yang memungkinkan pengguna untuk menonton konten video secara *online*. Beberapa *platform streaming* yang terkenal di Indonesia antara lain Netflix, iFlix, Vidio, dan Vision+.

Salah satu aspek penting dalam menilai kualitas pengalaman pengguna adalah analisis sentimen. Analisis sentimen merupakan teknik untuk mengekstrak dan memahami perasaan dan opini dari teks yang dituliskan oleh pengguna di media sosial, forum, atau *platform online* lainnya. Analisis sentimen dapat membantu platform streaming untuk memahami kebutuhan dan preferensi pengguna, sehingga dapat meningkatkan kualitas konten dan pengalaman pengguna.

Analisis sentimen adalah proses untuk menentukan sentimen atau pendapat yang terkandung dalam sebuah teks, apakah itu positif, negatif, atau netral. Dalam konteks ulasan Vision+ di Google Play Store, analisis sentimen digunakan untuk mengevaluasi pendapat pengguna terhadap aplikasi tersebut (Agustin & Risa Yuli, 2022).

Dalam konteks Vision+, analisis sentimen berbasis dapat membantu platform untuk memahami bagaimana pengguna merespons konten yang disajikan. Analisis sentimen berbasis aspek merupakan teknik analisis sentimen yang memeriksa perasaan dan opini pengguna terhadap aspek tertentu dari konten dan juga *user experience* (UX) dari pengguna platform streaming. Misalnya, dalam konteks film, aspek yang diperiksa dapat meliputi akting, alur cerita, efek visual, kemudahan dalam navigasi aplikasi, kestabilan dari aplikasi dan sebagainya .

Oleh karena itu, penelitian ini akan fokus pada analisis sentimen berbasis aspek terhadap konten dan eksperien penggunaan dari aplikasi *platform streaming* Vision+. Tujuan dari penelitian adalah untuk menganalisis sentimen ulasan Vision+ di Google Play Store menggunakan BERT. Dengan melakukan analisis sentimen ini, diharapkan dapat diperoleh wawasan tentang kepuasan pengguna, masalah yang mungkin ada, dan area yang perlu diperbaiki dalam aplikasi Vision+. Selain itu, skripsi ini juga bertujuan untuk mengevaluasi efektivitas BERT dalam melakukan analisis sentimen pada ulasan Vision+ di Google Play Store (Kaira & Fitria Milani, 2022).

* 1. **Rumusan Masalah**

Berdasarkan tujuan penelitian yang telah disebutkan sebelumnya, maka rumusan masalah yang akan dijawab dalam penelitian ini adalah sebagai berikut:

1. Bagaimana melakukan analisis sentimen terhadap konten Vision+ menggunakan *Bidirectional Encoder Representations from Transformers* (BERT)?
2. Bagaimana respon pengguna terhadap aspek-aspek tertentu dari konten Vision+ dalam analisis sentimen berbasis aspek?
3. Apa saja faktor-faktor yang mempengaruhi respon pengguna terhadap konten Vision+ dalam analisis sentimen berbasis aspek?
4. Bagaimana penerapan analisis sentimen berbasis aspek dapat membantu meningkatkan kualitas konten dan pengalaman pengguna di Vision+?

Dengan menjawab rumusan masalah di atas, diharapkan penelitian ini dapat memberikan wawasan dan pemahaman yang lebih mendalam tentang analisis sentimen berbasis aspek dalam konteks *platform streaming*, khususnya Vision+.

* 1. **Tujuan Penulisan**

Tujuan dari penelitian analisis sentimen adalah sebagai berikut:

1. Untuk melakukan analisis sentimen terhadap konten Vision+ dengan menggunakan metode algoritma *Bidirectional Encoder Representations from Transformers* (BERT).
2. Untuk memahami respon pengguna terhadap respon positive dan negative dari konten Vision+.
3. Untuk mengukur tingkat akurasi metode algoritma Bidirectional Encoder Representations from Transformers (BERT).
4. Untuk mengevaluasi penerapan analisis sentimen dalam meningkatkan kualitas konten dan pengalaman pengguna di Vision+.

Dengan tujuan tersebut, penelitian ini diharapkan dapat memberikan kontribusi dalam meningkatkan pemahaman tentang penggunaan teknik analisis sentimen berbasis aspek untuk meningkatkan kualitas konten dan pengalaman pengguna di platform streaming. Selain itu, penelitian ini juga diharapkan dapat memberikan rekomendasi dan masukan bagi *platform* Vision+ dalam mengoptimalkan konten dan pengalaman pengguna.

* 1. **Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat akademis, yaitu sebagai referensi atau sumber penelitian bagi peneliti lain yang tertarik dalam bidang analisis sentimen, khususnya dalam konteks platform streaming.
2. Manfaat praktis, yaitu sebagai masukan atau rekomendasi bagi *platform* Vision+ dalam meningkatkan kualitas konten dan pengalaman pengguna.
3. Manfaat bagi pengguna platform Vision+, yaitu dalam meningkatkan kualitas layanan dan pengalaman pengguna dalam menikmati konten di Vision+.
4. Manfaat bagi industri *platform streaming*, yaitu sebagai wawasan dalam menerapkan teknik analisis sentimen berbasis aspek untuk meningkatkan kualitas layanan dan pengalaman pengguna di platform streaming.
5. Manfaat dari tim internal Vision+ yaitu bisa dijadikan masukan kepada tim product dan engineering untuk melakukan improvement dari aplikasi vision+ agar lebih berkembang dan fitur aplikasi lebih relevan dengan kebutuhan user saat ini
   1. **Batasan Penelitian**

Pada subbab ini peneliti menyatakan Batasan serta ruang lingkup yang menjadi titik fokus dari penelitian tugas akhir yang dilakukan peneliti. Yaitu diantaranya;

1. Mengklasifikasikan ulasan positif. netral dan negatif.
2. Penelitian ini menggunakan data ulasan dari google play yang menggunakan bahasa Indonesia.
3. Data yang diambil menggunakan ulasan dari pengguna aplikasi Vision+ yang didapat dengan menggunakan package *library* google play scraper pada Python.
4. Penelitian ini menggunakan data yang diambil dari ulasan aplikasi Vision+ dengan Teknik *scraping* mencapai 8000 ulasan.
   1. **Metodologi Penelitian**

Pada subbab ini peneliti menggunakan python sebagai alat utama dalam mengelola analisis sentiment berikut ini adalah tahapanya:

1. Koleksi Data

Data yang digunakan dalam penelitian ini akan dikumpulkan dari review google play store dari platform streaming Vision+ menggunakan teknik web scraping. Python memiliki beberapa *library web scraping* yang dapat digunakan, seperti BeautifulSoup , Scrapy dan Google play *scraper*.Untuk penelitian ini penulis menggunakan Google play scraper untuk memperoleh data mentah.

1. Pra-Pemrosesan Data

Data yang telah dikumpulkan akan melalui proses pra-pemrosesan, seperti menghilangkan duplikasi data, membersihkan data dari karakter-karakter yang tidak diperlukan, serta memfilter data yang tidak relevan dengan penelitian.

1. Analisis Sentimen

Untuk melakukan analisis sentimen, penelitian ini akan menggunakan teknik Natural Language Processing (NLP) dengan Python. Terdapat beberapa *library* NLP yang dapat digunakan, seperti NLTK dan SpaCy. Teknik ini akan digunakan untuk memisahkan kalimat-kalimat dalam data ke dalam token-token dan kemudian mengklasifikasikan setiap token ke dalam aspek yang sesuai.

1. Pembangunan Model Analisis Sentimen

Model analisis sentimen akan dibangun dengan menggunakan algoritma *Machine Learning*. Python memiliki banyak *library Machine Learning* yang dapat digunakan, seperti Scikit-Learn dan TensorFlow. Model ini akan dilatih dengan menggunakan data yang telah dikumpulkan dan telah melalui proses pra-pemrosesan, serta dilakukan validasi dengan menggunakan teknik *k-fold cross-validation*.

1. Evaluasi Model dan Visualisasi Hasil

Hasil dari analisis sentimen berbasis aspek akan dievaluasi dengan menggunakan beberapa metrik, seperti akurasi, presisi, *recall*, dan *F1-score*. Selain itu, hasil dari analisis akan divisualisasikan menggunakan library plotting Python, seperti Matplotlib dan Seaborn, untuk mempermudah pemahaman hasil oleh pembaca.

* 1. **Sistematika Penulisan(Perbaikan Sistematika**

Sistematika penulisan dijelaskan secara garis besar mengenai apa yang terdapat

pada ball-bab yang ada. Adapun sistematika penulisan terdiri dari

1. BAB 1 PENDAHULUAN

Pada bab ini berisi uraian latar belakang masalah, rumusan masalah, Batasan masalah, tujuan penelitian, metode dan sistematika penulisan.

1. BAB 2 TINJAUAN PUSTAKA

Bab ini berisi tentang teori-teori yang menjadi dasar untuk melakukan penelitian dan penelitian-penelitian terdahulu yang berhubungan dengan permasalahan yang diteliti dan menjadi acuan.

1. BAB 3 METODE PENELITIAN

Bab ini berisi tentang tahapan proses dari klasifikasi dimulai dari Analisa kebutuhan data, pengambilan data, Data sampling, Pelabelan data dari ulasan aplikasi Vision+ pada Google Play Store, Melakukan *Pre-Processing, Database Splitting* sampe melakukan implementasi BERT dengan melakukan training data dan pada akhirnya evaluasi hasil training data

1. BAB 4 IMPLEMENTASI

Bab ini berisi tentang hasil pre-processing data, pelatihan data (*training*), dan pengujian data (*testing*) dalam analisis sentimen menggunakan metode *Bidirectional Encoder Representations From Transformers*

BAB 5 PENUTUP

Bab ini berisi kesimpulan dari seluruh hasil penelitian yang merupakan jawaban dari tujuan dan saran untuk pengembangan yang lebih baik di masa yang akan datang.

‘

**BAB 2**

**Tinjauan Pustaka**

* 1. **Aplikasi *Streaming***

Aplikasi streaming adalah program atau aplikasi yang memungkinkan pengguna untuk menonton atau mendengarkan konten multimedia seperti film, acara TV, musik, dan podcast secara online tanpa harus mengunduhnya terlebih dahulu. Aplikasi streaming ini menggunakan teknologi streaming untuk memutar konten multimedia secara langsung dan memungkinkan pengguna untuk menikmati konten ini hampir secara real time. Beberapa platform streaming yang terkenal di Indonesia antara lain Netflix, iFlix, Vidio, dan Vision+ (Rizki. A, 2020).

Aplikasi streaming memungkinkan pengguna untuk mengakses konten media dari berbagai perangkat seperti smartphone, tablet, laptop atau komputer, sehingga memudahkan Oleh karena itu, dapat disimpulkan bahwa aplikasi streaming adalah aplikasi yang menggunakan teknologi streaming untuk memudahkan pengguna menonton atau mendengarkan konten multimedia secara online melalui Internet. Aplikasi ini sangat populer dan banyak digunakan oleh masyarakat di seluruh dunia karena konten multimedianya mudah digunakan di berbagai perangkat dan fleksibel.Dari beberapa platform streaming di Indonesia salah satunya adalah Vision+.

* 1. **Vision+**

Vision+ adalah platform layanan streaming video on-demand (VOD) yang dikembangkan oleh MNC Group, salah satu perusahaan media terkemuka di Indonesia. Aplikasi ini menyediakan akses ke berbagai konten video seperti film, serial TV, dokumenter, dan program hiburan dari MNC Group dan mitra bisnisnya.

Dalam aplikasi Vision+, pengguna dapat menonton konten video kapan saja dan di mana saja dengan menggunakan perangkat mobile atau tablet yang terhubung ke internet. Aplikasi ini juga menyediakan fitur-fitur seperti fitur pencarian, kategori, rekomendasi, dan nonton offline (*download*). Aplikasi Vision+ tersedia untuk diunduh secara gratis di Google Play Store dan App Store. Namun, untuk menikmati layanan *streaming* video, pengguna harus berlangganan layanan berbayar yang disediakan oleh Vision+. Pada penelitian ini penulis ingin melalukan analisis sentimen pada ulasan Vision+ pada Google Play Store.

* 1. **Data Sampling**

Data sampling adalah teknik pengambilan sebagian data dari populasi untuk dijadikan sampel dalam penelitian. Teknik sampling digunakan untuk mengurangi biaya dan waktu dalam pengumpulan data, serta memudahkan analisis data. Terdapat beberapa jenis teknik *sampling*, antara lain:

1. *Probability Sampling*

Probability sampling adalah teknik pengambilan sampel yang memungkinkan setiap anggota populasi memiliki kesempatan yang sama untuk dipilih sebagai sampel. Beberapa jenis teknik probability sampling antara lain:

1. Simple random sampling adalah pengambilan sampel secara acak dari populasi.
2. Stratified random sampling: pengambilan sampel secara acak dari setiap strata dalam populasi.
3. Cluster sampling adalah pengambilan sampel dari kelompok-kelompok dalam populasi.
4. *Non-probability sampling*

Non-probability sampling adalah teknik pengambilan sampel yang tidak memungkinkan setiap anggota populasi memiliki kesempatan yang sama untuk dipilih sebagai sampel. Beberapa jenis teknik non-probability sampling antara lain:

1. *Convenience sampling*: pengambilan sampel yang paling mudah diakses.
2. *Quota sampling*: pengambilan sampel dengan memenuhi kuota tertentu untuk setiap kelompok dalam populasi.
3. *Purposive sampling*: pengambilan sampel dengan tujuan tertentu.

Pemilihan teknik *sampling* yang tepat sangat penting dalam penelitian karena dapat mempengaruhi hasil penelitian. Oleh karena itu, peneliti harus mempertimbangkan karakteristik populasi dan tujuan penelitian dalam memilih teknik sampling yang sesuai (Sugiyono, 2016).

* 1. **Analisis Sentimen**

Analisis sentimen adalah suatu proses untuk mengelola, mengekstrak, dan memahami data teks yang tidak terstruktur. Tujuan dari proses ini adalah untuk mengambil informasi sentimen yang terkandung dalam kalimat pendapat atau opini tentang suatu topik. Metode pemrosesan ini dapat diterapkan pada semua bidang opini. Ada bidang hukum, sosial, ekonomi, dan politik di antaranya.

Analisis sentimen mempelajari tentang pendapat, perilaku, dan emosi seseorang terhadap suatu entitas yg bisa mewakili suatu individu, acara atau suatu topik eksklusif. Analisis sentimen mencakup deteksi, analisis, dan penilaian keadaan pikiran manusia terhadap aneka macam insiden, dilema, layanan atau minat lainnya. Lebih tepatnya, bidang ini bertujuan buat menggali pendapat, sentimen serta emosi berdasarkan pengamatan orang-orang yg mampu didapatkan melalui tulisan, aktualisasi diri paras, ucapan, musik, gerakan, dan lain sebagainya.

Tujuan awal dari analisis sentimen itu sendiri adalah untuk mengetahui sifat data respon dengan mengidentifikasi sentimen yang mereka ungkapkan dan kemudian mengklasifikasikan polaritasnya. Beberapa pendapat mengenai analisis sentimen bisa diambil kesimpulan bahwa analisis sentimen ialah suatu proses untuk memilih sentimen atau opini berasal seorang yg umumnya di wujudkan dalam bentuk teks dan mampu dikategorikan sebagai sentimen positif atau negatif (Yoel Julianto, 2014).Pada penelitian ini penulis menggunakan machine learning untuk melalukan proses analisis sentimennya.

* 1. **Machine Learning**

Machine learning adalah teknologi yang memungkinkan komputer untuk belajar dari data tanpa harus diprogram secara eksplisit. Pembelajaran mesin dikembangkan berdasarkan disiplin ilmu lainnya seperti statistika, matematika, dan data mining sehingga mesin dapat belajar dengan menganalisa data tanpa perlu di program ulang atau diperintah. Cara kerja machine learning meliputi pengumpulan data, eksplorasi data, pemilihan model atau teknik, memberikan pelatihan terhadap model yang dipilih, dan mengevaluasi hasil dari ML(Sari. R.P & Wibowo. A, 2019) .

Ada tiga jenis utama dari machine learning:

1. *Supervised learning*, di mana model belajar dari data yang diberi label dan melakukan prediksi untuk data baru.
2. *Unsupervised learning*, di mana model belajar dari data tanpa label dan mencoba untuk mengidentifikasi pola atau klasifikasi pada data tersebut.
3. *Reinforcement learning*, di mana model belajar melalui interaksi dengan lingkungannya dan mendapatkan reward atau hukuman berdasarkan tindakan yang diambil.

ML telah digunakan secara luas dalam berbagai aplikasi, termasuk pengenalan wajah, klasifikasi teks, deteksi spam, dan pengenalan suara. Peningkatan teknologi komputer dan perkembangan algoritma ML, seperti deep learning, telah memungkinkan mesin untuk menyelesaikan tugas yang lebih kompleks dan lebih presisi.

* 1. ***Deep Learning***

Deep learning merupakan cabang dari machine learning yang masih menjadi bagian dari kecerdasan buatan. Deep learning merupakan *neural network* yang terdiri dari bebeberapa network dan bersifat mendalam karena memiliki lebih banyak layer dibandingkan dengan dengan *neural network* biasanya (Santoso & Ariyanto, 2018). Deep learning bekerja untuk sehingga tidak hanya dapat memprediksi tetapi juga merepresentasikan data berdasarkan data yang diproses dengan benar, sehingga cocok untuk melakukan prediksi. Deep learning dapat dibagi kedalam tiga metode pendekatan yaitu *supervised, semi-supervised dan unsupervised learning*. Deep learning didukung Oleh banyak framework seperti TensorFlow, Torch dan Theano.Untuk membangun *deep learning* membutuhkan dataset yang didapat dari 3 metode pengambilan data yaitu menggunakan Teknik *Scraping* data, *Crawling* data dan Import dataset dari sumber lainnya seperti repository instansi dll. Penulis menggunakan metode *scraping* data untuk mengambil data ulasan dari Google Play Store.

* 1. ***Scraping* data**

Scraping data adalah proses pengambilan atau pengumpulan data dari internet, termasuk dari website, dokumen, atau sumber data lainnya. Proses scraping data dapat dilakukan secara manual atau menggunakan perangkat lunak khusus yang disebut web scraper. Data yang diambil dapat berupa teks, gambar, video, dan lainnya(Sari, R. P, 2019).

Scraping data dapat digunakan untuk berbagai keperluan, seperti riset, analisis, dan pengembangan bisnis.Terdapat beberapa teknik scraping data yang dapat digunakan, seperti copy paste, HTML parsing, DOM parsing, vertical aggregation, web scraping, xpath, dan text pattern machine. Namun, perlu diperhatikan bahwa proses scraping data harus dilakukan dengan hati-hati dan memperhatikan aturan privasi dan hak cipta yang berlaku. Dari beberapa Teknik scraping data salah satunya adalah web scraping yaitu pengambilan data ulasan dari Google Play Store.

* 1. ***Web Scraping***

*Web scraping* adalah teknik pengumpulan data melalui internet yang sering digunakan dalam kegiatan penambangan data. Menurut teori, *web scraping* merupakan metode pengumpulan data yang berbeda dengan penggunaan API (*Application Programming Interfaces*). Proses ini dimulai dengan penulisan kode program yang digunakan untuk melakukan permintaan data secara otomatis kepada server. Data yang diperoleh dapat diekstrak untuk menghasilkan informasi yang dibutuhkan.Dari proses web scraping menghasilkan sebuah dataset yang digunakan untuk diproses tahap selanjutnya(Sari, R. P, 2019).

* 1. **Dataset**

Dataset adalah sekumpulan data yang terdiri dari himpunan data yang berasal dari informasi sebelumnya yang dapat diubah menjadi informasi baru melalui metode pembelajaran yang diawasi.Di mana sistem prediksi menggunakannya sebagai referensi untuk mendukung keputusan.Pada penelitian ini Dataset didapat dari data ulasan dari Google Play Store dan menggunakan *library* Google Play Scraper dan disorting jumlah data yang diambil adalah 8000 data dan komentar yang diambil berbahasa indonesia.Dari dataset tersebut akan diolah pada tahap preprocessing dengan melalukan data cleaning dan text mining

* 1. ***Text Mining***

*Text mining* adalah proses penggalian informasi dalam bentuk teks, dimana sumber informasi umumnya berasal dari sebuah dokumen, dan tujuannya adalah menemukan kata-kata yang dapat mewakili isi dokumen tersebut untuk membuat analisis terkait. antara dokumen-dokumen tersebut.  
  
*Text mining* adalah teknik penambangan yang dilakukan oleh komputer yang bertujuan untuk mengumpulkan informasi atau hal-hal baru yang sebelumnya tidak diketahui serta mengambil kembali informasi yang tersebar secara otomatis yang diambil dari berbagai sumber data teks. Text mining juga dapat digunakan untuk klasifikasi. Metode ini menemukan pola yang menarik dalam kumpulan data tekstual. Penambangan data dan teks mirip. Perbedaannya terletak pada fakta bahwa metode ini dimaksudkan untuk menangani data terstruktur atau semi-terstruktur.

* 1. **Neural Network**

Neural network atau jaringan syaraf tiruan adalah salah satu teknologi mekanik.Pembelajaran populer dengan mensimulasikan mekanisme “pembelajaran”. Terinspirasi dari cara kerja sistem saraf manusia dan organisme hidup lainnya. Sistem saraf terdiri dari sel-sel yang disebut neuron. Neuron ini terhubung satu sama lain oleh akson dan dendrit. Sinapsis adalah hubungan antara akson dan dendrit.

*Neural Network* disebut memiliki arsitektur yang terdiri berbagai lapisan yang saling berhubungan terhubung satu sama lain. Lapisan antara *input layer* dan *output layer* disebut *hidden layer*, dan keluaran dari hidden layer disebut *hidden* unit. Istilah "*hidden*" digunakan karena unit-unit ini tidak langsung terlihat oleh dunia luar sebagai *input* dan *output*. Inti dari Neural Network adalah lapisan tersembunyi yang terdiri dari unit-unit tersembunyi. Setiap unit tersembunyi adalah unit saraf yang mengambil jumlah input berbobot dan menerapkan lapisan nonlinier.

Setiap unit pada tiap *layer* mengambil *input* dan *output* dari semua unit *layer*

dan hubungan antara setiap unit dari dua *layer* yang saling berdekatan sehingga semua *layer* saling terhubung. Tiap *hidden* unit akan menjumlahkan semua *input* unit yang ada.*Neural network* memiliki beberapa parameter, seperti *weights* W, bias b yang dipelajari Oleh gradient descent. Sedangkan *hyperparameters* adalah parameter yang dipilih Oleh desainer algoritma nilai optimal yang diatur pada sebuah serber ,bukan dengan sebuah *gradient descent* di *dataset training*. Hyperparameters mencakup *learning rate n, mini-batch size*, arsitektur model (jumlah layer, jumlah *hidden node* tiap layer, fungsi aktivasi yang dipilih) dan lain-lain.

*Gradient descent* (penurunan gradien) adalah sebuah algoritma optimasi yang digunakan ketika melatih sebuah model machine learning. Gradient descent hanya digunakan untuk menemukan nilai parameter fungsi (koefisien) yang meminimalkan biaya. Tujuan utama dari *gradient descent* adalah menemukan weight yang optimal dengan meminimalisasi *loss function*. Seberapa besar step yang dilakukan gradient descent ditentukan Oleh *learning rate* yang menunjukkan seberapa cepat atau lambat bergeraknya fungsi ke *weight* yang optimal. Terdapat tiga jenis *gradient descent* yang sering digunakan, yaitu *Batch Gradient Descent, Stochastic Gradient Descent*, dan *Mini -Batch Gradient Descent*.

1. *Batch gradient descent (BGD)* atau *vanilla gradient descent* menghitung *error* dari tiap contoh yang ada pada dataset training. proses ini seperti sebuah siklus yang disebut juga *training epoch*. Semua *cost* dari *training* pada daraset dihitung. Algoritma ini memiliki kelebihan yaitu dapat membuat jaringan syaraf tiruan lebih cepat, menghasilkan gradien kesalahan yang stabil, Akan tetapi gradien kesalahan yang stabil dapat pula membuat model tidak dapat menghasilkan kondisi konvergensi yang terbaik\_ Selain itu,seluruh training dataset harus berada dalam memori.
2. *Stochastic gradient descent (SGD)* adalah algoritma yang meminimalkan loss function dengan menghitung gradiennya setiap kali melakukan proses training. Algoritma ini disebut stokastik karena hanya memiliki Satu contoh acak dalam satu waktu, memindahkan weight-nyasehingga dapat meningkatkan performa dari contoh tersebut. Sehingga, SGD akan melakukan proses tersebut pada setiap contoh dataset Satu per Satu. *Weight* yang terus diperbarui memungkinkan untuk mendapatkan peningkatan yang cukup detail. Akan tetapi pembahaman yang terlalu sering dilakukan lebih mahal dan dapat menyebabkan *gradient* yang tidak rata karena *error rate* dapat naik turun.
3. *Mini batch gradient descent* menggabungkan konsep dari SGD dan BGD. Mini batch adalah pruses training sekumpulan dataset (umumnya 512 atau 1024) yang lebih sedikit dari *dataset* asli. Algoritma ini akan membagi *training dataset* menjadi batch-batch yang kecil dan memperbarui weight dari masing-masing batch.

Salah Satu algoritma optimasi yang sering digunakan adalah ADAM. ADAM atau *Adaptive Moment Estiimation Algorithm* (Kingma & Ba, 2015) menghitung estimasi momen dan menggunakannya untuk mengoptimmsi fungsi. Adam merupakan kombinasi dari dua algoritma yaitu *AdaGrad* dan *RMSProp*. AdaGrad akan mempertahankan learning rare per parameter yang meningkatkan peforma pada masalah dengan gradien yang renggan dan RMSProp juga mempertahankan learning rare per parameter yang diadaptasi berdasarkan rata-rata besaran gradien untuk *weight*

(seberapa cepat berubah). Algoritma ini akan menghitung rata-rata *eksperimensial weighted* bergerak kemudian mengkuadratkan *gradien* yang dihitung. Alih-alih mengadaptasi learning rate parameter berdasarkan rata-rata momen pertama seperti di RMS Prop, ADAM juga menggunakan rata-rata momen kedua. Dengan menggunakan model dan *dataset* yang besar, ADAM dapat secara efektif menyelesaikan masalah-masalah *deep learning* praktikal. Hal ini membuat Ada *optimizer* menjadi metode yang efisien secara komputasi, membutuhkan sedikit memori, tidak berbeda dengan gradien dengan skala diagonal, cocok untuk masalah

yang menggunakan data atau parameter yang besar.

* 1. ***Artificial Neural Network*(ANN)**

*Artificial neural network* (ANN) atau jaringan saraf tiruan adalah model matematis yang terinspirasi dari sistem saraf biologis. ANN terdiri dari jaringan neuron buatan (artificial neuron) yang terhubung secara paralel, dan mampu melakukan pemrosesan informasi melalui propagasi sinyal yang disebut "*forward pass*" dan "*backward pass*".

Setiap neuron dalam ANN menerima masukan (*input*) dari neuron lain atau dari luar, dan kemudian menghasilkan keluaran (*output*) berdasarkan suatu fungsi aktivasi yang dapat berbeda-beda, seperti fungsi *sigmoid* atau fungsi *ReLU*. Hubungan antara neuron-neuron dalam ANN dapat direpresentasikan oleh bobot (*weight*), yaitu nilai numerik yang menentukan seberapa besar pengaruh suatu neuron terhadap neuron lainnya.

Proses pembelajaran (learning) ANN terjadi melalui pengaturan bobot dengan tujuan untuk meminimalkan suatu fungsi *error*. Terdapat berbagai macam algoritma pembelajaran ANN, seperti *backpropagation, gradient descent, dan stochastic gradient descent*, yang digunakan untuk menentukan perubahan bobot yang optimal.

ANN digunakan dalam berbagai aplikasi seperti klasifikasi, prediksi, pengenalan pola, dan optimasi. Contoh penggunaan ANN antara lain dalam pengenalan wajah, deteksi objek, dan prediksi harga saham (Wasito Budi, 2013).

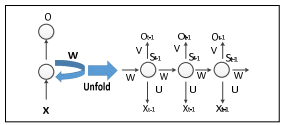
* 1. ***Recurrent Neural Network***

*Recurrent neural network* (RNN) atau jaringan saraf rekuren adalah model ANN yang dirancang untuk mengolah data sekuensial, seperti teks atau waktu. RNN memungkinkan informasi dari masa lalu untuk mempengaruhi pengambilan keputusan di masa sekarang melalui penggunaan sel memori (*memory cell*) dan *loop feedback* (kembali).

RNN terdiri dari serangkaian sel memori (*memory cell*) yang terhubung ke neuron-neuron lainnya. Setiap sel memori menghasilkan output berdasarkan input sekarang dan output sebelumnya. Output sebelumnya disimpan dalam sel memori, dan digunakan untuk mempengaruhi pengambilan keputusan selanjutnya.

Salah satu jenis RNN yang populer adalah LSTM (*Long Short-Term Memory*) yang memungkinkan sel memori untuk menyimpan informasi dalam jangka panjang dan jangka pendek. Hal ini memungkinkan LSTM untuk lebih efektif mengolah data sekuensial, terutama pada kasus di mana ada jarak waktu yang cukup jauh antara suatu kejadian dan kejadian berikutnya.

Penggunaan RNN terutama berfokus pada pengolahan data sekuensial seperti teks, waktu, dan suara. Beberapa aplikasi dari RNN antara lain terjemahan mesin, pengenalan suara, dan analisis sentimen.Berikut ini adalah gambaran dari Recurrent Neural Network (RNN).



Gambar 2.1 Recurrent Neural Network

Sumber : M Abdul Dwiyanto Suyudi, Esmeralda C. Djamal, Asri Maspupah (2021).

Dapat dilihat di gambar 2.1 bahwa looping dari RNN sebenarnya akan memproses input dari skala waktu 1 sampai t. RNN akan memproses data input satu per satu secara sekuensial dari huruf “h” sampai “l”, hidden layer pun akan melempar data menuju ke hidden layer pada skala waktu selanjutnya. Begitu seterusnya secara sekuensial.

Sifat dasar yang sekuensial ini menunjukkan bahwa RNN memang memiliki arsitektur yang didedikasikan untuk data berbentuk sequence dan list. Di beberapa tahun terakhir, RNN sudah cukup terbukti dalam menyelesaikan permasalahan seperti speech recognition, machine translation, sentiment analysis, image captioning dan masih banyak lagi.

* 1. ***Tokenizing***

*Tokenizing* adalah proses mengubah teks menjadi token atau kata-kata yang dapat diproses oleh komputer. *Tokenizing* adalah salah satu tahap awal dalam pemrosesan bahasa alami dan digunakan dalam berbagai aplikasi seperti mesin pencari, analisis sentimen, dan pemrosesan teks lainnya. *Tokenizing* dapat dilakukan dengan menggunakan algoritma tertentu seperti algoritma pemisahan kata, algoritma pemisahan frasa, atau algoritma pemisahan kata berdasarkan konteks (Salamah. N, 2019).

Terdapat beberapa jenis tokenizer yang dapat digunakan dalam pemrosesan bahasa alami, antara lain:

1. Tokenizing berdasarkan aturan (rule-based tokenizer): tokenizer ini menggunakan aturan tertentu untuk memisahkan kata-kata dalam teks. Aturan ini dapat berupa aturan pemisahan kata, aturan pemisahan frasa, atau aturan pemisahan kata berdasarkan konteks.
2. Tokenizing berbasis mesin (machine learning-based tokenizer): tokenizer ini menggunakan model pembelajaran mesin untuk memisahkan kata-kata dalam teks. Model ini dilatih dengan menggunakan data teks yang telah diberi label untuk memisahkan kata-kata dalam teks.
3. Tokenizing berbasis kamus (dictionary-based tokenizer): tokenizer ini menggunakan kamus kata-kata yang telah ditentukan sebelumnya untuk memisahkan kata-kata dalam teks. Kamus ini dapat berupa kamus kata-kata umum atau kamus kata-kata khusus.
   1. ***Case Folding***

Case folding adalah bentuk *preprocessing* yang sederhana namun efektif, meskipun sering diabaikan. Tujuan utamanya adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya karakter huruf dari 'a' hingga 'z' yang akan diterima, sedangkan karakter selain huruf akan dihilangkan dan dianggap sebagai pemisah. Terdapat beberapa pendekatan yang dapat digunakan dalam tahap case folding, dan metode yang dipilih tergantung pada tugas yang diberikan.

* 1. ***Data Cleansing***

*Data cleansing* atau pembersihan data adalah proses menghilangkan *noise* dan data yang tidak konsisten dalam sebuah dataset. Data yang tidak akurat, duplikat, tidak lengkap, tidak tepat, atau tidak relevan akan dihapus atau dimodifikasi agar dataset menjadi lebih akurat dan dapat diandalkan. Data cleansing merupakan salah satu langkah penting dalam pengambilan keputusan berbasis data yang berkualitas.

Beberapa hal yang perlu diperhatikan dalam melakukan *data cleansing* adalah peng-eksplorasian tentang isi data lebih jauh dan penyesuaian dengan kondisi atau kebutuhan real dari data.Meskipun tidak ada satu cara pasti untuk melakukan *data cleansing* karena prosesnya akan bervariasi tergantung pada dataset yang digunakan, namun terdapat beberapa langkah umum dalam proses *data cleansing*, seperti mengidentifikasi kesalahan data, mengubah, memperbarui, atau menghapus data yang salah, dan memperbaiki kesalahan seperti kesalahan penulisan dan penggabungan data dari beberapa sumber.

Data cleansing merupakan bagian penting dari proses manajemen data dan salah satu komponen inti dari persiapan data yang siap digunakan dalam aplikasi bisnis dan ilmu data. *Data cleansing* dapat dilakukan oleh analis dan insinyur kualitas data atau profesional manajemen data lainnya (Craig Stedman, 2022).

* 1. ***Stopwords Removal***

*Stopwords Removal* adalah komponen penting dalam pemrosesan bahasa alami untuk menghilangkan kata-kata umum yang tidak memiliki makna khusus dalam konteks tertentu. Dalam bahasa Indonesia, terdapat beberapa daftar stopwords yang tersedia bebas, namun belum ada kesepakatan mengenai daftar *stopwords* yang paling sesuai untuk berbagai tugas pemrosesan bahasa alami. Oleh karena itu, beberapa penelitian telah dilakukan untuk mengevaluasi pengaruh penghapusan *stopwords* pada berbagai tugas pemrosesan bahasa alami dalam bahasa Indonesia.

Beberapa referensi jurnal yang membahas penghapusan *stopwords* dalam bahasa Indonesia antara lain:

1. "Pengaruh Stopword Removal dan Stemming Terhadap Performa Klasifikasi Teks Komentar Kebijakan New Normal Menggunakan Algoritma LSTM" oleh Muhammad Nurfajri Rahman(Pradana, A. W., & Hayaty, R. 2019)
2. "Evaluasi Daftar Stopword Bahasa Indonesia" oleh Dian Purnama Sari dan Dwi Teguh Wijaya(Sari, D. P., & Wijaya, D. T. 2019)
3. "Pembentukan Daftar Stopword Menggunakan Term Based Random Sampling Pada Analisis Sentimen Dengan Metode *Naïve Bayes* (Studi Kasus: Kuliah Daring Di Masa Pandemi)" oleh Ahmad Agil Fauzi dan Agus Wibowo(Fauzi, A. A., & Wibowo, A. 2021)
4. "Analisis Performa Penggunaan *Stopwords* dan *Stemming* dalam Sentimen Analisis dengan Pendekatan Klasifikasi Naive Bayes" oleh Aditya Wiha Pradana dan Rizki Hayaty(Pradana, A. W., & Hayaty, R. 2019).
   1. ***Stemming***

*Stemming* adalah proses untuk memetakan berbagai variasi morfologikal dari kata menjadi bentuk dasar yang sama dalam bahasa Indonesia.Proses ini dilakukan untuk mencari kata dasar dari sebuah kata dan merupakan salah satu tahapan dalam *pre-processing. Stemming* memiliki pengaruh dalam meningkatkan performa *Information Retrieval* (IR) dengan mentransformasi kata-kata dalam sebuah dokumen teks ke kata dasarnya (Dony Yudha, 2014).

* 1. **Normalisasi Teks**

Normalisasi teks adalah proses pengolahan teks yang bertujuan untuk mengubah teks menjadi bentuk standar atau normal. Tujuan dari normalisasi teks adalah untuk menghilangkan variasi dan perbedaan dalam penulisan teks yang dapat mempengaruhi analisis atau pemrosesan lebih lanjut.Normalisasi teks dapat melibatkan penghilangan tanda baca, konversi huruf menjadi huruf kecil, penggantian kata-kata slang dengan kata-kata standar, dan lain sebagainya.Proses normalisasi teks memiliki peran penting dalam berbagai aplikasi seperti analisis sentimen, pemrosesan bahasa alami, dan penambangan teks. Dengan melakukan normalisasi teks, dapat meningkatkan kualitas dan akurasi hasil analisis serta mempermudah pemrosesan lebih lanjut.

Beberapa metode yang digunakan dalam normalisasi teks antara lain berbasis kamus, berbasis aturan, dan berbasis pembelajaran mesin. Metode berbasis kamus menggunakan kamus kata-kata yang telah ditentukan untuk menggantikan kata-kata slang atau tidak standar dengan kata-kata standar.Metode berbasis aturan menggunakan aturan-aturan yang telah ditentukan untuk mengubah teks menjadi bentuk standar. Sedangkan metode berbasis pembelajaran mesin menggunakan algoritma pembelajaran mesin untuk mempelajari pola-pola dalam teks dan melakukan normalisasi berdasarkan pola-pola tersebut (Prasetija, Z.R.N.S., Romadhony, A., & Setiawan, E.B, 2022).

* 1. ***Dataset Splitting***

*Dataset splitting* adalah suatu teknik yang digunakan untuk mengevaluasi performa model. Dalam metode ini, dataset dibagi menjadi beberapa bagian dengan proporsi tertentu, yaitu bagian untuk *training* data, *validation* data, dan *testing* data. Dataset training digunakan untuk melatih model, sementara dataset validasi digunakan untuk mengurangi overfitting yang sering terjadi pada jaringan saraf tiruan. Dataset testing digunakan sebagai uji akhir untuk melihat keakuratan model yang telah dilatih menggunakan *dataset* *training*.

* 1. ***Bidirectional Encoder Representations from Transformers* (BERT)**

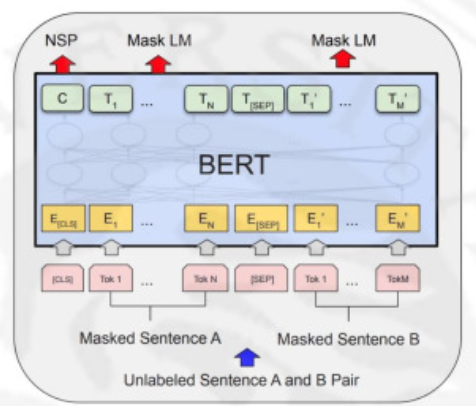
*Bidirectional Encoder Representations from Transformers* atau disingkat BERT adalah suatu model representasi bahasa terlatih yang dikembangkan Oleh para peneliti di Google AI *Language* pada tabun 2018. BERT merupakan sebuah teknik *open source* yang dikembangkan berdasarkan teknik *deep learning* dan berbagai metode seperti *semi-supervised learning*, ELMO, OpenAI *Transformer* dan *Transformer*. Sesuai dengan namanya, BERT menggunakan *Transformer*. *Transformer* adalah mekanisme yang hubungan kontekstual antara kata-kata dalam teks (Vaswani et al. 2017). *Transformer* dapat memahami dan mengkonversi yang diperoleh dengan mekanisme yang bernama *self-attention mechanism*. *Self-atttention* *mechanism* adalah cara *Transformer* untuk mengubah "pemahaman" kata terkait lainnya menjadi kata-kata yang akan diproses dengan mekanismenya. Terdapat dua mekanisme pada *Transformer* yaitu:

1. *Encoder*

*Encoder* memiliki fungsi untuk membaca keseluruhan teks masukan sekaligus.*encoder* terdiri dari stack N = 6 layers yang identik. Setiap layer memiliki dua *sub-layer* yaitu *self-attention layer* dan *feed-forward neural network*. Dengan *self-attention layer, encoder* dapat membantu *node* untuk tidak hanya fokus kepada kata yang sedang dilihat tetapi juga untuk mendapatkan konteks semantik darikata tersebut.

1. *Decoder*

*Decoder* berfungsi untuk menghasilkan urutan *output* yang berupa prediksi.*Decoder* juga terdiri dari stack N = 6 layers yang identik. Setiap layer terdiri dari dua *sub laye*r Seperti yang ada pada *encoder* dengan tambahan *attention layer* diantara dua *layers* tersebut untuk membantu *node* Saat ini mendapatkan *key content* yang membutuhkan *attention* (Vaswani et al. 2017) dengan melakukan *multi-head* *attention* pada *output* dari *encoder*. Sama dengan di *encoder*, *self-attention layer* di *encoder* membuat setiap posisi di *decoder* dapat menangani semua sebelumnya dan posisi saat itu.



Gambar 2.2. Arsitektur Model BERT

Sumber : Souradip Chakraborty.,Ekaba Bisong., (2020).

(Bab 2 harus harus dijelaskan data cleansing tokenization)

BERT di-*training* dengan *dataset*, *dataset* harus disesuaikan dengan representasi input yang dapat diterima Oleh BERT. Oleh karena itu dibutuhkan sebuah *tokenizer* yang berrtujuan untuk melakukan tokenisasi pada kalimat-kalimat dan menghasilkan *input* yang sesuai. Hal ini dilakukan karena BERT menggunakan *vocabulary* yang spesifik yang mana tergantung dengan model apa yang dipakai.*Vocabulary* dibuat dengan menggunakan sebuah model *WordPiece. Vocabulary* terdiri dari semua karakter dan tambahan sekitar 30 kata-kata yang sering digunakan Serta sub kata yang sering muncul pada corpus yang model latih. BERT menerima panjang yang tetap dan sama untuk setiap inputnya. BERT telah menentukan panjang urutan kalimat maksimumnya, yaitu 512 karena encoder pada Transformer hanya menghasilkan output dengan dimensi 512 saja. Jika panjang kalimat lebih dari panjang maksimum yang telah ditentukan, kalimat akan dikurangi (*truncate*).Sedangkan jika panjang urutan kalimat kurang dari panjang maksimum yang sudah ditentukan, kalimat akan ditambah dengan padding. Akan tetapi, panjang maksimum juga dapat ditentukan sendiri tergantung dengan kebutuhan dan tugas yang akan dilakukan. Proses mempersiapkan kalimat menjadi representasi input pada BERT dilakukan Oleh tokenizer, dengan langkah-langkah sebagai berikut:

1. Setiap kalimat ditokenisasi menjadi per kata atau sub kata menggunakan

*Wordpiece*. Untuk melakukan tokenisasi pada sebuah kata, *tokenizer* akanmemeriksa apakah tiap kata pada kalimat terdapat pada *vocabulary*. Jika tidak ada,*tokenizer* akan memecah kata menjadi *sub-sub* kata yang kemungkinan kemunculannya pada *vocabulary* paling besar. Jika *tokenizer* juga tidak menemukan *sub* kata pada vocabulary, kata tersebut dilrcah menjadi per karakternya. Akan tetapi, jika semua kata diubah menjadi sub kata atau karakter individual, akan terjadi *overload* Kata-kata yang tidak ada pada *vocabulary* akan diganti dengan *token* [UNKI] atau *unknown*. Namun jika semua kata diubah menjadi token tersebut, banyak informasi yang akan hilang\_ Oleh kalena itu, kata-kata dapat dipecah menjadi sub kata dengan simbol BERT melakukan ini karena dua hal yaitu pertama untuk mempercepat *processing* dan mengurangi jumlah *parameters* yang harus dilatih, dan kedua untuk mengatasi masalah *out-of-vocabulary*.

1. Setiap kalimat diberi token-token khusus yaitu [CLSI di awal kalimat dan [SEPI di

akhir kalimat. *Token* [CIS] menjadi indikator sebuah kalimat sekaligus

sentimennya Saat klasifikasi sentimen dilakukan. *Token* [SEP]adalah token yang

digunakan untuk memisahkan kalimat satu dengan kalimat selanjutnya. Kalimat

yang sudah diberi *token* khusus ini menjadi *token embeddings*.

1. Setelah itu kalimat-kalimat disesuaikan dengan panjang maksimum yang telah ditentukan dengan mengurangi atau memberi *padding* dengan *token* khusus [PAD].
2. Kemudian setiap kalimat dicocokkan dengan bilangan unik atau ID sesuai dengan *vocabulary* dan bilangan unik tersebut disimpan sebagai *token* id. Bilangan unik atau ID ini didapatkan ketika model di-*rraining*, yang mana membuat setiap kata,*sub* kata, dan karakter pada *vocabulary* memiliki bilangan uniknya masing-masing.ID tersebut diperoleh berdasarkan indeks kata pada vocabulary karena pada *vocabulary* disusun berdasarkan tingkat kemunculannya. Kata dan *sub* kata sendiri harus dikonversi menjadi ID-nya karena model *pre-trained* BERT hanya dapat mengerti ID dari *token*.
3. *Sentence embedding* diberi pada kalimat untuk membedakan kalimat pertama dan kalimat kedua atau *padding*. Tahap ini dapat dilakukan dengan memberi angka 1 pada kalimat pertama dan memberi angka 0 pada *padding*. *Tokenizer* dapat mengetahui mana ~~kalimat~~ dan mana *padding* dengan melihat *token* [SEPI] yang berfungsi memisahkan antara dua kalimat.
4. *Positional embedding* juga ditambahkan untuk tiap *token* untuk menunjukkan isi tiap kata pada kalimat\_tahap ini dilakukan karena BERT tidak mengetahui posisi dari tiap kata. Sehingga walaupun pada awal kalimat terdapat suatu kata dan di akhir kalimat juga terdapat kata yang sama tetapi memiliki makna yang berbeda, BERT tidak akan memperlakukan *token* tersebut dengan makna yang sama. BERT akan menerima unltan kata-kata atau sebuah kalimat sebagai input yangakan terus melalui tumpukan *encoder*. Tiap *encoder* mengaplikasikan *self-attention* dan memberikan output melalui *feed-forward network* yang kemudian dilanjutkan Oleh *encoder* selanjutnya. Pada penelitian ini, model yang dipilih adalah BERT-*base* sehingga proses ini berlanjut sebanyak dua belas Setelah melewati semua *encoder*, tiap *token* per posisi memberikan output vektor dengan ukuran *hidden size* yaitu 768 pada BERT-base. Untuk proses analisis sentimen, output yang diperhatikan adalah output dari pertama yaitu token [CLSI. Vektor tersebut digunakan sebagai input untuk classifier. BERT dapat mencapai hasil yang baik hanya dengan menggunakan neural network tunggal sebagai classifieFnya (Devlin et al., 2019). *Layer* yang digunakan untuk klasifikasi adalah *fully connected neural network* dengan fungsi *sofrmax*. Sehingga *output* dari BERT yang digunakan untukklasifikasi berasal vektor *token* [CLSI] karena *token* [CIS] dianggap melakukan pengumpulan rata-rata data.S token kata untuk mendapatkan vektor dari kalimat. *Layer* terakhir pada *classifier layer* menghasilkan *logits*. *Logits* adalah *outpu*t yang berupa prediksi probabilitas kasar dari kalimat yang akan diklasifikasikan. *Softmax* akan mengubah *logits* tersebut menjadi probabilitas dengan mengambil eksponen dari tiap nilai logit sehingga total probabilitasnya adalah tepat sehingga nilai probabilitas akan berada di antara 0 atau angka positif.
   1. **IndoNLU**

IndoNLU (Indonesian Natural Language Understanding) adalah sebuah komunitas yang mengembangkan sumber daya untuk melatih, menilai, dan menganalisis sistem pemahaman bahasa alami (natural language understanding) dalam bahasa Indonesia. Salah satu produk dari komunitas IndoNLU adalah IndoBenchmark.

IndoBenchmark merupakan sekumpulan tugas (tasks) yang dirancang untuk menguji kemampuan pemahaman bahasa alami sistem. Terdapat dua belas tugas dalam IndoBenchmark, yang meliputi berbagai tingkat kompleksitas, mulai dari klasifikasi kalimat tunggal hingga pelabelan urutan pasangan kalimat. Setiap tugas memiliki kumpulan data yang berbeda, baik dari segi domain maupun gaya penulisan.

Dalam penelitian ini , penggunaan IndoBenchmark difungsikan sebagai sarana untuk melakukan evaluasi dan perbandingan kinerja berbagai sistem pemahaman bahasa alami yang ada. Dalam konteks penelitian tersebut, IndoBenchmark digunakan sebagai benchmark, yaitu sebagai standar acuan untuk mengukur kemampuan sistem pemahaman bahasa alami yang dikembangkan oleh peneliti. Dengan menggunakan IndoBenchmark, peneliti dapat menguji dan membandingkan sistem yang berbeda-beda dan mengevaluasi kinerja mereka dalam tugas-tugas yang beragam.

Penelitian yang telah menggunakan IndoBenchmark antara lain termasuk dalam bidang pengembangan model pemahaman bahasa alami (NLU), analisis sentimen, pemeriksaan kesalahan tata bahasa, dan pemodelan struktur bahasa. Dengan menggunakan IndoBenchmark, para peneliti dapat menguji dan membandingkan kinerja berbagai model pemahaman bahasa alami dalam konteks bahasa Indonesia, serta melihat kemajuan dan tantangan yang ada dalam pengembangan teknologi pemahaman bahasa alami untuk bahasa tersebut.

* 1. **IndoBERT-BASE**

IndoBERT-BASE adalah model berbasis BERT (*Bidirectional Encoder Representations from Transformers*) yang dikembangkan oleh IndoNLU untuk bahasa Indonesia. Model ini dilatih menggunakan dataset indo4B yang terdiri dari empat miliar kata dalam bahasa Indonesia dan memiliki ukuran data teks sebesar dua puluh gigabia.

Ada dua varian model IndoBERT, yaitu IndoBERT-BASE dan IndoBERT-LARGE. Fokus pada IndoBERT-BASE, model ini memiliki hiperparameter yang mengikuti pengaturan pre-training BERT. Model ini memiliki total 124,5 juta parameter, terdiri dari 12 *layer*, 12 *head*, dan ukuran embedding sebesar 768. *Hidden size* model ini adalah 728.

IndoBERT-BASE dilatih menggunakan metode pre-training kontekstual. Dalam tahap pre-training, model mempelajari representasi kata-kata dalam teks bahasa Indonesia menggunakan pendekatan transformer. Dalam proses ini, model dipaparkan pada tugas-tugas seperti prediksi kata yang dihilangkan (*masked word prediction*) dan prediksi kesinambungan kalimat (*next sentence prediction*).

Dengan menggunakan IndoBERT-BASE, IndoNLU dapat menyediakan sebuah model yang telah dilatih secara canggih untuk pemahaman bahasa alami dalam bahasa Indonesia. Model ini dapat digunakan dalam berbagai tugas pemrosesan bahasa alami, termasuk pengenalan entitas, klasifikasi teks, analisis sentimen, dan tugas-tugas lainnya yang memerlukan pemahaman konteks bahasa Indonesia (Putra, H. K., Bijaksana, M. A., Ade Romadhony., 2021).

* 1. ***Seaborn***

*Seaborn* adalah sebuah library Python untuk visualisasi data yang dibangun di atas *matplotlib*. *Library* ini memiliki lebih banyak fungsi untuk visualisasi data dan lebih mudah digunakan dibandingkan dengan *matplotlib*. *Seaborn* bertujuan agar visualisasi menjadi bagian penting dari proses eksplorasi dan pemahaman data. *Seaborn* berorientasi pada fungsi *plotting* berdasarkan dataset yang beroperasi pada dataframe dan array yang berisi seluruh dataset yang secara internal melakukan pemetaan semantik dan agregasi statistik yang diperlukan untuk menghasilkan grafik informatif(Muhammad Zacky Asy'ari, 2020).

* 1. ***Matplotlib***

Matplotlib adalah sebuah library Python yang paling banyak digunakan untuk visualisasi data. Library ini memiliki tiga layer, dan yang paling mudah digunakan adalah scripting layer yang dapat diakses dengan mengimpor matplotlib.pyplot. Matplotlib sendiri dirancang agar dapat digunakan selayaknya menggunakan MATLAB. Library ini dapat digunakan di banyak tools, salah satunya Jupyter Notebook, yang populer digunakan untuk proyek data science. Matplotlib memiliki banyak fungsi perintah yang membuat matplotlib dapat bekerja seperti Matlab, seperti membuat gambar, membuat area plot dalam gambar, menambah label di plot dan lainnya. Matplotlib juga dapat digunakan untuk membuat berbagai jenis grafik, seperti line plot, scatter plot, bar plot, dan histogram. Untuk menginstall matplotlib, dapat dilakukan dengan menggunakan pip atau conda(Srimulia, 2022).

* 1. ***Wordcloud***

*Wordcloud* adalah representasi visual dari kumpulan kata-kata dalam sebuah teks. *Wordcloud* menampilkan kata-kata yang paling sering muncul dalam teks dengan ukuran yang lebih besar dan menampilkan kata-kata yang kurang sering muncul dengan ukuran yang lebih kecil.*Wordcloud* sering digunakan dalam analisis teks untuk memberikan gambaran visual tentang topik atau tema yang paling sering muncul dalam teks.

Untuk membuat wordcloud, terdapat beberapa langkah yang perlu dilakukan, antara lain:

1. Mengumpulkan teks yang akan digunakan untuk membuat wordcloud.
2. Membersihkan teks dari karakter-karakter yang tidak diperlukan seperti tanda baca dan angka.
3. Mengubah semua kata menjadi huruf kecil agar tidak terjadi duplikasi kata yang sama dengan huruf besar.
4. Menghapus kata-kata yang tidak relevan atau tidak penting seperti kata penghubung dan kata umum.
5. Menghitung frekuensi kemunculan kata dalam teks.
6. Menampilkan kata-kata dalam bentuk visual dengan ukuran yang berbeda-beda sesuai dengan frekuensi kemunculan kata tersebut.

*Wordcloud* dapat digunakan untuk berbagai tujuan seperti analisis sentimen, analisis topik, dan analisis opini. Dalam analisis sentimen, wordcloud dapat digunakan untuk menampilkan kata-kata yang paling sering muncul dalam teks positif atau negatif. Dalam analisis topik, wordcloud dapat digunakan untuk menampilkan kata-kata yang paling sering muncul dalam teks yang berkaitan dengan topik tertentu. Sedangkan dalam analisis opini, *wordcloud* dapat digunakan untuk menampilkan kata-kata yang paling sering muncul dalam teks yang berkaitan dengan opini atau pandangan tertentu (Ahmad Arif Samudro, 2019).

* 1. ***Confusion Matrix***

*Confusion Matrix* adalah sebagai suatu alat yang memiliki fungsi untuk melakukan analisis bahwa *classifier* tersebut baik dalam mengenali *tuple* dari kelas yang berbeda. nilai dari *True-Positive* dan *True-Negative* memberikan informasi ketika *classifier* dalam melakukan klasifikasi data bernilai benar, sedangkan *False-Positive* dan *False-Negative* memberikan informasi ketika *classifier* salah dalam melakukan klasifikasi data.

Tabel 2.1 *Confusion Matrix*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | Predicted Class |  |  |
|  |  |  |  |  |  |
|  |  | Positive | Neutral | Negative |  |
|  |  |  |  |  |  |
| True | Positive |  |  |  |  |
| Class |  | True Positive | False Neutral | False Negative |  |
|  |  |  |  |  |
|  |  | (TP) | (FNt) | (FN) |  |
|  |  |  |  |  |  |
|  | Neutral | False Positive | True Netral | False Negative |  |
|  |  | (FP) | (TNt) | (FN) |  |
|  |  |  |  |  |  |
|  | Negative |  |  |  |  |
|  |  | False Positive | False Neutral | True Negative |  |
|  |  | (FP) | (FNt) | (TN) |  |
|  |  |  |  |  |  |

Keenam istilah tersebut adalah :

* *True Positive* (TP), merupakan data positif yang memiliki prediksi benar
* *True Negative* (TN), merupakan data negatif yang memiliki benar
* T*rue Neutral* (TNt) adalah kalimat memiliki sentimen netral dan hasil prediksinya menunjukkan nilai sentimen netral.
* F*alse Neutral* (FNt) adalah kalimat yang memiliki sentimen positif atau negatif namun hasil prediksinya menunjukkan sentimen entral
* *False Positive* (FP), merupakan data netral atau negatif namun diprediksi sebagai data positif
* *False Negative* (FN), merupakan data positif atau netral namun diprediksi sebagai data negatif

Beberapa hal yang dapat dilakukan oleh *Confusion Matrix* dalam hal perhitungan

1. *Accuracy*

*Accuracy* memiliki rasio prediksi benar (positif dan negatif) dengan keseluruhan data. nilai *accuracy* diperoleh dengan persamaan berikut :

(2.1)

1. *Precision*

*Precision* memberikan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Nilai *precision* diperoleh dengan persamaan berikut :

(2.2)

1. *Recall*

*Recall* memberikan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Nilai *Recall* diperoleh dengan persamaan berikut :

(2.3)

1. *F1 Score*

*F1 Score* merupakan perbandingan rata-rata presisi dan *recall* yang dibibitkan. Nilai *F1 Score* dapat diperoleh dengan persamaan berikut :

(2.4)

* 1. **Python**

*Python* adalah bahasa pemrograman tingkat tinggi yang banyak digunakan akhir-akhir ini. dengan filosofi dimana kode mudah dibaca dan sintaksnya mengizinkan *programmer* untuk mengekspresikan konsep dalam kode baris yang lebih sedikit daripada di bahasa pemrograman C. kontruksi bahasa memungkinkan pengguna untuk menulis program secara jelas baik dalam skala kecil maupun besar.

Fitur terpenting dalam *python* adalah mendukung beberapa paradigma pemrograman, termasuk berorientasi objek, pemograman atau prosedural imperatif dan fungsional gaya. *Python* mendukung sistem tipe dinamis dan manajemen memori otomatis dan memiliki besar dan perpustakaan standar yang komprehensif. Penerjemah *python* tersedia untuk banyak sistem operasi.

* 1. **Google Colaboratory**

Google Colab atau Google Colaboratory, adalah sebuah *executable document* yang dapat digunakan untuk menyimpan, menulis, serta membagikanprogram yang telah ditulis melalui Google Drive. *Software* ini pada dasarnya serupa dengan Jupyter Notebook gratis berbentuk *cloud* yang dijalankan menggunakan *browser*, seperti Mozilla Firefox dan Google Chrome. Google Colabmemungkinkan user untuk menjalankan kode Python tanpa perlu melakukan proses instalasi dan setup lainnya karena semua keperluan *setting* dan *adjustment* akan diserahkan ke cloud (Glints.com, 2021).

Google Colaboratory dapat menulis dan mengeksekusi Python di *browser* dengan beberapa keuntungan yaitu tidak memerlukan konfigurasi, akses gratis ke GPU, berbagi dengan mudah serta dapat menggabungkan kode yang dapat dijalankan dan *rich text* dalam satu dokumen, beserta gambar, HTML, LaTeX, dan lainnya (*colab.research.google.com*). Dengan sifat yang fleksibel Google Colab dapat membuat *notebook* Colab oleh *user* itu sendiri, *notebook* tersebut akan disimpan di akun Google Drive pengguna. Pengguna dapat dengan mudah membagikan *notebook* Colab dengan pengguna lainnya, sehingga dapat saling memberikan komentar di *notebook* pengguna satu sama lainnya atau bahkan mengeditnya.

* 1. **Penelitian Terkait**

Penelitian ini dilakukan dengan mempelajari penelitian terdahulu yang memiliki keterkaitan tentang analisis sentimen. Beberapa penelitian terdahulu yang digunakan sebagai dasar pembelajaran pada Tabel 2.1

Tabel 2.2 Penelitian Terkait

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Nama | Tahun | Judul | Hasil |
| 1 | Cindy Alifia  Putri,  Adiwijaya,  Said Al Faraby | 2020 | Analisis Sentimen  Review Film  Berbahasa Inggris  dengan Pendekatan  Bidirectional Encoder  Representations from  Transformers | Akurasi yang didapatkan  dari analisis sentimen  menggunakan metode  BERT sebesar 73%.  dataset yang digunakan  berbentuk dokumen.  BERT-base terbukti |
| 2 | Raden Mas  Rizi Wahyu  Panca Kusuma  Atmaja, Wily  Yustanti | 2020 | Analisis Sentimen  Customer Review  Aplikasi Ruang Guru  dengan Metode  BERT(Bidirectional  Encoder  Representations from  Transformers) | Akurasi yang didapatkan  dari analisis sentimen  menggunakan metode  BERT sebesar 99%  dengan nilai F1 Score  sebesar 98.9%.  menggunakan 5437  testing data dimana 5254  komentar positif, 16  netral dan 167 komentar  negatif. |

Tabel 2.2 Penelitian Terkait ( Lanjutan )

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | **Nama** | **Tahun** | **Judul** | **Hasil** |
| 3 | Muhammad  Mahrus Zain,  Rizky  Nathamael  Simbolon,  Harlem  Sulung, Zaidan  Anwar | 2021 | Analisis Sentimen  Pendapat Masyarakat  Mengenai Vaksin  Covid-19 Pada Media  Sosial Twitter dengan  Robustly Optimized  BERT Pretraining  Approach | Akurasi yang didapatkan  dari analisis sentimen  menggunakan metode  RoBERT sebesar 95%.  dimana rata-rata hasil  akurasi prediksi pada  label positif 84%, Netral  97% dan Negatif 93% |

Berdasarkan penelitian diatas bisa disimpulkan kekurangan dan kelebihan penelitian sebelumnya yang diatasi atau digunakan dalam penelitian ini:

1. Penelitian oleh Cindy Alifia Putri, Adiwijaya, dan Said Al Faraby (2020) tentang analisis sentimen review film berbahasa Inggris menggunakan metode Bidirectional Encoder Representations from Transformers (BERT):

Kekurangan: Penelitian ini menggunakan dataset yang berbentuk dokumen tanpa menyebutkan jumlah data yang digunakan. Informasi tersebut penting untuk mengevaluasi sejauh mana hasil penelitian dapat diterapkan pada skala yang lebih besar.

Kelebihan: Metode BERT terbukti memberikan akurasi sebesar 73% dalam analisis sentimen. Hasil ini dapat menjadi dasar untuk melihat potensi penggunaan BERT dalam penelitian analisis sentimen pada bahasa Inggris.

1. Penelitian oleh Raden Mas Rizi Wahyu Panca Kusuma Atmaja dan Wily Yustanti (2020) tentang analisis sentimen customer review aplikasi Ruang Guru menggunakan metode BERT:

Kekurangan: Jumlah data yang digunakan dalam penelitian ini tidak dijelaskan secara rinci, hanya menyebutkan terdapat 5437 data pengujian dengan komentar positif, netral, dan negatif. Informasi lebih lanjut tentang bagaimana data ini dikumpulkan dan diragamkan dapat membantu memahami kehandalan hasil penelitian.

Kelebihan: Metode BERT memberikan hasil akurasi yang sangat tinggi, mencapai 99%, dengan nilai F1 Score sebesar 98.9%. Hal ini menunjukkan potensi BERT dalam menganalisis sentimen pada review pelanggan.

1. Penelitian oleh Muhammad Mahrus Zain, Rizky Nathamael Simbolon, Harlem Sulung, dan Zaidan Anwar (2021) tentang analisis sentimen pendapat masyarakat mengenai vaksin COVID-19 pada media sosial Twitter dengan pendekatan RoBERT:

Kekurangan: Penelitian ini tidak menyediakan informasi tentang ukuran dataset yang digunakan. Jumlah data yang cukup besar dapat memberikan keandalan yang lebih baik pada hasil analisis sentimen.

Kelebihan: Metode RoBERT memberikan akurasi sebesar 95% dalam analisis sentimen, dengan rata-rata akurasi prediksi yang tinggi pada label positif (84%), netral (97%), dan negatif (93%). Hal ini menunjukkan kemampuan RoBERT dalam memahami sentimen masyarakat terhadap vaksin COVID-19 di media sosial Twitter.

Dalam penelitian ini, kekurangan penelitian sebelumnya, seperti ketidakjelasan jumlah data dan kehandalan hasil, diatasi dengan menyediakan informasi yang lebih rinci dan menggunakan pendekatan yang lebih kuat. Alasan penggunaan pendekatan baru ini mungkin termasuk meningkatkan akurasi, mengatasi masalah yang dihadapi oleh pendekatan sebelumnya, atau menyesuaikan dengan konteks penelitian yang berbeda.

**Bab 3**

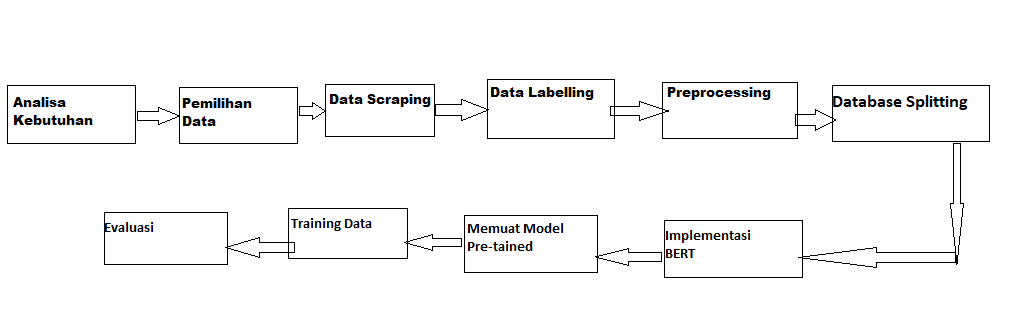
**Metodologi Penelitian**

* 1. **Deskripsi Penelitian**

Penelitian ini bertujuan untuk melakukan analisis sentimen pada ulasan pengguna Vision+ di Google Play Store menggunakan metode algoritma *Bidirectional Encoder Representations from Transformers* (BERT). Penelitian ini mengangkat masalah dalam memahami sentimen pengguna terhadap aplikasi Vision+ dan ingin mencari hasil yang dapat memberikan wawasan tentang kepuasan pengguna serta mencari pola umum dalam ulasan tersebut.

* 1. **Tahapan Penelitian**

Tahapan penelitian analisis sentimen yang dilakukan melalui komentar di Google Playstore pada aplikasi Vision+ untuk melihat bagaimana sentimen atau pendapat masyarakat mengenai aplikasi tersebut. Analisis sentimen ini menggunakan metode BERT untuk mengklasifikasikan data komentar yang mengandung kata Vision+ ke dalam tiga kategori, yaitu positif, netral, dan negatif. Diagram proses dari analisis sentimen dapat dilihat pada Gambar 3.1.

Gambar 3.1. Tahapan Penelitian

Langkah-langkah penelitian yang dilakukan dalam analisis sentimen pada ulasan Vision+ di Google Play Store menggunakan metode BERT dapat meliputi:

Dengan langkah-langkah di atas, penelitian ini bertujuan untuk memberikan pemahaman yang lebih baik tentang sentimen pengguna terhadap aplikasi Vision+ melalui analisis ulasan di Google Play Store menggunakan metode BERT.

* 1. **Analisa Kebutuhan**

Analisa kebutuhan adalah proses identifikasi, pemahaman, dan penentuan kebutuhan yang harus dipenuhi dalam sebuah proyek, penelitian, atau pengembangan produk. Hal ini melibatkan mengumpulkan informasi tentang kebutuhan pengguna, pemangku kepentingan, dan tujuan yang ingin dicapai. Analisa kebutuhan bertujuan untuk memastikan bahwa solusi yang dihasilkan memenuhi kebutuhan yang relevan dan bermanfaat.

Dalam konteks penelitian Analisis Sentimen pada Ulasan Vision+ di Google Play Store menggunakan metode algoritma Bidirectional Encoder Representations from Transformers (BERT), analisa kebutuhan mencakup beberapa aspek, antara lain:

1. Tujuan Penelitian: Menentukan tujuan yang ingin dicapai melalui analisis sentimen pada ulasan Vision+. Tujuan ini dapat berupa pemahaman tentang sentimen pengguna, identifikasi masalah, atau peningkatan pengalaman pengguna.
2. Jenis Data: Menentukan jenis data yang diperlukan untuk analisis sentimen. Dalam hal ini, data yang dibutuhkan adalah ulasan atau komentar pengguna Vision+ di Google Play Store.
3. Sumber Data: Mengidentifikasi sumber data yang akan digunakan. Pada kasus ini, sumber data adalah Google Play Store. Data ulasan pengguna dapat diakses melalui platform tersebut.
4. Skala Data: Menentukan skala data yang diperlukan untuk mendapatkan sampel yang representatif. Misalnya, menentukan jumlah ulasan yang perlu dikumpulkan untuk analisis yang akurat.
5. Kriteria Pemilihan Data: Menetapkan kriteria untuk memilih data yang relevan. Kriteria ini dapat mencakup rating ulasan, keberagaman pendapat, atau kejelasan dalam mengungkapkan sentimen.
6. Alat atau Metode: Memilih alat atau metode yang tepat untuk menganalisis sentimen. Dalam hal ini, metode BERT digunakan untuk menganalisis sentimen pada ulasan Vision+.

Analisa kebutuhan ini membantu memastikan bahwa penelitian dilakukan dengan fokus yang jelas, data yang relevan diperoleh, dan metode yang sesuai digunakan untuk mencapai tujuan penelitian. Dengan melakukan analisa kebutuhan yang baik, peneliti dapat mengarahkan langkah-langkah penelitian dengan lebih efektif dan menghasilkan hasil yang bermanfaat.

* 1. **Pemilihan Data**

Pemilihan data adalah proses memilih dan menentukan data yang akan digunakan dalam analisis sentimen. Pemilihan data yang tepat dapat sangat mempengaruhi hasil analisis sentimen yang dilakukan. Berikut adalah beberapa faktor yang perlu dipertimbangkan dalam pemilihan data untuk analisis sentimen:

1. Tujuan Analisis: Tentukan tujuan analisis sentimen Anda. Apakah Anda ingin menganalisis sentimen terhadap produk tertentu, merek, topik tertentu atau memahami respons publik terhadap suatu peristiwa
2. Sumber Data: Identifikasi sumber data yang relevan dengan tujuan analisis Anda. Sumber data ini dapat berupa media sosial, *platform e-commerce*, forum diskusi, atau platform lainnya di mana orang berbagi pendapat dan ulasan. Pastikan sumber data yang Anda pilih sesuai dengan konteks analisis Anda.
3. Jangkauan Waktu: Tentukan rentang waktu data yang akan Anda analisis. Apakah Anda ingin menganalisis sentimen dalam jangka waktu tertentu atau memantau perubahan sentimen seiring berjalannya waktu? Rentang waktu yang dipilih dapat mempengaruhi gambaran keseluruhan sentimen yang ditemukan.
4. Volume Data: Pertimbangkan jumlah data yang Anda butuhkan untuk analisis sentimen Anda. Volume data dapat berkisar dari sejumlah besar sampel hingga sampel yang lebih kecil tergantung pada tujuan Anda. Pastikan data yang Anda pilih mewakili sampel yang cukup untuk menghasilkan hasil yang valid.
5. Relevansi Konteks: Pastikan data yang Anda pilih relevan dengan konteks analisis Anda. Misalnya, jika Anda menganalisis sentimen terhadap sebuah produk, pastikan data yang Anda gunakan berasal dari pengguna produk tersebut, bukan dari pengguna produk serupa.
6. Pemrosesan dan Filterisasi: Setelah memilih data awal, Anda mungkin perlu melakukan pemrosesan dan filterisasi tambahan untuk membersihkan data dari konten yang tidak relevan atau noise. Anda dapat menggunakan metode seperti penghapusan tautan, penghilangan duplikat, pembersihan teks, dan filterisasi berdasarkan kata kunci untuk memperoleh data yang lebih bersih dan relevan.
   1. **Metode *Data Sampling***

Metode data sampling yang digunakan dalam penelitian ini adalah non-probability sampling dengan teknik purposive sampling. Sampel diambil berdasarkan kriteria tertentu, yaitu pengguna media sosial yang telah memposting ulasan tentang produk tertentu dalam periode waktu tertentu. Metode ini dipilih karena memungkinkan peneliti untuk memilih sampel yang sesuai dengan tujuan penelitian dan meminimalkan bias dalam pengambilan sampel.

Alasan Analisis Sentimen 8000 Merepresentasikan 5 Juta User adalah penelitian sebelumnya menunjukkan bahwa jumlah sampel yang cukup untuk merepresentasikan populasi adalah sekitar 1000 sampai 5000 yang bersumber dari jurnal Indra Kusumajati Susanto dengan judul Analisis Sentimen dan Topic Modelling Pada Pembelajaran Online di Indonesia Melalui Twitteryang diterbitkan pada tahun 2022 (Indrija Kusumajati Susanto, 2022). Oleh karena itu, dengan mengambil sampel sebanyak 8000 ulasan dari pengguna media sosial yang memenuhi kriteria sampel, penelitian ini diharapkan dapat merepresentasikan pandangan pengguna media sosial yang lebih luas terhadap produk tertentu, bahkan hingga mencapai 5 juta user.

* 1. ***Scrapping***

*Scrapping* atau pengambilan data dilakukan dengan metode *scraping* data di platform Google Play Store menggunakan *python script* dengan *library google* *play scraper*. *Scraping* dilakukan untuk mendapatkan komentar pengguna Vision+ yang tertulis pada platform Google Play Store seperti pada contoh berikut:

#scrape jumlah ulasan yang diinginkan

from google\_play\_scraper import Sort, reviews

result, continuation\_token = reviews(

'com.zte.iptvclient.android.idmnc',

lang='id',

Kode di atas memanggil fungsi reviews() dengan beberapa parameter. Parameter pertama adalah package name dari aplikasi yang ingin kita ambil ulasannya. Pada contoh ini, package name yang digunakan adalah 'com.zte.iptvclient.android.idmnc'.Parameter lang digunakan untuk menentukan bahasa ulasan yang ingin diambil. Di sini, kita mengatur bahasa menjadi 'id' yang berarti bahasa Indonesia.

* 1. **Data *Labelling***

Data *labelling* adalah proses menandai atau memberi label pada data untuk mengidentifikasi dan mengkategorikan informasi yang terkandung di dalamnya. Data *labelling* umumnya digunakan dalam konteks pembelajaran mesin dan kecerdasan buatan, di mana data yang dilabeli digunakan sebagai input untuk melatih algoritma atau model untuk mengenali pola, klasifikasi, atau memahami data yang serupa di masa depan.

Dalam proses data *labelling*, seorang penelabel atau anotator manusia menandai data dengan label yang sesuai berdasarkan aturan atau panduan yang telah ditentukan sebelumnya. Label ini dapat berupa kategori, atribut, tanda-tanda, atau anotasi lainnya yang relevan dengan tujuan yang ingin dicapai. Pada penelitian ini penulis membuat data *labelling* untuk memisahkan data komentar yang bersifat positif, netral dan negatif berdasarkan dari jumlah *rating* berikut ini adalah *sample code* dari proses data *labelling*

def to\_sentiment(rating)

rating = int(rating)

if rating <= 2:

return 0

elif rating == 3:

return 1

else:

return 2

df['category'] = df["score"].apply(to\_sentiment)

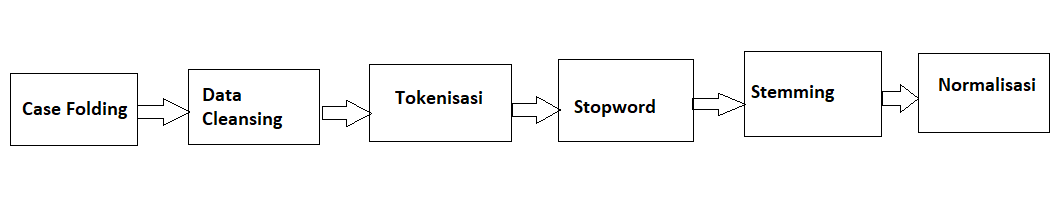
Pada kode diatas mendefinisikan fungsi to\_sentiment() dengan parameter rating. Pertama, kita mengonversi nilai rating menjadi tipe data *integer* menggunakan fungsi *int*(). Kemudian, kita menggunakan struktur kondisional untuk mengkategorikan sentimen berdasarkan *rating*.

1. Jika rating kurang dari atau sama dengan 2, maka ulasan tersebut dikategorikan sebagai sentimen negatif, dan fungsi mengembalikan nilai 0.
2. Jika rating sama dengan 3, maka ulasan tersebut dikategorikan sebagai sentimen netral, dan fungsi mengembalikan nilai 1.
3. Jika rating lebih dari 3, maka ulasan tersebut dikategorikan sebagai sentimen positif, dan fungsi mengembalikan nilai 2.

Di baris ini, fungsi to\_sentiment() diterapkan pada kolom "score" dalam DataFrame df menggunakan metode apply(). Hasilnya disimpan dalam kolom baru dengan nama "category". Dengan demikian, setiap nilai dalam kolom "score" akan digunakan sebagai argumen untuk fungsi to\_sentiment(), dan hasilnya akan disimpan dalam kolom "category".

* 1. ***Pre-processing***

*Pre-processing* dilakukan untuk membersihkan data komentar, tahap ini membuat data yang tidak baik atau belum siap diolah menjadi data yang siap diubah. tahap ini terdiri dari beberapa tahapan yaitu *case folding*, *data cleaning*, tokenisasi, *stopword*, *stemming* dan normalisasi kata tidak baku. Diagram proses dari *pre-processing* dapat dilihat pada Gambar 3.2.



Gambar 3.2 Proses *Preprocessing*

* + 1. ***Case Folding***

*Case folding* merupakan tahap merubah semua huruf besar yang ada pada dokumen (*uppercase*) menjadi huruf kecil (*lowercase*). case folding dilakukan menggunakan fungsi *lower()* yang ada di *library python*. Perintah untuk melakukan *case folding* adalah sebagai berikut:

def clean\_review(text):

text = text.lower()

Tabel 3.1 Contoh Hasil *case folding*

|  |  |
| --- | --- |
| Komentar | Hasil Case folding |
| Versi yg skrg byk bugnya. Ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. Parah bgt dah... | versi yg skrg byk bugnya. ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. parah bgt dah... |
| Parah nih aplikasi udh buffering Mulu, terus mau nonton live tv emang harus beli paket dulu yah? lawak sih klo emng gitu pelit bener gak kayak aplikasi video | parah nih aplikasi udh buffering mulu, terus mau nonton live tv emang harus beli paket dulu yah? lawak sih klo emng gitu pelit bener gak kayak aplikasi video |
| Kenapa sekarang app tidak dapat di gunakan sama sekali. Sinyal cuma muter-muter buffering terus | kenapa sekarang app tidak dapat di gunakan sama sekali. sinyal cuma muter-muter buffering terus |
| Lemot banget, saat sy tulis ulasan ini pun film nya tidak bisa diputar, padahal sudah berlangganan premium , lalu suka tiba2 putus ditengah film, mengecewakan untuk ukuran aplikasi berbayar | lemot banget, saat sy tulis ulasan ini pun film nya tidak bisa diputar, padahal sudah berlangganan premium , lalu suka tiba2 putus ditengah film, mengecewakan untuk ukuran aplikasi berbayar |

Tabel tersebut adalah contoh hasil dari proses case folding yang dilakukan pada kolom "Komentar" dalam suatu dataframe. Case folding dilakukan dengan menggunakan fungsi lower() yang tersedia di library Python. Tujuan dari case folding adalah untuk mengubah semua huruf kapital (uppercase) menjadi huruf kecil (lowercase) dalam teks.

Dalam tabel tersebut, terdapat dua kolom yaitu "Komentar" dan "Hasil Case folding". Kolom "Komentar" berisi teks komentar yang merupakan contoh input sebelum dilakukan case folding. Kolom "Hasil Case folding" berisi hasil dari proses case folding, di mana semua huruf besar pada komentar diubah menjadi huruf kecil.

Contoh pada baris pertama, komentar awalnya adalah "Versi yg skrg byk bugnya. Ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. Parah bgt dah...". Setelah proses case folding, komentar tersebut berubah menjadi "versi yg skrg byk bugnya. ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. parah bgt dah...".

Proses case folding diulang pada setiap baris dalam kolom "Komentar", sehingga semua huruf kapital dalam setiap komentar diubah menjadi huruf kecil dalam kolom "Hasil Case folding". Dengan demikian, output dari proses case folding adalah teks yang seragam dalam hal penggunaan huruf kecil, yang memudahkan analisis dan pemrosesan teks lebih lanjut.

* + 1. **Data *Cleansing***

*Data Cleansing* merupakan tahap menghilangkan angka, simbol, URL, at (@) hashtag (#), spasi berlebih, tanda baca, *emoji*, dan pengulangan karakter yang ada pada kalimat *review*. Data *cleansing* merupakan tahap penting dalam pemrosesan teks yang bertujuan untuk membersihkan teks dari berbagai elemen yang tidak relevan atau mengganggu, seperti angka, simbol, URL, tanda baca, emoji, dan pengulangan karakter. Berikut ini adalah potongan kode dari fungsi Data *cleansing* :

text = re.sub(r'\n', ' ', text) text = emoji.demojize(text)

text = re.sub(r"([xX;:]'?[dDpPvVoO3)(])", ' ', text)

text = re.sub(r"(https?:\/\/(?:www\.|(?!www))[a-zA-Z0-9][a-zA-Z0-9-]+[a-zA-Z0-9]\.[^\s]{2,}|www\.[a-zA-Z0-9][a-zA-Z0-9-]+[a-zA-Z0-9]\.[^\s]{2,}|https?:\/\/(?:www\.|(?!www))[a-zA-Z0-9]+\.[^\s]{2,}|www\.[a-zA-Z0-9]+\.[^\s]{2,})", "", text)

text = re.sub(r"@[^\s]+[\s]?", ' ', text) text = re.sub(r'#(\S+)', r'\1', text)

text = re.sub('[^a-zA-Z,.?!]+',' ',text) text = repeatcharClean(text)

text = re.sub('[ ]+',' ',text)

Kode yang diberikan menggunakan modul re pada Python untuk melakukan serangkaian penggantian string pada teks yang diberikan. Berikut adalah penjelasan dari setiap baris kode tersebut:

1. text = re.sub(r'\n', ' ', text): Baris ini mengganti semua karakter newline (\n) pada teks dengan karakter spasi.
2. text = emoji.demojize(text): Baris ini mengganti semua emoji pada teks dengan representasi teks yang sesuai.
3. text = re.sub(r"([xX;:]'?[dDpPvVoO3)(])", ' ', text): Baris ini mengganti semua emotikon pada teks dengan karakter spasi.
4. text = re.sub(r"(https?:\/\/(?:www\.|(?!www))[a-zA-Z0-9][a-zA-Z0-9-]+[a-zA-Z0-9]\.[^\s]{2,}|www\.[a-zA-Z0-9][a-zA-Z0-9-]+[a-zA-Z0-9]\.[^\s]{2,}|https?:\/\/(?:www\.|(?!www))[a-zA-Z0-9]+\.[^\s]{2,}|www\.[a-zA-Z0-9]+\.[^\s]{2,})", "", text): Baris ini menghapus semua URL dari teks.
5. text = re.sub(r"@[^\s]+[\s]?", ' ', text): Baris ini menghapus semua mention (kata-kata yang diawali dengan @) dari teks.
6. text = re.sub(r'#(\S+)', r'\1', text): Baris ini menghapus simbol # dari hashtag pada teks.
7. text = re.sub('[^a-zA-Z,.?!]+',' ',text): Baris ini menghapus semua karakter non-alphabet dari teks kecuali koma, titik, tanda tanya, dan tanda seru.
8. text = repeatcharClean(text): Baris ini memanggil fungsi repeatcharClean yang tidak didefinisikan dalam kode yang diberikan.
9. text = re.sub('[ ]+',' ',text): Baris ini mengganti semua spasi berturut-turut pada teks dengan satu spasi.

Secara keseluruhan, kode tersebut melakukan serangkaian operasi pembersihan teks untuk memproses teks agar dapat dianalisis lebih lanjut.Dari fungsi diatas menghasil data *cleansing* yang akan ditampilkan di tabel 3.2

Tabel 3.2 Contoh Hasil Data *Cleansing*

|  |  |
| --- | --- |
| Komentar | Hasil Data *Cleansing* |
| Versi yg skrg byk bugnya. Ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. Parah bgt dah... | versi yg skrg byk bugnya ndk bs streaming video dengan lancer walaupun sdh beli paket premium parah bgt dah |
| Parah nih aplikasi udh buffering Mulu, terus mau nonton live tv emang harus beli paket dulu yah? lawak sih klo emng gitu pelit bener gak kayak aplikasi video | parah nih aplikasi udh buffering mulu terus mau nonton live tv emang harus beli paket dulu yah lawak sih klo emng gitu pelit bener gak kayak aplikasi video |
| Kenapa sekarang app tidak dapat di gunakan sama sekali. Sinyal cuma muter-muter buffering terus | kenapa sekarang app tidak dapat di gunakan sama sekali sinyal cuma muter buffering terus |
| Lemot banget, saat sy tulis ulasan ini pun film nya tidak bisa diputar, padahal sudah berlangganan premium , lalu suka tiba2 putus ditengah film, mengecewakan untuk ukuran aplikasi berbayar | lemot banget saat sy tulis ulasan ini pun film nya tidak bisa diputar padahal sudah berlangganan premium lalu suka tiba2 putus ditengah film mengecewakan untuk ukuran aplikasi berbayar |

Contoh pada baris pertama, komentar awalnya adalah " Versi yg skrg byk bugnya. Ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. Parah bgt dah...". Setelah proses data cleansing, komentar tersebut berubah menjadi " versi yg skrg byk bugnya ndk bs streaming video dengan lancer walaupun sdh beli paket premium parah bgt dah".Bisa dilihat bagian characternya tanda “.” Menjadi hilang karena telah dilakukan proses data cleansing. Dengan demikian, output dari proses data cleansing adalah teks yang tidak memiliki character, yang memudahkan analisis dan pemrosesan teks lebih lanjut.

* + 1. **Tokenisasi**

Tokenisasi merupakan proses memecah teks menjadi unit-unit yang lebih kecil, seperti kata-kata atau token. Pada penelitian ini, tokenisasi dilakukan dengan menggunakan fungsi word\_tokenize dari library NLTK.

Proses tokenisasi penting dalam penelitian ini karena memiliki beberapa efek yang berdampak pada analisis teks:

* 1. Memecah teks menjadi unit-unit yang lebih kecil: Tokenisasi memisahkan teks menjadi kata-kata atau token-token terpisah. Ini memungkinkan analisis lebih lanjut pada tingkat kata-kata, seperti penghitungan frekuensi kata atau deteksi kata kunci.
  2. Membuang karakter yang tidak diinginkan: Dalam tokenisasi, karakter-karakter tertentu seperti tanda baca atau spasi yang berlebihan dapat dihapus. Hal ini membantu membersihkan teks dan memfokuskan analisis hanya pada kata-kata yang relevan.
  3. Membuat dasar untuk analisis teks lebih lanjut: Setelah dilakukan tokenisasi, data teks dapat diolah lebih lanjut, misalnya dengan melakukan stemming atau lemmatisasi untuk mendapatkan bentuk dasar kata-kata.

Dalam tabel yang diberikan, "Hasil Tokenisasi" menunjukkan teks yang telah melalui proses tokenisasi. Setiap kata dalam teks dipisahkan dan ditampilkan secara terpisah sebagai token. Hal ini mempermudah analisis teks lebih lanjut, seperti penghitungan frekuensi kata atau klasifikasi sentimen berdasarkan kata-kata yang muncul.

Dalam penelitian yang menggunakan analisis teks, tokenisasi membantu dalam mempersiapkan data teks sehingga dapat diolah lebih lanjut. Tokenisasi menjadi langkah awal yang penting untuk memahami dan menganalisis teks dengan lebih efektif.tahapan ini dijalankan menggunakan “*word\_tokenize*” dari library NLTK. Perintah untuk melakukan tokenisasi adalah sebagai berikut

def normalize\_review(text):

list\_text = word\_tokenize(text)

# ubah bahasa alay

list\_text = [normalize\_word\_dict[term] if term in normalize\_word\_dict else term for term in list\_text]

Kode yang diberikan merupakan sebuah fungsi Python dengan nama normalize\_review yang menerima sebuah parameter teks. Fungsi ini melakukan normalisasi pada teks dengan cara memisahkan setiap kata pada teks menggunakan fungsi word\_tokenize dari modul nltk. Selanjutnya, fungsi ini melakukan normalisasi pada kata-kata yang terdapat pada teks dengan cara mengganti kata-kata alay atau slang dengan kata-kata yang lebih umum menggunakan kamus normalize\_word\_dict. Kamus ini berisi pasangan kata-kata alay dan kata-kata umum yang sesuai. Fungsi ini mengganti setiap kata pada teks dengan kata yang sesuai pada kamus jika kata tersebut terdapat pada kamus. Jika kata tersebut tidak terdapat pada kamus, maka kata tersebut tidak diubah.

Dalam bahasa Indonesia, fungsi ini melakukan normalisasi pada teks dengan cara memisahkan setiap kata pada teks dan mengganti kata-kata alay atau slang dengan kata-kata yang lebih umum menggunakan kamus normalize\_word\_dict. Kamus ini berisi pasangan kata-kata alay dan kata-kata umum yang sesuai. Fungsi ini mengganti setiap kata pada teks dengan kata yang sesuai pada kamus jika kata tersebut terdapat pada kamus. Jika kata tersebut tidak terdapat pada kamus, maka kata tersebut tidak diubah.

Tabel 3.3 Contoh Hasil Tokenisasi

|  |  |
| --- | --- |
| Komentar | Hasil Tokenisasi |
| Versi yg skrg byk bugnya. Ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. Parah bgt dah... | “versi” “yg” “skrg” “byk” “bugnya” “ndk” “bs” “streaming” “video” “dengan” “lancer” “walaupun” “sdh” “beli” “paket” “premium” “parah” “bgt” “dah” |
| Parah nih aplikasi udh buffering Mulu, terus mau nonton live tv emang harus beli paket dulu yah? lawak sih klo emng gitu pelit bener gak kayak aplikasi video | “parah” “nih” “aplikasi” “udh” “buffering” “mulu” “terus” “mau” “nonton” “live” “tv” “emang” “harus” “beli” “paket” “dulu” “yah” “lawak” “sih” “klo” “emng” “gitu” “pelit” “bener” “gak” “kayak” “aplikasi” “video” |
| Kenapa sekarang app tidak dapat di gunakan sama sekali. Sinyal cuma muter-muter buffering terus | “kenapa” “sekarang” “app”  “tidak” “dapat” “di” “ gunakan” “sama” “sekali” “sinyal” “cuma" “muter” “buffering” “terus” |

Tabel 3.3 Contoh Hasil Tokenisasi (Lanjutan)

|  |  |
| --- | --- |
| Komentar | Hasil Tokenisasi |
| Lemot banget, saat sy tulis ulasan ini pun film nya tidak bisa diputar, padahal sudah berlangganan premium , lalu suka tiba2 putus ditengah film, mengecewakan untuk ukuran aplikasi berbayar | “lemot” “banget” “saat” “sy” “tulis” “ulasan” “ini” “pun” “film” “nya” “tidak” “bisa” “diputar” “padahal” “sudah” “berlangganan” “premium” “lalu” “suka” “tiba2” “putus” “ditengah” “film” “mengecewakan” “untuk” “ukuran” “aplikasi” “berbayar” |

* + 1. ***Stemming***

Pada tahap ini dilakukan untuk mengubah kata berimbuhan yang terdapat pada kata seperti pada awalan, sisipan atau akhiran kata menjadi kata dasar. Untuk melakukan *stemming* menggunakan *library* Sastrawi seperti pada perintah berikut:

factory = StemmerFactory()

stemmer = factory.create\_stemmer()

list\_text = [stemmer.stem(word) for word in list\_text]

list\_text = [word for word in list\_text if word not in list\_stopwords]

Perintah di atas akan memanggil “*StemmerFactory*” dari Sastrawi untuk menjalankan proses *stemming*. Setelah dipanggil, proses akan memeriksa tiap kata. Jika kata tersebut memiliki imbuhan maka imbuhan tersebut akan dihapus.

Perintah di atas akan memanggil “*StemmerFactory*” dari Sastrawi untuk menjalankan proses *stemming*. Setelah dipanggil, proses akan memeriksa tiap kata. Jika kata tersebut memiliki imbuhan maka imbuhan tersebut akan dihapus.

* + 1. ***Stopwords Removal***

Pada tahap ini digunakan untuk menghilangkan kata-kata yang umumnya sangat sering muncul pada suatu dokumen dan tidak memiliki arti khusus. Proses ini menggunakan *library stopwords* Indonesia yang disediakan oleh NLTK dan menambahkan kamus yang dibuat oleh Tala. Untuk melakukan *stopword* menggunakan perintah berikut:

additional\_stop = [] #['ya', 'deh', 'dia']

default\_stop = pd.read\_csv(stopword\_path, na, mes=['stopwords'])

default\_stop = default\_stop.stopwords.to\_list()

list\_stopwords = []

list\_stopwords = default\_stop + additional\_stop

default\_stop = default\_stop.stopwords.to\_list()

Kodingan diatas digunakan untuk memproses daftar kata-kata stopword dalam bahasa tertentu. Variabel additional\_stop adalah list kosong yang dapat digunakan untuk menambahkan kata-kata stopword tambahan. Variabel default\_stop adalah daftar kata-kata stopword default yang dibaca dari file CSV yang diidentifikasi oleh stopword\_path. Fungsi pd.read\_csv() digunakan untuk membaca file CSV ke dalam objek DataFrame Pandas. Fungsi ini memiliki beberapa parameter seperti filepath\_or\_buffer, sep, header, index\_col, usecols, dan na\_values

Setelah membaca file CSV, variabel default\_stop diisi dengan daftar kata-kata stopword default yang telah dibaca dari file CSV. Kemudian, variabel list\_stopwords diisi dengan gabungan dari daftar kata-kata stopword default dan kata-kata stopword tambahan yang telah didefinisikan sebelumnya.

Namun, terdapat kesalahan sintaksis pada kodingan ini, yaitu parameter na pada fungsi pd.read\_csv() tidak didefinisikan dengan nilai yang sesuai. Selain itu, pada baris terakhir, variabel default\_stop mencoba untuk mengakses atribut stopwords setelah sebelumnya diubah menjadi list, yang menyebabkan kesalaha. Dari pemrosesan kode diatas menghasilkan output yang akan dijelaskan pada tabel 3.4 yang bersumber dari Tala, F.Z.

Tabel 3.4 Daftar *Stopwords*

Sumber. Tala, F.Z. (2003)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ada | Anda | Atau | Belum | Dimana |
|  |  |  |  |  |
| Adalah | Antara | Bagaimanapun | Bisa | DLL |
|  |  |  |  |  |
| Agak | Anu | Bagi | Boleh | Dsb |
|  |  |  |  |  |
| Akan | Apakah | Bahwa | Dahulu | Dst |
|  |  |  |  |  |
| Amat | Apalagi | Begitu | Dalam | Dua |
|  |  |  |  |  |

* + 1. **Normalisasi Kata Tidak Baku**

Pada tahap ini dataset yang mengandung kata tidak baku diubah menjadi kata yang baku dan yang sesuai dengan ejaan. Normalisasi dilakukan karena masih terdapat banyak kata *slang* yang akan mempengaruhi program. Untuk melakukan normalisasi menggunakan perintah berikut:

kamus\_alay = pd.read\_csv(kamus\_alay\_path)

normalize\_word\_dict = {}

for index, row in kamus\_alay.iterrows():

if row[0] not in normalize\_word\_dict:

normalize\_word\_dict[row[0]] = row[1]

kamus\_alay = pd.read\_csv(kamus\_alay\_path)

normalize\_word\_dict = {}

for index, row in kamus\_alay.iterrows():

if row[0] not in normalize\_word\_ict:

normalize\_word\_dict[row[0]] = row[1]

Setelah membaca file CSV, kodingan menginisialisasi dictionary kosong normalize\_word\_dict. Kemudian, kodingan melakukan iterasi pada setiap baris dari DataFrame kamus\_alay menggunakan fungsi iterrows(). Pada setiap iterasi, kodingan memeriksa apakah kata pada kolom pertama (indeks 0) sudah ada di dalam dictionary normalize\_word\_dict. Jika belum, kodingan menambahkan kata tersebut ke dalam dictionary dengan nilai yang sesuai pada kolom kedua. Tujuan dari kodingan ini adalah untuk mengubah kata-kata alay menjadi kata-kata normal yang lebih mudah dipahami. Kamus kata-kata alay dapat digunakan untuk memetakan kata-kata alay ke dalam kata-kata normal yang sesuai. Dalam pemrosesan bahasa alami, kamus kata-kata alay dapat digunakan untuk meningkatkan kinerja model pembelajaran mesin dalam tugas-tugas seperti klasifikasi teks dan analisis sentimen.

Setelah proses tokenisasi selesai, contoh kalimat yang telah melalui proses tokenisasi akan terlihat pada tabel 3.5

Tabel 3.5 Contoh Hasil Normalisasi

|  |  |
| --- | --- |
| Komentar | Hasil Normalisasi Tidak Baku |
| “versi” “yg” “skrg” “byk” “bugnya” “ndk” “bs” “streaming” “video” “dengan” “lancer” “walaupun” “sdh” “beli” “paket” “premium” “parah” “bgt” “dah” | “versi” “yang” “sekarang” “banyak” “bugnya” “tidak” “bisa” “streaming” “video” “dengan” “lancar” “walaupun” “sudah” “beli” “paket” “premium” “parah” “banget” “sudah” |
| “parah” “nih” “aplikasi” “udh” “buffering” “mulu” “terus” “mau” “nonton” “live” “tv” “emang” “harus” “beli” “paket” “dulu” “yah” “lawak” “sih” “klo” “emng” “gitu” “pelit” “bener” “gak” “kayak” “aplikasi” “video” | “parah” “nih” “aplikasi” “udh” “buffering” “mulu” “terus” “mau” “nonton” “live” “tv” “emang” “harus” “beli” “paket” “dulu” “iya” “lawak” “sih” “kalau” “memang” “gitu” “pelit” “bener” “gak” “kayak” “aplikasi” “video” |
| “kenapa” “sekarang” “app”  “tidak” “dapat” “di” “ gunakan” “sama” “sekali” “sinyal” “cuma" “muter” “buffering” “terus” | “kenapa” “sekarang” “app”  “tidak” “dapat” “digunakan” “sama” “sekali” “sinyal” “cuma" “memutar” “buffering” “terus” |
| “lemot” “banget” “saat” “sy” “tulis” “ulasan” “ini” “pun” “film” “nya” “tidak” “bisa” “diputar” “padahal” “sudah” “berlangganan” “premium” “lalu” “suka” “tiba2” “putus” “ditengah” “film” “mengecewakan” “untuk” “ukuran” “aplikasi” “berbayar” | “lemot” “banget” “saat” “saya” “tulis” “ulasan” “ini” “pun” “film” “nya” “tidak” “bisa” “diputar” “padahal” “sudah” “berlangganan” “premium” “lalu” “suka” “tiba-tiba” “putus” “ditengah” “film” “mengecewakan” “untuk” “ukuran” “aplikasi” “berbayar” |

* 1. ***Dataset Splitting***

Data selanjutnya dibagi menjadi data *training*, *validation* dan *testing*. Sebanyak 8000 data yang ada, data dibagi dengan komposisi 80% atau sebanyak 5599 data *training*, 13,4 % atau sebanyak 1608 data validasi, dan 6.6% atau sebanyak 794 data *testing*. Data *training* digunakan untuk melatih dan mengelola algoritma dalam melakukan klasifikasi dari analis sentimen. Data validasi digunakan untuk meminimalisir *overfitting* yang sering terjadi pada *Neural Network* , *Overfitting* adalah suatukedaan dimana data yang digunakan pada data training memiliki prediksi yang terlalu baik, namun prediksinya buruk pada data testing. Sedangkan, data *testing* akan digunakan sebagai penguji seberapa baik metode BERT dapat mengklasifikasi sebuah data.

train\_set, val\_set = train\_test\_split(df\_v2, test\_size=0.3, stratify=df\_v2.category, random\_state=1)

val\_set, test\_set = train\_test\_split(val\_set, test\_size=0.33, stratify=val\_set.category, random\_state=1)

Kodingan tersebut merupakan implementasi dari fungsi train\_test\_split() dari library scikit-learn pada Python. Fungsi ini digunakan untuk membagi dataset menjadi dua subset, yaitu training set dan test set. Pada kodingan tersebut, dataset df\_v2 dibagi menjadi tiga subset, yaitu train\_set, val\_set, dan test\_set.

Pertama, dataset df\_v2 dibagi menjadi train\_set dan val\_set dengan rasio 70:30 menggunakan fungsi train\_test\_split(). Parameter test\_size diatur sebesar 0.3 untuk membagi dataset menjadi 30% untuk val\_set. Parameter stratify diatur dengan nilai df\_v2.category untuk memastikan bahwa proporsi kategori pada dataset train\_set dan val\_set sama. Parameter random\_state diatur dengan nilai 1 untuk memastikan hasil yang sama setiap kali fungsi ini dijalankan.

Selanjutnya, subset val\_set dibagi menjadi val\_set dan test\_set dengan rasio 67:33 menggunakan fungsi train\_test\_split(). Parameter test\_size diatur sebesar 0.33 untuk membagi dataset menjadi 33% untuk test\_set. Parameter stratify diatur dengan nilai val\_set.category untuk memastikan bahwa proporsi kategori pada dataset val\_set dan test\_set sama. Parameter random\_state diatur dengan nilai 1 untuk memastikan hasil yang sama setiap kali fungsi ini dijalankan.

* 1. **Implementasi BERT**

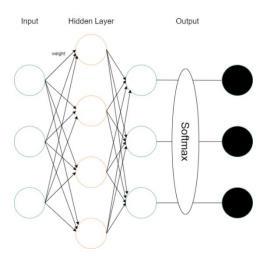
Pada penelitian ini menggunakan Model *pre-trained* IndoBERT yang merupakan arsitektur yang secara khusus dilatih menggunakan data *corpus* bahasa Indonesia. Data set yang digunakan mencapai 4 milyar kata, baik bahasa informal maupun formal dengan 12 korpus bahasa Indonesia yang berbeda. Dataset ini kemudian dilatih dengan arsitektur BERT standard yang memiliki 12 *transformers layers.*

Model ini dipilih dalam penelitian ini karena model dari IndoBERT tersebut mendukung 104 bahasa termasuk Bahasa Indonesia. Pada penelitian ini menggunakan *library transformers* yang telah disediakan oleh *HuggingFace* yang berisi sejumlahbesar model terlatih dimana dapat digunakan untuk melakukan berbagai macam tugas seperti klasifikasi, ekstraksi informasi, tanya jawab, *summarization*, translasi, *text generation* sebagai bantuan untuk proses yang panjang. Dua perpustakaanpembelajaran mendalam yang paling populer adalah *PyTorch* dan *TensorFlow*. *Transformers* menggunakan kedua *library* ini untuk membuat model pembelajaranmesin yang andal. Kemudian melakukan penyesuaian dataset sesuai format agar dapat dilakukan pelatihan atau finetuning. Proses penyusunan kalimat yang akan dihasilkan input pada BERT akan dilakukan oleh tokenizer, dengan tahapan-tahapan berikut :

1. Semua Kalimat ditokenisasi dan dipecah menjadi sebuah kata menggunakan WordPiece. *Tokenizer* akan memeriksa apakah setiap kata dalam kalimat tersebut berada di dalam kosa kata atau di luar kosa kata (OOV). *Tokenizer* akan memecah kata menjadi sub-kata dengan kemungkinan kemunculan paling signifikan dalam kosakata. Jika *tokenizer* tidak menemukan sub-kata dalam kosakata, kata tersebut akan dipecah menjadi per-karakternya. Kata-kata ini dipecah menjadi sub-kata atau karakter dengan simbol ##. Namun jika kata yang panjang diubah menjadi sub-kata atau karakter individu, akan terjadi kelebihan beban atau *overload*, dan kata tersebut akan diganti dengan token [UNK] atau *unknown*. BERT melakukan ini karena dua alasan, pertama untuk mempercepat pemrosesan dan mengurangi jumlahnya parameter yang harus dilatih. Kedua adalah untuk mengatasi masalah *out-of-vocabulary*.
2. Tahap selanjutnya adalah *token embedding*, yaitu penambahan *token* [CLS] di awal kalimat sebagai indikator awal kalimat saat dilakukan klasifikasi sentimen. *Token* [SEP] di akhir kalimat berfungsi untuk memisahkan satu kalimat dari kalimat berikutnya.
3. Tahap selanjutnya adalah menyesuaikan panjang kalimat dengan panjang maksimal yang telah ditentukan. Pada tahap ini dilakukan pengurangan token jika kalimat melebihi panjang maksimum atau menambahkan token [PAD] jika kalimat kurang dari panjang maksimum. *Encoder* pada *transformer* hanya menghasilkan output dengan dimensi 512 saja. Jikapanjang hukuman lebih dari panjang maksimal yang telah ditentukan,Efeknya akan dikurangi (truncate). Jika panjang rangkaian kalimat kurang dari panjang maksimal yang telah ditentukan, maka kalimat akan ditambah dengan *padding*.
4. Tahap selanjutnya adalah *token* disesuaikan dengan nomor unik atau id token sesuai dengan *vocabulary*. model *pre-trained* BERT hanya dapat menerima ID token. Dalam model IndoBERT, token [PAD] memiliki token ID 0, token [UNK] memiliki token ID 1, token [CLS] memiliki token ID 2, dan token [SEP] memiliki token ID 3. Begitu juga token-token lain dalam *vocabulary* masing-masing sudah memiliki token mereka sendiri.Pemberian *attention mask* berfungsi untuk membedakan antara token yang harus diperhatikan dan tidak perlu diperhatikan. *Attention mask* akan menunjukkan posisi indeks yang diberi *padding* sehingga model tidak perlu memproses token tersebut. Tahap ini dapat dilakukan dengan memberikan angka 1 pada token yang harus diperhatikan dan memberikan angka 0 pada *padding*.

|  |  |  |
| --- | --- | --- |
|  | Tabel 3.6 Contoh Hasil Tahapan Tokenizer | |
|  |  |  |
| Tahap |  | Hasil |
|  |  |  |
| Tahap 1 |  | Versi yg skrg byk bugnya. Ndk bs streaming video dengan lancar, walaupun sdh beli paket premium. Parah bgt dah... |
|  |  |  |
| Tahap 2 |  | [CLS], versi yg skrg byk bugnya ndk bs streaming video dengan lancar walaupun sdh beli paket premium parah bgt dah, [SEP] |
| Tahap 3 |  | [CLS], versi yang sekarang banyak bugnya tidak bisa streaming video dengan lancar walaupun sudah beli paket premium parah banget dah, [SEP], [PAD], |
|  |  | [PAD],[PAD], [PAD] |
|  |  |  |
| Tahap 4 |  | 2, 1841, 1339, 4844, 3, 0, 0, 0,0  1,1,1,1,1, 0,0,0,0  1,2,3,4,5,6,7,8,9 |
|  |  |  |

BERT akan menerima urutan kata-kata atau sebuah kalimat sebagai *input* yang akan terus melalui tumpukan *encoder*. Tiap *encoder* mengaplikasikan self-*attention* dan memberikan *output* melalui *feed-forward network* yang kemudiandilanjutkan oleh *encoder* selanjutnya. Pada penelitian ini, model yang dipilih adalah BERTBASE sehingga proses ini berlanjut sebanyak dua belas kali. Setelah melewati semua *encoder*, tiap token per posisi memberikan *output* berupa vektor dengan *hidden size* yaitu 768 pada BERTBASE. Untuk proses analisis sentimen, *output* yangdiperhatikan adalah *output* dari posisi pertama yaitu token [CLS]. Vektor tersebut digunakan sebagai input untuk *classifier*. BERT dapat mencapai hasil yang baik hanya dengan menggunakan *neural network* tunggal sebagai classifier-nya (Devlin et al., 2019). *Layer* yang digunakan untuk klasifikasi adalah *fully connected neural network* dengan fungsi *softmax*. Sehingga *output* dari BERT yang digunakan untukklasifikasi berasal vektor token [CLS] karena token [CLS] dianggap melakukan pengumpulan rata-rata atas token kata untuk mendapatkan vektor dari kalimat. *Layer* terakhir pada *classifier layer* menghasilkan *logits*. *Logits* adalah *output* yangberupa prediksi probabilitas kasar dari kalimat yang akan diklasifikasikan. *Softmax* akan mengubah *logits* tersebut menjadi probabilitas dengan mengambil eksponen dari tiap nilai *logit* sehingga total probabilitasnya adalah tepat 1. Sehingga nilai probabilitas akan berada di antara 0 atau angka positif.



Gambar 3.3 Ilustrasi *Layer* untuk Analisis Sentimen

Misalkan terdapat vektor logits = yang harus dikonversi ke distribusi probabilitas dari sebuah komentar agar dapat diklasifikasi ke dalam kategori positif, negatif, dan netral. Maka, langkah untuk mendapatkan probabilitas yaitu:

1. Hitung seluruh eksponensial dari setiap elemen pada vektor.

(3.1)

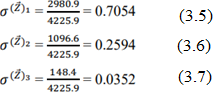
(3.2)

(3.3)

1. Normalisasi nilai dengan menjumlahkan semua eksponensial.

(3.4)

1. Bagi eksponensial dari setiap elemen dengan normalisasi untuk Memperoleh output softmax dari tiap elemen yang ada.



Untuk memastikan apakah hasil prediksi tiap dari tiap probabilitas adalah 1, maka semua probabilitas dijumlahkan

(3.8)

Hasil prediksi probabilitas tersebut akan menunjukkan probabilitas dari komentar ke kategori sentimen. Berarti komentar tersebut memiliki probabilitas sebesar 0.7054 kepada kategori sentimen positif, sedangkan terhadap sentimen negatif sebesar 0.2594, dan sentimen netral sebesar 0.0352. Hal ini juga dilakukan kepada semua komentar yang ada pada dataset. Pada penelitian ini terdapat 3 kategori sentimen yaitu positif, netral, dan negatif. Selain itu, dibutuhkan pula dropout untuk membuat jaringan tidak mengalami *overfitting*. *Overfitting* adalah suatu masalah yang timbul ketika data di-training, loss berkurang dan akurasi meningkat. Tetapi akurasi yang diperoleh saat testing tidak meningkat atau terus menurun.

Setelah BERT *di-pre-train* dengan dataset yang ada, maka dilakukan *training*. Sebelum melakukan *training*, dataset terlebih dahulu dibagi menjadi tigajenis yaitu dataset untuk *training*, validasi, dan *testing*. Karena dataset yang digunakan cukup banyak, memasukkan dataset ke dalam memori sekaligus dapat membuat memori berhenti bekerja dan memperlambat program. Oleh karena itu

dibutuhkan sebuah data loader untuk membantu menghemat memori dan meningkatkan kecepatan selama training. *DataLoader* merupakan fungsi pada *PyTorch* yang berperan sebagai iterator. *DataLoader* akan menggabungkan datasetdan sample untuk iterasi. Karena dataset dibagi menjadi tiga, maka terdapat juga 3 data *loader* yaitu data *loader* untuk *training*, data *loader* untuk validasi, dan data *loader* untuk *testing*. Seperti pada paper aslinya (Devlin et al., 2019), *optimizer* yang digunakan adalah AdamW yang bertujuan untuk mengoreksi *weight* dari kalimat.

* 1. **Memuat Model *Pre-trained***

Penelitian ini menggunakan *pre-trained* model indobenchmark/indobert-base-p1 dari indoNLU.*Pre-trained* model dari indoNLU diakses melalui Hugging-Face. Model *pre-trained* untuk penelitian ini menggunakan kelas model *BertForSequenceClasssification* dari *library transformers* milik *HuggingFace*

tokenizer = BertTokenizer.from\_pretrained('indobenchmark/indobert-base-p1')

config = BertConfig.from\_pretrained('indobenchmark/indobert-base-p1')

Kodingan tersebut merupakan implementasi dari library transformers pada Python untuk menggunakan model IndoBERT. Pada kodingan tersebut, dilakukan dua hal yaitu:

1. Membuat tokenizer dengan menggunakan fungsi BertTokenizer.from\_pretrained('indobenchmark/indobert-base-p1'). Fungsi ini digunakan untuk membuat tokenizer yang akan digunakan untuk memproses teks pada dataset. Tokenizer ini akan mengubah teks menjadi token-token yang dapat diproses oleh model. Parameter 'indobenchmark/indobert-base-p1' digunakan untuk memilih model IndoBERT yang akan digunakan.
2. Membuat konfigurasi model dengan menggunakan fungsi BertConfig.from\_pretrained('indobenchmark/indobert-base-p1'). Fungsi ini digunakan untuk membuat konfigurasi model yang akan digunakan. Konfigurasi ini akan menentukan arsitektur model, jumlah layer, jumlah head, ukuran embedding, dan lain-lain. Parameter 'indobenchmark/indobert-base-p1' digunakan untuk memilih model IndoBERT yang akan digunakan.

Dengan menggunakan tokenizer dan konfigurasi model yang sesuai, kita dapat memproses teks pada dataset dan menggunakan model IndoBERT untuk melakukan tugas-tugas pemrosesan bahasa alami seperti klasifikasi teks, analisis sentimen, dan lain-lain.

* 1. ***Training***

*Training* didapatkan dari proses pembagian data sebelumnya yaitu *train data*, *validation data* dan *testing data*. *Data training* di *load* sesuai *hyperparameter tuning* yang telah ditetapkan sebelumnya yaitu *batch size* sebesar 32, *batch size* adalah jumlah *sample* yang dimasukkan ke dalam *network* sebelum *weight* disesuaikan. Jumlah *worker* sebanyak 16, maksimal dari panjang *sequence* adalah 512 dimana 512 adalah dimensi panjang token yang bisa diterima oleh BERT. Selama proses pelatihan, akan dilakukan pemantauan terhadap *loss* dan akurasi dari *data training* dan data validasi pada setiap *epoch*. Variasi *epoch* yang digunakan yaitu 5 dikarenakan Semakin besar batch size maka semakin lama waktu yang dibutuhkan untuk menyelesaikan satu batch, *epoch* adalah jumlah berapa kali jaringan melihat seluruh *dataset*. Proses pelatihan akan dijalankan menggunakan Google Collab. Data latih menggunakan menggunakan *learning rate* 3e-6, *learning rate* menentukan seberapa banyak *weight* pada *neural network* yang akan diubah. Pada tahapan sebelumnya, didapatkan hasil pencarian parameter yang optimal. Hasil pencarian ini berguna untuk menyesuaikan penyetelan parameter. Perintah untuk melakukan *training* menggunakan perintah berikut :

for epoch in range(n\_epochs):

model.train()

torch.set\_grad\_enabled(True)

total\_train\_loss = 0

list\_hyp\_train, list\_label = [], []

train\_pbar = tqdm(train\_loader, leave=True, total=len(train\_loader))

Dari kodingan diatas berikut ini adalah penjelasan fungsi dari tiap function yang ada di kode tersebut :

for epoch in range(n\_epochs): adalah loop yang digunakan untuk melatih model dalam beberapa epoch. n\_epochs adalah jumlah epoch yang ditentukan sebelumnya.

model.train() digunakan untuk mengatur model dalam mode pelatihan.

torch.set\_grad\_enabled(True) mengaktifkan perhitungan gradien saat melatih model.

total\_train\_loss = 0 digunakan untuk menginisialisasi total loss pelatihan.

list\_hyp\_train dan list\_label adalah list kosong yang akan digunakan untuk menyimpan hasil prediksi dan label pada setiap batch pelatihan.

train\_pbar = tqdm(train\_loader, leave=True, total=len(train\_loader)) digunakan untuk membuat progress bar saat melatih model dengan menggunakan train\_loader sebagai iterator.

.

Dengan menggunakan kodingan tersebut, model akan dilatih dalam beberapa epoch dengan menghitung loss pada setiap batch pelatihan. Selain itu, dataset val\_set akan dibagi menjadi val\_set dan test\_set untuk digunakan dalam evaluasi model.

* 1. **Visualisasi Frekuensi Data**

Pada tahap ini akan menampilkan frekuensi kata yang paling sering muncul pada dataset yang ada dengan batasan maksimal 20 kata. Untuk menampilkan hasil frekuensi data menggunakan perintah seperti berikut:

plt.style.use('default')

sns.set(style='ticks', palette='Set2')

mpl.rcParams['axes.titlesize'] = 20

mpl.rcParams['axes.titlepad'] = 20

f, ax1 = plt.subplots(1,figsize=(15,5))

sns.barplot(x='word', y='freq', data=corpus\_freq, ax=ax1)

ax1.set\_title('Word Frequency in Train Data')

ax1.tick\_params(axis='x', rotation=45)

Kode diatas digunakan digunakan untuk membuat visualisasi word cloud berdasarkan frekuensi data. Berikut penjelasan dari setiap baris kode:

plt.style.use('default')

Baris ini mengatur gaya plot menjadi gaya default.

sns.set(style='ticks', palette='Set2')

Baris ini mengatur gaya plot seaborn menjadi 'ticks' dan palet warna menjadi 'Set2'.

mpl.rcParams['axes.titlesize'] = 20

mpl.rcParams['axes.titlepad'] = 20

Baris ini mengatur ukuran judul dan jarak judul pada sumbu x dan y.

f, ax1 = plt.subplots(1,figsize=(15,5))

Baris ini membuat sebuah figure dengan ukuran 15x5.

sns.barplot(x='word', y='freq', data=corpus\_freq, ax=ax1)

Baris ini membuat bar plot menggunakan seaborn dengan sumbu x adalah 'word', sumbu y adalah 'freq', dan data yang digunakan adalah corpus\_freq. Plot ini akan ditampilkan pada ax1.

ax1.set\_title('Word Frequency in Train Data')

Baris ini mengatur judul pada ax1 menjadi 'Word Frequency in Train Data'.

ax1.tick\_params(axis='x', rotation=45)

Baris ini mengatur rotasi label pada sumbu x sebesar 45 derajat.

plt.show()

Baris ini menampilkan plot yang telah dibuat.

Kode tersebut akan menghasilkan sebuah bar plot yang menampilkan frekuensi kata-kata dalam data train. Plot ini menggunakan gaya default, gaya plot seaborn 'ticks', dan palet warna 'Set2'. Judul plot adalah 'Word Frequency in Train Data', dan label pada sumbu x akan diputar sebesar 45 derajat. Plot tersebut akan ditampilkan dengan ukuran 15x5.

* 1. **Visualisasi *Wordcloud***

Data yang ditampilkan pada *word cloud* adalah data dalam bentuk 20 kata yang paling sering muncul ada pada *dataset*. Proses visualisasi data *word cloud* dilakukan menggunakan perintah seperti berikut ini:

wordcloud = WordCloud(max\_font\_size=50, max\_words=100, background\_color="white").generate(corpus)

plt.figure()

plt.imshow(wordcloud, interpolation="bilinear")

plt.axis("off")

plt.show()

Kode diatas digunakan untuk membuat visualisasi Word Cloud. Berikut adalah penjelasan dari setiap baris kode tersebut:

wordcloud = WordCloud(max\_font\_size=50, max\_words=100, background\_color="white").generate(corpus)

Baris ini menciptakan objek WordCloud dengan parameter yang telah ditentukan. max\_font\_size menetapkan ukuran font maksimum untuk kata-kata dalam Word Cloud, max\_words menetapkan jumlah maksimum kata yang akan ditampilkan, dan background\_color menetapkan warna latar belakang dari Word Cloud. Variabel corpus adalah teks masukan yang digunakan untuk membuat Word Cloud.

plt.figure()

plt.imshow(wordcloud, interpolation="bilinear")

plt.axis("off")

plt.show()

Baris-baris ini menciptakan sebuah gambar (figure), menampilkan gambar Word Cloud menggunakan imshow(), menghilangkan label sumbu (axis) dengan axis("off"), dan akhirnya menampilkan Word Cloud menggunakan show().

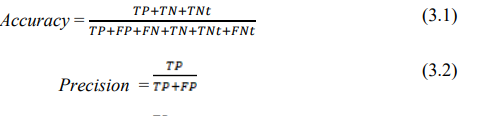
Kode di atas akan menghasilkan visualisasi Word Cloud berdasarkan teks masukan (corpus). Word Cloud akan menampilkan kata-kata yang paling sering muncul dalam teks tersebut, dengan ukuran kata yang lebih besar menunjukkan frekuensi yang lebih tinggi.

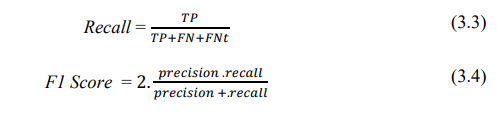
* 1. **Evaluasi**

Tahap evaluasi bertujuan untuk mendapatkan prediksi dari performa model menggunakan *confusion matrix* yang didapatkan dari tahap training sebelumnya*.* seperti yang ditunjukkan pada Tabel 3.7. (Epoch Harus di masukin di bab 2).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | Predicted Class |  |  |
|  |  |  |  |  |  |
|  |  | Positive | Neutral | Negative |  |
|  |  |  |  |  |  |
| True | Positive |  |  |  |  |
| Class |  | True Positive | False Neutral | False Negative |  |
|  |  |  |  |  |
|  |  | (TP) | (FNt) | (FN) |  |
|  |  |  |  |  |  |
|  | Neutral | False Positive | True Netral | False Negative |  |
|  |  | (FP) | (TNt) | (FN) |  |
|  |  |  |  |  |  |

Nilai *akurasi, presisi, recall,* dan *F1-score* juga dapat diperoleh setelah mendapatkan nilai *confusion matrix*. Akurasi adalah perbandingan prediksi Benar (Positif, Netral, dan Negatif) dengan keseluruhan data. *Precision, recall,* dan *F1-score* dibagi dengan jumlah kelas. Sedangkan *F1-score* merupakan perbandinganberbobot dari rata-rata *presisi* dan *recall*. Pada penelitian ini menggunakan fungsi *classification\_report()* dari *library Sklearn* untuk menghitung proses evaluasi.Rumus perhitungan untuk mendapatkan *akurasi, presisi, recall,* dan *F1-score* ditunjukkan pada persamaan berikut :





Proses uji coba dengan *confusion matrix* dilakukan menggunakan perintah berikut :

def show\_confusion\_matrix(confusion\_matrix):

hmap = sns.heatmap(confusion\_matrix, annot=True, fmt="d", cmap="Blues")

hmap.yaxis.set\_ticklabels(hmap.yaxis.get\_ticklabels(), rotation=0, ha='right')

hmap.xaxis.set\_ticklabels(hmap.xaxis.get\_ticklabels(), rotation=30, ha='right')

plt.ylabel('True sentiment')

plt.xlabel('Predicted sentiment');

Hasil dari *confusion matrix* yang ada pada penelitian ini menunjukkan bahwa model memiliki kekurangan dalam hal memprediksi sentimen netral, namun untuk sentimen positif dan negatif sudah cukup baik.

Setelah dilakukan *confusion matrix* maka langkah selanjutnya yaitu menghitung *accuracy, precision, recall,* dan *F1-Score* menggunakan fungsi *classification\_report()* dan sklearn. untuk menampilkan hasil perhitunganmenggunakan perintah seperti berikut :

cm = confusion\_matrix(test\_real, test\_pred)

df\_cm = pd.DataFrame(cm, index=['positive', 'neutral', 'negative'], columns=['positive', 'neutral', 'negative'])

show\_confusion\_matrix(df\_cm)

Kode ditas digunakan untuk menghasilkan confusion matrix. Berikut penjelasan dari setiap baris kode:

cm = confusion\_matrix(test\_real, test\_pred)

Baris ini menghitung confusion matrix berdasarkan prediksi (test\_pred) dan nilai sebenarnya (test\_real). Confusion matrix adalah tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan nilai prediksi dengan nilai sebenarnya.

df\_cm = pd.DataFrame(cm, index=['positive', 'neutral', 'negative'], columns=['positive', 'neutral', 'negative'])

Baris ini membuat DataFrame menggunakan confusion matrix yang telah dihitung sebelumnya. DataFrame ini akan memiliki indeks 'positive', 'neutral', dan 'negative', serta kolom 'positive', 'neutral', dan 'negative'.

show\_confusion\_matrix(df\_cm)

Baris ini memanggil fungsi show\_confusion\_matrix dengan argumen DataFrame df\_cm. Fungsi ini akan menampilkan confusion matrix dalam bentuk visual.

Kode tersebut akan menghasilkan confusion matrix berdasarkan prediksi dan nilai sebenarnya. Confusion matrix ini akan ditampilkan dalam bentuk visual menggunakan DataFrame. Indeks dan kolom DataFrame akan diberi label 'positive', 'neutral', dan 'negative'.

**Bab 4**

**HASIL DAN PEMBAHASAN**

* 1. **Evaluasi Hasil**

Pada penelitian ini, penulis training data menggunakan 8000 data dikarenakan dengan menggunakan jumlah data tersebut membuat *accuracy* sentimen. Untuk mengkategorikan data positif, data netral dan data negative penulis melakukan *data labelling*.Setelah itu penulis melakukan data preprocessing untuk membersihkan symbol atau emoji di kolum review text di *dataframe.*Tahap selanjutnya yaitu implementasi BERT untuk memproses data testing menjadi data prediction.Penulis menggunakan *n\_epoch* sebanyak 4 kali dikarenakan dengan melaukan processing epoch sebanyak 4 kali menghasilkan tingkat akurasi yang cukup baik. Setelah dilakukan tahap implementasi BERT maka tahap selanjutnya yaitu tahap evaluasi. Sebanyak 792 data *testing* yang dilakukan dari proses pengujian menggunakan model BERT menghasilkan prediksi komentar *true positive* sebanyak 418 data, *true neutral* sebanyak 71 data, dan *true negative* sebanyak 303 data. Hasil prediksi beberapa data *testing* dapat dilihat pada tabel-tabel berikut.

Tabel 4.1 *True Positive*

|  |  |  |
| --- | --- | --- |
| Komentar | Sentimen | Prediksi |
| good aplikasi . jangan ragu untuk mendownload aplikasi ini . semua saluran ada disini . tingkatkan terus kualitasnya . mantap | Positive | Positive |
| lumayan buat cadangan menonton tv disaat mati listrik , tapi masih bisa menonton di vision , senang banget | Positive | Positive |
| puas dengan vision tidak mengecewakan , fitur lengkap dan mudah penggunaannya | Positive | Positive |
| request supaya app bisa berjalan di background dong min seperti youtube . menghemat baterai kan keren | Positive | Positive |
| dari banyak aplikasi untuk menonton siaran tv , vision terbaik menurut sy.bs menonton siaran ulang nya dan siaran langsung tentu nya dan gratis | Positive | Positive |

Tabel diatas merupakan contoh dari hasil *true positive* yang dihasilkan dari *testing.* *True positive* merupakan kalimat positif yang diprediksi positif oleh model.

Contoh *True Neutral*

|  |  |  |
| --- | --- | --- |
| Komentar | Sentimen | Prediksi |
| kenapa bintang , karena kalua menonton champion atau pertandingan bola yang sekiranya big match ypasti susah masuk ke chenel tv tersebut | Neutral | Neutral |
| maaf , . saya kasih bintang dulu , soalnya mau cek dulu antv lancar apa enggak , soalnya kalo saya download apps tv lain selalu antv enggak bisa lancar alias hilang . kalo lancar saya kasih bintang . | Neutral | Neutral |
| loding masih terlalu berat , mohon tampilan di tv android dibuat memudahkan untuk ganti chanel , agar tidak harus keluar masuk saat akan ganti chanel tv . semoga kedepan loading lebih ringan dan fitur memudahkan pengguna | Neutral | Neutral |
| bisa ditayangkan di tv android saya mohon tingkatkan lagi kedepannya dari kualitas dan lainnya | Neutral | Neutral |
| membutuhkan koneksi yang benar kuat untuk bisa menonton film | Neutral | Neutral |

Tabel diatas merupakan contoh dari hasil *true neutral* yang dihasilkan dari *testing.* *True neutral* merupakan kalimat netral yang diprediksi netral oleh model.

Contoh *True Negative*

|  |  |  |
| --- | --- | --- |
| Komentar | Sentimen | Prediksi |
| katanya sudah di perbaiki tapi masih saja kayak tadi kalau begini gua enggak suka lagi dukung timnas .. mengapai di dukung lihat mereka main saja enggak bisa. | Negative | Negative |
| sudah di perbarui masih saja enggak bisa di buka. | Negative | Negative |
| aplikasi busuk enggak bisa dibuka ! uninstall saja | Negative | Negative |
| sangat buruk susah banget buat berlangganan lagi tolong di perbaiki kembali sistematis pelayanannya | Negative | Negative |
| sudah kecewa banget , ku kira bisa menonton siaran ulang di tv nasional yang terlewat malah error , tolong diperbaiki lah , kalau aplikasi enggak berguna buat apa orang tertarik beli premium nya ? | Negative | Negative |

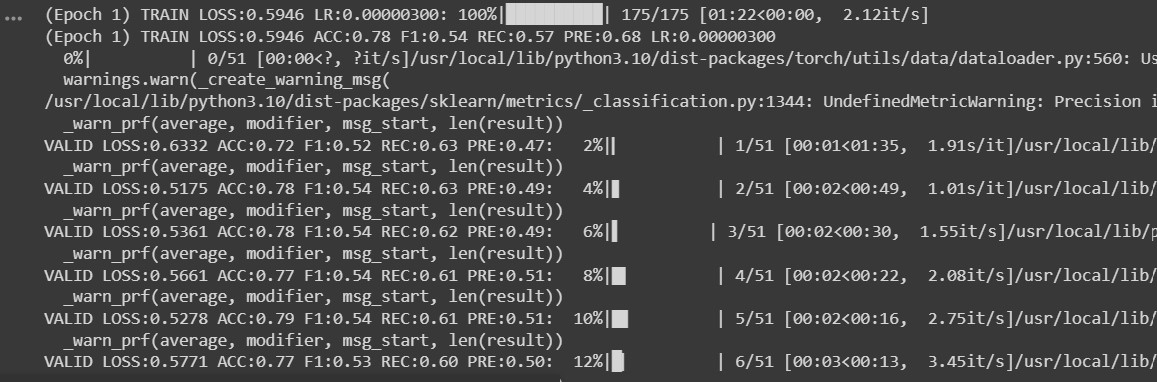
Tabel diatas merupakan contoh dari hasil *true negative* yang dihasilkan dari *testing.*

*True negative* merupakan kalimat negatif yang diprediksi negatif oleh model.

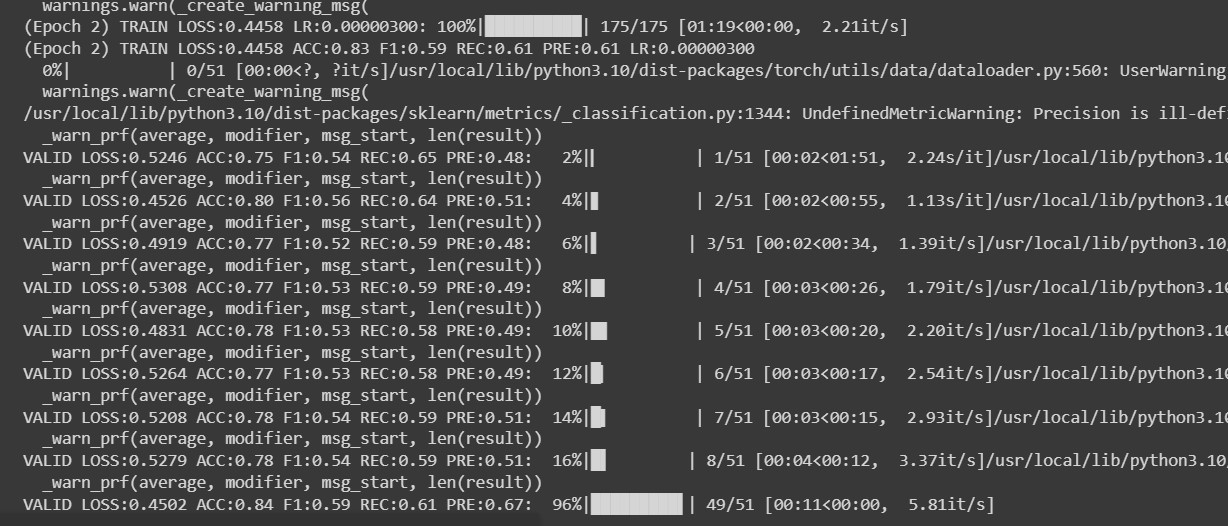
* 1. **Uji Akurasi**

Tahap ini akan membahas tentang hasil uji coba model yang telah dilatih. Pada tahap ini akan ditampilkan hasil pelatihan model dan hasil uji coba model dengan data uji yang akan ditunjukkan dengan *confusion matrix* untuk mengukur seberapa banyak model berhasil memprediksi dengan benar semua sentimen pada *data testing*.

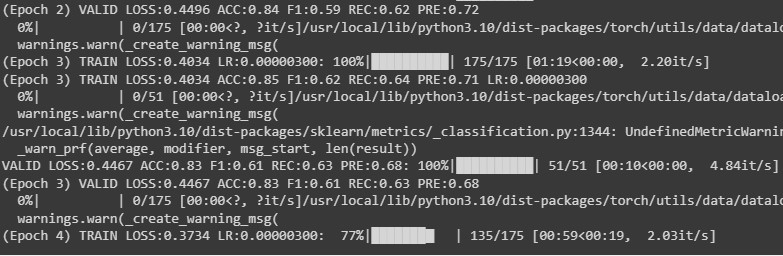
Kemudian akan ditampilkan juga *classification report* yang akan menunjukkan skor *accuracy, precision, recall,* dan *f1-score*. Proses *training* dan evaluasi dapat dilihat pada gambar 4.1.



Gambar 4.1 Proses Uji Akurasi Epoch 1

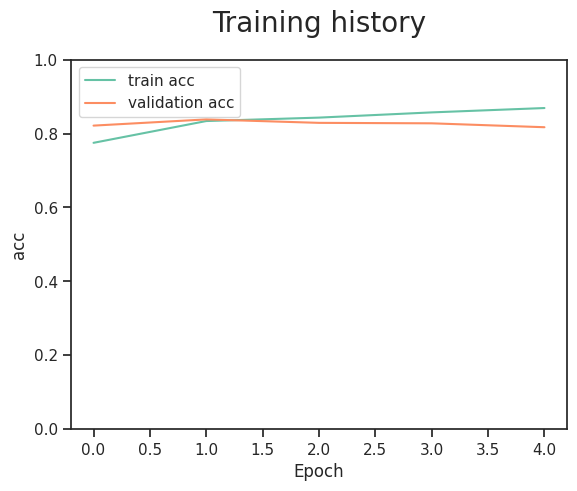


Gambar 4.1 Proses Uji Akurasi Epoch 2



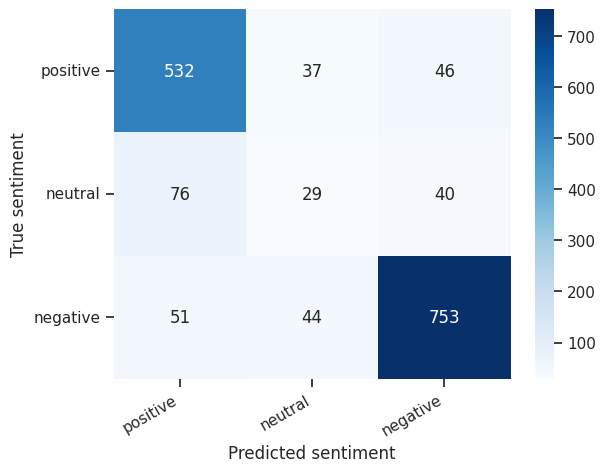
Gambar 4.3 Proses Uji Akurasi Epoch 3

Setelah melalui proses uji coba dengan perulangan per-*epoch*, hasil uji coba akan disimpan. Kemudian akan dibuat *learning curve* sebagai visualisasi perbandingan antara akurasi yang didapat dari *training* dan validasi.



Gambar 4.2 Learning Curve

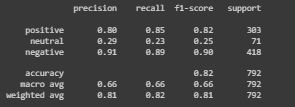
Berdasarkan di hasil learning curve diatas bisa disimpulkan bahwa pada saat epoch pertama sampai epoch keempat mengalami kenaikan dari train acc dari < 0.8 sampai diatas 0.8. Namun pada saat validation acc mengalmi penurunan sejak epoch 1 menjadi < 0.8. Pada gambar 4.2 menghasilkan learing curve yang sesuai atau *Good Fit*. *Good fit* diidentifikasi oleh training dapat beradaptasi ketika diberi dataset baru.Hasil data *training* dan validasi yang telah diuji menunjukkan bahwa data *training* mempunyai hasil akurasi yang lebih tinggi dari pada hasil akurasi data validasi. Kemudian akan dilakukan uji coba kurasi terhadap 396 dataset yang ada menggunakan *confusion matrix* yang hasilnya dapat dilihat pada gambar 4.3



Gambar 4.3 *Confusion Matrix*

Confusion matrix yang didapatkan pada penelitian ini menghasilkan 532 t*rue positive*, *True positive* merupakan kalimat positif yang diprediksi positif olehmodel. 127 *false positive*, *false positive* merupakan kalimat positif yang diprediksi netral atau negatif oleh model. 29 *true neutral, true neutral* merupakan kalimat netral yang diprediksi netral oleh model. 81 *false neutral*, *false neutral* merupakan kalimat netral yang diprediksi positif atau negatif oleh model. 102 *true negative*, *true negative* merupakan kalimat negatif yang diprediksi negatif oleh model. 753 *false negative, false negative* merupakan kalimat negatif yang diprediksi posiitifatau netral oleh model.

Setelah melakukan proses confusion matrix lalu defined amenghasilkan perhitungan *accuracy, precision, recall,* dan *F1-Score* yang ada pada penelitian ini ada pada gambar 4.4



Gambar 4.4 Hasil Evaluasi

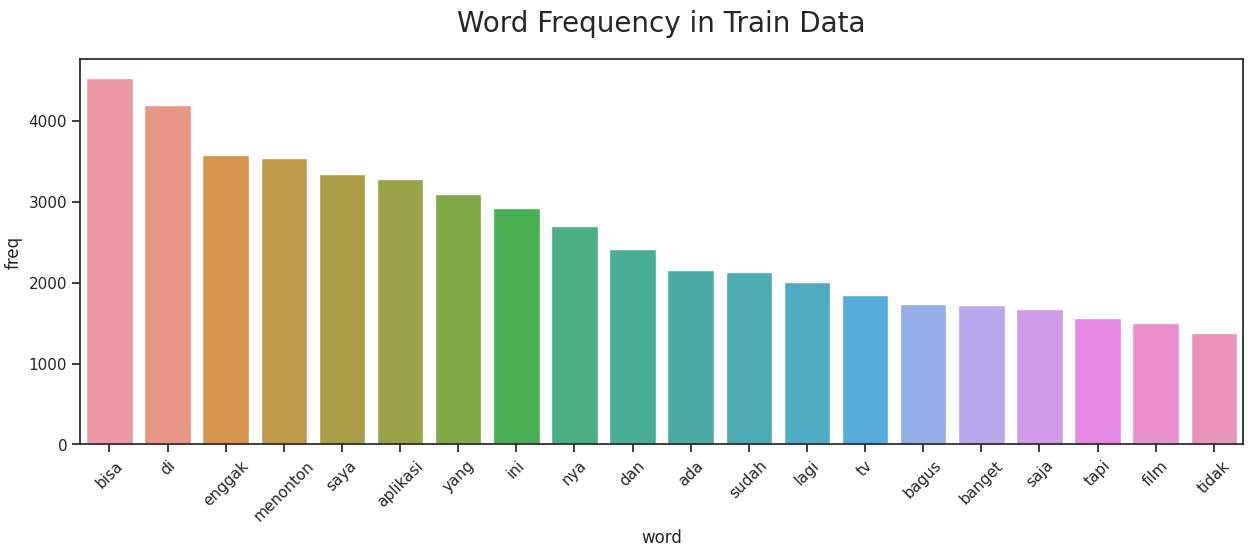
Dari hasil uji coba akurasi yang telah dilakukan, didapatkan hasil akurasi dari metode BERT seperti berikut:

1. Akurasi (Accuracy): Akurasi adalah persentase keseluruhan prediksi yang benar. Untuk menghitungnya, Anda membagi jumlah prediksi yang benar dengan total jumlah prediksi dan mengalikannya dengan 100. Dalam penelitian ini didapat akurasi adalah 80%.
2. Precision: Precision (presisi) mengukur sejauh mana prediksi positif yang dilakukan benar. Ini menggambarkan persentase prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Precision dapat dihitung dengan membagi jumlah prediksi benar positif dengan total jumlah prediksi positif. Anda memberikan nilai precision untuk tiga kelas yang berbeda. Setelah diproses diapat hasil precision kategori positif adalah 80%, untuk kelas netral 29%, dan untuk kelas negatif 91%.
3. Recall: Recall (pemulihan) mengukur sejauh mana prediksi benar positif dibandingkan dengan keseluruhan data yang sebenarnya positif. Ini menggambarkan persentase prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Recall dapat dihitung dengan membagi jumlah prediksi benar positif dengan total jumlah data yang benar positif. Anda memberikan nilai recall untuk tiga kelas yang berbeda. Didapat kelas positif recall adalah 85%, untuk kelas netral 23%, dan untuk kelas negatif 89%.
4. F1-score: F1-score adalah ukuran rata-rata harmonik antara presisi dan recall. Ini memberikan keseluruhan skor yang menggabungkan kedua metrik tersebut. F1-score dapat dihitung dengan menggunakan rumus: 2 \* (precision \* recall) / (precision + recall). Anda memberikan nilai F1-score untuk tiga kelas yang berbeda. Didapat, kelas positif F1-score adalah 82%, untuk kelas netral 25%, dan untuk kelas negatif 81%.

Berdasarkan hasil *accuracy, precision, recall,* dan *F1-Score* diatas bisa dilihat bahwa accuracy darid data negative sangat tinggi dibanding data positif dan negatif. Hal ini bisa terjadi karena data support yang berlabel negatif memiliki jumlah yang lebih banyak dibanding data support berlabel positif dan netral. Pada data netral memiliki tingkat akurasi yang sangat kecil dikarenakan data support yang digunakan jumlahnya sedikit sehingga mempengaruhi tingkat akurasi dari analisis sentimen. Namun

* 1. **Visualisasi Frekuensi Data**

Pada tahap ini akan menampilkan frekuensi kata yang paling sering muncul pada dataset yang ada dengan batasan maksimal 20 kata. Setelah *running* *process* kode visualisasi frekuensi data didapatkan hasil pada gambar 4.5 sebagi berikut:



Gambar 4.5 Visualisasi Frekuensi Data

Dari hasil analisis frekuensi kata dalam dataset, ditemukan bahwa kata "bisa" memiliki frekuensi yang tinggi, dengan jumlah frekuensi di atas 4000 dan menyentuh 5000. Sementara itu, kata "tidak" memiliki jumlah frekuensi 1000. Kata-kata paling sering muncul dalam dataset adalah "saya", "yang", "ini", "dan", "ada", "lagi", "tv", "bagus", "banget", "sudah", "aplikasi", "tapi", "film", "enggakm", dan "nya".

Visualisasi grafik batang membantu memperjelas temuan tersebut dengan menampilkan frekuensi kata-kata paling sering muncul dalam dataset secara visual. Dari visualisasi tersebut, dapat dilihat bahwa kata "saya" memiliki frekuensi tertinggi, diikuti oleh kata "yang" dan "ini". Selain itu, visualisasi grafik batang juga membantu membandingkan frekuensi kata-kata yang berbeda secara langsung.

Berdasarkan hasil analisis, dapat disimpulkan bahwa kata-kata yang paling sering muncul dalam dataset adalah kata-kata yang umum digunakan dalam bahasa sehari-hari. Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa kata-kata umum cenderung memiliki frekuensi yang tinggi dalam dataset.

* 1. **Visualisasi *Word Cloud***

Pada tahap ini akan menampilkan frekuensi kata yang paling sering muncul pada dataset yang ada dengan batasan maksimal 20 kata. Setelah *running* *process* kode visualisasi wordcloud didapatkan hasil pada gambar 4.6 sebagi berikut :



Gambar 4.6 *Wordcloud*

Berdasarkan hasil dari *wordcloud* ada beberapa poin kesimpulan yang didapat sebagai berikut:

1. Frekuensi kata yang tinggi: Dalam hasil analisis frekuensi kata dalam dataset, kata "bisa" memiliki frekuensi yang tinggi, dengan jumlah frekuensi di atas 4000 dan menyentuh 5000. Hal ini menunjukkan bahwa kata "bisa" sering muncul dalam dataset dan memiliki pengaruh yang signifikan dalam konteks yang sedang dianalisis.
2. Kata-kata yang paling sering muncul: Kata-kata paling sering muncul dalam dataset adalah "saya", "yang", "ini", "dan", "ada", "lagi", "tv", "bagus", "banget", "sudah", "aplikasi", "tapi", "film", "enggak", dan "nya". Kata-kata ini dapat memberikan gambaran tentang topik atau tema yang dominan dalam dataset.
3. Visualisasi yang jelas dan mudah dipahami: Visualisasi wordcloud dapat memberikan gambaran visual yang mudah dipahami tentang kata-kata yang paling sering muncul dalam dataset. Ukuran kata dalam *wordcloud* menunjukkan seberapa sering kata tersebut muncul dalam dataset. Semakin besar ukuran kata, semakin sering kata tersebut muncul. Dengan demikian, visualisasi wordcloud dapat membantu pengguna untuk dengan cepat mengidentifikasi kata-kata yang paling penting atau dominan dalam dataset.
4. Kesesuaian dengan tujuan analisis: Evaluasi hasil visualisasi wordcloud juga perlu mempertimbangkan kesesuaian dengan tujuan analisis. Apakah visualisasi wordcloud berhasil menggambarkan frekuensi kata yang paling sering muncul dalam dataset dengan jelas dan sesuai dengan konteks analisis yang sedang dilakukan.

Dengan mempertimbangkan kriteria-kriteria di atas, evaluasi hasil visualisasi *wordcloud* dapat memberikan pemahaman yang lebih baik tentang kata-kata yang paling sering muncul dalam *dataset* dan membantu dalam analisis teks yang lebih mendalam.

* 1. **Ketidakakuratan Hasil Testing**

Pada tahap evaluasi yang didapatkan dari testing menghasilkan confusion matrix 14 nilai *false positive* dimana sentimen bernilai netral tetapi diprediksi positif, 26 nilai *false positive* dimana sentimen bernilai negatif tetapi diprediksi positif. Kemudian terdapat 1 *false neutral* dimana sentimen bernilai positif tetapi diprediksi netral, 3 *false neutral* dimana sentimen bernilai negatif tetapi diprediksi netral. Terakhir, 23 *false negative* dimana sentimen bernilai negatiff tetapi diprediksi positif dan 11 *false negative* dimana sentimen bernilai netral tetapi diprediksi negatif.

Hal ini dapat terjadi dikarenakan model menilai komentar berdasarkan pola dari data training sehingga prediksi model bisa berbeda dengan labelisasi.

Contoh hasil *False Positive True*

|  |  |  |
| --- | --- | --- |
| Komentar | Sentimen | Prediksi |
|  |  |  |
| film keren keren | Netral | Positif |
|  |  |  |
| tolong langgan biasa bayar pakai gopay punya | Netral | Positif |
| gopay tolong sudah kasih bintang full |  |  |
|  |  |  |
| kasih star kouta putar film aneh kalau kouta | Netral | Positif |
| flashnya jalan kouta maxsteam bagaimana |  |  |
|  |  |  |

Tabel diatas merupakan contoh dari nilai *false positive* dimana sentimen bernilai netral tetapi diprediksi positif oleh model. Hal ini dapat terjadi dikarenakan model menilai komentar berdasarkan pola dari data training sehingga prediksi model bisa berbeda dengan labelisasi.

Contoh hasil *False Positive False*

|  |  |  |  |
| --- | --- | --- | --- |
| Komentar |  | Sentimen | Prediksi |
|  |  |  |  |
| aplikasi segi kualitas gambar | enggak lama | Negatif | Positif |
| premium jangkau sayang update film film baru | |  |  |
| film action film indonesia baru bom | |  |  |
|  |  |  |  |
| sudah bayar |  | Negatif | Positif |
|  |  |  |  |
| aplikasi bagus menonton film | saja masang wifi | Negatif | Positif |
| tau dapat Vision+ gratis pasang wifinya | |  |  |
| keren |  |  |  |
|  |  |  |  |

Tabel diatas merupakan contoh dari nlai *false positive* dimana sentimen bernilai negatif tetapi diprediksi positf oleh model. Hal ini dapat terjadi dikarenakan model menilai komentar berdasarkan pola dari data training sehingga prediksi model bisa berbeda dengan labelisasi.

Contoh Hasil *False Neutral True*

|  |  |  |
| --- | --- | --- |
| Komentar | Sentimen | Prediksi |
|  |  |  |
| putar set top box smart tv subtitle indonesia putar | Positif | Netral |
| smartphone sedia bahasa indonesia |  |  |
|  |  |  |

Tabel diatas merupakan contoh dari nlai *false neutral* dimana sentimen bernilai positif tetapi diprediksi netral oleh model. Hal ini dapat terjadi dikarenakan model menilai komentar berdasarkan pola dari data training sehingga prediksi model bisa berbeda dengan labelisasi.

Contoh hasil *False Neutral False*

|  |  |  |
| --- | --- | --- |
| Komentar | Sentimen | Prediksi |
|  |  |  |
| demon slayer season update | Negatif | Netral |
|  |  |  |
| film jangan | Negatif | Netral |
|  |  |  |
| buka aplikasi tampil muat putar | Negatif | Netral |
|  |  |  |

Tabel diatas merupakan contoh dari nlai *false neutral* dimana sentimen bernilai negatif tetapi diprediksi netral oleh model. Hal ini dapat terjadi dikarenakan model menilai komentar berdasarkan pola dari data training sehingga prediksi model bisa berbeda dengan labelisasi.

Contoh hasil *False Negative True*

|  |  |  |
| --- | --- | --- |
| Komentar | Sentimen | Prediksi |
|  |  |  |
| putar set top box smart tv subtitle indonesia putar | Positif | Netral |
| smartphone sedia bahasa indonesia |  |  |
|  |  |  |
| loading home screen mood menonton | Positif | Netral |
|  |  |  |
| kasih bintang download putar | Positif | Netral |
|  |  |  |

Tabel diatas merupakan contoh dari nlai *false negative* dimana sentimen bernilai positif tetapi diprediksi netral oleh model. Hal ini dapat terjadi dikarenakan model

menilai komentar berdasarkan pola dari data training sehingga prediksi model bisa berbeda dengan labelisasi.

Contoh hasil *False Negative False*

|  |  |  |  |
| --- | --- | --- | --- |
| Komentar |  | Sentimen | Prediksi |
|  | |  |  |
| sudah beli paket kuota maxtream tetap kuota | | Netral | Negatif |
| reguler kepake sampai ambil pulsa habis sudah | |  |  |
| buka tetap pakai kuota reguler |  |  |  |
|  |  |  |  |
| film kadang putar terus loading | suara putus putus | Netral | Negatif |
| eror |  |  |  |
|  |  |  |  |
| bagus kadang sinyal suka lama | buka aplikasi | Netral | Negatif |
| lancar |  |  |  |
|  |  |  |  |

Tabel diatas merupakan contoh dari nlai *false negative* dimana sentimen bernilai netral tetapi diprediksi negatif oleh model. Hal ini dapat terjadi dikarenakan model menilai komentar berdasarkan pola dari data training sehingga prediksi model bisa berbeda dengan labelisasi.

**Bab 5**

**PENUTUP**

* 1. **Kesimpulan**

Penelitian analisis sentimen Aplikasi Vision+ pada Android menggunakan metode BERT atau *Bidirectional Encoder Representations from Transformers* dengan bahasa pemrograman *Python* untuk mengklasifikasikankomentar ke dalam sentimen positif, netral, dan negatif telah berhasil dilakukan. Berdasarkan hasil pengujian yang telah dilakukan, didapatkan hasil akurasi sebesar 82% dari jumlah data komentar sebanyak 8000 data dengan data latih sebanyak 5599 data, data validasi sebanyak 1608 data, dan data uji sebanyak 792 data. Berdasarkan hasil analisis sentimen yang dilakukan menghasilkan akurasi persentase sentimen pengguna Vision+ pada positif sebesar 82%, netral sebesar 23% dan negatif sebesar 90%.

* 1. **Saran**

Adapun saran yang diberikan pada penelitian ini untuk pengembangan lebih lanjut, pengembangan yang dilakukan dapat melakukan pengambilan dataset yang lebih banyak, proses pelabelan dengan user yang lebih memahami kosakata Bahasa Indonesia, dan menggunakan model *Bidirectional Encoder Representations* (BERT) yang lain agar dapat melihat perbedaan hasil akurasi yang didapatkan dan dapat menggunakan *learning rate* yang berbeda.

**Daftar Pustaka**

Putri, C.A., & A.S. (2021). Analisis Sentimen Review Film Berbahasa Inggris dengan Pendekatan Bidirectional Encoder Representations from Transformers. JEISBI.

Asy'ari, M. Z. (2020). Berkenalan Dengan Seaborn Python | Belajar Mudah Visualisasi Data Bagian 0. <https://auftechnique.com/berkenalan-dengan-seaborn-python/>.

Tala, F. Z. (2023). A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. M.Sc. *Thesis, Appendix, pp,* 39-46

Sari, R. P., & Wibowo, A. (2019). Perbandingan Algoritma Machine Learning dan Deep Learning pada Klasifikasi Citra. Jurnal Ilmiah Teknologi Informasi Asia, 13(2), 1-10.

Samudro A. A. (2019). "Normalisasi Teks Bahasa Indonesia pada Media Sosial Berdasarkan FastText Embeddings." Tugas Akhir. Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Institut Teknologi Sepuluh Nopember, Surabaya.

Prasetija, Z.R.N.S., Romadhony, A., & Setiawan, E.B. (2022). Analisis Pengaruh Normalisasi Teks pada Klasifikasi Sentimen Ulasan Produk Kecantikan. e-Proceeding of Engineering, Vol.9, No.3, halaman 1769. ISSN: 2355-9365.

Stedman, C. (2022). "What is data cleansing?" TechTarget. Dalam DEFINITION. Diakses pada 8 Agustus 2023 dari <https://www.techtarget.com/searchdatamanagement/definition/data-scrubbing>.

Pradana, A. W., & Hayaty, R. (2019). "Pengaruh Stopword Removal dan Stemming Terhadap Performa Klasifikasi Teks Komentar Kebijakan New Normal Menggunakan Algoritma LSTM" oleh Muhammad Nurfajri Rahman.

Sari, D. P., & Wijaya, D. T. (2019). "Evaluasi Daftar Stopword Bahasa Indonesia" oleh Dian Purnama Sari dan Dwi Teguh Wijaya.

Putra, H. K., Bijaksana, M. A., Ade Romadhony., (2021). Deteksi Penggunaan Kalimat Abusive Pada Teks Bahasa Indonesia Menggunakan Metode IndoBERT. [PDF] Available at: https://openlibrary.telkomuniversity.ac.id/pustaka/files/167739/jurnal\_eproc/deteksi-penggunaan-kalimat-abusive-pada-teks-bahasa-indonesia-menggunakan-metode-indobert.pdf

Fauzi, A. A., & Wibowo, A. (2021). "Pembentukan Daftar Stopword Menggunakan Term Based Random Sampling Pada Analisis Sentimen Dengan Metode Naïve Bayes (Studi Kasus: Kuliah Daring Di Masa Pandemi)" oleh Ahmad Agil Fauzi dan Agus Wibowo.

Pradana, A. W., & Hayaty, R. (2019). "Analisis Performa Penggunaan Stopwords dan Stemming dalam Sentimen Analisis dengan Pendekatan Klasifikasi Naive Bayes" oleh Aditya Wiha Pradana dan Rizki Hayaty.

Salamah, N. (2019). Penerapan Tokenisasi Kalimat dan Metode TF (Term Frequency) pada Peringkas Teks. Skripsi. UIN Syarif Hidayatullah Jakarta.

Lestiyanto, D. Y. (2014). Stemming Bahasa Indonesia Sebagai Media Belajar Siswa Sekolah Menggunakan Algoritma Porter. Semarang: Universitas Dian Nuswantoro.

Kingma, D.P., & Ba, J. (2015). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kaira, Fitria M. (2022). Implementasi Model Bahasa OpenAI GPT-3 Untuk Aplikasi Text Content Generator Berbasis Web dan Aplikasi Mobile. Fakultas Teknik, Universitas Lampung

Agustin, Risa Yuli. (2022). Analisis Sentimen Ulasan Aplikasi Vlog Now (Vn) Pada Google Play Store Menggunakan Algoritma Naive Bayes. S1 - Sarjana thesis, Universitas AMIKOM Yogyakarta.

Rizki, A. (2020). Preferensi Kaum Milenial Terhadap Aplikasi Digital Streaming (Studi Pada Followers Media Sosial Twitter @drakorfess\_). Universitas Muhammadiyah Makassar.

Julianto, Y., & Setiabudi, D.H. (2014). Analisis Sentimen Ulasan Restoran Menggunakan Metode Support Vector Machine. Universitas Kristen Petra.

Wasito, B. (2013). Kajian Penerapan Artificial Neural Network (Ann) Untuk Memprediksi Harga Saham Mustika Ratu Dengan Metode Support Vector Machine (SVM) Dan Multi Layer Perceptron (MLP). Kwik Kian Gie Business School.

Susanto, I. K. (2021). Analisis Sentimen dan Topic Modelling Pada Pembelajaran Online di Indonesia Melalui Twitter. JOINTECS (Journal of Information Technology and Computer Science), 6(2), 85-98. e-ISSN:2541-6448 p-ISSN:2541-3619. http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs/article/view/1843

Santoso, A., & Ariyanto, M.H. (2018). Implementasi Algoritma Deep Learning untuk Memprediksi Klasifikasi Gambar Menggunakan Convolutional Neural Network. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 2(3), 718-723.

Lian, A.H.S. (2020). Implementasi Sentimen Analisis Tanggapan Masyarakat Terhadap Pemangunan di Kota Pontianak. JUSTIN.

Zain, M.M., & R.S. (2018). Analisis Sentimen Pendapat Masyarakat Mengenai Vaksin Covid-19 Pada Media Sosial Twitter dengan RO-BERT Pretraining Approach. Komputer Terapan, 2.

Atmajaya, R., & W.Y. (2021). Anlisis Sentimen Customer Review Aplikasi Ruang Guru Dengan Menggunakan Metode BERT (Biderectional Encoder Representations From Transformers). JEISBI.

Putra, E.A. (2021). Sentiment Analysis Using Bidirectional Encoder Representations from Transformers (BERT). Diambil dari https://medium.com/@eza.a.putra/implementasi-bert-untuk-analisis-sentimen-terhadap-ulasan-aplikasi-flip-berbahasa-indonesia-557d691e0440 diakses pada tanggal 27 April 2022.

Srimulia. (2022). Mengenal Matplotlib Untuk Melakukan Visualisasi Data, Simak Penjelasan Berikut ini! - IDMETAFORA. https://idmetafora.com/news/read/1254/Mengenal-Matplotlib-Untuk-Melakukan-Visualisasi-Data-Simak-Penjelasan-Berikut-ini.html

Han, J., & Kamber, M. (2006). Data mining: Concepts and techniques. Morgan Kaufmann.

Sugiyono. (2016). Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D. Bandung: Alfabeta.

Rianto, A. (2018). Perancangan Aplikasi Streaming Film Berbasis Android. Jurnal Teknik Informatika dan Sistem Informasi, 4(2), 106-113.

Nugroho, K.S. (2019). Confusion Matrix untuk Evaluasi Model pada Supervised Learning. https://ksnugroho.medium.com/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f.

Hastie, T., Tibshirani, R., & Friedman, J. (2017). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

Jordan, M.I., & Mitchell, T.M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Sembodo, E., Agustina, L., & Tandyono, R. (2016). Web scraping: Crawling data.

M Abdul Dwiyanto Suyudi, Esmeralda C. Djamal, Asri Maspupah (2021). Prediksi Harga Saham menggunakan Metode Recurrent Neural Network.Universitas Jenderal Ahmad Yani.

Souradip Chakraborty.,Ekaba Bisong., (2020). BioMedBERT: A Pre-trained Biomedical Language Model for QA and IR. Conference: Proceedings of the 28th International Conference on Computational Linguistics.

**Lampiran**

1. **Evaluasi Hasil Uji Training Data**

|  |  |  |
| --- | --- | --- |
| review\_text | category | prediction |
| tak kasih bintang lima walau siaran mnc grup di acak tetap bisa menonton di hp lebih jernih dan banyak channel tv nya bagus | positive | positive |
| kenapa enggak bisa menonton padahal aku sudah beli paket premium dan premium sport tapi tetap enggak bisa menonton tulisannya malah segera langganan v premium sport beli sekarang dan nikmati konten ini | neutral | negative |
| seru banget ya dirumah saja jadi enggak bakal bosan lagi ditemani vision . | positive | positive |
| sangat membantu bisa untuk menonton siaran ulang tanpa hambatan | positive | positive |
| super keren menonton streaming film n tv bisa kapan saja dan dimana saja | positive | positive |
| bismillah semoga dapat smarthphonenya . oh ya kak tayngin drakor dong di situ sudah ada tha heirs saran saja bertambah drakor yang banyak ya supaya pecinta drakor bisa lihat drakor di sini . terima kasih kak semoga sukses | positive | positive |
| keluar sendiri atau aplikasi tidak menanggapi . jadi saya copot . cukup satu saja nn buat ini aplikasi . | negative | negative |
| salut banget sama yang bikin apps ini keren . enggak cuma satu channel tv saja tapi bisa buat berbagai channel tv apalagi bisa menonton film dan drama korea . thanks | positive | positive |
| terimakasih aplikasinya sangat berjalan lancar dapat gambar hd dan gratis . saya jadi bisa melihat pertandingan indonesia . terbaik untuk kalian . dan untuk kawan yang lain yang bermasalah dengan jaringan cek dulu kuota dan jaringan kalian kita kan di beri gratis jangan ngeluh dong . lu bayar kalau bermasalah di aplikasi baru ngeluh ini aplikasinya bagus gratis jaringan lu jelek ngeluh jangan di aplikasi ya . | positive | positive |
| vision bagus semoga kedepan nya stabil enggak lambat . dan channel semakin banyak | positive | positive |
| gem filim vision saya ini suka karena ini juga mengigat kan saya tentang penampakan asli di sekolahan sdn bendeorejo tapi tidak apa ? bukan dekat sama cewek saja dekdekan dekat dekat sama hantu ada suara itu juga dekdekan j.daffa y. love salsabila mengapa suka suka sekali sama gem horror dan filim horror lagi ada di indonesia mantab sekali dan bagus sekali | positive | positive |
| maaf bintang nya sedikit dulu ya soalnya enggak support buat android tv aplikasi nya enggak ada di mi tv | neutral | neutral |
| saya mohon maaf selama ini saya kurang teliti mengerti cara memasang aplikasi ini ternyata memang aplikasi anda yang paling mantap diantara aplikasi yang lain.sekali lagi mohon maaf n sukses terus menyertai anda kedepannya amin . | positive | positive |
| kenapa ya enggak bisa dibuka aplikasinya .. ? ? kembali ke menu utama terus .. ? ? | neutral | negative |
| mohon maaf apkalsi vison ini beberapa hari tiga sampai empat hari saja bisa habis itu bisa login bisa lihat vidio nya setelah itu tidak bisa lihat tv nya itu terus bagaimana ya . | neutral | neutral |
| aplikasi ini sangat bagus dan tidak ngelek tapi kenapa episode upin ipin yang bajak laut enggak ada padahal episode bajaklaut terakhir | positive | positive |
| aku suka sekali dengan aplikasi ini terima kasih ya yang buat aplikasi ini karena tv daya rusak dan aku install aplikasi terus menonton maksi | positive | positive |
| bagus banget aku suka karena ini tuh menonton nya tenang enggak digangguin iklan banyak | positive | positive |
| kenapa ya pas saya mau menonton kenapa di kunci dan kenapa mau play video nya kok lama tolong kasih tutorialnya dong kepada pihak mnc now ? | neutral | negative |
| enggak bisa mutar film juga kalo download terjadi kesalahan server . terima kasih | neutral | negative |
| percuma sudah beli paket niat buat dukung timnas taunya aplikasinya jelek keluar sendiri terus enggak bisa di putar setiap klik pertandingan timnas pasti langsung keluar dari aplikasi | negative | negative |
| kenapa ya sudah aktivin paket spot tv tetapi beberapa siaran yang termasuk dalam paket itu tidak bisa dibuka ada tulisan e .smart card out of date . mohon bantuannya | negative | neutral |
| apknya sudah enak tapi kurang update film nya min contohnya amanah wali season belum ada diapk ini | positive | positive |
| kenapa aplikasi vision setelah di install di android box .pas kalau mau di buka selulu nutup sendiri.mungkin ada solusi nya .. mohon info nya | positive | neutral |
| selalu puas dengan vision menonton tv jadi lebih mudah | positive | positive |
| kecewa sekali setelah mengupdate yang baru ini .. mau menonton siaran enggak bisa .lola sekali tidak berpungsi sama sekali | negative | negative |
| saya sudah bayar tapi untuk konten the flash masih terkunci katanya harus ubah ke premium dan saya sudah ubah dan saya coba bayar lagi tetap tidak bisa edit terakhir kalinya rating tanggal april sekarang sudah mei tetap tidak bisa di buka | negative | negative |
| wah betah deh bisa menonton banyak film di aplikasi ini mantap pokoknya tingkatkan biar lebih mantepp lagi | positive | positive |
| kok saya isi ulang paket hari menunggu biasanya cepat tolong di proses dan tolong channel wbtvnya jangan di ganti hits ya terima kasih | negative | neutral |
| aplikasi yang mantap sekarang kalo menonton film enggak usah bingung lagi karena sudah vision | positive | positive |
| aplikasi nya bagus banget tapi ada sedikit kekurangan karena saat mau menonton sinetron yang dulu enggak bisa ke buka dan pakai bayar padahal aplikasi nya sudah bagus banget | neutral | positive |
| huh hp nya bohongan saya setia juga dari awal relese saja enggak dapat kebohongan publik mnc now yang diiklan setinggan saja . buktinya saya dari jawa tengah saja tidak dapat hadiah handphone nya padahal stay tune dari awal | negative | negative |
| saya suka aplikasi ini . boleh saya usul untuk translate atau menterjemahkan subtitle film bahasa bisakah huruf nya di perbesar menurut saya kurang besar atau bisa di atur besar kecil huruf nya . terima kasih atas perhatianya | positive | positive |
| aplikasi busuk banyak masalahnya rencana gua mau langganan eh malah di buat kecewa ! selalu gagal loading . proses pengambilan data gagal ! auto gua uninstal ! | negative | negative |
| keren bos aplikasi ini tonton tayangan ulang dengan mudah | positive | positive |
| aplikasi sangat bagus .. siaran tv nya juga lengkap .. bisa menonton siaran tv yang ketinggalan hingga hari .. dijamin enggak menyesal download ini .. | positive | positive |
| vision ples sangat membantu saya kalau saya pergi keluar kota bisa menonton film india .kalau dirumah saya pasang indovison sudah tahun . | positive | positive |
| mantap. ! chanelnya lengkap .. kualitas hd loading cepat banyak filmya juga . kalo bisa sih ditambah chanel fox sama hbo biar lebih mantap . sukses selalu buat yang buat aplikasi ini . | positive | positive |
| bagus banget channel lokal nya lengkap ditambah channel premium nya cocok buat yang mau menonton tv lewat hp untuk kedepannya lebih di perbanyak lagi channel nya | positive | positive |
| saya baru unduh aplikasinya tapi kenapa tidak bisa dipakai ya terus bacaannya chanel ini tidak tersedia dalam kategori ini maksudnya bagaimana ? | negative | neutral |
| bagus banget aplikasinya jadi suka menonton tv di hp daripada lihat tv langsung hehe . lebih ditingkatkan lagi kedepannya ya | positive | positive |
| aplikasi nya bagus lengkap dan enggak seru . tapi satu saran saya penggunaan datanya diperkecil . terlalu memakan banyak data kuota gue cepat habis . oke sekian | positive | positive |
| dari sekian banyak apktv yang prnah saya install apktv ini yang sangat cocok bagus . saya kasih bintang karena apktv nya benar sangat puas . | positive | positive |
| kok suara ada tapi gambar kotak | positive | neutral |
| bagus banget bos .tapi sayang tidak kompli t. mana film film.dan serial rcti . mnctv . dll . tolong lah dilengkapi biar yang on tidak bosan . nontonya itu mulu . tolonglah untuk dilengkapi.bos | positive | positive |
| kok enggak bisa ya masuk di akun nya malah keluar terus tolong sih ini bagaimana ya | negative | negative |
| aplikasi nya bagus cocok untuk yang enggak mau ribet lihat tayangan kesayangan . semoga kedepannya lebih baik lagi ya | positive | positive |
| halo . harus berlangganan indovison mnc vision dulu kah untuk bisa menikmati chanel sport bein fox sport dll atau bisa langsung beli paketnya vision ? | neutral | positive |
| mudah mudahan beruntung soalnya mau main game hp nya kentang sering lag dan sering malu juga kalok dilihat kawan kawan . terima kasih semoga bermanfaat | positive | positive |
| aplikasi yang menghibur saya saat sedang istirahat kerja menonton tv sekarang enggak perlu kabel lagi . sekarang bisa lihat di mana mana . | positive | positive |
| aplikasi enggak jelas sudah sinyal bagus tapi gambarnya enggak ada apalikasi tholol | negative | negative |
| sebetulnya saya download aplikasi ini karena chanel yang banyak dan bagus .. tapi setelah saya gunakan ternyata aplikasi ini sering log out sendiri padahal saya sedang menonton .. dan saya harus bolak balik login lagi . selain itu untuk masuk aplikasi lama sekali loadingnya sering saya harus hapus data dan cache . ini benaran bikin saya enggak nyaman baru kali ini pakai aplikasi yang dipakenya setengah .. rugi saya sudah berlangganan | negative | negative |
| aplikasi ini banyak manfaatnya terutama tentang update info terkini ya . semoga beruntung dapat smartphone ! | positive | positive |
| aku turunin bintangnya karena ujung nya bayar juga . padahal cuma menonton live tv saja . | neutral | negative |
| bagus aplikasinya .. tingkat kan lagi semoga mnc now makin maju dan seru serta bermanfaat buat semuanya | positive | positive |
| tidak bisa masuk selalu muncul now terus ! di iphone saya juga enggak bisa | negative | negative |
| saya sudah install dapat hadiah nya enggak kalo enggak saya percuma dong install aplikasi ini penuh penuhin memory | positive | negative |
| aku kasih bintang lima karena keren nih .. bisa menonton tv streaming film kesukaan di satu aplikasi jadi enggak ribet .. | positive | positive |
| bagus sekali yang jadi bisa streaming di mana mana enggak perlu nunggu lihat tv | positive | positive |
| bagus bisa lihat semua film and sinetron yang jaman dahulu tapi . agak eror . tolong di cek and dj perbaiki | neutral | positive |
| ini nih aplikasi yang bagus bisa menonton semua tayangan film drama dan lainnya koneksinya juga cepat mantap deh | positive | positive |
| mau nonoton aff saja enggak bisa padahal sudah bayar | negative | negative |
| bersyukur masih ada layanan tv online yang channelnya lengkap dan murah.semoga kedepannya semakin baik lagi untuk memudahkan orang yang install aplikasi ini | positive | positive |
| suka banget deh soalnya lengkap banget mau menonton apa saja ada terima kasih banyak vision semoga makin bagus lagi ini aplikasinya | positive | positive |
| kenapa cuma loading saja lama ke lamaan ada tulisan oppss serve lagi tinggi coba lagi nanti | negative | negative |
| bagus sih tapi sayang cuma bisa kebuka bentar selepas itu enggak ada gambar | positive | positive |
| baru tau ada aplikasi ini sangat menguntungkan karena bisa menonton film yang enggak sempat ditonton saya kasih bintang lima | positive | positive |
| keren banget nih aplikasi banyak banget film film yang aku suka di aplikasi ini . tingkatkan terus | positive | positive |
| favorit banget sama aplikasi ini komplit semua channel tv khususnya saya yang jarang ke ruang tv dan cuma dikamar saja ini salah satu aplikasi untuk bisa menikmati acara tv lainnya | positive | positive |
| setelah kemarin sulit banget berlangganan akhirnya berhasil langganan setelah jam percobaan dan berhasil mencoba aplikasinya overall channelnya oke sih banyak karena incernya memang tv channel tapi aplikasi versi tvnya kacau banget berat ganti channel harus back ke menu channel jadi enggak seamless pengalaman nontonnya buka aplikasinya sendiri saja lama dan berat buka streaming tv sering blank worst e erience banget buat aplikasi tvnya masih banyak yang perlu ditingkatkan untuk bisa dapat | negative | positive |
| pandemi bikin saya tidak ada melakukan ibadah di gereja untung ada aplikasi vision yang nyediain acara kerohanian juga . mantap lah ini terimakasih . | positive | positive |
| aplikasi ini bagus dan cocok banget buat yang suka menonton tv atau film dimanapun dan kapanpun lewat hp pas lagi streamingnya juga lancar . | positive | positive |
| okay cukup menyenangkan menggunakan aplikasi ini yang sukanya menonton korea ini rekomended banget | positive | positive |
| baik sangat disaat buka channel lancar tapi sayang banyak channel yang hilang dari free terlalu rumit untuk membeli channel tv seharusnya dipermudah lagi transaksinya . | neutral | negative |
| sudah setahun langganan tapi agak nyesal sih . gambarnya di tv burik . tidak hd . di hp masih lumayan tapi di tv berasa nonto kwalitas p | negative | positive |
| aplikasinya suka performa oke biaya langganan murah streaming lancar channelnya banyak sejauh ini oke dan tidak ada kendala . | positive | positive |
| bagus sih cuman aku bingung nanti beli paket nya bagaimana waktu habis paket nya tanggal febuari tahun nanti nah itu tinggal beberapa bulan lagi nih bagaimana cara ngisi paket nya nih biar berlangganan | positive | positive |
| aplikasi sangat bermanfaat bisa buat streaming tv favorit lengkap statiun tv nya ada banyak film juga | positive | positive |
| puas banget sama aplikasi nya vision mudah untuk menonton film yang aku ingin kan intinya puas banget sama aplikasi nya | positive | positive |
| kecewa saya beli premium tapi kuota saya yang disedot lalu apa guna nya premium | negative | negative |
| selalu sering eror kalau lihat tayangan ulang di mnctv.nyi loro kidul yang selalu eror padahal itu film yang gue tunggu tunggu.soalnya tayangnya mlm bngt.jdi gue sering lihat tayangan ulangnya pagi tapi kenapa sekarang sering eror.perbaiki tolong min . kalau enggak eror eror lagi gue kasih bintang full | negative | negative |
| trimakasih mnc now dengan aplikasi ini sangat membantu dan menghibur di waktu luang .. bisa juga dilihat saat ketinggalan siarannya .. sangat membantu | positive | positive |
| menonton streaming kok servernya down kan goblok sekali .. kembalikan menit hidup saya yang terbuang buat pasang aplikasi jelek ini ! un install and goodbye . | negative | negative |
| assalamu allaikum alhamdulillah terimakasih mnc now saya suka banget dengan aplikasinya bagus dan bermanpaat.dan juga film filmnya sangat bagus dan bisa di putar sesuai episodenya . i like mnc now . | positive | positive |
| kemarin mencoba download vision karena direkomendasikan teman.ternyata benar vision ini sangat memberi hiburan untuk saya .. apalagi saat sedang dilanda macet yang lama tapi saya masih bisa menikmati acara tv favorite saya . suka banget pokoknya sama vision | positive | positive |
| misi kak kok saya daftar akun ke facebook tapi kok gbisa mohon di koreksi | neutral | positive |
| kok setiap buat menonton bola timnas selalu enggak bisa ya muter muter terus apakah di acak .. kalau untuk saluran yang lain bagus | positive | neutral |
| koo setelah di upgrade ke versi terbaru enggak ada siaran antv ya di saluran tv nasionalnyaa | positive | negative |
| saya sangat suka dngan alikasi ini bisa menonton tv di mana saja kapan saja dngan cara mudah | positive | positive |
| vision ok juga agak ringan tidak perlu internet kencang saya bisa menonton dengan lancar walaupun internet pas an terus tidak menguras kuota . | positive | positive |
| boleh ditambah lagi siaran tv daerah nya ? agar yang di luar daerah asal juga bisa melihatnya | positive | positive |
| oo iya min kalau update pakai fitur hemat kuota dong . hehe . soalnya kadang pas jaringan buruk untuk akses vidio di mnc now jadi susah . saya pakai telkomsel . kadang sinyal agak terganggu enggak bisa menonton . | positive | positive |
| sering keluar mnc now terhenti padahal saya sudah keluar aplikasi .. dan lagi sering enggak bisa di buka saya mesti uninstall dan download ulang .. ini saya baru uninstall dan download lagi | neutral | negative |
| lag terus min kenepa ya ? padahal jaringan stabil | positive | negative |
| apakah aplikasi ini berbayar ? karena saya tidak bisa menonton apa .. hindari sentuhan blablabla begitu infonya selalu . saya tambah bintangnya nanti kalau lancar free menonton tv | negative | negative |
| sangat bagus wajib donwload ini . menyesal kalau enggak donwload sekarang sudah bisa menonton antv dan chanel lainnya . terima kasih | positive | positive |
| aplikasinya sangat simpel dan mudah untuk di pakai | positive | positive |
| punya tv dirumah semua trsambung keinternet enggak ada pun yang bisa konex menonton aff ane coba pakai hp pakai kuota kartu prabayar tetap juga enggak konex pakai aplikasi inipasti pas aff selesai terkonexi dengan baik enggak ada masalah deh jadi ganti pakai aplikasi v di o tanpa bufering dan enggak pernah ngadat jwbannya saya sudah matiin modem brkali pakai modem jaringan kabel sama saja bufering coba pakai jaringan hp lte sama juga buferering silahkan dicoba sendiri saja besok menonton diaff sekarang aff deh habis lancar jaya kembali | negative | negative |
| padahal paket full mau menonton tv luar negeri lelet kurang puas .. ? ! | positive | negative |
| aplikasinya bagus yang paling aku suka loadingnya enggak lama dan enggak ribet . dikembangkan terus ya min biar tambah nyaman | positive | positive |
| kok kalo menonton biasa nya hitam begitu ya enggak ada gambar nya sudah nunggu beberapa menit tetap saja hitam | neutral | negative |
| filmnya di update dong yang tahun lama kalo bisa enggak usah diputar sih .. munculin film film yang baru baru dong . dan jangan dikit dikit unnable tapi connect mulu tiap menonton . baru sekian detik menonton deh ngebug . durasi jam bisa sampai besok baru kelar nontonnya | neutral | negative |
| seru banget dwnload aplikasi ini bisa menonton tv dari hp.nntn film keren banget pokoknya sudah begitu bagi hadiahh lagi moga saya beruntung ya cinta lah pokoknya sama aplikasi ini . | positive | positive |
| bagus banget aplikasi nya tapi tolong dong episode ih seremnya di tambah lagi yang di bagian serial p | positive | positive |
| saya sudah berlangganan paket premium hari ribu rupiah tapi setelah update ke versi terbaru saya tidak bisa menonton channel premium . tolong segera diperbaiki ! | negative | negative |
| banyak salurannya makin enggak bosan bisa menonton walaupun ketinggalan hari keren sih ini | positive | positive |
| wah senng sekali bisa memberikan b karena lama lama makin hancur pelayanan pengaduan via wa saja lambat bilang ada kunjungan ternyata hoax apliksi juga sering error | negative | negative |
| aplikasi sampah .. masuk keluar lagi .. masuk lagi keluar lagi . aplikasi enggak guna ! | negative | negative |
| alhamdulillah semakin baik setelah diperbarui . saran saja tambahkan fitur resolusi video dong biar kita hemat internet saat menonton . dan juga fitur skip atau fitur download begitu . supaya kita sebagai pengguna puas untuk menggunakan aplikasi ini . thank | positive | positive |
| aku kasih bntang dulu ya .. coal nya jaringan nya lelet .. nanti kalau aku menonton film nya lancar seperti biasa aku tambahkan deh .. tolong dong admin diperbaiki jaringan nya . | negative | neutral |
| aplikasi yang bagus .. bisa menonton acara tv meskipun di kost saya enggak ada tv.terimakasih mnc now . | positive | positive |
| tolong perbaiki bug secepatnya . karena bug yang enggak bikin nyaman ketika menonton | negative | neutral |
| entah kenapa menonton di tv di hp android channel channel olahraga pada hilang semua .. kaya fox sport soccer channel mnc sport bein sport .. padahal di paket tv saya nambahin paket bein sport .. tapi kenapa ya di android nya tidak ada .. saya kecewa | negative | negative |
| apa sih baru juga beli paket premium malah enggak bisa masuk logout sendiri terus enggak bisa masuk | negative | negative |
| awal beli paket cling saluran ada semua pas beli ke xnya tvn movie thrill celestial classic hilang semua padahal sebelum beli paket cling tvn movie selalu ada kenapa ya ? sudah begitu sekarang susah login lagi ke appnya | neutral | negative |
| aku suka sekali sama vision tapi kenapa sekarang aku buka tidak bisa ya . padahal sinyal bguss | positive | positive |
| woyy gua sudah berlangganan tapi kok enggak bisa menonton acara live rugilah | negative | negative |
| ini benaran enggak sih hadiahnya kalo kita install aplikasinya terus daptar pakai email langsung dapat hadiahnya enggak | positive | negative |
| sekarang jadi jelek kwalitasnya pengin menonton badminton.udah beli paket premium.eeh gambar buram.sering loading . padahal di aplikasi lain lancar lancar saja buruk pokoknya mending menghilangkan ini aplikasi | negative | negative |
| mungkin bisa ditingkatkan untuk penggunaan yang di pc . tetapi overall sudah baik semoga semakin mantap . terimakasih | positive | positive |
| halo namaku santy susah banget daftarnya tolong di bantu nanti kasih bintang lima deh pleas tapi caranya bagaimana emailnya sudah benar dan saya juga belum punya facebook soalnya aku masih kls | neutral | positive |
| saya sudah beli paket tan di vision yang premium tetapi sampai sekarang saya tidak bisa meng akes film yang saya tonton . dan saya di paksa harus beli paket lagi . padahal saya sudah langganan paket bulan . kalau saya disuruh beli lagi buang uang saja kalau begitu .. tolong min jawab . jangan rugikan konsumen . ! | negative | negative |
| waahh mantap apknya lancar jaya tidak sama dengan aplikasi lain saya suka sekali nunggu supaya bisa full screen semoga bisa tambah bagus lagi terimakasih banyak | positive | positive |
| setiap pembaruan menjadi lebih baik namun ada kekurangan di balik kelebihan tersebut | positive | positive |
| rating khusus app ini paling buruk .. aplikasi eror terus sudah bakan mau lihat chanel nasional enggak mau kebuka | positive | negative |
| sangat kecewa saya sudah melakukan pembaayaran untuk bulan tapi tidak bisa singkron dengan perangkat resiver kak vision ketwrangan nya id sudah ada . | negative | negative |
| sangat irit apknya murah meriah bisa menonton sepuasnya .lanjutkan ! | positive | positive |
| dulu sih sudah bagus sekarang kok makin jelek saja nih masa sudah pakai wifi masih loading dan enggak mau keputar videonya sih semoga kedepannya lebih baik lagi ya | negative | negative |
| acara yang ditampilkan dapat dipilih sesuai dengan hobi saya jadi saya bisa menonton acara yang saya suka dengan mudah | positive | positive |
| oke terima kasih kak cara nya alhamdulillah pas saya coba berhasil dan btw aplikasinnya bagus banget | positive | positive |
| aplikasi bagus sih buat menonton ulang yang ketinggalan tontonan yang kemarin .tapi kenapa akhir ini agak susah masuk aplikasi nya selalu keluar sendiri juga | positive | neutral |
| tolong untuk sinyal dari pusat vision tolong di perbaiki lagi soalnya sinyal yang sudah kuat masih tetap tidak bisa menonton di vision | negative | neutral |
| vision memang baru bagi saya seperti bisa membawa tv kemana saja dan kapan saja sambil apa saja siip pokoknya | positive | positive |
| semenjak ada aplikasi ini aku lebih sering menonton acara favorit disini semoga makin di tingkatkan kualitas nya | positive | positive |
| keren sangatt berguna banget . banyak fiturnyaa juga jadi enggak menyesal download aplikasi ini . thank you min ! semoga makin sukses ya ! ! | positive | positive |
| aplikasi nya sangatt mantap saya jadi bisa menonton tv di hp walaupun mati listrik hehe tapi sayang nya layar hp saya cuman inchi . jadi semoga saja saya bisa beruntung dan mendapatkan hp dari mnc now | positive | positive |
| aplikasinya sangatt bagus membantu banget . enggak menyesal download aplikasi ini . film filmnya juga banyak banget jadi enggak bosan deh kalau dirumah . thank you | positive | positive |
| aslinya top markotop mnc now ini cuman kalau menonton sinetron kesukaan saya yai itu kembalinya kian santang di langit pajajaran durasinya enggak sampai habis tolong dong di perbaiki durasinya sampai habis padahal sudah terpotong sama sinetron jafar tolong ya tambah durasinya sampai habis nanti saya tambah satu bintang lagi | positive | neutral |
| setelah download aplikasi nya wah bagus juga jadi bisa menonton tv semua chanel tanpa harus mke tv . di perjalanan juga bisa menonton terima kasih vision | positive | positive |
| kalau disini suka menonton film hiburan . enggak tau kenapa suka banget dan enak kalau pakai aplikasi ini enggak mecet macetan dan sangat jelas pula . | positive | positive |
| mantap ! chenelnya lengkap dan gambarnya kelihatan jelas . tingkatkan lagi biar makin wah . sukses selalu ! | positive | positive |
| setiap muter film sampai tengah pasti macet . di ulang macet terus padahal jaringan bagus | negative | negative |
| mantap bro aplikasi sangat mudah dibuka dan semua yang disajikan puas vision memang mantab .. tab.tab.tab | positive | positive |
| saya sudah bayar kenapa ketika saya buka siaran live film selalu muncul pasan upgrade apakah aplikasi ini penipu ? dibuat untuk menipu uang tidak seberapa bagi saya yang bikin saya marah karena saya marasa ditipu | negative | negative |
| woahh paket komplit ini aplikasi menonton tv series film semua bisa dalam satu aplikasi | positive | positive |
| mantap banget kak aplikasi enggak menyesal saya download hehe tingkatkan lagi ya | positive | positive |
| pengalaman seperti menonton tv sungguhan tetapi kenapa tidak bisa melihat pertandingan bola pada saat live malah di alihkan acara lain ? | positive | negative |
| aplikasinya sangat bagus buat menonton tv dan streaming .. | positive | positive |
| bagus banget simple . tapi tolong upgrade berikut nya di buat nya biar bisa full di layar . kan sekarang sudah banyak hp yang layar seperti itu . | positive | positive |
| sejauh ini masih cukup bagus aplikasinya fiturnya juga banyak gapang dipahami juga saran saya supaya lebih dikembangkan lagi biar bisa di download tuh film | positive | positive |
| sangat jelk sinyal bagus . notip ermintaan ke server sedang penuh . menyesal sangat dalam baru download minta kemudian saya uninstal.sungguh kesal | negative | negative |
| kok antv nya hilang ya ? sudah di update masih enggak muncul juga antv nya padahal suka banget sama aplikasi ini .. lagi eror apa memang enggak ada lagi siaran antv ? | positive | neutral |
| saya kira ini seru ternyata tidak .. saya kira langsung menonton ternyata harus pakai email tolong dong jangan pakai itu langsung menonton saja biar banyak yang suka aplikasi ini . | negative | positive |
| aku suka aplikasi vision sangat simpel lancar banget pakainya thank you so much | positive | positive |
| saya sudah menonton beberapa film di mnc now tapi saya belum dapat handphone . solusinya bagaimana ? untuk aplikasinya sudah lumayan loadingnya enggak terlalu lama hanya saja saya belum dapat handphone nya . tolong diulas kembali | positive | positive |
| enaknya itu bisa lihat siaran ulang tv berkali kali . anti lambat club | positive | positive |
| hari ini bete banget lagi enak menonton flim tayangan ulang baru menonton menit kok flim nya berhenti saya berpikir apa karena wifi nya lambat pas saya lihat bagus kok wifi .. makin binggung dong kan jadi enggak enak menonton nya . tolong dong di memperbaiki lagi . ini apalikasih yang paling bagus banget enggak perlu ribet cari chanel dan ada tayangan ulang nya . | positive | negative |
| bagus sudah ok bangetz lah .. garafik jgk ok. sinyal zuuz . ok lah pokok nya . aku suka bgtz . terus di kembangkan . sukses selalau mnctv | positive | positive |
| wah mnc now mantap sekali bisa buat menonton film apa saja aku jadi enggak susah kalau mau menonton film yang ke tinggalan pokoknya ndk rugi download aplikasi nih . deh buat mnc now salamdarilombok | positive | positive |
| mnc now memang top.ini lah apps tv android yang ku tunggu dan saya bisa menonton afc live .. top deh mnc now .. maju terus dan tingkatkan lagi | positive | positive |
| dari tdi mau live streaming antv sama net kok enggak bisa terus ya padahal yang lain bisa . kenapa ? dan selalu ada tulisan unable tapi connect sudah di update juga dan sudah coba log out dan login lagi tapi tetap enggak bisa semoga bisa cepat balik lagi kaya dulu.plis bagaimana ini ? ! padahal sudah nyaman banget sama apknya | neutral | negative |
| di hubungkn kak alat wifi donggle enggak sudah gambar di lyr tv hanya suara yang ada kalau ada kerusakan dari alat enggak mungkin karena lihat youtube bagus ajjh saya pakai aplikasi ini untuk lihat tv karena jaringn antena tv di wilayah saya jelek harus pakai parabola . tlg ditingkatkan lagi diperbaiki | negative | negative |
| versi lama lebih bagus bisa menonton tv tiongkok rrc . | negative | positive |
| vision bagus dan saya suka. dan tapi ada salah satu acara tayangan film yang saya suka di salah satu stasiun televisi. tetapi selalu tidak menyelesaikan pemutaran film nya dan ini membuat saya jadi sebal. | positive | positive |
| bagus sih aplikasinya .. kontennya beragam dan banyak variasi film yang recommended . | positive | positive |
| aplikasi berat dan susah .. | negative | negative |
| channel nya masih kurang lengkap sbsin kbs world indonesia enggak ada padahal kan channel itu channel indonesia juga sbsin fta di palapa sedangkan kbs world ini diacak biss di telkom si dua channel itu channel faforit ku tolong tambah ya mnc . terimakasih . | neutral | neutral |
| lagi di rumah sakit umum adanya televisinya dibuat bersama jadi aku pakai earphone nyalain hp menonton streaming televisi di aplikasi vision plus | positive | positive |
| tak kasih bintang saja . masa menonton film buffer mulu down padahal resolusi deh pakai yang paling kecil p. giliran tak coba menonton di youtube n iflix lncar jaya p p no buffer . apalah ini aplikasi | negative | negative |
| sorry karena apknya masih ngebug aku uninstall dulu . nanti kalo apknya sudah diperbaikin aku download lagi | negative | negative |
| saya kasih satu bintang karena setiap saya nonto serial kok videonya lomot padahal paketa masih banyak sinyal bagus mungkin ini aplikasi butuh pengapdedtan ke lebih baik . | positive | negative |
| oke sudah normal mantap sudah lancar aplikasinya semoga aplikasinya ke depanya makin baik .. | positive | positive |
| thanks buat teman yang rekomendasikan install app ini . bisa jadi teman menonton dikala macet . | positive | positive |
| kenapa aplikasi vision tidak bisa pakai layar transmisi ? | neutral | negative |
| bagus tapi kok lama lama enggak bisa di buka sih ! tolong dong di perbaiki hehe | positive | positive |
| tidak bisa di buka karena meminta di upgrade tapi setelah di upgrade tetap tidak bisa di buka | negative | negative |
| sebenarnya saya mau kasih bintang cuma mslah di aplikasi ini tolong ya min perbaiki mslah rotasi otomatisnya .. kalau sudah pasti saya kasih bintang oke | positive | neutral |
| aplikasi ini bagus banyak pilihan channelnya dengan harga paket yang terjangkau sudah enggak susah mau menonton film favorit .. | positive | positive |
| untuk e erience sangat kurang apalagi di tv android loading suka stuck haru keluar masuk akun terlebih dahulu tingkatkan kualitas lagi | neutral | negative |
| kenapa vision tidak siarkan sea games payah tidak di siarkan di vision pelit padahal sudah beli paket langganan tahun | positive | negative |
| biasanya kalau menonton tv selalu pakai tv yang besar sekarang mah sudah simpel menonton dimana saja bisa cuma bawa hp yang install aplikasi vision disini sudah lengkap semua ada keren pokoknya | positive | positive |
| suka bangetlah pokoknya sama aplikasi ini bagus semua channel tv saja ada . gambarnya juga bagus . pertahankan ya min . | positive | positive |
| enggak ada indosiar jadi enggak bisa menonton panggilan kalo bisa ditambahkan sama chenel tv yang lain . semangat developer ! tolong indosiar ditambahkan pls . | positive | negative |
| ini enggak ada usaha perbaikan apa ya padahal sinyal bagus tapi tetap saja enggak bisa buat menonton motogp | negative | negative |
| awalnya kali saya buka aplikasinya lancar saja tetapi sekarang aplikasinya enggak bisa dibuka . what happened ? apa mungkin saya harus uninstall padahal saya senang dengan fitur nya . | neutral | negative |
| suka sama aplikasi ini tapi kadang suka error kalo menonton di net tv kadang net tv jadi channel mnc shop | positive | positive |
| sebenar nya apilikasi ini harus daftar tapi saya enggak tau email saya dan telepon aku maaf saya kasih bintang dulu ya kalau sudah enggak daftar dan tentang email baru saya kasih bintang ok bye | negative | negative |
| keren banget aplikasi nya bisa menonton tv kalo kita lagi di perjalanan | positive | positive |
| kalau enggak bisa menonton di televisi gara gara rebutan sama adek jadi bisa menonton disini . terima kasih om sudah buat aplikasi berguna kayak begini . | positive | positive |
| terimakasih banyak vision dengan ini aku bisa lihat jadwal favorit aku jadi enggak bakal ketinggalan . ah senang banget pokoknya . wajib download | positive | positive |
| selalu saja seperti ini tuh kenapa episode dunia terbalik lompat dari ep langsung ke sama ep ke ? dan juga ep sampai error network no file . tuh tim terkait kayak enggak ada niat upload . mohon di perbaiki | negative | negative |
| parah nih aplikasi sudah buffering mulu terus mau menonton live tv memang harus beli paket dulu ya ? lawak sih kalau memang begitu pelit benar enggak kayak aplikasi video . | negative | negative |
| aplikasi sampah dibuka loading muter saja habis itu keluar sendiri .. padahal buka youtube lancar saja | negative | negative |
| katanya banyak chanel.tp kok saya lihat cuma dikit.apalagi banyak chanel yang hilang . seblmnya masih ada.d rumah kan saya juga berlangganan skligus.tp di hp kok cuma dikit chanelnya .. | neutral | neutral |
| kenapa pas mau berlangganan registrasi berhasil tapi gagal saat mengirim email itu kenapa ya min ? mohon segera balasannya saya kasih bintang dulu | negative | negative |
| lelet pas masuk ke menu tv bahkan bisa error hitam . | negative | negative |
| kenapa channel music jadi harus bayar . ksell ih jadi enggak bisa menonton lagi .. | negative | negative |
| pas di stel kok muncul playback failed di meminta retry ga bisa juga . padahal di bulan mei kemarin baik saja . sekarang sudah ok. tolong dukungan buat web os nya ya . di apps web os belum support . api nya dikembangkan lagi ya . | positive | positive |
| asli saya suka dengan aplikasi ini karena bisa menonton tv yang blman ada di tv aku dan aku bisa sambil menonton tv di tengah perjalanan menuju pulang kerumah visiont semoga berjaya menuju kesuksesan walaupun ada sedikit retry mohon di perbaiki terima kasih | positive | positive |
| jelek banget masih asik menonton tiba tiba eror kasih code code kenapa sih aplikasinya | negative | negative |
| tolong dong tayangin liga champion dan liga inggris . jangn di acak juga di aplikasi mnc now | positive | negative |
| live tv tapi kalau menonton live moto gp selalu buffering yang pelit lah indonesia ini ditambah sponsor telkomsel makin depan mulai enggak puas lagi menonton pakai paket saja pun susah nya meminta ampun | negative | negative |
| saya sudah download vision connect id pelanggan . sudah bayar bulanan juga untuk bulan depan tapi belum dapat gratis all channel ya selama sebulan . | positive | neutral |
| mau masuk tapi kow tulisanya silahkan cek koneksi anda.padahal jaringannya masih bagus | neutral | negative |
| kok setiap mau menonton suka bermasalah sistem eror padahal jaringan saya bagus saja . mohon diperbaiki agar kedepannya orang bisa lebih nyaman dan nontonnya tidak ada gangguan lagi . | negative | neutral |
| masa sudah daftar bukannya dapat mamen malah kuota cepat habis padahal sudah log ini dan sudah daftar enggak dapat dapat sudah kai begitu enggak dapat dapat | negative | negative |
| sangat sangat burik aplikasi nya para enginer vision lihat enggak sih komentar diaplikasi ini hampir semuanya negatif ayo belajar dari aplikasi sebelah tapi sampai saat ini aplikasi masih begitu saja kalo menonton live tv hilang mulu . | negative | negative |
| aplikasi yang sangat bermanfaat untuk menunjang terbentuknya masyarakat yang berwawasan tinggi . terima kasih banyak mnc now semoga tambah sukses dan terus berkembang fitur fiturnya | positive | positive |
| mantab mudah di gunakan dan tidak buffering di tunggu fitur fitur menarik lainnya | positive | positive |
| kenapa kalo lagi menonton vidio suka berenti sendiri sih padahal sinyal bagus | negative | negative |
| jangan buang kuota download aplikasi ini . aplikasi paling lama loadingnya . untung saya pakai wifi . pakai wifi saja susah loading loh . aware ! | negative | positive |
| kenapa pas di tengah episode videonya buram ya padahal sudah bagus tuh sinetronnya sinyalnya pun juga bagus tapi videonya buram | negative | negative |
| kenapa tidak ada gambar ketika di konesikan ke smart tv | negative | negative |
| resolusi tolong tambahkan minimal hd p di channel channel mnc group doang gapapa biar rasio gambar enggak . sama loading beranda tolong dipercepat suka force close juga . aplikasi streaming tetangga sudah p kok jangan mau kalah . | positive | negative |
| aplikasi black screen terus padahal internet aku sudah bagus pakai biznet mbps tapi pas buka ini aplikasi di tv black screen mulu aneh | negative | negative |
| bagus jelas jadi kalau pas tv gangguan atau pas lagi diluar rumah bisa langsung buka vision | positive | positive |
| kenapa permintaan server sedang tinggi padahal jaringan full.mohon di perbaiki.supaya sinyal rendah tetap bagus gambarnya.jangan hanya promosi doang .. ! saya kecewa | positive | negative |
| aplikasi yang sangat membantu kita untuk tidak ketinggalan tontonan walaupun tidak sempat menonton di televisi . banyak pilihan channelnya meskipun belum selengkap di televisi . tolong terus perbaiki fitur nya supaya para pengguna bisa lebih nyaman memakai aplikasi ini . | positive | positive |
| sudah berulang di update tetap saja tidak bisa di buka aplikasi vision nya.sampek sudah di install lagi tetap saja tidak bisa di buka | negative | negative |
| aduh kok ada lagi sih iklan tiap mau pindah channel ? sangat mengganggu .. susah mau pindah pindah chanel belum lagi masing masing channel nya sudah banyak iklan . payah ! | negative | negative |
| bagus banget wajib download dengan adanya aplikasi ini saya jadi bisa jonton tv dimana saja terus enggak ketinggalan serial kesayangan aku karena bisa diputar ulang terima kasih vision semoga tambah maju | positive | positive |
| saya sangat puas dangan aplikasi ini . tapi tolong dek perbaiki sedikit . pertama kenapa saat aku menonton tv video nya kecil.yang ke dua kenapa cuman otomatis saja . padahal aku ingin lebih hd . | positive | neutral |
| enggak bisa di pakai aplikasinya . sering nge crash | negative | negative |
| saran untuk vision semoga ada versi untuk semua smart tv jadi menonton nya hanya tinggal digeser pakai remot tv saja . kalau pakai mouse rib t. dan semoga vision bisa kerjasama dengan semua provider jadi ada paket kusus nya begitu seperti aplikasi sebelah . | positive | positive |
| sudah tersedia di android tv tapi ui nya masih kurang nyaman terasa berat mbps masa enggak sanggup loading . semoga kedepannya semakin ringan cepat kayak aplikasi sebelah vidio rcti dan mudah digunakan agar bisa berganti channel tanpa keluar dari yang sedang ditonton . | positive | positive |
| mana nih kok saya belum dapat hape nya ya min padahal sudah memakai ini aplikasi boleh tiap hari menonton terus malah semua episodenya epen cupen sudah saya tonton . parah lah | neutral | negative |
| untuk aplikasinya lumayan bagus tapi kok ketika memilih audio bahasa indonesia di channel fox family movies tidak ada suaranya ya ? mohon diperbaiki ya audionya | negative | positive |
| chanel tvn tiba tiba enggak ada . di home ada tapi enggak bisa di buka . sudah di uninstal terus di install ulang tetap saja begitu . perasaan kemarin ada chanel tvn balikin lagi kenapa .. ! ! ? | negative | negative |
| aplikasinya bagus enggak kecewa saya download ini aplikasi . saya kasih bintang deh sebagai bentuk apresiasi karena telah berinovasi mengeluarkan aplikasi ini | positive | positive |
| sudah aktifkan premium sport tapi masih utuh enggak bisa dibuat menonton sudah bayar padahal mohon tingkatkan lagi kualitasnya | negative | negative |
| bagus banget tapi agak telat ya mohon di perbaiki kan kata nya live tapi pas aku lihat tv kok agak beda ya | neutral | positive |
| gambar nya kuarang jernih . enggak kayak video tingkatkan lagi | negative | positive |
| ini kenapa aplikasinya enggak bisa dibuka deh loading di logo vision plus mulu kesel | negative | negative |
| mantap langganan mnc vision bisa dipakai untuk gadget .. | positive | positive |
| lumayanlah buat menonton film iseng iseng cuman sayang film dan fitur kurang lengkap perbaikan lagi dong untuk filmnya di tambahkan lagi biar enggak bosan para pengguna nya rekomendasi banget deh | positive | positive |
| bagus cuma kadang tiba tv luar hilang semua . harus install ulang | positive | positive |
| tolong dong cenel islami nya di tambah . biar bisa menonton kajian dengan mudah | positive | positive |
| tolong perbaiki saat play catch up baru beberapa menit langsung buffer | neutral | neutral |
| awal awal bagus tapi lama lama jadi enggak jelas seperti keluar sendiri padahal jaringan nya bagus terus ada tulisan server sedang tinggi kalah sama aplikasi tv lainnya | negative | negative |
| iklan dari mnc now nya enggak keluar terus lama nunggu nya . sbnrny sih bagus ya .. bisa menonton film dimana saja kapan pun bisa .. salah nya pas buka app nya eh malah enggak nongol itu iklan .. kan setiap buka app kan ada iklan nya tuh .. nah itu yang buat lama nunggu | neutral | positive |
| secara keseluruhan suka sama aplikasinya buat orang awam enggak terlalu kagok . kedepannya ditambah program yang lebih menarik . | positive | positive |
| apps nya berat di tv loading lama . enggak kayak apps sebelah . scan scan terus masa sehari bisa kali scan . mohon diperbaiki . kalo dibandinh video mola netflix agak jauh . pr nya agak banyak | negative | neutral |
| aplikasi ini sangat jelek banget . eh admin bodoh aplikasi ini error hapus saja cuiih . semua nya jangan download aplikasi ini jelek banget dan error banget . enggak usah download | negative | negative |
| mantap deh aplikasinya . bis menemukan channel sebanyak ini cuma dalam aplikasi . bis dibawa kemana saja . terus enggak lelet . moga saja kedepannya bisa bertambah terus fitur fitur kerennya . | positive | positive |
| habis beli paket premium malah langsung log out dan tidak bisa login kembali | negative | negative |
| sebener nya aku suka menggunakan aplikasi mnc now ini tapi kenapa enggak bisa di buka saat mengulan kuota aplikasi gratis dari im . malah paket utama aku yang habis . saat sedang menonton loding terus . | negative | neutral |
| saya di rumah pakai wifi indi home .. tapi tetap enggak bisa saya sudah hps aplksi saya dounlowd ulang tetap enggak bisa selalu eror .. pakai data pun sama .. tolong dong perbaikannya | negative | negative |
| sebanyak aplikasi tv yang saya download baru kali ini dapat yang bagus lengkap ! | positive | positive |
| kebanyakan bufering menonton timnas nya mungkin bisa di perbaiki lagi ? ? | negative | negative |
| kenapa tayangan premium harus langganan kemarin gratis saya kecewa | negative | negative |
| setiap menonton ikatan cinta kok cuma muter saja pas sudah bisa juga masih muter . tolong di perbaiki dong | negative | negative |
| sangat membantu sekali karena dengan app vision plus dan signal yang kuat tentunya bisa update .. semua berita dan siaran dimanapun berada | positive | positive |
| asik banget pakai ini .. bcs channelnya banyak banget dan daftar filmnua juga banyak | positive | positive |
| aplikasi ini sangat berguna untuk menonton acara yang ketinggalan ! | positive | positive |
| saya suka aplikasinya membantu tapi mohon dibantu pencarian channel nya biar lebih gampang terimakasih | positive | positive |
| bagus lah aplikasi ini login sangat mudah dan bisa diakses kapanpun jadi enggak bakal ketinggalan kartun favoritku | positive | positive |
| apa apaan deh lagi seru seru nya menonton terkadang suka terputus sendiri unable tapi connect ajg memang belum juga menit menonton deh error saja | negative | negative |
| jangan sampai enggak donlod lah praktis lebih mobile juga buat yang sibuk tapi tetap mau menonton tivi | positive | positive |
| mau menonton film terbaik ? vision tempatnya . aplikasi ini menyajikan beragam pilihan film berkualitas . keren . | positive | positive |
| saya ingin menikmati open all channel selama bulan tetapi tiap kali akan connect selalu gagal dan ada pemberitahuan bahwa id pelanggan sudah ada ! mohon penjelasannya terima kasih | positive | negative |
| bagus cuma seiring bertambahnya user server semakin lambat low respon streaming | positive | positive |
| senang banget sama aplikasi vision sekarang aku bisa lihat acra mnc dimanapun dan kapanpun . pokoknya best banget | positive | positive |
| jadi jelek sekarang mah.masa streaming tv bisa mau buka dunia terbalkk malah ada tulisan internet maslah saya kan pakai wifi .di update dong | negative | negative |
| nice enggak perlu susah payah bawa ditv dikosan aplikasi ini bisa menemani disaat aku gabut hahahah | positive | positive |
| channel tv di vision benar benar lengkap banget deh menonton tv jadi makin gampang | positive | positive |
| aplikasi di update kok makin enggak bisa ya ? | negative | negative |
| aplikasi nyh sekarang jadi jelek dulu bagus sekarang jelek suka ngeleg padahal sinyal nyh bagus . dan kadang suka enggak bisa login padahal sinyal bagus momon tolong perbaiki dong aplikasi nyh | negative | negative |
| kenapa aplikasinya keluar sendiri dan sudah bolak balik donwload ulang tetap saja sama padahal mau nontonn film yang ada disini | negative | negative |
| aplikasi sering error sering mengalami gangguan seperti tiba tiba keluar di saat menonton film padahal sudah berlangganan membuat tidak nyaman | negative | negative |
| saya sangat suka sekali aplikasi vision karena sangat membantu saya menonton tv secara mudah cepat | positive | positive |
| chanelnya tidak sama dengan mnc play media padahal saya sudah langganan selama sekitar tahun mnc play media sebelumnya bisa konek semua tv.mohon bantuannya | negative | negative |
| tidak mau ganti aplikasi masih suka banget .belum bisa membuka lagi sudah di edit bahkan ganti sandi tetap tidak bisa.mohon bantuan nya gan . | positive | negative |
| haha saya bisa nilai aplikasi ini dari iklan di tv bahwa yang menang hadiah itu itu saja sedang kan kita yang baru susah dapat nya | negative | positive |
| bagus banget . penggunaannya juga enggak ribet .. sip deh pokoknya . dan mudah an dapat smartphone gratis dari mncnow | positive | positive |
| bagus top tingkatkan terus pertahankan kuwalitas nya . | positive | positive |
| aku menonton filmnya kok kayak ngebug begitu ya suaranya saja yang dengar tapi visualnya enggak ada apa karena sinyal internet mnc ku yang lelet ? setting auto tetap saja .. | negative | neutral |
| apa n nih aplikasi lambat parah dikit masuk ke akun mu . hadeh . mana loadingnya lamaa banget durasi menit bisa sampai jam karena loadingnya . | negative | negative |
| sudah bagus dan kalau bisa untuk pengaturan kwalitas video bisa dikunci ketika pindah channel lain maksudnya tidak selalu menyetel lagi ketika berganti channel lain.misal memilih kwalitas audio maka setelah pindah channel lain kwalitas videonya bisa tetap audio tidak berubah.kecuali memilih otomatis memang bisa pindah sendiri kwalitas nya.hanya masukkan supaya user tidak bolak balik menyetel pengaturan kwalitas videonya.good job teruslah berkarya | positive | positive |
| saya sudah langganan acara tvnya dan sudah lama konek kak hp tapi sekarang saya buka acaranya hilang kenapa bisa begini .. bolak balik harus telepon ke operatornya .. bikin saya malas dan mau ta putus langganan .. tolong solusinya secepatnya | negative | negative |
| untuk promo event sportnya lebih di buat menarik biarpun sudah akan tayang di tv reguler .. terimakasih | positive | positive |
| ini aplikasi kok jadi error.dulu lacar dan sekarang malah enggak bisa dipakai diluar negara padahal kemarin masih bisa pakai.dan data internet masih banyak . | negative | negative |
| tolong di perbaiki lagi masa pas aplikasi di buka kluar sendiri tnpa di pencet.dlu aplikasi nya sngt bagus sekarang jlek sekali | negative | negative |
| mantap.selain aplikasinya yang simpl enggak bertele tele.pokonya sip.mudah an juga bruntung dapat hp baru | positive | positive |
| dulu pernah pakai app ini bagus n lancar karena ganti hp kentang jadi ku copot . ee.skrng deh ganti hp pas download app ini kok cuma muter doang.auto copot | negative | negative |
| apakah kalo mau menangin hadiah harus daftar berlangganan dulu kalo iya apalah daya aku hanya seorang penganggur yang berharap bisa menangin hp bagus supaya saya bisa rekomendasi kan ke teman saya soal aplikasi ini . | positive | neutral |
| bermanfaat banget aplikasi nya kita bisa lihat acara ditv yang ketinggalan dan bisa dilihat kapan saja sewaktu luang kita | positive | positive |
| .nyoba mana tau saya dapat hp soalnya hp saya sudah tak layak pakai . .mantap tadi malam saya nonoton film karton orang purba enggak ada sedikit pun laq video nya nyaman sekali kasih film naruto dong min saya hobi menonton naruto tapi bahasanya yang indonesia ya | positive | positive |
| sumpah aplikasinya enggak hanya di hp ataupun di android tv berat banget jadi pas buka ngefreze lama banget | negative | positive |
| mau streaming film apa pun di vision banyak banget pilihannya | positive | positive |
| percuma pakai langganan nonto indonesia vs thailand saja enggak bis | negative | negative |
| bagus banget aplikasinya terbantu banget buat streaming dan menonton film favorit saat ketinggalan semoga kualitasnya ditingkatkan lagi | positive | positive |
| wah akhirnya ada juga aplikasi ini . dari dulu mencari yang seperti ini . keren .. tapi saya kasih bintang dulu ya . satu bintang lagi kalau saya dapat hadiah . kalau saya dapat hadiah . saya kasih bintang . sukses selalu untuk mnc | positive | positive |
| knapa enggak bisa di buka aplikasinya .. semalam enak saja itu sekarang malah meminta email ponsel padahal semlam sudah | negative | negative |
| live tidak sesuai dengan tv .masa di tv main bola sudah ada .di vision masih sinetron .kan enggak sama | negative | negative |
| mantap juragan biar di jalan dimanapun kapanpun jadi kalo pas bosan langsung streaming tv nya lewat aplikasi ini enggak ribet tapi tolobg di baguslan lagi graviknya hehe suwun suwun | positive | positive |
| nickelodeon bermasalah hanya setelah tahun baru setelah update januari sama saja enggak ada bahasa indonesia nya padahal sebelum update aman saja | negative | negative |
| keren sekarang menonton streaming film dan menonton tv pun bisa dimana saja dan kapan saja puas | positive | positive |
| aplikasi lambat padahal jaringan enggak | negative | negative |
| aplikasi ini jenjang keren selalu berkarya untuk negeri . salam mnc . | positive | positive |
| saya kurangi jadi bintang . karena apa karena buka aplikasi ya kok ada tulisan cek koneksi internet padahal jaringan internet saya baik saja . buka aplikasi lain saja bisa . kenapa yang ini enggak bisa .. ? | negative | negative |
| so far bagus banyak pilihan filmnya .. tolong film anak anak nya diperbanyak .. | positive | positive |
| nih aplikasi kenapa sih setiap mau dibuka suka keluar sendiri padahal jaringan bagus . sudah pakai wifi juga tetap saja . hp juga enggak ada kendala apa apa . tolong dong di perbaiki . ini sangat sangat menggangu . | negative | negative |
| saya sangat menyukai aplikasi ini banyak pilihan . pokok nya top deh untuk kwalitas gambar nya tolong di tingkatkan lagi biar bertambah jernih . bagus lah | positive | positive |
| aneh.aku enggak bisa daftar.kecewa banget .. padahal kode nya sudah dimasukin tapi enggak bisa gagal terus . | negative | negative |
| tolong diperbaiki saya kalau buka aplikasi ini pakai konekai dari aplikasi vpn selalu tidak bisa terbuka dianggap offline aliast tidak terhubung internet terutama bagian profil sama live streaming padahal di aplikasi streaming app lain masih bisa buat streaming live tv saya asli kecewa . | neutral | negative |
| seminggu deh pakai aplikasi ini bagus kok aplikasi nya buat menonton tv dsb suka deh | positive | positive |
| dari kemarin susah amat di buka lagi gangguan kah ? | negative | negative |
| maaf kok pas mau buka film enggak bisa ada tulisan security errornya | negative | neutral |
| nice nih cuman kurang kumplit mau menonton lagi yang film stake land nya enggak ada | positive | positive |
| aku beri bintang dulu ya soalx masih blum lemkap cennelx tolong dong kak di tambah sctv dan indosiar karena di keluarga ku pada pakai tv vision katax bagus tapi kurang sdikit lagi kalau yang cennel di pasang pasti lebih bagus dan amat bagus sekali . rasa kurang lengkap | neutral | neutral |
| untuk paket datanya bagaimana ? apakah ada paket buat vision biar enggak makan kouta data utama yang tinggi .. apakah ada kerja sama dngan telkomsel atau yang lain ? biar bisa dipaketin data enggak makan kouta utama | positive | neutral |
| vision bagus ahmad gue jadi bisa menonton film naruto tv saya rusak . aplikasi ini sangat membantu sekali . | positive | positive |
| mempermudah kami untuk menonton ketika sedang dalam perjalanan ataupun bepergian | positive | positive |
| vision sangat bagus membantu kita agar kita ketinggalan film favorit dan bagusnya bisa digunakan dipc juga | positive | positive |
| mohon maaf kok punya saya pass menonton nya disuruh beli.paket premium padahal saya buka aplikasi pakai jaringan wifi bukan data saluler memang aplikasi nya enggak bisa ditonton pakai wifi kah mohon info nya ? ? | negative | negative |
| ini kenapa setelah saya memakai aplikasi vision sinyal ditv saya malah hilang total yakali digital nya yang rusak kan enggak mungkin saya baru bulan pakai digital optus kak vision dan baik saja tapi setelah saya install aplikasi ini kok sinyal nya hilang semua padahal antena nya juga baru diganti astaga ! tolong penjelasannya aplikasi ini pantas mendapatkan bintang | negative | negative |
| sekarang parah. enggak kayak dulu . | negative | negative |
| sangat sangat membantu dengan adanya aplikasi mnc now buat saya yang kerja di lapangan bisa streaming menonton film kapan pun dimana pun . tks ya | positive | positive |
| aplikasi vision sangat bagus jadi bisa ngulang lagi kalau ketinggalan acara kesukaan . | positive | positive |
| kadang lelet semoga kedepan lebih lancar dan mudah di akses | positive | positive |
| mohon maaf saya kasih satu bintang karena sudah beberapa hari gagal koneksi padahal pakai wifi tapi tetap saja masih gagal koneksi mohon di perbaiki | negative | negative |
| senang sekali bisa menonton drakor lewat streaming langsung semoga tayangan drakornya lanjut terus ya saya suka.saya suka . | positive | positive |
| ayok pecinta film merapat karena dengan aplikasi vision enggak perlu lagi menonton ke bioskop . cukup rebahan dikamar | positive | positive |
| gue suka menonton animasi di semua mnc grup.bgus .tpi memang hdiah smartphone.nya betulan .. sampai tanggal berapa | positive | positive |
| bagus aplikasinya bermanfaat supaya enggak boring di jalan bisa buka aplikasi ini sambil stream tv . | positive | positive |
| sering buffer mohon diperbaiki servernya padahal internet saya speedtest mbps | neutral | negative |
| scam katanya enggak ada iklan ternyata ada iklan padahal sudah beli paket premium dan bukanya lama . | negative | negative |
| kalau dibuka tidak mau buka dengan alasan server sibuk tingkat tinggi .apa itu maksudnya .sy ingin lihat malah tidak bisa dilihat macet | positive | negative |
| saya coba pakai yang premium tapi enggak ada bedanya bahkan chanel net tv ada tapi pas di buka isinya rcti bahkan suka error .. | negative | negative |
| kenapa selalu gagal ngedaftar sih sudah enth berapa kali ku daftar enggak bisa bisa .knp sih .mohon lah di perbaiki .saya mendaftar nya melalui fb .knp enggak bisa .pake email waktu itu enggak bisa juga | negative | negative |
| saya sudah beli paket melalui sms dan sukses tetapi mengapa saya buka aplikasi nya tetap tidak bisa menonton siaran premium nya . saya sudah dirugikan saya mohon untuk segera diperbaiki masalah ini . saya sangat kecewa . | negative | negative |
| sangat bagus tapi tolong ya jangan ada lage yang dtolat atau terima | positive | positive |
| aku kecewa dengan aplikasi vision saya sangat suka menonton channel antv tetapi channel tersebut hilang dari aplikasi visionnya tolong dong antv masukin lagi | negative | negative |
| saya baru install terus mau menonton tv tidak bisa selalu ada pemberitahuan cek koneksi mu padahal jaringan bagus hampir full kecewa banget ditambah kuota data banyak | negative | negative |
| maaf ya min saya kecewa jadi saya turunkan bintang nya karena sekarang kok susah lihat siaran antv kenapa ya min sudah di perbarui log out sudah restart sudah jringan bagud juga yang lain bisa mutar antv saja yang enggak bisa bilangnya enggak ada koneksi internet | negative | negative |
| kenapa haru beli paket premium tolong di update lagi dong buat gratis kan sudah memakai data masak masih di suruh bayar buat beli paket sangat boros | negative | negative |
| saya waktu mau pilih canelnya selalu ada tulisan periksa koneksi anda padahal saya memakai wi fi dan saya coba pakai data seluler juga masih ada tulisan periksa koneksi jadi saya enggak bisa lihat apa di aplikasinya bagaimana caranya biar bisa nontonnya ? ? maksudnya periksa koneksi anda itu apa padahal saya pakai wi fi dan saya coba pakai data masih tidak bisa jadi caranya itu bagaimana . tolong di jawab . | negative | negative |
| kenapa sih ini tv pertama tama bagus pas sudah berapah hari eror .trus harus hapus aplikasi terus di download lagi baru bisa benar . tolong dong jangan kaya begini | negative | negative |
| aplikasi nya bagus banget ! ! maaf ya kakak ini dulu sudah saya dowload tapi di uninstal karena memo penuu jadi aku dowload lagi karena setiap aku mau menonton drama india saja dimlm hari enggak boleh sama mak katanya jangan ganggu lagi menonton ikatan cinta eh jadi curhat kan pokoknya aplikasi ini bagus nya pakai banget ! | positive | positive |
| aplikasinya bagus kemarin di tv enggak bisa lihat disney channel sekarang bisa mantap ! | positive | positive |
| kenapa ya setiap menonton tv on demand setelah dua menitan layar selalu kembali lagi ke awal .. apa harus ada pengaturan tertentu .. hape saya oppo a | negative | neutral |
| vision adalah aplikasi streming tv pulang buruk yang pernah saya coba . kuota pulang boros hai untuk menonton beberapa video banyak sekali iklan terlalu sering muncul notifikasi . bang yang mau aplikasi ini gunakan jaringan enggak dengan kuota ratusan giga baru bisa menikmati aplikasi ini . semoga beruntung ! | negative | positive |
| bagus banget aplikasih nya suka bisa menonton kapanpun dan dimanapun | positive | positive |
| aku suka banget deh aplikasi ini karena aku kalau di rumah rebutan remot tv jadi aku menonton tv di vision | positive | positive |
| developer tolong di perbaiki lagi vision plus versi android servernya sangat tidak stabil dibanding versi website . | negative | negative |
| aplikasi apa padahal lihat motogp sudah live juga di tv trans masak kok tayangan nya beda sama yang di tv .. masak suruh mbayar khusus menonton motogp .. kecuali kalau tidk di tayangin di tv baru saya lihat di aplikasi . mohon maaf sebelumnya . | negative | negative |
| kata film jagoan bola go go go bakal tayang lagi tapi sampai sekarang enggak ada saya mohon banget plis plis plis adakan lagi flim nya plis lah puaskan hati saya sekian dari saya terimakasi | positive | positive |
| aplikasi ini sangat bagus sekali ayo cepat download enggak bakal menyesal . terima kasih mnc now | positive | positive |
| enggak bisa masuk keluar mulu dari aplikasinya tiap dibuka | negative | negative |