

## ANALISIS TINGKAT KEBERMANFAATAN MYPERTAMINA MENGGUNAKAN K-MEANS CLUSTERING

Kharisma <sup>1)</sup>, Ulfi Saidata Aesy <sup>2)</sup>

<sup>1,2)</sup> Sistem Informasi Universitas Jenderal Achmad Yani Yogyakarta  
email : [kharisma.anoe@gmail.com](mailto:kharisma.anoe@gmail.com)<sup>1)</sup>, [ulfiaesy@gmail.com](mailto:ulfiaesy@gmail.com)<sup>2)</sup>

### Abstraksi

Pertamina, menerapkan digitalisasi Stasiun Pengisian Bahan Bakar Umum (SPBU) melalui aplikasi mypertamina yang bekerja sama dengan linkaja. Meskipun aplikasi tersebut menawarkan beberapa program yang dapat mempermudah pelanggan, ternyata masyarakat banyak mengomentari aplikasi tersebut sehingga menyebabkan topik mypertamina di twitter menjadi trending topik pada bulan Juli. Akan tetapi komentar tersebut tidak seluruhnya positif. Komentar yang terlalu banyak tentang aplikasi mypertamina ini, menyebabkan sulitnya berbagai kalangan untuk menyimpulkan tingkat kebermanfaatan dari aplikasi tersebut, termasuk Pertamina. Sehingga diperlukan penelitian untuk menganalisis tingkat kebermanfaatan aplikasi mypertamina dari banyaknya komentar masyarakat di twitter. Oleh karena itu, penelitian ini mengumpulkan data berupa tweet masyarakat dari twitter tentang aplikasi mypertamina. Data tersebut kemudian dibersihkan dan dihitung kedekatan aktanya menggunakan TF-IDF. Setelah itu data dikelompokkan dengan menggunakan K-Means. Dengan menggunakan coherence score, kluster terbaik ada di kluster 2. Kluster 0 berisi kata keluhan masyarakat terhadap aplikasi my Pertamina. Kluster 1 berisi kata umum terkait aplikasi Pertamina. Berdasarkan analisis hasil kluster yang diperoleh, maka aplikasi mypertamina mendapatkan banyak keluhan dari masyarakat yang cenderung berpendapat bahwa aplikasi mypertamina tidak bermanfaat untuk masyarakat.

### Kata Kunci :

*mypertamina, datamining, klustering, K-Means*

### Abstract

*Pertamina, implementing digitization of Public Fuel Filling Stations (SPBU) through the mypertamina application in collaboration with linkaja. Even though the application offers several programs that can make it easier for customers, it turns out that a lot of people comment on the application, causing the topic myertamina on Twitter to become a trending topic in July. However, these comments were not entirely positive. Too many comments about the mypertamina application make it difficult for various groups to conclude the level of usefulness of the application, including Pertamina. So research is needed to analyze the level of usefulness of the mypertamina application from the many public comments on Twitter. Therefore, this study collects data in the form of community tweets from Twitter about the mypertamina application. The data is then cleaned and the closeness of the deed is calculated using the TF-IDF. After that the data is grouped using K-Means. By using the coherence score, the best cluster is in cluster 2. Cluster 0 contains the words people complain about the my Pertamina application. Cluster 1 contains common words related to Pertamina applications. Based on the analysis of the cluster results obtained, the mypertamina application gets a lot of complaints from the public who tend to think that the mypertamina application is not useful for the community.*

### Keywords :

*mypertamina, datamining, clustering, K-Means*

### Pendahuluan

Pertamina adalah perusahaan milik negara yang bergerak di bidang energi meliputi minyak, gas, energi baru dan terbarukan serta penyumbang devisa terbesar di Indonesia [1]. Saat ini Pertamina bekerjasama dengan linkaja [2] telah melakukan digitalisasi Stasiun Pengisian Bahan Bakar Umum (SPBU) melalui aplikasi MyPertamina. Aplikasi tersebut bisa di download melalui *Google Playstore* dan *Apps Store*. MyPertamina menawarkan berbagai program yang diharapkan mempermudah pelanggan [3] dan pemberian poin untuk ditukarkan reward dari PT Pertamina (Persero), diresmikan pada bulan Agustus 2017.

Berbagai fungsi dari aplikasi MyPertamina diantaranya : 1. Saat ini *cashless* MyPertamina menggunakan dompet digital LinkAja. Pengguna harus menghubungkan akun MyPertamina dengan LinkAja untuk melakukan pembayaran, saldo Linkaja akan otomatis terpotong setelah pengguna melakukan pembelian BBM. 2. Fitur *Charging Station* pada aplikasi MyPertamina berfungsi untuk mengetahui informasi Stasiun Pengisian Kendaraan Listrik Umum. 3. Info riwayat pembelian dan point yang bisa ditukarkan dengan *merchandise*. 4. Fitur pemesanan BBM untuk diantar ke rumah, dll [4][5]. Per 1 Juli 2022 PT Pertamina mewajibkan pengguna BBM Peralite dan Solar Subsidi untuk

mendaftarkan kendaraannya pada website MyPertamina. Hal tersebut dilakukan agar subsidi BBM tepat sasaran. Kebijakan tersebut diberlakukan karena anggaran subsidi BBM terus mengalami pembengkakan. Menurut Kepala Badan Kebijakan Kementrian Keuangan Febrio Nathan Kacaribu bahwa yang menikmati hampir 80% dari total seluruh penggunaan rumah tangga perbulan adalah 60% masyarakat yang tergolong terkaya, sedangkan 40% nya adalah masyarakat yang terba wah[6]. Pendaftaran ini bisa dilakukan secara online melalui laman MyPertamina atau *offline* dengan mendatangi tempat pendaftaran di area-area yang telah ditentukan. Ada beberapa berkas yang perlu disiapkan, yaitu: foto KTP, foto diri, foto STNK depan dan belakang, foto KIR, foto kendaraan semua sisi, dan foto nomorpolisi kendaraan. Rupanya kebijakan pendaftaran kendaraan untuk pembelian pertalite dan solar subsidi ini tidak serta merta diterima baik oleh publik meskipun berbagai fitur yang disediakan untuk menarik pelanggan seperti telah disebutkan di atas. Aplikasi MyPertamina terpantau hanya mendapatkan nilai 1,3. Rata-rata pengguna juga mengeluhkan aplikasi itu lambat dan menyusahkan. Bukan hanya di *Google Play*, warganet juga menggunjingkan MyPertamina di Twitter. Sehingga perlu adanya analisis dari komentar masyarakat di twitter terkait tingkat kebermanfaatan dari aplikasi MyPertamina.

### Tinjauan Pustaka

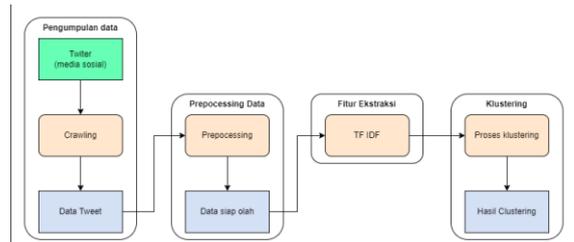
Analisis dilakukan menggunakan teknik pengelompokan data mining yaitu k-means klustering. Teknik ini merupakan teknik paling sederhana untuk melakukan pengelompokan [7]. Metode K-means ini telah banyak digunakan oleh beberapa penelitian untuk melakukan pengelompokan. Salah satunya untuk mengelompokkan data berdasarkan tingkat pemasokan barang dan penjualan hasil produksi dari perusahaan untuk mengetahui keinginan dari konsumennya di pasar [8]. K-Means juga pernah digunakan untuk melakukan pengelompokan hasil penjualan suatu perusahaan yang laku dan tidak laku dipasaran. Dari hasil penelitian tersebut, diperoleh 4 produk yang laku dan 4 produk yang tidak laku [9]. Sedangkan pada penelitian ini akan melakukan analisis dari hasil pengelompokan pendapat masyarakat terhadap aplikasi MyPertamina.

Data yang akan digunakan dalam analisis ini adalah data dari twitter dengan menggunakan teknik *crawling*. *Crawling* data pada twitter yang dimaksud adalah proses mengumpulkan data dari server twitter dengan bantuan Application Programming Interface (API) baik user maupun data tweet [10]. Data yang telah diperoleh tersebut akan dilakukan pembersihan data. Kemudian data setelah *preprocessing* akan dihitung Term-Frequency Inverse Document Frequency (TF-IDF). TF-IDF adalah pembobotan terhadap kata melalui beberapa teknik *preprocessing* untuk mencari kata yang penting dari dokumen [11].

Setelah itu dilakukan proses pengelompokan menggunakan *K-Means*. Hasil dari pengelompokan akan dianalisis. Dengan adanya analisis ini, diharapkan masyarakat dapat mengetahui hasil pengelompokan pendapat masyarakat terkait kebermanfaatan aplikasi mypertamina.

### Metode Penelitian

Penelitian ini menggunakan data dari hasil *crawling* dari salah satu media sosial yaitu twitter dengan menggunakan hashtag tertentu. Tahapan penelitian dapat dilihat pada gambar 1.



Gambar 1. Tahapan penelitian

#### a. Pengumpulan data

Pengumpulan data yang dimaksud adalah proses mengambil data. Pada penelitian ini data diambil dari Twitter yang merupakan salah satu media sosial yang sering digunakan oleh masyarakat untuk menyampaikan ide atau pendapat terhadap suatu permasalahan yang sedang diperbincangkan. Pengambilan data pada sosial media twitter ini menggunakan teknik *crawling*.

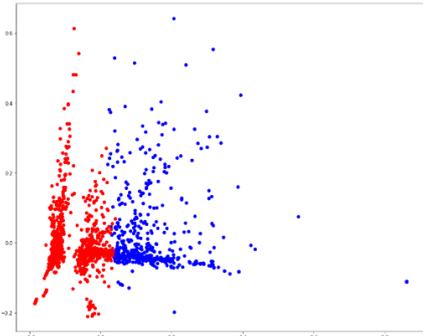
#### b. Preprocessing data

*Preprocessing* data yang dimaksud adalah membersihkan data mulai dari menghapus data yang tidak lengkap, terdapat data error, terdapat ketidakkonsistenan data [12], duplikasi data, serta *typografi* [13]. Data tweet yang diperoleh pada penelitian ini akan dibersihkan dan diseragamkan untuk mendapatkan hasil yang akurat untuk klustering. Tahap ini terdiri dari *tokenizing*, *stopword*, dan *stemming* [14]. Dimana *tokenizing* adalah tahap untuk memecah *text* dari kalimat menjadi token atau kumpulan kata dengan menghilangkan tanda baca dan membuat seluruh huruf menjadi huruf kecil. *Stopword* ini digunakan untuk mengurangi jumlah kata dengan melakukan penghapusan dan menyaring kata. Pada *stemming*, kata yang berlebihan diubah menjadi kata asli tanpa imbuhan, sisipan, awalan, akhiran, kombinasi.

#### c. Fitur ekstraksi

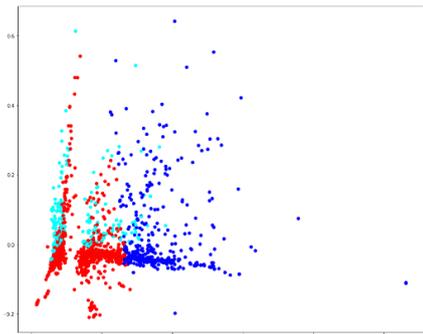
Fitur ekstraksi merupakan proses untuk melakukan perubahan format teks yang formatnya tidak terstruktur menjadi teks yang terstruktur [15]. Tahap ini adalah mencari kedekatan token dalam *corpus* [16]. Pada penelitian ini TF\_IDF digunakan untuk memilih fitur yang merupakan hasil riskasan





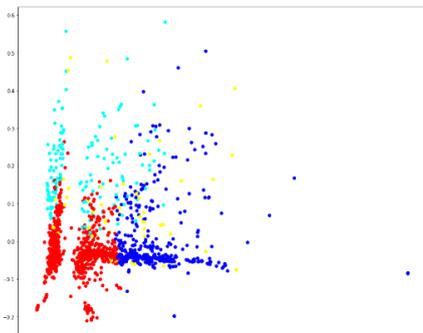
Gambar 5. Hasil kluster 2

Gambar 6 menunjukkan kluster dibagi menjadi 3, yaitu kluster 0, kluster 1, dan kluster 2.



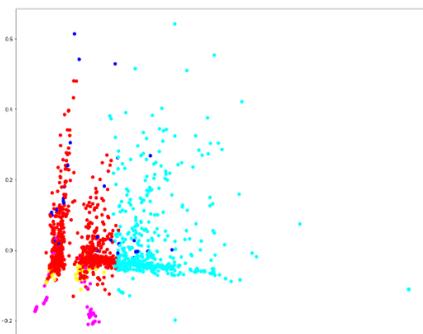
Gambar 6. Hasil kluster 3

Pada gambar 7, kluster dibagi menjadi 4, yaitu kluster 0, kluster 1, kluster 2, dan kluster 3.



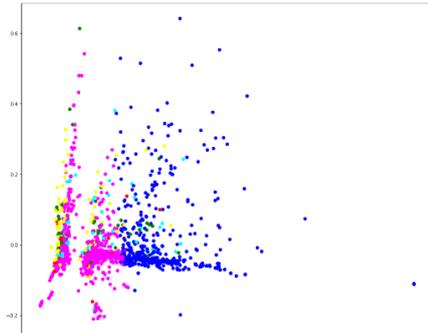
Gambar 7. Hasil kluster 4

Gambar 8 menunjukkan pembagian kluster menjadi 5 kelompok kluster.



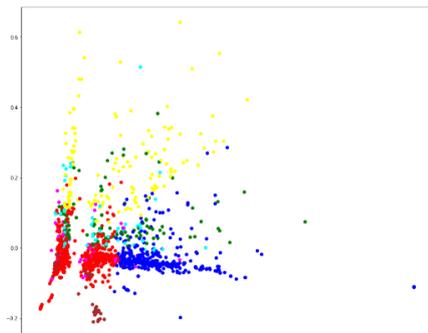
Gambar 8. Hasil kluster 5

Gambar 9 menunjukkan kluster dibagi sebanyak 6 kluster.



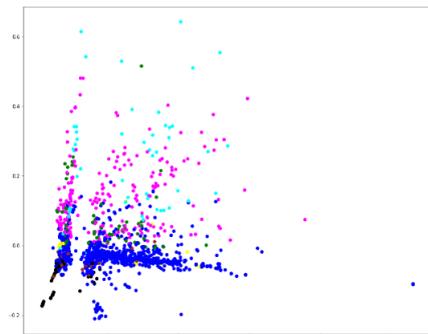
Gambar 9. Hasil kluster 6

Pada gambar 10, kluster dibagi menjadi 7 kluster.



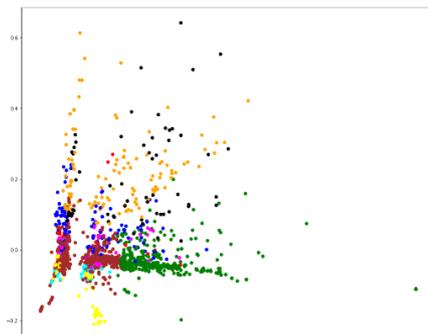
Gambar 10. Hasil kluster 7

Gambar 11 menunjukkan pembagian kluster menjadi 8 kluster.



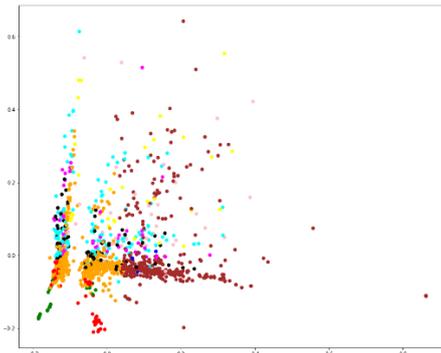
Gambar 11. Hasil kluster 8

Pada gambar 9, kluster dibagi menjadi 9 kluster.



Gambar 12. Hasil kluster 9

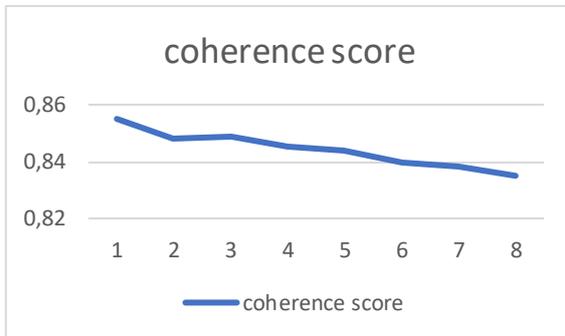
Gambar 13 menunjukkan kluster dibagi menjadi 9 kluster.



Gambar 13. Hasil kluster 10

Dari kluster di atas, dapat disimpulkan bahwa semakin banyak pembagian kluster pada topik mypertamina, hasil data berkelompok juga semakin tidak bisa terukur. Sehingga diperlukan suatu cara untuk menentukan kluster terbaik dari penelitian ini yaitu menggunakan *coherence score* untuk mengetahui *elbow* dari setiap kluster. Semakin jelas *elbow* yang terbentuk, maka semakin baik kluster itu.

Berdasarkan pengujian menggunakan *coherence score* diperoleh bahwa *elbow* terlihat pada kluster 2. Gambar 14 menunjukkan hasil pengujian menggunakan *coherence score*.



Gambar 14. Pengujian Coherence Score

Cluster 0	Cluster 1
my Pertamina unfaedah	my Pertamina unfaedah
rakyat	aplikasi
aplikasi	beli
susah	rakyat
ribet	link
ruwet	kamipercayaact
ras	my Pertamina
mudah	brunei
pertamina	pertamina
sulit	bbm
negara	pake
up	jokowi
woy	harga
bijak	rp
my	atur
bocor	spbu
tagar	adi
pake	pakai
ledak	pagi
my Pertamina	
kamipercayaact	
repot	

Gambar 15. Sample hasil cluster

Dari gambar 15, dapat dilihat pada kluster 0 adalah kelompok menurut masyarakat aplikasi

my Pertamina tidak bermanfaat dan banyak kekurangan. Sehingga pada kelompok kata kluster 0 berisi keluhan masyarakat tentang aplikasi my Pertamina, misal susah, ribet, ruwet, sulit, repot dan lain-lain. Sedangkan pada kluster 1 lebih kepada pembahasan yang umum dan tidak membahas terkait keberuntungan aplikasi my Pertamina. Sehingga dapat disimpulkan bahwa masyarakat menganggap aplikasi my Pertamina tidak bermanfaat.

### Kesimpulan dan Saran

Komentar masyarakat di twitter terkait aplikasi my Pertamina sempat menjadi trending topik pada bulan Juli 2022. Data tersebut diambil dan dibersihkan untuk dilakukan analisis menggunakan *K-Means*. Kluster dilakukan dari kluster 2 hingga 9. Seluruh kluster diujikan dengan menggunakan *coherence score*. Hasil dari pengujian tersebut diperoleh *elbow* terbaik pada kluster 2. Dari kluster 2 tersebut menghasilkan kelompok kata yang berbeda. Kluster 0 adalah kelompok kata keluhan masyarakat terhadap aplikasi my Pertamina. Kluster 1 membahas aplikasi my Pertamina secara umum. Rencana penelitian ini akan melakukan filtering pada kluster 0, dimana kluster 0 berisi keluhan masyarakat tentang aplikasi My Pertamina. Filtering tersebut akan melakukan filter komentar yang berhubungan dengan kesulitannya menggunakan aplikasi My Pertamina. Dari hasil filter tersebut, akan diperoleh data dari masyarakat terkait proses bisnis Pertamina dalam menggunakan aplikasi my Pertamina. Data tersebut dapat dilakukan penelitian untuk membuat usulan proses bisnis dan UI/UX baru aplikasi my Pertamina.

### Daftar Pustaka

- [1] N. K. Hikmawati, "Analisis Kualitas Layanan My Pertamina Menggunakan Pendekatan e-GovQual pada Beberapa Kota Percobaan," *J. Manaj. Inform.*, vol. 12, no. 2, pp. 100–111, Oct. 2022, doi: 10.34010/JAMIKA.V12I2.7977.
- [2] A. A. Sinurat et al., "My Pertamina Application To Increase Consumer Engagement," *Int. J. Bus. Rev. (The Jobs Rev.)*, vol. 5, no. 1, pp. 65–74, Aug. 2022, doi: 10.17509/TJR.V5I1.48470.
- [3] F. M. Oktaviana, D. Wijayanto, and T. Wahyudi, "PENGARUH SOCIAL MARKETING CAMPAIGN TERHADAP KEPUTUSAN KONSUMEN BERTRANSAKSI MENGGUNAKAN APP MYPERTAMINA DI PONTIANAK," *J. TIN Univ. Tanjungpura*, vol. 5, no. 1, p. 23, Feb. 2021, Accessed: Nov. 15, 2022. [Online]. Available: <https://jurnal.un tan.ac.id/index.php/jtinUNTAN/article/view/44772>.
- [4] G. Keramidas, N. Voros, and M. Hübner, "Components and services for IoT platforms: Paving the way for IoT standards," *Components Serv. IoT Platforms Paving W. IoT Stand.*, pp. 1–383, Jan. 2016, doi: 10.1007/978-3-319-42304-3.
- [5] I. Nilasari, "Fungsi Aplikasi My Pertamina Sesuai Fitur yang Ada, Penting Dipahami!," 2022.

- <https://www.harapanrakyat.com/2022/07/fungsi-aplikasi-mypertamina/> (accessed Nov. 16, 2022).
- [6] E. Syahputra, "Pertamina Buka Pendaftaran BBM Subsidi Per 1 Juli 2022," Jun. 29, 2022. <https://www.cnbcindonesia.com/news/20220629165209-4-351500/pertamina-buka-pendaftaran-bbm-subsidi-per-1-juli-2022> (accessed Nov. 16, 2022).
- [7] Noviyanto, "Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian," *Paradig. – J. Inform. dan Komput.*, vol. 22, no. 2, 2020.
- [8] M. Stmik and T. Dharma, "Analisis Data Mining Pada Strategi Penjualan Produk PT Aquasolve Sanaria Dengan Menggunakan Metode K-Means Clustering," *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 2, no. 1, pp. 32–41, Jan. 2019, doi: 10.53513/JSK.V2I1.87.
- [9] F. M. Huda and R. Helilintar, "Implementasi Data Mining Pada Hasil Penjualan Barang Menggunakan Metode K-Means Clustering," *Pros. SEMNAS INOTEK (Seminar Nas. Inov. Teknol.)*, vol. 4, no. 2, pp. 61–65, Sep. 2020, doi: 10.29407/INOTEK.V4I2.121.
- [10] J. Eka Sembodo, E. Budi Setiawan, and Z. Abdurahman Baizal, "Data Crawling Otomatis pada Twitter," pp. 11–16, Sep. 2016, doi: 10.21108/INDOSC.2016.111.
- [11] M. A. Rofiqi, A. C. Fauzan, A. P. Agustin, A. A. Saputra, and H. D. Fahma, "Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query," *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 1, no. 2, pp. 58–64, Dec. 2019, doi: 10.28926/ILKOMNIKA.V1I2.18.
- [12] A. Supoyo and P. T. Prasetyaningrum, "Analisis Data Mining Untuk Memprediksi Lama Perawatan Pasien Covid-19 Di DIY," *Bianglala Inform.*, vol. 10, no. 1, pp. 21–29, Mar. 2022, doi: 10.31294/BI.V10I1.11890.
- [13] H. Susanto and S. Sudiyatno, "Data mining untuk memprediksi prestasi siswa berdasarkan sosial ekonomi, motivasi, kedisiplinan dan prestasi masa lalu," *J. Pendidik. Vokasi*, vol. 4, no. 2, Jun. 2014, doi: 10.21831/JPV.V4I2.2547.
- [14] A. T. J. Harjanta, "Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining," *J. Inform. Upgris*, vol. 1, no. 1 Juni, 2015, doi: 10.26877/JIU.V1I1.
- [15] P. D. Negeri, B. Bohong, H. Di, M. Sosial, and A. Budiman, "Berita Bohong (Hoax) di Media Sosial dan Pembentukan Opini Publik," *Pusat Penelitian Badan Keahlian DPR RI*, vol. IX, no. 01, pp. 2009–2012, 2017.
- [16] P. W. Cahyo and M. Habibi, "Clustering followers of influencers accounts based on likes and comments on Instagram Platform," no. x, pp. 1–10, 2020, doi: 10.22146/ijccs.xxxx.
- [17] N. Komang *et al.*, "Seleksi Fitur Bobot Kata dengan Metode TFIDF untuk Ringkasan Bahasa Indonesia," *J. Ilm. Merpati (Menara Penelit. Akad. Teknol. Informasi)*, vol. 6, no. 2, pp. 119–126, Aug. 2018, doi: 10.24843/JIM.2018.V06.I02.P06.
- [18] I. Romli, R. Firana Puspita Dewi, and K. Kunci, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS UNTUK KLASIFIKASI PENYAKIT ISPA," *Indones. J. Bus. Intell.*, vol. 4, no. 1, pp. 10–15, Jun. 2021, doi: 10.21927/IJUBI.V4I1.1727.