

## Analisis Sentimen Opini Terhadap Vaksin Covid-19 pada Media Sosial Twitter Menggunakan *Support Vector Machine* dan *Naive Bayes*

Frizka Fitriana<sup>1\*</sup>, Ema Utami<sup>2</sup>, Hanif Al Fatta<sup>3</sup>  
<sup>1,2,3</sup> Magister Teknik Informatika, Universitas Amikom Yogyakarta  
\*email: [Frizka.1304@students.amikom.ac.id](mailto:Frizka.1304@students.amikom.ac.id)

DOI: <https://doi.org/10.31603/komtika.v5i1.5185>

Received: 21-06-2021, Revised: 01-07-2021, Accepted: 10-07-2021

### ABSTRACT

The corona virus outbreak, commonly referred to as COVID-19, has been officially designated a global pandemic by the World Health Organization (WHO). To minimize the impact caused by the virus, one of the right steps is to develop a vaccine, however, with the vaccination for the Indonesian people, it is controversial so that it invites many people to give an opinion assessment, but the limited space makes it difficult for the public to express their opinion, because Therefore, people choose social media as a place to channel public opinion. Support vector machine algorithm has better performance in terms of accuracy, precision and recall with values of 90.47%, 90.23%, 90.78% with performance values on the Bayes algorithm, namely 88.64%, 87.32%, 88, 13%, with a difference of 1.83% accuracy, 2.91% precision and 2.65% recall, while for time the Naive Bayes algorithm has a better performance level with a value of 8.1 seconds and the Support vector machine algorithm gets a time speed of 11 seconds with a difference of 2, 9 seconds. With the results of sentiment analysis neutral 8.76%, negative 42.92% and positive 48.32% for Bayes and neutral 10.56%, negative 41.28% and positive 48.16% for SVM..

**Keywords:** Covid-19, Vaccines, Twitter, Support Vector Machine, Naïve Bayes.

### ABSTRAK

Wabah virus korona yang biasa disebut dengan COVID-19 ditetapkan secara resmi sebagai pandemi global oleh World Health Organization (WHO). Untuk meminimalisir dampak yang disebabkan oleh virus salah satu langkah yang tepat adalah dengan mengembangkan vaksin. Akan tetapi dengan adanya vaksinasi untuk masyarakat Indonesia tersebut menimbulkan kontroversial sehingga mengundang banyak kalangan untuk memberikan penilaian opininya. Keterbatasan tempat membuat masyarakat sulit dalam menyampaikan opininya. Oleh karena itu masyarakat memilih sosial media sebagai tempat untuk menyalurkan opini masyarakat. Dalam penelitian ini dilakukan perbandingan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) untuk melakukan analisa sentimen tentang vaksin Covid – 19. Berdasarkan hasil penelitian diperoleh bahwa algoritma SVM memiliki performa lebih baik pada bagian akurasi, presisi dan *recall* dengan nilai 90,47%, 90,23%, 90,78% dan performa pada algoritma *Naive Bayes* adalah 88,64%, 87,32%, 88,13%, dengan selisih akurasi 1,83%, presisi 2,91% dan *recall* 2,65%. Sedangkan untuk waktu algoritma *Naive Bayes* memiliki tingkat performa lebih baik dengan nilai 8,1 detik dibandingkan SVM dengan kecepatan waktu 11 detik. Hasil sentimen analisis netral diperoleh 8,76%, negatif 42,92% dan positif 48,32% untuk *Naive Bayes* dan netral 10,56%, negatif 41,28% dan positif 48,16% untuk SVM.

**Kata-kata kunci:** Covid-19, Vaksin, Twitter, Support Vector Machine, Naïve Bayes.

### PENDAHULUAN

Wabah virus korona yang biasa disebut dengan COVID -19 ditetapkan secara resmi sebagai pandemi global oleh *World Health Organization* (WHO) pada tanggal 11 Maret 2020. Pusat penyebaran virus tersebut dimulai pada akhir tahun 2019 lalu, dan sampai saat ini sudah

menyebarkan ke berbagai negara-negara dengan tingkat penularan mencapai nilai 40 juta lebih kasus dan jumlah kematian mencapai 1 juta lebih hingga Oktober 2020. Efek lanjutan dari COVID -19 ini berpotensi membawa tantangan besar bagi sistem kesehatan dunia dan memiliki konsekuensi yang luas pada ekonomi global jika penyebaran virus tidak dikendalikan secara efektif [1].

Melihat pesatnya penyebaran COVID -19 dan dampak yang akan ditimbulkan apabila tidak secepatnya diselesaikan, salah satu solusi untuk mengurangi tingkat penyebaran virus yaitu dengan pembuatan vaksin. Vaksin akan melindungi masyarakat terhadap penyebaran virus dan juga dapat menurunkan tingkat penyebaran virus tersebut. Dalam menyikapi hal tersebut, Pemerintah Indonesia juga turut aktif dalam rencana kegiatan vaksinasi yang akan diberikan kepada masyarakatnya. Untuk mempercepat penanganan COVID -19, Pemerintah Republik Indonesia telah mengeluarkan vaksinasi untuk masyarakat Indonesia pada akhir tahun 2020 [2].

Dengan adanya vaksinasi untuk masyarakat Indonesia tersebut dinilai kontroversial sehingga mengundang banyak kalangan untuk memberikan pendapat dan opininya. Masyarakat bebas memberikan respon dan opininya di berbagai media. Sosial media yang paling banyak digunakan masyarakat untuk menyampaikan pendapat dan opininya. Berdasarkan data dari Global Digital Statistic “*Digital, Social & Mobile in 2019*” di *We Are Social* (2019), 2019 merupakan tahun dengan jumlah pengguna media sosial dengan jumlah 150 juta pengguna. Twitter menjadi media sosial yang memiliki pengguna aktif paling banyak di Indonesia dengan jumlah mencapai 52% dari seluruh pengguna sosial media di Indonesia [3].

Sosial media merupakan tempat yang dikenal untuk mencurahkan segala hal oleh semua masyarakat. Salah satu sosial media yang dikenal oleh masyarakat adalah Twitter yang digunakan masyarakat untuk membagikan hal yang dirasakan oleh penggunanya. Melalui postingan pada Twitter masyarakat dapat membagikan dan mendapatkan informasi mengenai hal apapun. Dengan memanfaatkan data dari media sosial Twitter, dapat dilakukan analisis pendapat dan opini masyarakat terhadap vaksin melalui analisis sentimen dengan cara melakukan klasifikasi pendapat dan opini kedalam 2 kelas yaitu negatif dan positif.

Analisis sentimen merupakan suatu metode untuk mengklasifikasikan tingkat emosional seseorang baik netral, positif ataupun negatif [4]. Salah satu penerapan analisis sentimen dilakukan pada Twitter, dimana sistem akan secara otomatis mengambil data lewat postingan atau tweet dari pengguna dan sistem akan melakukan klasifikasi untuk menilai apakah postingan tersebut mengandung kalimat netral, positif atau negatif [5]. Berdasarkan klasifikasi tersebut, pengguna dapat mengetahui dan menilai tentang suatu topik tertentu serta dapat membuat keputusan [6]. Metode Naïve Bayes dan *Support Vector Machine* (SVM) digunakan untuk melakukan analisis sentimen melalui komentar pada postingan sosial media instagram terkait pada kasus COVID-19 pada postingan World Health Organization (WHO) mengenai Covid-19. Hasil penelitian diperoleh bahwa dengan menggunakan algoritma Naïve Bayes mendapatkan akurasi 78,02% dan algoritma SVM mendapatkan nilai akurasi 80,23%. Melalui optimasi *Particle Swarm Optimization* (PSO), akurasi metode SVM meningkat menjadi 81,16% dan akurasi Naïve Bayes menjadi 79,07% [7].

Untuk Mengetahui opini positif atau negatif suatu peristiwa atau isu dapat dilakukan analisis sentimen data Twitter dengan kata kunci ‘Vaccine Covid-19’ menggunakan metode

Naïve Bayes. Hasil pengukuran sentimen dengan lebih dari 3,4 ribu *tweet* negatif, lebih dari 2,4 ribu *tweet* positif, dan sisa 301 *tweet* dianggap netral selama periode seminggu pada minggu kedua dan ketiga Januari 2021. Analisis pada periode tersebut menunjukkan 39% sentimen positif, 56% sentimen negatif, dan 1% sentimen menguntungkan. Pendapat negatif pada periode tersebut muncul karena masyarakat tidak percaya bahwa vaksin itu aman [8]. Penelitian analisis sentimen masyarakat mengenai pendapat tentang COVID-19 melalui media sosial Twitter juga diteliti menggunakan KNN serta Naïve Bayes. Dengan menggunakan KNN diperoleh akurasi sebesar 85% sedangkan untuk nilai akurasi Naïve Bayes mendapatkan nilai 82%. Melalui hasil tersebut algoritma KNN terlihat lebih baik dalam melakukan analisis sentimen untuk mengklasifikasi komentar pada media sosial Twitter jika dibandingkan dengan algoritma *Naïve Bayes* [9]. Selain itu algoritma *Naïve Bayes* juga digunakan untuk analisis sentimen menggunakan *twitter* API yang terhubung dengan RStudio. Dan hasil yang didapat dari penelitian analisis sentimen ini yaitu pada data *tweet* Maret 2020, sentimen positif 51% dan sentimen negatif 49% dengan tingkat akurasi 0,7586, spesifisitas 0,6667, prevalensi 0,5862. Sedangkan data *tweet* Juni 2020 menunjukkan, 59% sentimen negatif dan 41% sentimen positif. Dengan tingkat akurasi 0,6486, spesifisitas 0,6111, prevalensi 0,5135 [10].

Algoritma Naïve Bayes dikenal mampu melakukan klasifikasi dengan baik berdasarkan penalaran probabilitas. Algoritma SVM juga merupakan salah algoritma *machine learning* yang dikenal cukup baik dalam melakukan klasifikasi berdasarkan pembobotan yang diproses dalam algoritma tersebut. Dalam penelitian ini dilakukan perbandingan model algoritma SVM serta *Naïve Bayes* untuk melakukan analisis sentimen pada postingan Twitter berdasarkan kata kunci vaksin Covid -19.

## METODE

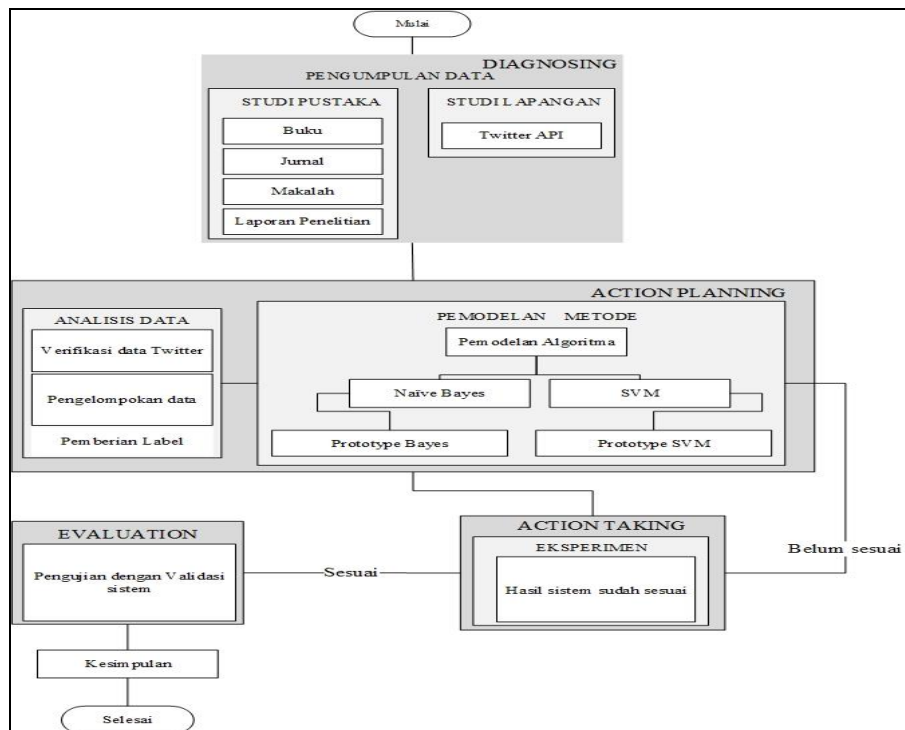
Tahapan penelitian yang dilakukan dalam penelitian ini disajikan dalam Gambar 1 yang terdiri dari beberapa tahap meliputi pengumpulan data, analisis data, pemodelan dan evaluasi. Data yang digunakan dalam penelitian ini diperoleh dari data postingan Twitter berdasarkan kata kunci vaksin Covid -19. Interval waktu data Twitter diambil dari data tahun 2021 berkisar 1000 data melalui kata vaksin Covid-19 yang diuji menggunakan algoritma *Naïve Bayes* dan SVM untuk mendapatkan nilai performa tertinggi. Dalam penelitian ini dilakukan perbandingan tingkat akurasi, presisi, recall, dan waktu training dengan melakukan analisis sentimen.

Dalam tahapan analisis data dilakukan verifikasi data Twitter, pengelompokan data dan pemberian label. Untuk menentukan tingkat performa pada algoritma *Naïve Bayes* dan SVM. Analisis dilakukan melalui proses verifikasi data Twitter, pengelompokan data dan pemberian label. Tahap *preprocessing* meliputi *case folding* dengan merubah teks dalam bentuk baku sebuah dokumen (*lowercase*), *tokenizing* yaitu tahap melakukan pemotongan string input berdasarkan tiap kata yang menyusunnya, *filtering* yang merupakan tahap mengambil kata-kata penting dari hasil *token* dengan menggunakan algoritma *stoplist* (membuang kata kurang penting) atau *wordlist* (menyimpan kata penting), dan *stemming* untuk melakukan sortir dengan tujuan memperoleh indeks yang berbeda dari berbagai dokumen.

Dalam tahap *action planning* dilakukan dengan membuat pemodelan data menggunakan algoritma algoritma Naïve Bayes dan SVM. Selanjutnya dilakukan tahap *action taking*, yaitu

dengan melakukan eksperimen pengujian tingkat performa melalui pengujian serta membandingkan performa kedua metode melalui parameter pengujian akurasi, presisi, *recall* dan waktu training untuk mendapatkan nilai performa terbaik antara kedua algoritma.

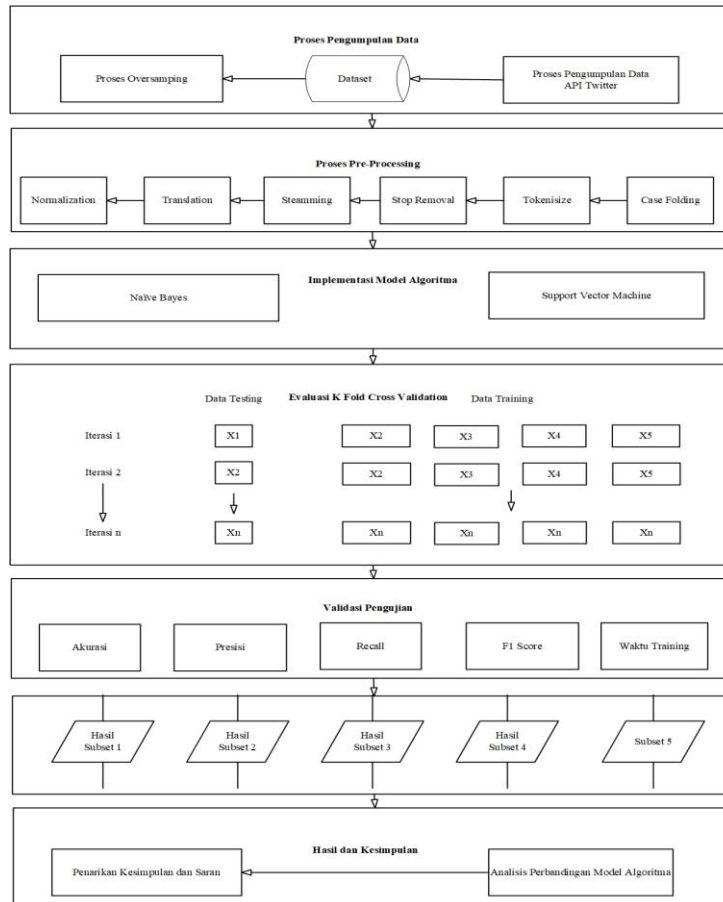
Tahap *evaluation* merupakan tahapan untuk melakukan evaluasi hasil perbandingan dua metode yang digunakan yaitu *Naïve Bayes* dan SVM. Dalam tahap ini dihitung hasil yang sesuai untuk mendapatkan prosentase performa pada masing-masing metode sehingga dapat diperoleh kesimpulan metode yang sesuai untuk melakukan analisis sentimen data Twitter melalui kata vaksin Covid-19.



Gambar 1. Metode Penelitian

Alur analisis sentimen vaksin Covid-19 menggunakan algoritma *Naïve Bayes* dan SVM disajikan dalam Gambar 2. Proses pengumpulan data didapatkan melalui API Twitter untuk mengambil data postingan melalui kata vaksin Covid-19. Setelah data diperoleh maka data tersebut menjadi *dataset* awal untuk dilakukan proses *over sampling* yang membentuk data ke dalam pola yang telah ditentukan dan format yang sesuai untuk proses *preprocessing*. Selanjutnya dilakukan *preprocessing* yaitu dengan melakukan sortir ke dalam proses *case folding*, *tokenize*, *stop removal*, *stemming*, *translation* dan *normalization*. Data yang telah dilakukan *preprocessing* selanjutnya siap untuk diimplementasikan ke dalam algoritma *Naïve Bayes* dan SVM menggunakan bahasa pemrograman PHP 7.4 dan JavaScript 1.8.5 serta *database* dengan MySQL. Data yang menjadi *input* dibagi menjadi data *training* dan data *testing*. Dalam tahap evaluasi *K-Fold Cross Validation* dilakukan validasi untuk pengujian algoritma, dimana *dataset* diatur melalui skenario 5 *subset* dan di poses sebanyak 5 kali perulangan untuk mendapatkan validasi pengujian. Pada bagian akhir dilakukan validasi pengujian terhadap model yang dibentuk untuk mengukur performa algoritma melalui

parameter akurasi, presisi, *recall* dan waktu training sehingga dapat dibuat kesimpulan performa algoritma yang sesuai untuk analisis sentimen vaksin Covid – 19



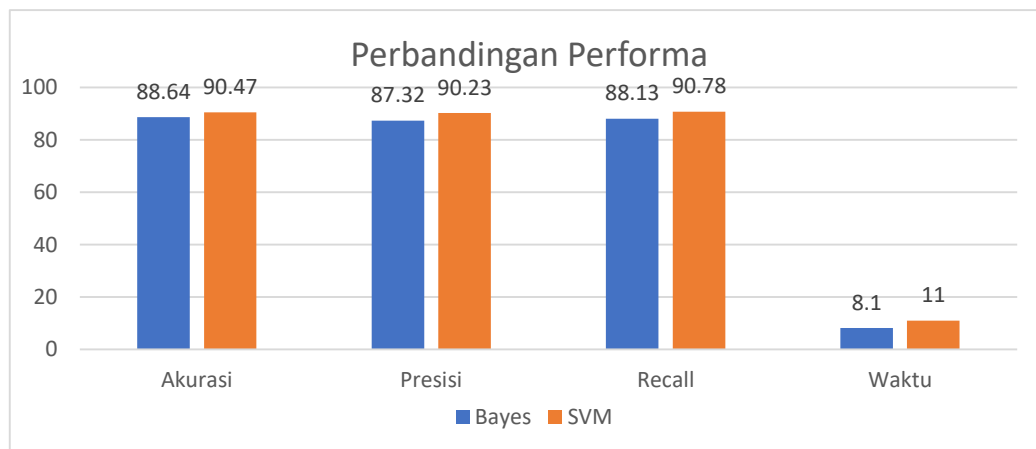
Gambar 2. Alur Analisis Sentimen Algoritma

## HASIL DAN PEMBAHASAN

Berdasarkan hasil pengujian, algoritma *Naïve Bayes* dan SVM memiliki hasil tingkat performa yang berbeda seperti disajikan dalam Tabel 1. Hasil yang diperoleh dari Tabel 1 selanjutnya dilakukan perbandingan performa dari kedua algoritma yang digunakan seperti disajikan pada Gambar 3. Dalam hal ini, algoritma SVM terlihat lebih unggul dalam hal akurasi, presisi dan *recall* dan algoritma *Naïve Bayes* lebih unggul dalam hal eksekusi waktu *training*. Performa SVM untuk akurasi, presisi dan *recall* lebih tinggi dibandingkan *Naïve Bayes* karena proses pembobotan pada algoritma SVM lebih detail dan melakukan perulangan secara terus menerus untuk menentukan nilai bobot yang tepat dibandingkan algoritma *Naïve Bayes* yang hanya melakukan satu kali perulangan untuk menentukan bobot yang tepat. Hal tersebut yang menyebabkan *Naïve Bayes* lebih unggul dari segi waktu *training* dibandingkan SVM.

Tabel 1. Perbandingan Performa

Algoritma	Performa			
	Akurasi (%)	Presisi (%)	Recall (%)	Waktu (detik)
Bayes	88,64	87,32	88,13	8,1
SVM	90,47	90,23	90,78	11



Gambar 3. Perbandingan Hasil Performa

Melalui hasil tersebut, sistem dapat melakukan klasifikasi analisis sentimen opini vaksin Covid – 19 yang meliputi opini netral, negatif dan positif. Nilai masing – masing opini dengan menggunakan algoritma *Naïve Bayes* dan SVM disajikan dalam Tabel 2. Berdasarkan hasil yang diperoleh maka dapat dinyatakan bahwa opini positif tentang vaksin Covid – 19 di Twitter lebih banyak dibandingkan opini negatif maupun yang netral.

Tabel 2. Perbandingan Sentimen Analisis

Algoritma	Performa		
	Netral (%)	Negatif (%)	Positif (%)
Bayes	8,76	42,92	48,32
SVM	10,56	41,28	48,16

## KESIMPULAN

Dalam penelitian ini melakukan perbandingan algoritma *Naïve Bayes* dan SVM untuk melakukan analisis sentimen opini vaksin Covid – 19 pada media Twitter. Berdasarkan hasil yang telah diperoleh dapat dinyatakan bahwa SVM mempunyai performa yang lebih bagus dibandingkan *Naïve Bayes*. Namun dalam segi waktu, algoritma *Naïve Bayes* memerlukan waktu yang lebih sedikit dibandingkan SVM. Opini vaksin Covid – 19 di media Twitter memberikan opini positif yang artinya bahwa penerimaan vaksin Covid – 19 lebih besar dibanding penolakannya. Untuk meningkatkan performa dapat ditingkatkan melalui optimasi pada kedua algoritma SVM dan *Naïve Bayes*, seperti menggunakan *particle swarm optimization*.

## DAFTAR PUSTAKA

- [1] World Health Organization (WHO). 2021. Coronavirus disease (COVID-19) pandemic <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>, diakses 25 Januari 2021
- [2] Nareza, M. (2021). Jangan Termakan Isu, Ini Fakta Penting Vaksin COVID-19. <https://www.alodokter.com/>, diakses 25 Januari 2021

- [3] Our World In Data. 2021. Coronavirus (COVID-19) Vaccinations. <https://ourworldindata.org/covid-vaccinations>, diakses 25 Januari 2021
- [4] Mubarok, M.S., Adiwijaya, Aldhi, M.D., 2016. Aspect-based Sentiment Analysis to Review Products Using Naïve Bayes. 2nd International Conference on Mathematics: Pure, Applied and Computation (ICOMPAC)
- [5] Tripathy, A., Agrawal, A. and Rath, S.K., 2016. Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, 57, pp.117-126
- [6] M. Adriani, J. Asian, B. Nazief, and H. E. Williams, Stemming Indonesian : A Confix-Stripping Approach, *ACM Transactions on Asian Language Information Processing (TALIP)*, vol. 6, no. 4. pp. 1–33, 2007.
- [7] Ratino, Noor, H., Sita A., Windu, G. 2020. “Sentimen Analisis Informasi Covid-19 menggunakan Support Vector Machine dan Naïve Bayes”. *Jurnal JUPITER* Vol. 12 No. 2 Bulan 10 Tahun 2020, Hal 1 - 11
- [8] Muhammad, A., Agus, A, Fauziah, H. 2020. “ Analisis sentimen vaksin COVID-19 di Indonesia menggunakan Algoritma Naïve Bayes”. *Konferensi Tahunan tentang Ilmu Komputer dan Teknologi Rekayasa (AC2SET)*. doi: 10.1088 / 1757-899X / 1088/1/012045.
- [9] Ardianne, L., F., Rima, D., R., Nia, A., F. 2021. “Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter”. *Jurnal Dinda*. Vol.1x No.1 (2021) Feb 2021.
- [10] Cahyo, P., Nisa, H. 2020. “Analisis Sentimen Covid-19 Sebagai Pandemi Media Sosial”. *Jurnal Internasional Sistem Informasi & Teknologi* . Vol. 4, Tidak. 1, ( 2020), hlm.509-517



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

---