Bagaimana Masyarakat Menyikapi Pembelajaran Tatap Muka: Analisis Komentar Masyarakat pada Media Sosial Youtube Menggunakan Algoritma Deep Learning Sekuensial dan LDA

Fawwaz Zaini Ahmad#1, Muhammad Fauzi Satria Arifandy#2, Muhammad Rasyad Caesarardhi#3, Nur Aini Rakhmawati#4

#Departemen Sistem Informasi, Institut Teknologi Sepuluh Nopember  
Kampus ITS Sukolilo, Surabaya, Jawa Timur

1fawwaz.18052@mhs.its.ac.id

2arifandy.18052@mhs.its.ac.id

3caesarardhi.19052@mhs.its.ac.id

4nur.aini@is.its.ac.id

Abstrak— Saat Pandemi Covid-19 memasuki tahun ke-2, ada beberapa pertanyaan yang mengganjal di antara kita, apakah kita harus terus belajar online, atau berkompromi dengan Pandemi dan membuka kembali sekolah kita. Saat vaksinasi dimulai di seluruh dunia, pertanyaan 'apakah kita harus segera membuka kembali sekolah kita?' menjadi semakin keras. Tujuan dari penelitian ini adalah menganalisis sentiment sebagian masyarakat Indonesia, dalam hal ini pengguna Youtube, dalam hal pembelajaran tatap muka. Data yang digunakan dalam penelitian ini adalah komentar dari sembilan video youtube yang berkaitan dengan pembelajaran tatap muka. Persiapan data meliputi *upsampling, casefolding, cleansing, labelling, dan tokenizing.* Algoritma yang digunakan adalah pemodelan data LSTM dengan pengoptimasi ADAM Setelah pemodelan, setiap model dievaluasi dan berhasil memperoleh akurasi tertinggi sebesar 78%. Setelah dilakukan permodelan, model dengan akurasi tertinggi digunakan untuk memberi label secara otomatis pada komentar yang belum memiliki label. Langkah terakhir dari penelitian ini adalah melakukan permodelan topik terhadap setiap sentiment, guna mencari tau setiap sentiment berkaitan dengan topik apa.

Kata kunci— LSTM, LDA, Youtube, Analisis Sentimen

1. PENDAHULUAN

Pada 31 Desember 2019 dunia digemparkan oleh pelaporan pertama kali kasus novel Corona virus atau yang sekarang disebut dengan COVID-19 di Wuhan, China [1]. Berdasarkan data dari WHO terdapat 14.348.858 kasus terkonfirmasi di seluruh dunia pada Juli 2020, jumlah ini dapat dipastikan terus bertambah sampai detik ini. Pandemi ini terjadi di seluruh belahan dunia, tak terkecuali Indonesia. Menurut data Satgas COVID-19 di Indonesia tercatat terdapat 132.816 kasus sampai pertengahan Agustus 2020.

Pandemi COVID-19 mengubah banyak hal dalam hidup masyarakat saat ini. Perlu dilakukan modifikasi perilaku, supaya penyebaran COVID-19 dapat dibatasi sehingga sistem kesehatan maupun pendidikan tidak terbebani. Untuk itu, saat ini pemerintah bergerak menuju penerapan Pembelajaran Tatap Muka. pelaksanaan pembelajaran tatap muka disesuaikan dengan level PPKM sesuai Instruksi Menteri Dalam Negeri Nomor 41 Tahun 2021 Tentang Pemberlakuan Pembatasan Kegiatan Masyarakat Level 3, Level 2, dan Level 1, Serta Mengoptimalkan Posko Penanganan Corona Virus Disease 2019 di Tingkat Desa dan Kelurahan Untuk Pengendalian Penyebaran Corona Virus Disease 2019 [2]. Kebijakan Pemerintah Indonesia ini ditandai dengan usaha untuk mengembalikan aktivitas masyarakat pada bidang pendidikan menjadi “normal” kembali dengan tetap meminta perguruan tinggi dan masyarakat untuk memperhatikan protokol kesehatan yang berlaku.

Demi keberhasilan kebijakan, penting bagi pemerintah dan stakeholder terkait untuk dapat memahami sikap masyarakat terhadap kebijakan tersebut dan mengembangkan strategi komunikasi yang efektif supaya kebijakan dapat dipahami dan dilaksanakan oleh masyarakat. Ini penting sekali terutama untuk kebijakan Pembelajaran Tatap Muka, dimana kepatuhan masyarakat merupakan kunci keberhasilannya. Untuk dapat memahami sikap dan tingkat pemahaman masyarakat terhadap kebijakan ini, terdapat berbagai sumber yang dapat dipergunakan. Salah satu sumber yang penting saat ini adalah media sosial Youtube.

Berdasarkan laporan data dari Hootsuite (We Are Social) pada 2020 Youtube merupakan platform media sosial yang paling banyak dipakai oleh masyarakat Indonesia yakni 88% dari jumlah populasi. Pengguna media sosial terbiasa untuk mengutarakan opininya melalui komentar pada video Youtube yang menarik perhatiannya. Sentimen atau opini dari media sosial ini memberikan informasi yang paling up to date dan inklusif, hal tersebut dikarenakan maraknya media sosial yang memberikan hambatan rendah untuk memposting sebuah pesan [3]. Komentar-komentar masyarakat terutama pada video berita yang berkaitan dengan Pembelajaran Tatap Muka adalah sebuah bentuk penyaluran aspirasi bagi masyarakat dan fungsi kontrol kepada pemerintah. Selain itu, opini masyarakat ini juga merupakan hasil dari pemikiran mereka yang melandasi tindakan dalam menjalankan aktivitas sehari-hari.

Jika pada penelitian yang dilakukan sebelumnya telah dibahas mengenai sentimen terkait sekolah daring pada Twitter dan menghasilkan kesimpulan bahwa kebijakan yang ditetapkan pemerintah terkait sekolah daring memiliki beberapa kekurangan. Kekurangan dari kebijakan ini ialah workload yang cenderung membuat siswa lebih mudah lelah, ditambah dengan metode pengajaran yang belum efektif membuat siswa memahami materi yang diberikan sekolah akan tetapi untuk kelebihannya berupa masyarakat setuju atas kebijakan yang diberikan dan merasakan dampak positif akibat kebijakan tersebut [4]. Maka pada penelitian ini penulis akan berfokus pada pengembangan model dan pengaplikasiannya pada data baru untuk analisis komentar Youtube yang berkaitan dengan kebijakan Pembelajaran Tatap Muka yang diterapkan untuk dilakukan analisis sentimen, klasifikasi multikelas, dan memproses dengan metode Natural Language Processing (NLP) serta Sequential Deep Learning. Komentar yang diberikan masyarakat ini dapat diklasifikan lebih lanjut, dilakukan pengamatan terhadap sentimen, dan dianalisis dari segi kebahasaan yang digunakan.

1. TINJAUAN PUSTAKA

Pada bagian ini akan dijabarkan beberapa penelitian sebelumnya yang dijadikan acuan dalam pengerjaan penelitian ini serta berisi dasar teori sebagai penunjang.

2.1 Studi Literatur

Peneliti di berbagai belahan dunia telah mencoba untuk memahami berbagai aspek dari pandemi COVID-19. Sejak penyakit COVID-19 diidentifikasi pada akhir 2019, banyak studi dari berbagai macam bidang telah dipublikasikan. Salah satu contohnya adalah penelitian pengolahan bahasa alami yang menganalisis COVID-19 di media sosial. J. Samuel et al. [5] melakukan analisis sentimen terhadap kebijakan pembukaan kembali Amerika Serikat. Penelitian ini menggunakan dataset dari twitter sebanyak 293,597 tweet yang didapat dengan cara crawling.

Selain itu Ghasiya et al. [6] melakukan permodelan topik menggunakan Top2Vec dan analisis sentimen menggunakan algoritma RoBERTa terhadap berita terkait COVID-19 di 4 negara, yakni Britania Raya, India, Jepang, dan Korea Selatan. Pada penelitian tersebut dikumpulkan 102,278 artikel, setelah dikumpulkan artikel tersebut kemudian permodelan topik, serta dilakukan analisis sentimen terhadap headline dari artikel tersebut. Setelah selesai dilakukan permodelan topik dan analisis sentimen, hasil dari dua proses tersebut dilakukan analisis mendalam guna memahami sifat dalam pembuatan berita dari keempat negara tersebut.

Kemudian Oktavianto dan Persada [7] melakukan analisis sentimen menggunakan *ELK Stack* dan Python, terhadap persepsi publik terhadap pembelajaran daring. Penelitian ini menggunakan 133,417 tweet yang didapat dengan cara crawling.

Berdasarkan tiga penelitian tersebut, kami mengidentifikasi adanya research gap dalam penelitian pengolahan bahasa alami yang berkaitan dengan pandemi COVID-19. Meskipun permodelan topik sudah sering dilakukan, belum ada permodelan topik yang spesifik terhadap komentar masyarakat seputar pembelajaran tatap muka di Indonesia. Selain itu, belum ada juga penelitian analisis sentimen seputar topik pembelajaran tatap muka pada media sosial. Kedua hal tersebut adalah topik yang akan diperdalam pada penelitian ini.

2.2 Analisis Sentimen

Analisis sentimen adalah salah satu cabang dari Pengolahan Bahasa Alami. Analisis sentimen pada komentar berfokus pada analisis orientasi sentimen dari korpus komentar tersebut, yang menunjukkan bahwa pengguna mengekspresikan sentimen positif, negatif, atau netral terhadap produk atau peristiwa [8]. Cara yang dapat dilakukan untuk menghasilkan polaritas sentimen dapat diraih dengan melakukan klasifikasi multikelas yang merupakan bagian dari tugas sebuah model supervised learning dalam mengklasifikasikan sesuatu menjadi lebih dari dua jenis kelas [9].

2.3 Permodelan Topik

Pemodelan topik merupakan salah satu cabang dalam Pengolahan Bahasa Alami yang digunakan untuk menemukan sebuah topik pada dokumen. Salah satu metode pemodelan topik yang populer adalah Latent Dirichlet Allocation LDA. LDA adalah model probabilistik generatif dari korpus. Kata-kata dengan probabilitas tertinggi di setiap topik biasanya memberikan gambaran yang baik tentang topik apa yang dapat dikatakan probabilitas dari LDA [10]. Pada penelitian ini LDA akan diterapkan untuk menemukan topik dominan pada komentar media sosial mengenai pembelajaran tatap muka.

2.4 Long Short-Term Memory

*Long Short-Term Memory* merupakan varian dari *Recurrent Neural Network* (RNN), yang diciptakan oleh Hocreiter et al [11]. pada tahun 1997 guna mengatasi kekurangan yang dimiliki RNN . LSTM terdiri dari tiga *gate* yaitu, *forget gate*, *input gate*,dan *output gate* [12].

![Diagram

Description automatically generated]()

Gambar. 1 Cara kerja LSTM

Pada gambar. 1 pada *forget gate* dihilangkan informasi dari kondisi saat ini dengan mengaplikasikan fungsi sigmoid pada kondisi langkah sebelumnya yang disembunyikan dan input saat ini dengan bobot *forget gate*  dan bias *forget gate* yang ditunjukkan dalam persamaan di bawah ini.

(2.1)

Sedangkan *input gate* mengontrol laju informasi baru ke sel memori. *Input gate* bekerja menurut persamaan di bawah ini.

(2.2)

Kemudian *output gate* mengontrol informasi yang dibutuhkan untuk aktivasi output dari sel memori. *Output gate* bekerja menurut persamaan ini.

(2.3)

dan merupakan bobot dari *input gate* dan *output gate*. Sedangkan dan meupakan bias dari *input gate* dan *output gate*. Kemudian setelah melalui *gate*, pada persamaan di bawah ini kondisi sel saat ini dari LSTM diperbarui dengan *forget gate* dan *input gate* yang diaplikasikan pada kondisi sel sebelumnya dan fungsi tanh dari dan . dan merupakan bobot dan bias dari kondisi sel saat ini.

(2.4)

Kondisi output dihitung sebagaimana persamaan di bawah ini, dengan *output gate* dan fungsi aktivasi tanh diaplikasikan pada kondisi sel saat ini yang kemudian diberikan pada memori sel selanjutnya dan juga pada *layer* selanjutnya.

(2.5)

[1]

Algoritma ini dapat mengekstrak dan mengingat pola data yang mendasari ketergantungan data jangka panjang pada data – data hasil klasifikasi menggunakan alogoritma RNN. LSTM sendiri terdiri dari berbagai bentuk seperti Traditional LSTM (T-LSTM), Peephole LSTM, dan Convolutional LSTM (Conv-LSTM). Masing – masing bentuk memiliki arsitekturnya tersendiri [12].

1. METODOLOGI

Metodologi penelitian ini dimulai dari pengumpulan data dari Youtube, kemudian dilakukan pra-proses data seperti *casefolding, cleansing,* pembuatan kategori dan *labelling*, dan *tokenizing*. Lalu dilakukan *upsampling*, *splitting*, implementasi *word embedding*, pembuatan model *Sequential Deep Learning*, analisa evaluasi performa model, implementasi model pada data baru, dan evaluasi sentimen serta kebahasaan.

3.1 Pengumpulan Data

Pengumpulan data diambil dari komentar Youtube sebanyak X post yang berkaitan dengan kebijakan New Normal. Pengumpulan data dilakukan dengan bahasa pemrograman Python modifikasi yang memanfaatkan library Selenium Webdriver dan dilakukan selama X hari.

3.2 Pra-proses Data

Sebelum dilakukan training data dengan model yang akan dibuat, data mentah yang didapatkan dari scraping akan diproses terlebih dahulu untuk memperoleh hasil yang maksimal dari *training* dan *test* pada model. Dalam melakukan pra-pemrosesan data diuraikan pada sub bab berikut.

3.2.1 Casefolding

Pada tahap ini dilakukan proses pengubahan semua huruf pada data yang diperoleh menjadi huruf kecil. Perubahan yang dilakukan hanya 26 huruf yaitu dari a sampai z, selain karakter tersebut akan dihilangkan. Tahap ini dilakukan dengan Python tanpa library apapun.

3.2.2 Cleansing

Pada langkah ini, angka, tanda baca atau simbol, karakter kosong, serta baris data yang terpisah dengan baris baru atau enter pada satu kolom akan dihilangkan atau dihapus. Karakter khusus harus dihilangkan untuk menghilangkan perbedaan dalam pemrosesan polarisasi. Langkah ini dilakukan dengan Python, khususnya *regular expression (regex)* saat menghapus angka.

3.2.2 Pembuatan Kategori dan Labelling

Pada tahap ini dilakukan pembuatan daftar kategori dan label untuk menganalisis perilaku masyarakat dan kebahasaan. Nantinya terdapat empat kategori yang pada label. Label ini akan menjadi acuan untuk *labelling* pada dataset yang digunakan untuk pelatihan (*training*).

3.2.3 *Tokenizing*

Pada tahap ini dilakukan proses tokenisasi yang merupakan proses untuk memisahkan teks menjadi kata, frasa, simbol, atau elemen yang bermakna lainnya yang biasa disebut token. Tujuan dari proses ini adalah untuk ,mengeksplor kata-kata pada sebuah kalimat. Dalam penelitian ini, proses tokeniasi dilakukan dengan memisahkan kalimat menjadi kata-kata untuk dilakukan analisa lebih lanjut.

3.3 Upsampling

Pada tahap ini dilakukan *upsampling* dikarenakan dataset memiliki *imbalanced* *classes*. *Upsampling* dilakukan dengan membuat suatu fungsi yang berulang terus menerus sehingga mengambil sampel dengan *replacement*, dari kelas minoritas hingga kelas yang ukurannya sama dengan mayoritas.

3.4 Splitting

Pada tahap ini dilakukan proses pemisahan data menjadi dataset *training* dan dataset *test*. Dataset *training* digunakan untuk pelatihan model, sementara dataset *test* digunakan untuk pengujian model dan performa yang diukur melalui skor evaluasi.

3.5 Implementasi Word Embedding

*Word embedding* merupakan jenis representasi istilah yang memungkinkan kata-kata menggunakan makna serupa mempunyai representasi yang serupa.Seperti contohnya kata ‘tidak dan ‘nggak’ mempunyai arti yang sama. Implementasi *Word Embedding* model dilakukan dengan tujuan untuk memetakan kata menjadi vektor yang digunakan untuk banyak macam tugas *Natural Languange Processing*, dimana dalam penelitian ini yakni untuk diproses dalam arsitektur LSTM. Implementasi *Word Embedding* dalam penelitian ini memakai *library word2vec* dan pembuatan model memakai *genism* menggunakan dataset Wikipedia.

3.6 Pembuatan Model Long Term Short Memory (LSTM)

Pada tahap ini dilakukan training data dengan model *Long Term Short Memory* (LSTM) dan model pembanding Bidirectional Long Short Term Memory (Bi-LSTM) menggunakan Python dengan library Keras Tensorflow.

3.7 Analisa Evaluasi Performa Model

Pada tahap ini dilakukan proses pelatihan dataset *training* dengan model LSTM yang telah dibuat, kemudian dilakukan prediksi dengan data *test* yang belum pernah dipelajari oleh model yang telah dibuat. Kemudian dilakukan evaluasi terhadap model LSTM tersebut dengan metriks ROC-AUC, *Hamming Loss*, dan *F1-Score*. Nilai pada metriks tersebut berada diantara 0 hingga 1. Jika nilai ROC-AUC dan *F1-Score* mendekati 1 maka model akan semakin bagus. Sebaliknya jika *Hamming Loss* mendekati 0 maka model juga dikatakan semakin bagus.

*3.8 Permodelan Topik*

Setelah dilakukan evaluasi, model yang memiliki performa paling baik dipilih untuk selanjutnya digunakan untuk memberi label pada komentar yang belum diberi label pada tahap 3.2.2. Setelah dilakukan pelabelan secara otomatis, hasil pelabelan ini akan digabung dengan data yang dilabeli secara manual. Setelah penggabungan akan kembali dipisahkan berdasarkan sentimennya, yang kemudian masing-masing sentimen akan dilakukan permodelan topik. Permodelan topik ini bertujuan untuk mengetahui topik apa yang dibicarakan pada masing-masing sentimen.

1. PEMBAHASAN

Data yang berhasil dikumpulkan dari youtube adalah sebanyak 1759 komentar, data tersebut dikumpulkan dari sembilan video youtube yang berkaitan dengan pembelajaran tatap muka. Sebanyak 1759 komentar tersebut, 563 diantaranya dilakukan pelabelan secara manual. Komentar yang sudah diberi label ini lah yang akan digunakan untuk permodelan.

|  |  |  |
| --- | --- | --- |
| Judul Video | Tanggal Rilis | Jumlah Komentar yang Diambil |
| Kuliah Tatap Muka Dimulai Juli 2021 (Tribunnews) | 10 Maret 2021 | 39 |
| Sekolah Tatap Muka Dimulai Juli 2021, Ini Syaratnya (MetroTVnews) | 19 Maret 2021 | 180 |
| Curhat Anak SMP Ingin Sekolah Tatap Muka: Kangen Upacara - Coba-Coba Tatap Muka (Part 1)| (Mata Najwa) | 2 September 2021 | 120 |
| Sekolah di Jateng Siap Pembelajaran Tatap Muka (Ganjar Pranowo) | 17 Maret 2021 | 222 |
| Jokowi Targetkan Belajar Tatap Muka Juli 2021 (CNN Indonesia) | 24 Februari 2021 | 21 |
| Disebut Melarang Sekolah Tatap Muka, Gubernur Lampung Tantang Nadiem Makarim (Kompas TV) | 27 Agustus 2021 | 308 |
| Sekolah Tatap Muka Diizinkan Mulai Januari 2021 (CNN Indonesia) | 20 November 2020 | 272 |
| Januari 2021, Pemerintah Izinkan Sekolah Tatap Muka (CNN Indonesia) | 21 November 2020 | 224 |
| Sekolah Tatap Muka jadi Klaster Penyebaran COVID-19, 90 Siswa di Purbalingga Terinfeksi Corona (Kumparan) | 22 September 2021 | 378 |

Tabel. 1 Video Sumber Komentar

Setelah dilakukan pengumpulan data dan pelabelan secara manual, dilakukan upsampling. Upsampling diperlukan karena jumlah sentimen dari komentar yang sudah diberi label tidak seimbang.

|  |  |  |
| --- | --- | --- |
| Sentimen | Jumlah Sebelum Upsampling | Jumlah Setelah Upsampling |
| Positif | 210 | 245 |
| Netral | 245 | 245 |
| Negatif | 95 | 245 |

Tabel. 2 Upsampling

Setelah dilakukan penyeimbangan data dengan upsampling dilanjutkan dengan semua proses pra-pemrosesan data dan word-embedding yang sudah dijelaskan pada bab sebelumnya. Setelah selesai dilakukanlah permodelan dengan algoritma LSTM, permodelan dilakukan dua kali dengan parameter yang berbeda. Setelah permodelan selesai, meskipun kedua model menggunakan algoritma dan pengoptmasi yang sama, yakni LSTM dan ADAM, rupanya tetap ada perbedaan antara dua model yang dihasilkan.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| NO | *Learning Rate* | *Batch Size* | Akurasi | Skor ROC-AUC | Hamming Score | Hamming Loss |
| 1 | 0.001 | 32 | 80% | 81% | 74% | 14% |
| 2 | 0.01 | 64 | 84% | 84% | 78% | 14% |

Tabel. 3 Evaluasi Performa Model

Perbedaan yang sedikit signifikan juga terlihat dari confusional matrix kedua model tersebut. Dimana model dengan *batch size* dan *learning rate* lebih kecil lebih banyak false neutral, sedangkan model yang lainnya lebih banyak false positive. Bagaimanapun juga kedua model sudah cukup akurat dalam meramalkan yang bersentimen negatif.

A picture containing background pattern

Description automatically generated

Grafik. 1 *Confusional Matrix Model 1*

A picture containing application

Description automatically generated

Grafik. 2 *Confusional Matrix Model 2*

Berdasarkan hasil evaluasi model, maka digunakan model dua untuk melabeli secara otomatis komentar yang sudah dikumpulkan namun belum diberi label. Komentar ini nantinya akan digabungkan dengan komentar yang sudah diberi label secara manual untuk dilakukan permodelan topik.

|  |  |  |
| --- | --- | --- |
|  | Label | Jumlah |
| Label Manual | Positif | 210 |
| Netral | 245 |
| Negatif | 95 |
| Label Otomatis | Positif | 370 |
| Netral | 557 |
| Negatif | 271 |

Tabel. 4 Jumlah Hasil Klasifikasi

Setelah selesai dilakukan pelabelan secara otomatis dengan model yang sudah dibuat. Kita telah sampai pada tahap terakhir yakni permodelan topik. Permodelan ini akan dilakukan terpisah untuk setiap sentimen.

Grafik. 3 Hasil Permodelan Topik Sentimen Positif

Grafik. 4 Hasil Permodelan Topik Sentimen Positif

Grafik. 5 Hasil Permodelan Topik Sentimen Positif

Dari hasil permodelan topik diatas dapat dilihat bahwa setiap topik memiliki kecendungan terhadap topik yang berbeda-beda. Sentimen positif mayoritas berisi doa dan harapan orang-orang yang menyambut pembelajaran tatap muka. Netral cenderung berisi komentar membangun. Sedangkan negatif berisi komentar orang-orang yang pesimistis terhadap pembelajaran tatap muka.

1. KESIMPULAN

Dari penelitian ini dapat ditarik beberapa kesimpulan, pertama parameter terbaik untuk digunakan dalam melakukan analisis sentiment menggunakan algoritma LSTM adalah pengoptimasi ADAM, dengan *batch size* 64, dan learning rate sebesar 0.01. Ditemukan juga bahwa jumlah epoch tidak begitu berpengaruh, karena akurasi sudah konvergen pada kisarah epoch ke 40.

Dari sisi kebahasaan, dapat kita lihat bahwa sentiment pengguna Youtube sudah mengarah pada netral dan positif, bukan pada sentimen negatif. Hal ini dapat mengindikasikan bahwa masyarakat sudah antusias untuk menghadapi pembelajaran tatap muka. Adapun dari hasil permodelan topik dapat kita lihat bahwa pada sentimen positif mayoritas berisi doa dan harapan orang-orang yang menyambut pembelajaran tatap muka. Netral cenderung berisi komentar membangun. Sedangkan negatif berisi komentar orang-orang yang pesimistis terhadap pembelajaran tatap muka. Komentar-komentar pesimistis ini kebanyakan berasal dari masyarakat yang di daerahnya sudah pernah ada wacana pembelajaran tatap muka dan tidak jadi. Dari hasil penelitian ini diharapkan dapat menjadi referensi pemerintah dalam menyusun strategi peningkatan animo masyarakat menjelang pembelajaran tatap muka.

Referensi

[1] M. Malta, A. W. Rimoin, and S. A. Strathdee, “The coronavirus 2019-nCoV epidemic: Is hindsight 20/20?,” *EClinicalMedicine*, vol. 20. Lancet Publishing Group, Mar. 01, 2020. doi: 10.1016/j.eclinm.2020.100289.

[2] “PENYELENGGARAAN PEMBELAJARAN TATAP MUKA TAHUN AKADEMIK 2021/2022 – Direktorat Jenderal Pendidikan Tinggi Kementerian Pendidikan dan Kebudayaan Republik Indonesia.” https://dikti.kemdikbud.go.id/pengumuman/penyelenggaraan-pembelajaran-tatap-muka-tahun-akademik-2021-2022/ (accessed Sep. 23, 2021).

[3] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, “A survey of sentiment analysis in social media,” *Knowledge and Information Systems*, vol. 60, no. 2, pp. 617–663, Aug. 2019, doi: 10.1007/s10115-018-1236-4.

[4] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 1, Apr. 2021, doi: 10.28932/jutisi.v7i1.3216.

[5] J. Samuel *et al.*, “Feeling Positive about Reopening? New Normal Scenarios from COVID-19 US Reopen Sentiment Analytics,” *IEEE Access*, vol. 8, pp. 142173–142190, 2020, doi: 10.1109/ACCESS.2020.3013933.

[6] P. Ghasiya and K. Okamura, “Investigating COVID-19 News across Four Nations: A Topic Modeling and Sentiment Analysis Approach,” *IEEE Access*, vol. 9, pp. 36645–36656, 2021, doi: 10.1109/ACCESS.2021.3062875.

[7] A. Oktavianto and S. F. Persada, “PersepsiPublikTentangPembelajaran Daring di Indonesia: Studi Menggunakan ELK Stackdan Python untuk Analisis Sentimen di Twitter,” *JURNAL TEKNIK ITS*, vol. 9, no. 2, 2020.

[8] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, “Sentiment analysis of comment texts based on BiLSTM,” *IEEE Access*, vol. 7, no. c, pp. 51522–51532, 2019, doi: 10.1109/ACCESS.2019.2909919.

[9] D. A. Nurdeni, I. Budi, and A. B. Santoso, “Sentiment Analysis on Covid19 Vaccines in Indonesia: From the Perspective of Sinovac and Pfizer,” *3rd 2021 East Indonesia Conference on Computer and Information Technology, EIConCIT 2021*, no. April, pp. 122–127, 2021, doi: 10.1109/EIConCIT50028.2021.9431852.

[10] H. Jelodar *et al.*, “Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey,” *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15169–15211, 2019, doi: 10.1007/s11042-018-6894-4.