**ANALISIS SENTIMEN PENGGUNA YOUTUBE MENGENAI VIRUS *PNEUMOLIA* DENGAN METODE *Vader Lexicon***

**Rafi Hardiansyah**1)**, Meyla Syahrani Utami2**)**, Hendri**3)

1) 2) 3) Teknik Informatika, Fakultas Sains Teknologi, Universitas Bina Darma

Jl. Jendral Ahmad Yani No.3, 9/10 Ulu, Kecamatan Seberang Ulu I, Kota Palembang.

1. [rafiardiansyah@gmail.com](mailto:rafiardiansyah@gmail.com)

2)  [meylasyahrani1@gmail.com](mailto:%20meylasyahrani1@gmail.com)

3)  [hendrineknag@gmail.com](mailto:%20hendrineknag@gmail.com)

**Abstrak**

Era digital telah membuka jalan bagi masyarakat untuk berbagi pendapat dan perasaan mereka tentang berbagai topik. Salah satu topik yang sering dibahas adalah pandemi virus pneumonia. Salah satu platform yang sering digunakan untuk berbagi pendapat adalah YouTube. Di YouTube, pengguna dapat mengekspresikan sentimen mereka melalui komentar. Komentar ini dapat memberikan gambaran tentang persepsi publik terhadap pandemi.Penelitian ini melakukan analisis dengan mengumpulkan komentar dari video YouTube yang berjudul “Seberapa baha Wabah Pneumonia Misterius dari TiongkokIni Penjelasan Pengamat”. Data yang berhasil terhimpun dari komentar tersebut kemudian menjadi fokus utama analisis dalam penelitian ini menggunakan bahasa pemrograman Python dan memanfaatkan library nltk, tweet-preprocessor, textblob, dilakukan beberapa langkah, termasuk Translate, Cleansing, Case Folding, Tokenizing, Stopword, dan Stemming.mayoritas komentar memberi respon netral 99(52.7%) sedangkan respon positif 54(29.3%) dan negatif 33(18.1%). Mayoritas pengguna YouTube tampaknya memiliki pandangan yang netral terhadap virus pneumonia. Namun, jumlah komentar positif hampir dua kali lipat dari komentar negatif. Hal ini menunjukkan bahwa ada lebih banyak pengguna YouTube yang memiliki persepsi positif atau mendukung upaya penanganan virus pneumonia dibandingkan dengan mereka yang memiliki sentimen negatif. Meski demikian, masih ada sejumlah pengguna YouTube yang memberikan sentimen negatif. Hal ini menunjukkan bahwa masih ada tantangan dalam penanganan dan komunikasi informasi tentang virus pneumonia kepada publik.

**Kata Kunci :** *Analisis Sentimen*, *Vader Lexicion*, *Youtube*, *Newcovid*

**Abstract**

*The digital age has opened the way for people to share their opinions and feelings on a variety of topics. One of the topics that is frequently discussed is the pneumonia virus pandemic. One platform that is often used to share opinions is YouTube. On YouTube, users can express their feelings through comments. This commentary can provide an overview of the public's perception of the pandemic.The study carried out the analysis by collecting comments from a YouTube video entitled "Some of the mysterious pneumonia outbreaks from China are an observer's explanation". The data collected from the comments then became the main focus of the analysis in the study using the Python programming language and using the nltk library, tweet-preprocessor, textblob, performed several steps, including Translate, Cleansing, Case Folding, Tokenizing, Stopword, and Stemming.The majority of comments gave a neutral response of 99 (52.7%) while a positive response of 54(29.3%) and a negative response of 33(18.1%). The majority of YouTube users seem to have a neutral view of the pneumonia virus. However, the number of positive comments is almost twice that of negative comments. This suggests that there are more YouTube users who have positive perceptions or support efforts to treat the pneumonia virus than those who have negative feelings. It shows that there are still challenges in the handling and communication of information about the pneumonia virus to the public.*

***Keyword*** *: Sentiment Analysis, Vader Lexicion, Youtube Newcovid*

1. **PENDAHULUAN**

Era digital telah membuka jalan bagi masyarakat untuk berbagi pendapat dan perasaan mereka tentang berbagai topik. Salah satu topik yang sering dibahas adalah pandemi virus pneumonia. Salah satu platform yang sering digunakan untuk berbagi pendapat adalah YouTube. Di YouTube, pengguna dapat mengekspresikan sentimen mereka melalui komentar. Komentar ini dapat memberikan gambaran tentang persepsi publik terhadap pandemi.

Analisis sentimen adalah metode yang efektif untuk memahami persepsi publik terhadap suatu topik atau peristiwa. Dengan menganalisis sentimen, kita dapat memahami bagaimana opini publik berkembang dan berubah seiring waktu. Ini sangat penting dalam konteks pandemi virus pneumonia. Dengan memahami sentimen publik, kita dapat merumuskan strategi yang lebih efektif untuk mengatasi pandemi. Selain itu, analisis sentimen juga dapat membantu kita memahami bagaimana pandemi mempengaruhi masyarakat.

Virus pneumonia telah menjadi topik yang banyak dibicarakan di berbagai platform media sosial, termasuk YouTube. Dengan menganalisis sentimen pengguna YouTube, kita dapat memahami bagaimana persepsi publik terhadap virus pneumonia berkembang seiring waktu. Hal ini penting untuk membantu pihak berwenang dan organisasi kesehatan dalam merumuskan strategi komunikasi dan intervensi yang efektif. Analisis sentimen juga dapat membantu kita memahami bagaimana masyarakat merespons terhadap berbagai upaya penanganan pandemi. Selain itu, analisis sentimen juga dapat memberikan gambaran tentang bagaimana pandemi mempengaruhi kehidupan sehari-hari masyarakat.

Untuk melakukan analisis sentimen, kita menggunakan metode Vader Lexicon. Metode ini telah terbukti efektif dalam menganalisis sentimen dalam teks, termasuk komentar YouTube. Metode Vader Lexicon dapat mengidentifikasi sentimen positif, negatif, dan netral dalam teks. Selain itu, metode ini juga dapat memberikan skor sentimen yang mencerminkan intensitas sentimen dalam teks. Dengan menggunakan metode ini, kita dapat memahami bagaimana sentimen berubah seiring waktu.

Dengan menggunakan metode Vader Lexicon, kita dapat mengidentifikasi sentimen positif, negatif, dan netral dalam komentar dan memahami bagaimana sentimen ini berubah seiring waktu. Ini memberikan gambaran yang lebih baik tentang bagaimana masyarakat merespons dan bereaksi terhadap pandemi virus pneumonia. Analisis sentimen juga dapat membantu kita memahami bagaimana pandemi mempengaruhi kehidupan sehari-hari masyarakat. Selain itu, analisis sentimen juga dapat memberikan wawasan tentang bagaimana pandemi mempengaruhi persepsi masyarakat tentang kesehatan dan keselamatan mereka. Akhirnya, analisis sentimen dapat memberikan informasi berharga yang dapat digunakan untuk merumuskan strategi penanganan pandemi yang lebih efektif.

1. **METODOLOGI PENELITIAN**
2. **Diagram Alur Proses**

Penelitian ini melakukan analisis dengan mengumpulkan komentar dari video YouTube yang berjudul “Seberapa baha Wabah Pneumonia Misterius dari Tiongkok?Ini Penjelasan Pengamat”. Data yang berhasil terhimpun dari komentar tersebut kemudian menjadi fokus utama analisis dalam penelitian ini. Gambaran alur proses penelitian dapat dilihat pada Gambar 1.

klasifikasi

Pengumpulan data

Dataset

Preprocessing

Gambar 1. Diagram Alur Proses

Gambar tersebut menggambarkan secara keseluruhan langkah-langkah yang dilakukan dalam penelitian ini. Proses dimulai dengan mengumpulkan data berupa teks komentar dari YouTube. Data tersebut kemudian menjalani tahap preprocessing dan klasifikasi.

1. **Pengumpulan Data**

Penelitian ini mengumpulkan data dengan melakukan pengambilan komentar di YouTube melalui teknik scrapping pada video dari channel KOMPASTV. Proses scrapping dilakukan menggunakan website https://netlytic.org/. Data yang ingin kita kumpulkan kemudian kita copy videoidnya lalu dimasukkan pada website. Netlytic akan melakukan pengumpulan secara otomatis dan kemudian data hasilnya akan dapat kita download dengan format csv atau excel, data dapat kita ambil untuk proses selanjutnya.

1. ***Preprocessing***

*Preprocessing* data adalah praktik membersihkan, mengubah, dan mengatur ulang data mentah sebelum diproses dan dianalisis, yang juga dikenal sebagai persiapan data[1]. Pada tahap ini, menggunakan bahasa pemrograman Python dan memanfaatkan library nltk, tweet-preprocessor, textblob, dilakukan beberapa langkah, termasuk *Translate*, *Cleansing, Case Folding, Tokenizing, Stopword,* dan *Stemming*.

**Translate:** Data ditranslate ke bahasa Inggris menggunakan google sheets.

**Cleansing:** Data dibersihkan dari simbol, emoji, tanda baca, URL, dan angka.

**Case Folding:** Dataset dengan huruf kapital diubah menjadi huruf kecil.

**Tokenizing:** Teks atau kalimat dipisahkan menjadi bagian-bagian kecil, yang disebut "token".

**Stopword:** Kata hubung atau kata-kata tidak penting dalam kalimat dihilangkan.

**Stemming:** Kata-kata disederhanakan ke bentuk dasar agar dianggap sebagai bentuk yang sama.

1. **Klasifikasi**

Klasifikasi adalah suatu proses untuk memberikan label sentiment pada dataset. Data yang ditampung kemudian dilabeli dengan memanfaatkan polarity[2]. Pemrograman dilakukan dengan menggunakan bahasa Python, dan *labelling* diukur berdasarkan jumlah kata yang dikategorikan sebagai kata positif, negative dan netral. Label diberikan sesuai dengan kategori yang memiliki jumlah tertinggi.

Pada tahapan ini akan dilakukan proses menggunakan algoritma Vader Lexicon. Pelabelan data menggunakan empat sumber leksikal: VADER. VADER (Valence Aware Dictionary for Sentiment Reasoning) adalah model analisis sentimen yang menggabungkan leksikon dan berbasis aturan. VADER mempertimbangkan tanda baca dan huruf besar untuk meningkatkan akurasi nilai sentimen dalam sebuah teks. VADER tidak hanya mengklasifikasikan kategori sentimen negatif, netral, dan positif, tetapi juga membagi intensitas dari setiap kategori. Kerangka kerja VADER mencakup tiga tahap Preprocessing, yang terdiri dari tokenisasi, penghilangan stopword, dan penghilangan tanda baca; Peningkatan data dilakukan pada kata-kata seperti 'hebat', 'sangat', dan 'luar biasa' [3]. Data yang telah melalui tahap *scrapping* diproses dengan algoritma tersebut. Mesin akan mengenali pola dan mengklasifikasikan data ke dalam tiga kategori: positif, netral, dan negatif..

Penelitian ini menggunakan VADER dari paket SentimentIntensityAnalyzer di Python, yang memiliki nilai majemuk dengan rentang -1 dan 1, untuk menentukan jenis sentimen untuk setiap baris dalam dataset. Nilai gabungan yang dihasilkan kemudian dikategorikan menjadi tiga jenis, nilai <= -0.05 adalah negatif, nilai > -0.05 dan < 0.05 adalah netral, dan nilai >= 0.05 akan menghasilkan label positif.[4].

1. **HASIL DAN PEMBAHASAN**
2. **Scrapping**

Data dalam penelitian ini diperoleh dari komentar pengguna Youtube pada video di channel KOMPASTV dalam tayangan yang berjudul “Seberapa baha Wabah Pneumonia Misterius dari Tiongkok?Ini Penjelasan Pengamat|”. Dalam penelitian ini diambil sebanyak 186 komentar yang didapat dari video yang dipilih. Komentar diambil pada tanggal 28 Desember 2023. Hasil scrapping dipaparkan sebagai berikut :

|  |
| --- |
| **Komentar** |
| Kok masih pcy, liat tuh corona stelah ada perang Rusia vs ukraina hilang tuh.. Ini peramainan dunia. Saya pcy adanya penyakit tp seperti coruna dan ini terlalu dibesarkan |
| Cuan n cuan.. Bisnis pcr, swab dan vaksin Corona virus....sukses yg kaya makin kaya.. Yg susah makin stres dam gila...wkkkk |
| Semoga cukup di Tiongkok jgn menyebar,ke negara2 lain. |
| Itu kemungkinan karena virus nyamuk wolbachia |
| Bahaya untuk anak umur di bawah 3 tahun ðŸ˜¢ðŸ˜¢ðŸ˜¢ðŸ˜¢.. |

1. **Text Preprocessing**

Sebelum kita melakukan pengelolah data menggunakan vader lexicon, data harus kita translate karena libarary Vader lexicon hanya membaca data dengan bahasa inggris dan langkah preprosesing harus dijalani terlebih dahulu. Proses preprocessing diperlukan untuk menangani data yang tidak konsisten, sehingga data dapat dipersiapkan untuk proses selanjutnya[5].

Langkah-langkah dalam *preprocessing* melibatkan *Translate*, *Tokenizing*, Normalisasi, *Stemming*, dan penghapusan *Stopword*. Pada proses ini digunakan bahasa Pemrograman Python dengan library *textblob*, *tweet-preprocessor* dan *pandas* untuk menghapus karakter yang tidak diinginkan. Library nltk untuk membuat token dan menghapus stopword. Berikut hasil dari tahap text preprocessing yang telah dilakukan :

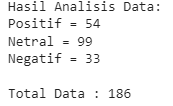
|  |  |
| --- | --- |
| **Komentar** | **Hasil Text Preprocessing** |
| Kok masih pcy, liat tuh corona stelah ada perang Rusia vs ukraina hilang tuh.. Ini peramainan dunia. Saya pcy adanya penyakit tp seperti coruna dan ini terlalu dibesarkan | how come its still pcy see corona after the russian vs ukraine war is gone this is a world tamper i am a disease like a coruna and this is too raised |
| Cuan n cuan.. Bisnis pcr, swab dan vaksin Corona virus....sukses yg kaya makin kaya.. Yg susah makin stres dam gila...wkkkk | cuan n cuan pcr business swab and vaccine corona virus success that is rich getting richer which is difficult to stress and crazy wkkkk |
| Semoga cukup di Tiongkok jgn menyebar,ke negara2 lain. | hopefully enough in china does not spread to other countries |
| Itu kemungkinan karena virus nyamuk wolbachia | thats probably because of the wolbachia mosquito virus |
| Bahaya untuk anak umur di bawah 3 tahun ðŸ˜¢ðŸ˜¢ðŸ˜¢ðŸ˜¢.. | danger for children under years old |

1. **Klasifikasi**

Data sentimen komentar video Youtube dilakukan proses klasifikasi menggunakan bahasa pemrograman Python yang telah dibagi menjadi file dengan kata positif dan kata negatif. Sistem yang dijalankan pada program tersebut yaitu dengan menghitung jumlah kata positif dan negatif dalam komentar lalu membandingkan untuk menentukan sentimen komentar. Berikut hasil klasifikasi yang telah dilakukan pada penelitian ini:

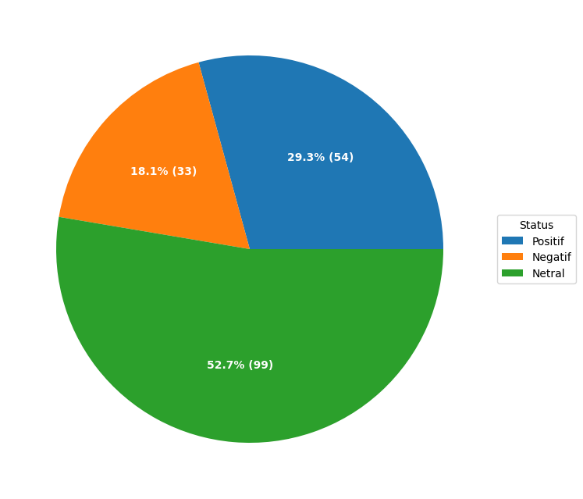
|  |  |  |
| --- | --- | --- |
| **Komentar** | **Hasil Text Preprocessing** | **Klasifikasi** |
| Kok masih pcy, liat tuh corona stelah ada perang Rusia vs ukraina hilang tuh.. Ini peramainan dunia. Saya pcy adanya penyakit tp seperti coruna dan ini terlalu dibesarkan | how come its still pcy see corona after the russian vs ukraine war is gone this is a world tamper i am a disease like a coruna and this is too raised | netral |
| Cuan n cuan.. Bisnis pcr, swab dan vaksin Corona virus....sukses yg kaya makin kaya.. Yg susah makin stres dam gila...wkkkk | cuan n cuan pcr business swab and vaccine corona virus success that is rich getting richer which is difficult to stress and crazy wkkkk | negatif |
| Semoga cukup di Tiongkok jgn menyebar,ke negara2 lain. | hopefully enough in china does not spread to other countries | negatif |
| Itu kemungkinan karena virus nyamuk wolbachia | thats probably because of the wolbachia mosquito virus | netral |
| Bahaya untuk anak umur di bawah 3 tahun ðŸ˜¢ðŸ˜¢ðŸ˜¢ðŸ˜¢.. | danger for children under years old | positif |

Penelitian ini dilakukan dengan 186 data yang diperoleh dari 1 video youtube di akun KOMPASTV dan telah melalui tahap preprocessing. Data tersebut dilakukan klasifikasi menggunakan Vader Lexicon. Berikut hasil klasifikasi data yag dilakukan dengan menggunakan bahasa pemrograman Python :



Gambar 1. Hasil Klasifikasi

Metode klasifikasi Vader Lexicon dapat kita liat juga penyebaran datanya melalui diagram pie sebagai berikut:



Gambar 2. Diagram Pie

Pada gambar diagram diatas dapat kita liat mayoritas komentar memberi respon netral(52.7%) sedangkan respon positif(29.3%) dan negatif(18.1%). Adapun kata yang sering muncul dapat kita buat menggunakan *wordcould* dengan hasil yang terdapat pada gambar.

****

Gambar 3. Wordcould

1. **KESIMPULAN**

Penelitian ini menganalisis sentimen pengguna YouTube terhadap berita munculnya penyakit baru bernama pneumonia menggunakan metode klasifikasi Vader Lexicon. Hasilnya menunjukkan bahwa dari total komentar yang dianalisis, 54 komentar memberikan sentimen positif, 99 komentar bersifat netral, dan 33 komentar memberikan sentimen negatif terhadap virus pneumonia.

Mayoritas pengguna YouTube tampaknya memiliki pandangan yang netral terhadap virus pneumonia. Namun, jumlah komentar positif hampir dua kali lipat dari komentar negatif. Hal ini menunjukkan bahwa ada lebih banyak pengguna YouTube yang memiliki persepsi positif atau mendukung upaya penanganan virus pneumonia dibandingkan dengan mereka yang memiliki sentimen negatif. Meski demikian, masih ada sejumlah pengguna YouTube yang memberikan sentimen negatif. Hal ini menunjukkan bahwa masih ada tantangan dalam penanganan dan komunikasi informasi tentang virus pneumonia kepada publik.

Penelitian ini menunjukkan pentingnya memahami sentimen publik terhadap isu kesehatan seperti virus pneumonia. Dengan memahami sentimen ini, para peneliti dan praktisi kesehatan masyarakat dapat merancang strategi komunikasi dan edukasi yang lebih efektif untuk meningkatkan pemahaman dan dukungan publik terhadap upaya penanganan virus pneumonia.

1. **REFERENSI**

[1] C. Fan, M. Chen, X. Wang, J. Wang, and B. Huang, “A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data,” *Front. Energy Res.*, vol. 9, no. March, pp. 1–17, 2021, doi: 10.3389/fenrg.2021.652801.

[2] S. F. Pane and J. Ramdan, “Pemodelan Machine Learning : Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Menggunakan Data Twitter,” *J. Sist. Cerdas*, vol. 5, no. 1, pp. 12–20, 2022, doi: 10.37396/jsc.v5i1.191.

[3] C. Dev, A. Ganguly, and H. Borkakoty, “Assamese VADER: A Sentiment Analysis Approach Using Modified VADER,” in *2021 International Conference on Intelligent Technologies (CONIT)*, 2021, pp. 1–5. doi: 10.1109/CONIT51480.2021.9498455.

[4] M. Hayaty and A. H. Pratama, “Performance of Lexical Resource and Manual Labeling on Long Short-Term Memory Model for Text Classification,” *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 1, pp. 74–84, 2023, doi: 10.26555/jiteki.v9i1.25375.

[5] S. W. Min, D. Komunikasi, and D. Komunikasi, “Comparative Evaluation of Lexicons in Performing,” *J. Adv. Comput. Technol. Appl.*, vol. 2, no. 1, pp. 14–20, 2020.