

Analisis Sentimen Penerapan Kurikulum Merdeka Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor Dengan Forward Selection

Wachid Darmawan^{*1}, Muhammad Faizal Kurniawan², WahyuSetianto³, Wim Hapsoro⁴

^{1,2,3,4} STMIK Widya Pratama Pekalongan

E-mail: ^{*1}wachiddw.stmikwp@gmail.com, ²faizal@stmik-wp.ac.id, ³kian@stmik-wp.ac.id, ⁴wimhapsoro06@gmail.com

Abstrak

Kurikulum yang di terapkan di Indonesia untuk kegiatan belajar mengajar selalu mengalami perubahan yang tujuannya untuk meningkatkan sistem pembelajaran yang ada. Kurikulum terbaru yang di terapkan pemerintah adalah kurikulum merdeka. Kurikulum merdeka ini akan serta merta menggantikan kurikulum 2013, walaupun saat ini masih ada sekolah yang menerapkan kurikulum 2013. Kurikulum yang baru diluncurkan dan diterapkan tahun ajaran 2022/2023 sudah banyak dirasakan dan banyak bermunculan opini yang bernada positif ataupun negatif di Twitter. Untuk mengetahui Analisis Sistemen dari opini yang ada, akan digunakan metode klasifikasi pada hasil opini masyarakat yang ada di Twitter. Data yang digunakan sebanyak 500 tweet, setelah dilakukan pembersihan data didapatkan data sebanyak 220 tweet, terdiri dari 110 data tweet positif dan 110 data tweet negatif. Metode klasifikasi yang digunakan untuk mengetahui Analisis Sentimen pada kurikulum Merdeka Belajar adalah metode K-Nearest Neighbor (K-NN) dengan Forward Selection (FS). Berdasarkan eksperimen yang sudah dilakukan metode K-NN akurasi sebesar 73.64%, sedangkan K-NN + FS akurasi meningkat menjadi 76.82%. berdasarkan perhitungan machine learning dihasilkan Sentimen Analisis bernada atau opini negatif yang banyak di tweetkan oleh masyarakat terkait penerapan Kurikulum Merdeka. Untuk penelitian selanjutnya bisa menggunakan algoritma klasifikasi lainya serta ada penambahan datasetnya.

Kata Kunci— kurikulum merdeka, k-nearest neighbor, analisis sentimen, Forward Selection

1. PENDAHULUAN

Penggunaan media sosial di Indonesia setiap tahunnya mengalami peningkatan, salah satu media sosial yang populer di Indonesia adalah Twitter. Berdasarkan data yang di himpun oleh We Are Social pengguna Twitter di Indonesia pada tahun 2022 mencapai 18,45 juta orang [1]. Dengan banyaknya jumlah pengguna twitter, tentu banyak opini yang bermunculan terkait dengan trending topik yang sedang dibicarakan. Dengan banyaknya opini tersebut, pasti ada opini yang positif maupun opini negatif [2].

Penyampaian opini terkait Kurikulum Merdeka cukup banyak disampaikan oleh masyarakat pengguna Twitter, karena kurikulum ini baru yang di terapkan pada Tahun Ajar 2022/2023. Penerapan Kurikulum Merdeka juga masih belum di terapkan kesemua sekolah, masih di lakukan tahap uji coba di sekolah-sekolah tertentu saja. Dengan adanya kurikulum baru pasti akan bermunculan kontroversi, salah satu kontroversi yang sering muncul di tweet adalah jam istirahat siswa berkurang karena banyaknya projek yang harus di kerjakan secara mandiri ataupun secara kelompok. Dengan adanya kontroversi tersebut banyak bermunculan opini terkait Kurikulum Merdeka yang disampaikan lewat Twitter.

Pengguna Twitter di Indonesia terbanyak ke lima di dunia [1], diperkirakan dalam sehari pengguna twitter melakukan tweet sebanyak 400 tweet [3]. Berdasarkan laporan yang dilakukan

oleh Reuters Institute Digital News Report tahun 2021 yang di kutip oleh katadata.co.id pengguna Twitter semakin tahun semakin banyak karena digunakan untuk mencari berita terbaru di bandingkan sosial media lainnya, seperti facebook, instagram, dll. Selain itu banyak pengguna yang melakukan debat atau berkomentar atau beropini dengan di buat utas (thread), yang di sosial media lain tidak ada.

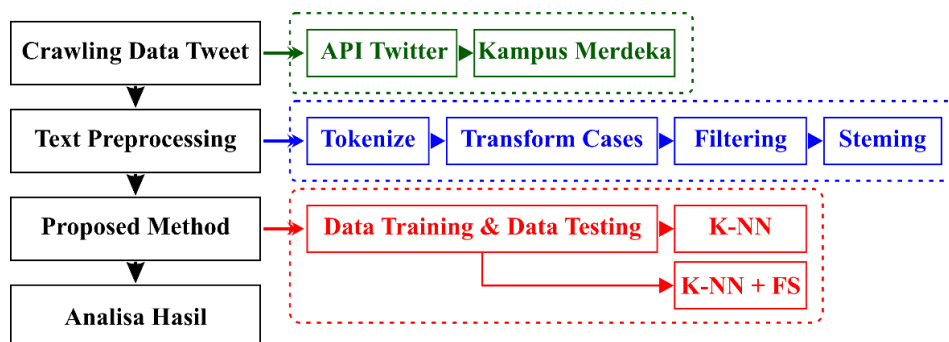
Penelitian tentang Analisis Sentimen yang menggunakan data dari pengguna Twitter sudah banyak dilakukan, seperti: Analisis Sentimen terhadap Program Merdeka Belajar Kampus Merdeka pada Twitter menggunakan *Support Vector Machine* (SVM). Penggunaan metode SVM dapat digunakan untuk memprediksi arah sentimen seseorang apakah positif atau negatif pada kegiatan PMM-BKM dengan nilai akurasi 95% [4]. Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan *Naïve Bayes*, *K-Nearest Neighbors* dan *Decision Tree*. Dari ketiga metode yang digunakan metode *Naïve Bayes* memiliki tingkat akurasi yang lebih baik yaitu 99,2% dengan menerapkan tiga subjek yaitu positif, negatif dan netral [5]. Analisis Sentimen Kurikulum 2013 pada Twitter menggunakan *Ensemble Feature* dan Metode *K-Nearest Neighbor*. Penerapan fitur *Ensemble Feature* pada metode *K-Nearest Neighbor* dapat meningkatkan akurasi sebesar 96% dari sebelumnya akurasi hanya 80% [2].

Menurut Han, Kamber dan Pei metode K-NN merupakan metode yang populer dalam klasifikasi [6], K-NN menggunakan teknik *lazy learning* dalam melakukan proses klasifikasi, berdasarkan jarak terdekat dan termasuk kedalam metode *supervised learning* [7]. K-NN melakukan klasifikasi melalui data pembelajaran dengan tingkat kemiripan yang paling dekat dengan objek pembelajarannya [8]. Menurut Bijalwan, et.al metode terbaik dalam melakukan *mining* pada dokument teks adalah K-NN dibandingkan *Naïve Bayes* atau Term Graph [9].

Dataset pada penelitian ini diperoleh dari opini yang ada di Twitter dengan melakukan teknik *Crawling* data, dengan kata kunci Kurikulum Merdeka. Setelah tahap *Crawling* data dilakukan tahap *preprocessing*, selanjutnya akan dilakukan tahap *proposed metode* atau menghitung akurasi dari metode K-NN dan K-NN + *forward selection*. *Forward selection* sebagai metode seleksi fitur pada dataset yang prinsip kinerjanya berdasarkan *regresi liner* [10]. Setelah tahap *proposed metode* selesai akan dilakukan analisis hasilnya berdasarkan perhitungan *machine learning*.

2. METODE PENELITIAN

Penelitian eksperimental tentang analisis sistimen penerapan Kurikulum Merdeka pada pengguna Twitter digunakan untuk mengetahui model yang diinginkan, berikut metode penelitian yang dilakukan seperti di gambar 1.



Gambar 1. Metode Penelitian

2.1. Crawling Data Tweet

Untuk mendapatkan data tweet tools yang digunakan adalah Rapidminer, teknik sampling yang digunakan *purposive sampling*, yaitu penentuan sampel berdasarkan tujuan yang diinginkan berdasarkan kriteria tertentu untuk proses penelitian [12]. Kata kunci yang digunakan dalam penentuan sampel adalah Kampus Merdeka dengan jumlah data sebanyak 500 *tweet*. Dari *tweet* dilakukan pembersihan dan pemberian label untuk *tweet* bernada positif atau negatif.

2.2. Text Preprocessing

Pada *text preprocessing* atau pemrosesan data hasil *crawling data tweet* diolah menggunakan beberapa operator, diantaranya: *Tokenize*, *Transform Cases*, *Filter* (*Tokens* dan *Stopwords*) dan *Stem*.

2.3. Proposed Method

Pada tahapan ini dilakukan untuk mengetahui tingkat akurasi dari metode yang digunakan, yaitu *K-Nearest Neighbor* dan *K-Nearest Neighbor* dengan *Forward Selection*. Dari hasil *crawling tweet* digunakan juga sebagai data *training* dan data *testing* dengan teknik *cross validation*.

2.4. Analisa Hasil

Untuk mengetahui analisis sentimen dari Kurikulum Merdeka akan menggunakan alat bantu Rapidminer 9.10. Hasil dari model yang sudah di tentukan akan dilakukan analisis sentimen dan melakukan evaluasi hasilnya.

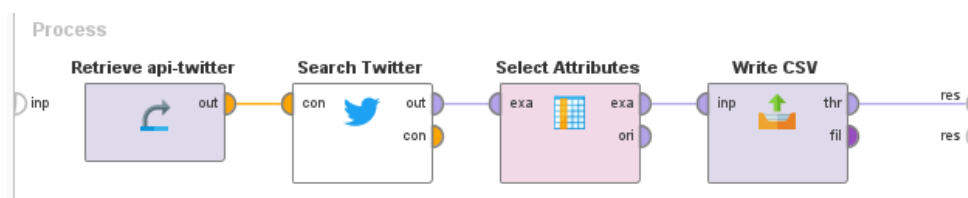
3. HASIL DAN PEMBAHASAN

3.1. Pembahasan

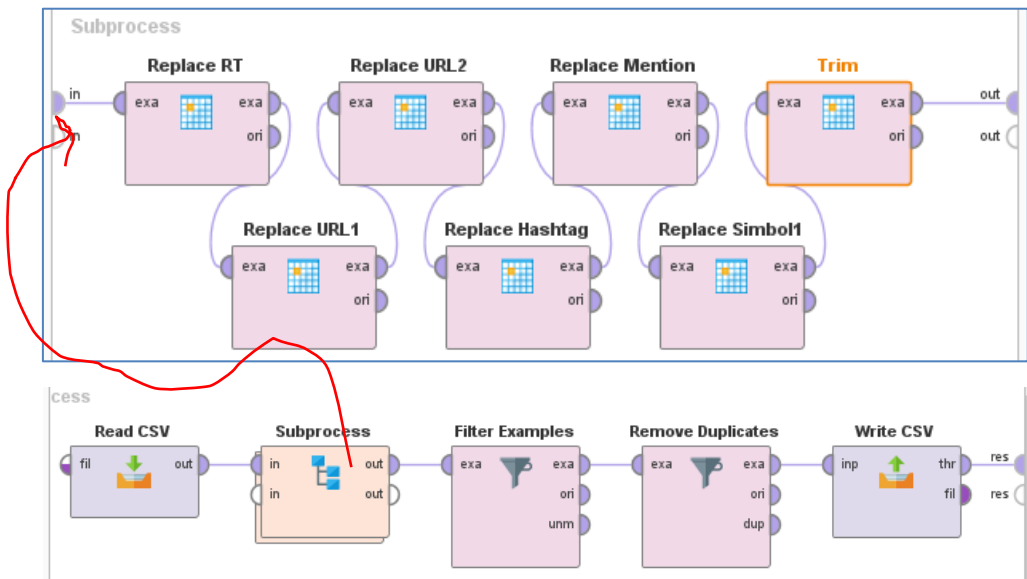
Penelitian ini menggunakan dataset hasil dari *crawling data tweet* sebanyak 500 *tweet*, setelah dataset terbentuk selanjutnya dilakukan pembersihan dan pelabelan antara *tweet* bernada positif maupun bernada negatif.

3.1.1. Crawling Data Tweet

Proses *crawling data tweet* menggunakan alat bantu Rapidminer 9.10. Untuk sampling data yang digunakan sebanyak 500 *tweet*. Setelah data didapatkan dilakukan pembersihan data dan pemberian label pada *tweet* yang bernada positif atau negatif. Untuk *tweet* bernada positif diambil data sebanyak 110 *tweet*, sedangkan *tweet* bernada negatif juga diambil sebanyak 110 *tweet*. Selanjutnya dilakukan pelabelan dari data yang sudah didapatkan.



Gambar 2. Crawling Data Tweet



Gambar 3. Perbersihan Crawling Data Tweet

3.1.2. Text Preprocessing

Setelah dataset hasil *crawling data tweet* sudah bersih dan sudah ada labelnya, selanjutnya di lakukan tahap *text preprocessing* pada dataset tersebut [11]. Pada tahap ini operator yang digunakan adalah:

- 1. *Tokenize* digunakan untuk memisahkan setiap kata-kata dari *tweet* yang ada.

Tabel 1. Contoh Hasil Tokenize

Sebelum Tokenize	Sesudah Tokenize
Kurikulum merdeka ini yang merdeka siapa ya kok aku merasa terjajah	Kurikulum merdeka ini yang merdeka siapa ya kok aku merasa terjajah
Itu keponakan2 aku pd suka metodenya. Aku jg ikut senang. Anak2 jd lbh happy wktu belajar kognitif sosialnya brkembang.	Itu keponakan2 aku pd suka metodenya. Aku jg ikut senang. Anak2 jd lbh happy wktu belajar kognitif sosialnya brkembang.

- 2. *Transform Cases* digunakan untuk menyamaratakan semua huruf yang digunakan.

Tabel 2. Contoh Hasil *Transform Cases*

Sebelum Tokenize	Sesudah Tokenize
nie gara gara kurikulum merdeka ak jd bljar kimia ALIAS AK GA SUKA	nie gara gara kurikulum merdeka ak jd bljar kimia alias ak ga suka
Ngga lagi mengutamakan NILAI akademis tapi menekankan penguatan KARAKTER ga ada PR belajar yg lebih menyenangkan	ngga lagi mengutamakan nilai akademis tapi menekankan penguatan karakter ga ada pr belajar yg lebih menyenangkan

3. *Filter (Tokens dan Stopwords)* digunakan untuk menghapus kata yang tidak penting, contoh: a, s, di, yang, dan, agar, ada dll.

Tabel 3. Contoh Kata-Kata Dalam *Filter Stopword*

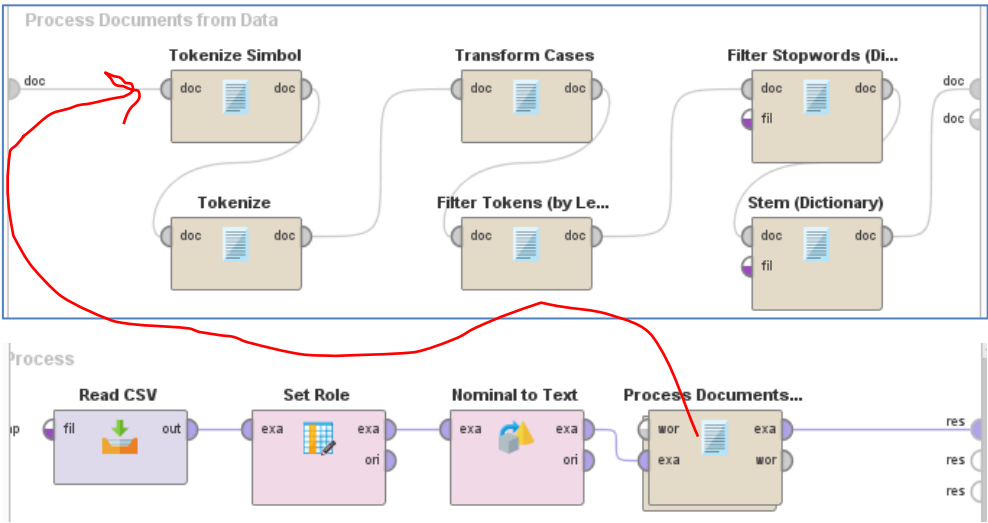
Filter Stopword						
apatah	bapak	cara	datang	empat	haruslah	ibaratnya
artinya	baru	caranya	dekat	enggak	harusnya	ibu
asal	bawah	cukup	demi	enggaknya	hendak	ikut
asalkan	beberapa	cukupkah	demikian	entah	hendaklah	ingat
...	...	...	...	...	...	...

4. *Stem* digunakan untuk mencari kata dasar (*root*) dari *tweet* yang ada.

Tabel 4. Contoh Hasil *Stem*

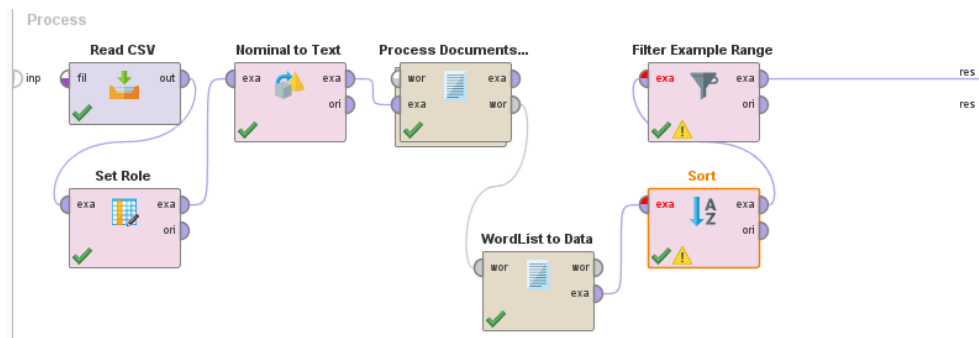
Sebelum Stem	Sesudah Stem
terancam	ancam
berbahaya	bahaya
tercepat	cepat
terdepan	depan
...	...

Berikut penggunaan operator-operator yang dibutuhkan dalam melakukan *text preprocessing* di Rapidminer 9.10, seperti berikut:



Gambar 4. *Text Preprocessing Crawling Data tweet*

Dari hasil *text preprocessing* yang sudah dilakukan akan muncul kata-kata yang sering di *tweetkan*, berikut visualisasi kata-kata tersebut dalam *Wordcloud*, *tweet* bernada positif atau negatif.



Gambar 5. *Visualisasi Wordcloud Pada Text Preprocessing*



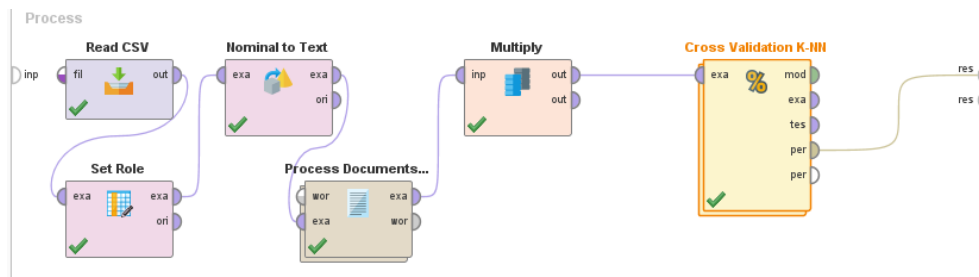
Gambar 6. Hasil Visualisasi *Wordcloud* Tweet Bernada Positif



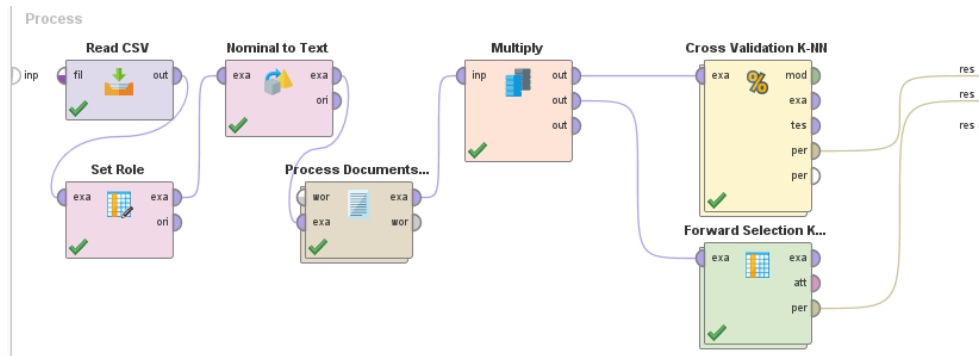
Gambar 7. Hasil Visualisasi *Wordcloud* Tweet Bernada Negatif

3.1.3. Proposed Method

Setelah melakukan *text preprocessing* selanjutnya melakukan komparasi perhitungan metode yang digunakan dengan alat bantu Rapidminer 9.10 untuk mengetahui akurasi yang dihasilkan dari kedua metode yang di usulkan menggunakan teknik *cross validation*.



Gambar 8. Proses Analisis Sentimen Menggunakan Metode KNN



Gambar 9. Proses Analisis Sentimen Menggunakan Metode KNN+FS

3.1.4. Analisa Hasil

Analisis sentimen pada Kurikulum Merdeka dengan menggunakan model K-NN dan K-NN+FS, hasilnya dapat di lihat pada gambar 10 dan 11 setelah melakukan eksperimen kedua metode yang diusulkan.

accuracy: 73.64% +/- 15.27% (micro average: 73.64%)

	true negatif	true positif	class precision
pred. negatif	87	35	71.31%
pred. positif	23	75	76.53%
class recall	79.09%	68.18%	

Gambar 10. Hasil Eksperimen Metode K-NN

Dari gambar 10 dapat dijelaskan bahwa Akurasi yang dihasilkan dari metode K-NN sebesar 73,64%, untuk *tweet* yang bernada positif *machine learning* memprediksi *tweet* positifnya sebesar 75, sedangkan *tweet* negatifnya 35, dengan nilai *recall* sebesar 68,18% dan *precisionnya* 76,53%. Sedangkan *tweet* yang bernada negatif *machine learning* memprediksi *tweet* positif sebesar 23, sedangkan *tweet* negatifnya 87, dengan nilai *recall* sebesar 79,09% dan *precisionnya* 71,31%.

accuracy: 76.82% +/- 6.93% (micro average: 76.82%)

	true negatif	true positif	class precision
pred. negatif	103	44	70.07%
pred. positif	7	66	90.41%
class recall	93.64%	60.00%	

Gambar 11. Hasil Eksperimen Metode K-NN + FS

Sedangkan pada gambar 11 dapat dijelaskan bahwa Akurasi yang dihasilkan dari metode K-NN + FS sebesar 76,82%, untuk *tweet* yang bernada positif *machine learning* memprediksi *tweet* positifnya sebesar 66, sedangkan *tweet* negatifnya 44, dengan nilai *recall* sebesar 60,00% dan *precision*nya 90,41%. Sedangkan *tweet* yang bernada negatif *machine learning* memprediksi *tweet* positifnya sebesar 7, sedangkan *tweet* negatifnya 103, dengan nilai *recall* sebesar 93,64% dan *precision*nya 70,07%.

Dari gambar 10 dan 11 dapat disimpulkan bahwa penambahan FS sebagai metode seleksi fitur pada metode K-NN meningkatkan akurasi. Sedangkan Analisis Sentimen penerapan Kurikulum Merdeka berdasarkan hasil *machine learning* dengan menggunakan metode K-NN maupun K-NN + FS mendapatkan respon negatif yang lebih kuat dibandingkan respon positifnya.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan, dapat ditarik kesimpulan, seperti berikut:

1. Berdasarkan penelitian eksperimen yang sudah dilakukan model *K-Nearest Neighbor* menghasilkan akurasi sebesar 73,64%.
2. Sedangkan model *K-Nearest Neighbor* dengan *Forward Selection* menghasilkan akurasi sebesar 73,64%.
3. Penambahan seleksi fitur, yaitu *Forward Selection* pada sebuah metode (*K-Nearest Neighbor*) dapat meningkatkan hasil akurasi.
4. Analisis Sentimen pada penerapan Kurikulum Merdeka banyak mendapatkan respon atau opini yang bernada negatif berdasarkan perhitungan *machine learning* yang sudah dilakukan.

DAFTAR PUSTAKA

- [1] M. A. Rizaty, "Pengguna Twitter di Indonesia Capai 18,45 Juta pada 2022," 10 Agustus 2022. [Online]. Available: <https://dataindonesia.id/digital/detail/pengguna-twitter-di-indonesia-capai-1845-juta-pada-2022>.
- [2] M. R. Irfan, M. A. Fauzi and Tibyani, "Analisis Sentimen Kurikulum 2013 pada Twitter menggunakan Ensemble Feature dan Metode K-Nearest Neighbor," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, pp. 3006-3014, 2018.
- [3] N. F. F. d. Silva, E. R. Hruschka and E. R. Hruschka, "Tweet sentiment analysis with classifier ensembles," Decision Support Systems, pp. 170-179, 2014.
- [4] N. Kholila, "Analisis Sentimen terhadap Program Merdeka Belajar Kampus Merdeka pada Twitter menggunakan Support Vector Machine (SVM)," ANTIVIRUS: Jurnal Ilmiah Teknik Informatika, pp. 252-261, 2021.

- [5] A. Rozaq, Y. Yunitasari, K. Sussolaikah, E. R. N. Sari and R. I. Syahputra, "Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan Naïve Bayes, K-Nearest Neighbors dan Decision Tree," *Jurnal Media Informatika Budidarma*, pp. 746-750, 2022.
- [6] J. Han, M. Kamber and J. Pei, *Data Mining Concepts and Techniques Third Edition*, Burlington, Middlesex County: Morgan Kaufmann Publishers, 2012.
- [7] D. Cahyanti, A. Rahmayani and S. A. Husniar, "Analisis performa metode Knn pada Dataset pasien pengidap Kanker Payudara," *Indonesian Journal of Data and Science*, pp. 39-43, 2020.
- [8] E. Hardiyanto and F. Rahutomo, "Studi Awal Klasifikasi Artikel Wikipedia Bahasa Indonesia Dengan Menggunakan Metoda K Nearest Neighbor," in *Prosiding Sentrinov (Seminar Nasional Terapan Riset Inovatif)*, Semarang, 2016.
- [9] V. Bijalwan, V. Kumar, P. Kumari and J. Pascual, "KNN Based Machine Learning Approach for Text and Document Mining," *International Journal of Database Theory and Application*, pp. 61-67, 2014.
- [10] R. Noori, A. R. Karbassi, A. M. Nia, D. Han, M. H. Z. Ashtiani, A. Farokhnia and M. G. Gousheh, "Assessment of Input Variables Determination on the SVM Model Performance Using PCA, Gamma Test, and Forward Selection Techniques for Monthly Stream Flow Prediction," *Journal of Hydrology*, vol. 401, no. 3-4, pp. 177-189, 2011.
- [11] A. P. Wibowo, W. Darmawan and N. Amalia, "Komparasi Metode Naïve Bayes dan K-Nearest Neighbor terhadap Analisis Sentimen Pengguna Aplikasi Pedulilindungi," *IC-Tech - Journal of Informatic and Computer Technology*, vol. 17, no. 1, pp. 18-23, 2022.
- [12] N. Indriantoro and B. Supomo, *Metodologi Penelitian Bisnis: Untuk Akuntansi dan Manajemen*, Yogyakarta: Penerbit BPFE, 2016.