

# Implementasi Algoritma K-Nearest Neighbour dalam Menganalisis Sentimen Terhadap Program Merdeka Belajar Kampus Merdeka (MBKM)

Dewi Sartika

Program Studi Teknik Informatika Fakultas Ilmu Komputer, Universitas Indo Global Mandiri,  
Palembang 30129, Sumatera Selatan, Indonesia  
Email: dewi.sartika@uigm.ac.id

**Abstract.** *K-Nearest Neighbor Algorithm Implementation in sentiment analysis towards Merdeka Belajar Kampus Merdeka (MBKM) Program.* Merdeka Belajar Kampus Merdeka (MBKM) is a program that supports students to improve their skills by having direct experience in the work environment to prepare for competition and a future career. MBKM program has been implemented by Indonesia's Ministry of Education, Culture, Research, and Technology (Kemendikbudristek) since 2020. Every policy needs to be evaluated; a simple evaluation can be done through sentiment analysis to determine public responses to the MBKM program. The results are used as suggestions for program improvement. Sentiment analysis is done by applying the Natural Language Processing (NLP) algorithm to process crawled data from Twitter, then classified using the K-NN Algorithm. Based on the results, the sentiment is neutral. This illustrates that people are only partially interested in the MBKM program policy. The accuracy of the classification model using the K-NN algorithm is 95%, and an F1-score value of 0.96 for the classification model with a ratio of 80% training data and 20% test data.

**Keywords:** MBKM, NLP, K-NN, F1-Score

**Abstrak.** Program Merdeka Belajar Kampus Merdeka (MBKM) merupakan suatu kebijakan dalam mendukung pemberian kebebasan terhadap mahasiswa untuk mengasah kemampuan dengan merasakan langsung pengalaman di dunia kerja sebagai bekal untuk menghadapi persaingan dan persiapan berkarir di masa mendatang. Program MBKM mulai diberlakukan oleh Kementerian Pendidikan Kebudayaan Riset dan Teknologi (Kemendikbudristek) Republik Indonesia sejak tahun 2020. Setiap kebijakan tentunya perlu dievaluasi, evaluasi sederhana dapat dilakukan melalui analisis sentimen untuk mengetahui tanggapan masyarakat mengenai program MBKM. Hasilnya digunakan sebagai saran perbaikan untuk pengembangan program. Analisis sentimen dilakukan dengan menerapkan algoritma Natural Language Processing (NLP) untuk memproses data hasil crawling dari Twitter, selanjutnya diklasifikasikan menggunakan algoritma K-NN. Berdasarkan hasil analisis diperoleh bahwa sentimen masyarakat bersifat netral. Hal ini menggambarkan bahwa masyarakat tidak sepenuhnya tertarik terhadap kebijakan program MBKM, sedangkan untuk tingkat akurasi model klasifikasi menggunakan algoritma K-NN sebesar 95% dan nilai F1-score sebesar 0,96 untuk model klasifikasi dengan perbandingan 80% data latih dan 20% data uji.

**Kata Kunci:** MBKM, NLP, K-NN, F1-Score

## 1. Pendahuluan

Program Merdeka Belajar Kampus Merdeka (MBKM) merupakan suatu kebijakan dalam mendukung pemberian kebebasan terhadap mahasiswa untuk mengasah kemampuan dengan merasakan langsung pengalaman di dunia kerja sebagai bekal untuk menghadapi persaingan dan persiapan berkarir di masa mendatang. Berdasarkan Peraturan Menteri Pendidikan dan Kebudayaan (Permendikbud) Nomor 3 Tahun 2020 Pasal 15 Ayat 1 dinyatakan bahwa pada Semester 5 Mahasiswa ditawarkan untuk mengikuti matakuliah di luar program studi, namun dalam lingkup Universitas. Sedangkan pada Semester 6 dan 7 terdapat 8 jenis kegiatan MBKM yang ditawarkan dan bersifat opsional.

Kegiatan MBKM dapat dilaksanakan melalui keikutsertaan program yang disediakan oleh Kemendikbudristek maupun program yang disediakan oleh internal Perguruan Tinggi (PT). Program MBKM yang disediakan oleh Kementerian Pendidikan Kebudayaan Riset dan Teknologi (Kemendikbudristek) dapat diikuti oleh setiap mahasiswa yang memenuhi syarat dan melalui tahapan seleksi. Selama menjalani program MBKM, setiap program mewajibkan mahasiswa untuk melaksanakan tugas spesifik sesuai dengan tema program yang diikuti. Berdasarkan rangkaian kegiatan dan evaluasi mahasiswa akan mengajukan pengakuan ke Program Studi untuk konversi SKS ke matakuliah yang bersesuaian.

Kebijakan program MBKM mulai diberlakukan sejak tahun 2020 hingga saat ini. Perguruan tinggi melakukan penyesuaian kurikulum yang berlaku sesuai dengan kebijakan program MBKM. Setiap kebijakan tentunya perlu dievaluasi, evaluasi sederhana dapat dilakukan melalui analisis sentimen. Analisis sentimen termasuk dalam bidang *Natural Language Processing* (NLP) [1]. Analisis sentimen merupakan analisis terhadap pernyataan tentang sesuatu, bisa berupa produk, event, fenomena maupun program yang mengindikasikan ekspresi positif, negatif, atau netral [2], [3]. Analisis sentimen dilakukan dengan memanfaatkan berbagai platform media sosial sebagai sumber data. Hal itu dikarenakan media sosial dijadikan sebagai tempat bertukar informasi serta berekspeksi dalam menyampaikan pendapat [4]. Data yang diperoleh dari sosial media berupa teks yang kemudian diekstraksi informasinya untuk memperoleh opini dari masyarakat terkait topik tertentu. Opini ini selanjutnya akan bernilai untuk pengambilan keputusan.

Salah satu platform media sosial yang dapat dijadikan dataset dalam analisis sentimen adalah Twitter. Indonesia merupakan negara ke lima yang memiliki jumlah pengguna Twitter terbanyak. Pada tahun 2022 jumlah pengguna Twitter di Indonesia mencapai 18,45 juta [5]. Berdasarkan peluang tersebut, maka akan dilakukan analisis sentimen masyarakat terkait program MBKM berdasarkan data Twitter. Pada penelitian ini dilakukan pengolahan data menggunakan algoritma NLP, kemudian dihitung nilai polaritas untuk menentukan label dari setiap data. Tahap akhir dilakukan klasifikasi menggunakan algoritma *K-Nearest Neighbour* (KNN). Algoritma ini dipilih karena memiliki tingkat akurasi yang tinggi [6]. Penentuan nilai  $k$  dapat dilakukan berdasarkan perolehan tingkat akurasi tertinggi. Tujuan dari penelitian ini untuk mengetahui tanggapan masyarakat mengenai program MBKM. Hasilnya digunakan sebagai saran perbaikan untuk Kemendikbudristek.

## **2. Tinjauan Pustaka**

### **2.1. Text Mining**

*Text mining* atau penambangan data merupakan proses pengekstraksian informasi yang diperoleh dari basis data berupa sekumpulan teks atau dokumen. Tantangan dalam melakukan *text mining* adalah tidak terlepas dari permasalahan ambiguitas [7]. Hal tersebut dikarenakan satu kata dapat memiliki banyak arti dan banyak kata dapat memiliki satu arti. *Text mining* dapat dimanfaatkan untuk merubah data teks yang tidak terstruktur menjadi terstruktur.

### **2.2. Analisis Sentimen**

Analisis sentimen merupakan salah satu bentuk dari penerapan *text mining*. Analisis sentimen memanfaatkan teknologi NLP untuk mengidentifikasi, memahami, dan mengevaluasi sentimen maupun opini yang diperoleh dari data sosial dengan tujuan untuk menentukan sentimen terhadap suatu topik tersebut termasuk pada kategori positif, negatif, atau netral. Selain tiga kategori tersebut sentimen juga dapat dikategorikan menggunakan skala lainnya seperti baik, sangat baik, jelek, dangat jelek dan sebagainya [8]. Hasil dari analisis sentimen dapat digunakan untuk menentukan tindak lanjut yang harus dilakukan dengan tujuan perbaikan untuk masa mendatang.

Terdapat dua teknik dalam melakukan analisis sentimen yaitu menggunakan teknik *machine learning method* dan *lexical based method* [9]. Pada *machine learning method*, sentimen diperoleh berdasarkan data latih yang telah diketahui labelnya untuk selanjutnya digunakan dalam memprediksi data uji. Sedangkan pada *lexical based method*, sentimen diperoleh berdasarkan kumpulan data yang saling terhubung dengan sentimen yang spesifik.

### 2.3. Natural Language Processing (NLP)

NLP merupakan algoritma yang digunakan untuk *pre-processing* data teks. NLP dilakukan sebelum melakukan tahapan selanjutnya, seperti klasifikasi, pengelompokan dan sebagainya. *Pre-processing* dilakukan untuk menghilangkan *noise* pada data, sekaligus membuat data lebih terstruktur [10]. Beberapa penelitian sejenis yang telah dilakukan sebelumnya melakukan tahapan data processing menggunakan NLP dengan tahapan yang berbeda. Pada penelitian yang dilakukan Arief et.al tahapan dari NLP yang dilakukan yaitu *filtering*, *case folding*, *word normalization*, *tokenization* dan *stop word removal* [11]. Penelitian yang dilakukan oleh Rani dan Nasib melakukan pembobotan TF-IDF pada tahapan NLP [12]. Penelitian yang dilakukan oleh F.M Javed et.al melakukan lemmatization dan object identification pada tahapan NLP [11]. Tahapan NLP yang dilakukan pada penelitian ini adalah *case folding*, *data cleansing*, *tokenization*, *word normalization*, *stop word removal*, dan *stamming*.

### 2.4. Machine Learning

Pembelajaran mesin bertujuan untuk membangkitkan suatu model secara otomatis berdasarkan sekumpulan data latih untuk pemrosesan data yang tidak dikenal [13]. Pembelajaran mesin memanfaatkan algoritma dan model matematis untuk menganalisis data dan memperoleh pola-pola yang kemudian digunakan untuk membuat prediksi atau keputusan dalam optimalisasi tugas yang diinginkan. Berdasarkan cara kerjanya pembelajaran mesin terdiri dari dua yaitu *supervised learning* dan *unsupervised learning*. *Supervised learning* akan membentuk suatu model berdasarkan sekumpulan data yang telah diketahui pasangan masukkan dan keluarannya, sedangkan *unsupervised learning* membentuk model berdasarkan data yang tidak terstruktur dan belum diketahui keluarannya [14].

### 2.5. K-Nearest Neighbour (KNN)

KNN adalah suatu algoritma pembelajaran mesin bekerja dengan cara menghitung jarak antara data uji dengan data latih. Data latih dengan jarak terdekat akan digunakan untuk menentukan prediksi dari data uji. Rumus perhitungan jarak yang dapat digunakan pada algoritma KNN yaitu *euclidean distance*. Persamaan dari perhitungan *euclidean distance* dapat dilihat pada Persamaan 1, diketahui  $p$  dan  $q$  adalah data yang akan dihitung jaraknya.

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (1)$$

Berdasarkan penelitian sebelumnya menyatakan dibandingkan algoritma Naive Bayes, KNN memiliki tingkat akurasi yang lebih baik dalam menganalisis kepuasan pengguna pada pembayaran digital di Indonesia [15]. Penentuan analisis sentimen menggunakan KNN dilakukan dengan menghitung nilai polaritas, jika suatu teks memiliki polaritas lebih besar dari 0 maka diklasifikasikan sebagai sentimen positif, jika sebaliknya maka negatif, dan jika sama dengan 0 maka dinyatakan netral. Persamaan yang digunakan untuk menghitung polaritas dapat diilustrasikan pada Persamaan 2–5.

$$\text{mean}, \bar{x} = \frac{\sum x}{n} \quad (2)$$

$$\text{median} = \frac{n+1}{2} \quad (3)$$

$$\text{average minimum} = \frac{(n-1) \min + \max}{n} \quad (4)$$

$$\text{average maximum} = \frac{\min + (n-1) \max}{n} \quad (5)$$

## 2.6. Perhitungan Tingkat Akurasi

Perhitungan tingkat akurasi digunakan untuk mengukur keberhasilan model klasifikasi dalam menentukan prediksi. Tingkat akurasi dihitung berdasarkan nilai *confusion matrix* [16]. Tabel *confusion matrix* dapat dilihat pada Tabel 1. Perhitungan tingkat akurasi, *precision*, dan *recall* menggunakan Persamaan 6-8, sedangkan untuk mengevaluasi model yang dihasilkan dengan menghitung *F1-score* menggunakan Persamaan 9 [17]. Model yang baik akan memiliki nilai *F1-score* mendekati nilai 1 [18].

**Tabel 1. Confusion Matrix**

Klasifikasi	Hasil Klasifikasi	
	Positif	Negatif
<i>True</i>	Jumlah Data <i>true</i> yang dipresiksi benar (TP)	Jumlah Data <i>false</i> yang dipresiksi benar (TN)
<i>False</i>	Jumlah Data <i>true</i> yang dipresiksi salah (FP)	Jumlah Data <i>false</i> yang dipresiksi salah(FN)

$$Akurasi = \frac{TP+TN}{TP+FP+TN+FN} \quad (6)$$

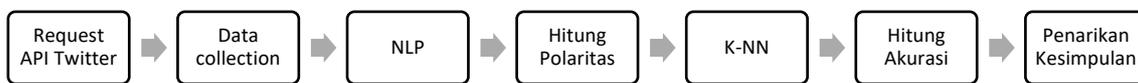
$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

$$F1 - score = 2x \frac{precision \times recall}{precision+recall} \quad (9)$$

## 3. Metodologi Penelitian

Tahapan – tahapan yang dilaksanakan pada penelitian dapat dilihat pada Gambar 1.



**Gambar 1. Tahapan Penelitian**

1. Request API Twitter: untuk memperoleh akses ke Twitter, perlu melakukan permintaan *Application Programming Interface* (API) kepada pihak Twitter. Setelah disetujui akan diperoleh kunci dan token yang bisa digunakan untuk tahapan selanjutnya.
2. *Data collection*: tahapan ini dilakukan untuk mengumpulkan data berupa *tweet* pengguna Twitter menggunakan *library* Tweepy. Data yang terkumpul diubah dalam bentuk *dataframe* menggunakan *library* Pandas.
3. Memproses data *tweet* menggunakan NLP: tahapan ini dilakukan dengan *library* yang tersedia oleh pihak ketiga yaitu NLTK dan Pipelines.
4. Hitung polaritas: tahapan ini dilakukan untuk memperoleh sentimen dari *tweet* pengguna berdasarkan nilai polaritas. Perhitungan polaritas dilakukan menggunakan *library* Textbolb.
5. Mengklasifikasi sentimen menggunakan K-NN: tahapan ini dilakukan untuk membuat model klasifikasi dengan  $k = 3$  dan *splitting* data dengan perbandingan 80:20 (*test size* = 0,2), 70:30 (*test size* = 0,3) dan 60:40 (*test size* = 0,4).
6. Hitung akurasi: tahap ini dilakukan perhitungan *confusion matrix* untuk mendapatkan nilai akurasi dan *F1-score* untuk mengetahui jumlah *splitting* data yang mana yang menghasilkan tingkat akurasi terbaik.
7. Penarikan kesimpulan: kesimpulan didapatkan berdasarkan hasil perhitungan akurasi dari model klasifikasi.

#### 4. Hasil dan Diskusi

Tahapan yang dilakukan setelah mendapat kunci dan token hasil dari *request* API dari laman pengembang Twitter, selanjutnya dilakukan pengumpulan data. Data *tweet* yang terkumpul selanjutnya diubah menjadi *dataframe* yang terdiri dari 4 kolom, yaitu waktu, id, *username*, teks. Tabel 2 menampilkan contoh data *tweet* yang diperoleh dari hasil *crawling* data. Diperoleh sebanyak 373 *tweet* yang berkaitan dengan MBKM.

**Tabel 2. Data Hasil Crawling Data**

Waktu	Id	Username	Teks
2023-04-01 14:06:29+00:00	1642166545681100805	COLLE [cm]	gaiis mau nanyaa, misal kita daftar mbkm pas sem ester 5 nanti nah itu tuh kita bakal mulai mbkmnya semester 5... <a href="https://t.co/7z19ficlwD">https://t.co/7z19ficlwD</a> 5 tuh mbkmnya semester 5
2023-03-30 05:19:48+00:00	1641309226428678146	UIN Raden Intan	Buku Panduan Penyelenggaraan MBKM UIN Rad en Intan Lampung <a href="https://t.co/sU5nFb4rW3">https://t.co/sU5nFb4rW3</a>
2023-03-26 14:28:20+00:00	1639997719925370880	02L	smt 5 bener2 smt paling kek babykk, gamau ngula ng lagi anjir gr2 mbkm 14 tm dimampetin jadi 3 m inggu kek wtf gilakk... <a href="https://t.co/2fDepLKiiA">https://t.co/2fDepLKiiA</a>
2023-03-25 15:31:50+00:00	1639651310281818114	haruharuhaha	RT @bintang4_ : Salam pak menteri dan jajaranny a, saya mau tanya terkait msib dan mbkm kampus zakat, itu kan seluruhnya full dikonversi 20 s...

Sebelum penerapan NLP, dilakukan penghapusan data redundansi yang ada didalam *dataframe*. Pada tahapan tersebut terdapat sebanyak 89 *tweet* redundansi yang sebagian besar merupakan *retweet* dari *tweet* orang lain. Sebanyak 284 data yang selanjutnya diproses melalui beberapa tahapan berikut:

1. Merubah seluruh huruf menjadi huruf kecil;
2. Melakukan *cleansing* data dengan menghapus simbol, *link*, dan emoji;
3. Melakukan tokenisasi untuk memperoleh kata dari kalimat;
4. Mengganti kata singkatan dengan kata sesuai Kamus Besar Bahasa Indonesia (KBBI);
5. Normalisasi data;
6. Menghapus *stop word* (yang tidak memiliki makna);
7. *Stemming* menggunakan *library* Sastrawi dan singkatan.

**Tabel 3. Data Penentuan Sentimen Tweet**

Waktu	Id	Username	Teks	Sentimen
2023-04-01 14:06:29+00:00	1642166545681100805	COLLE	gai nanyaa daftar mbkm pas semester 5 tuh mbkmnya semester 5	Netral
2023-03-30 05:19:48+00:00	1641309226428678146	UIN Raden Intan	buku pandu selenggara mbkm uin raden intan lampung	Netral
2023-03-26 14:28:20+00:00	1639997719925370880	02L	smt 5 benar smt kayak babykk enggak mau ngulang anjir gara-gara mbkm 14 tm dimampetin 3 minggu kayak wtf gilakk	Negatif
2023-03-25 15:31:50+00:00	1639651310281818114	haruharuhaha	rt salam menteri jajar kait msib mbkm kampus zakat full konversi 20 s	Positif

Setelah melalui tahapan NLP, selanjutnya dilakukan perhitungan polaritas untuk memperoleh sentimen dari masing-masing *tweet*. Rincian data setelah dihitung polaritas diperoleh sebanyak 264 *tweet* dinyatakan sebagai opini netral, 14 *tweet* dinyatakan sebagai opini positif dan 6 *tweet* lainnya dinyatakan sebagai opini negatif. Tabel 3 berisikan contoh data yang telah melalui algoritma NLP dan dilakukan analisis sentimen berdasarkan perhitungan polaritas.

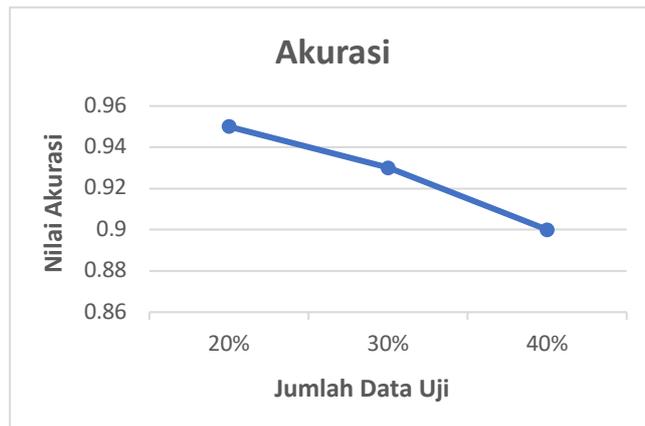
Pada penerapan NLP dan perhitungan polaritas terdapat *tweet* yang secara manusiawi dikelompokkan pada sentimen yang kurang tepat. Hal ini dikarenakan perlu adanya penambahan kata singkatan yang *up to date* pada *library* pada tahapan sebelum normalisasi data. Selain kata singkatan, beberapa *tweet* yang menggunakan bahasa daerah juga perlu dihapus atau disesuaikan

dengan penambahan *library* khusus bahasa daerah. Tabel 4 berisikan contoh *tweet* yang hasil analisis sentimennya kurang tepat.

**Tabel 4. Data Penentuan Sentimen Kurang Tepat**

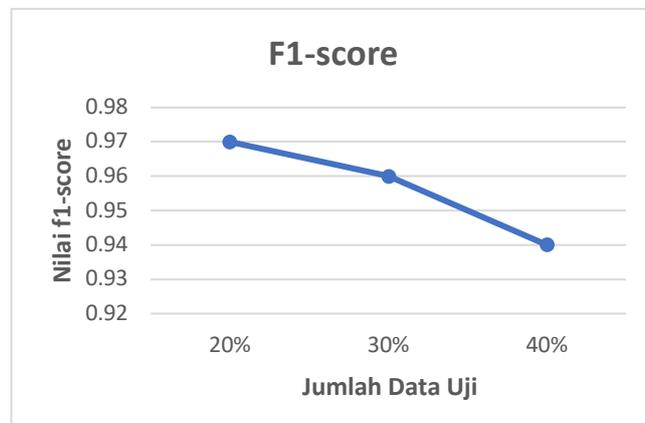
Waktu	Id	Username	Teks	Sentimen
2023-04-01 22:50:47+00:00	1642298491702775808	K Saadah	maaf mas menteri tapi iku mbkm ni bikin mules kemringet	Netral
2023-04-01 15:27:28+00:00	1642186927901401089	Anna	capek banget stresssssss stres mbkm serta printilannya ya allah plis asa pengen nyerah saja an	Netral
2023-03-30 15:30:03+00:00	1641462803612733441	Media Indonesia	mbkm nilai bangun ekosistem didik kualitas guru	Netral
2023-03-29 16:26:49+00:00	1641114698379837440	y	sing wajar wae kampus endi mbkm dibebanke semester 8 nek ora kampus	Netral

Selanjutnya dilakukan klasifikasi dengan membagi data secara acak dengan perbandingan data latih dan data uji sebesar 80:20 (*test size* = 0,2), 70:30 (*test size* = 0,3) dan 60:40 (*test size* = 0,4). Perhitungan tingkat akurasi dilakukan berdasarkan *confusion matrix*. Hasil yang diperoleh menyatakan bahwa *splitting* data dengan *test size* sebesar 0,2 memberikan tingkat akurasi sebesar 95% seperti tercantum pada Gambar 2.



**Gambar 2. Hasil Akurasi Model Klasifikasi K-NN**

Gambar 3 merepresentasikan hasil perhitungan *F1-score* dari masing-masing perbandingan. Pada grafik tersebut dinyatakan bahwa *F1-score* yang dihasilkan model klasifikasi dengan perbandingan 80% data latih dan 20% data uji paling tinggi yaitu sebesar 0.96.



**Gambar 3. F1-Score Model Klasifikasi K-NN**

## 5. Kesimpulan dan Saran

Berdasarkan data Twitter yang terkumpul terkait program MBKM diperoleh bahwa sentimen masyarakat bersifat netral. Hal ini menggambarkan bahwa masyarakat tidak sepenuhnya tertarik terhadap kebijakan program MBKM, sehingga perlu adanya evaluasi lebih lanjut sebagai dasar perbaikan guna keberlanjutan program dimasa mendatang. Sedangkan untuk akurasi yang dihasilkan oleh model klasifikasi menggunakan algoritma K-NN menyatakan bahwa dengan perbandingan 80% data latih dan 20% data uji untuk analisis sentimen tentang program MBKM memperoleh akurasi sebesar 95% dan nilai *F1-score* sebesar 0,96.

Saran untuk penelitian selanjutnya perlu menyusun *library* kosa kata singkatan yang lebih *up to date*. Hal ini bertujuan untuk meningkatkan akurasi dalam perhitungan polaritas. Sehingga hasil analisis sentimen dari data teks yang diperoleh dari Twitter lebih tepat.

## 6. Ucapan Terima Kasih

Penulis menghaturkan ucapan terimakasih kepada Universitas Indo Global Mandiri atas dukungan yang telah diberikan selama ini.

## Referensi

- [1] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 8, no. 4, Jul. 2018, doi: 10.1002/widm.1253.
- [2] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment texts based on BiLSTM," *IEEE Access*, vol. 7, pp. 51522–51532, 2019, doi: 10.1109/ACCESS.2019.2909919.
- [3] D. Alita, S. Priyanta, and N. Rokhman, "Analysis of Emoticon and Sarcasm Effect on Sentiment Analysis of Indonesian Language on Twitter", *Journal of Information Systems Engineering and Business Intelligence*, vol. 5, no. 2, pp. 100, Oct. 2019, doi: 10.20473/jisebi.5.2.100-109.
- [4] H. Saif, Y. He, and Alani, "Semantic Sentiment Analysis of Twitter International Semantic Web Conference, vol. 11, 2012. doi: 10.1007/978-3-642-35176-1.
- [5] M. A. Rizaty, "Pengguna Twitter di Indonesia Capai 18,45 Juta pada 2022" *DataIndonesia.id*, Aug. 10, 2022.
- [6] B. Gnana Priya, "Emoji Based Sentiment Analysis Using Knn B.Gnana Priya" 2019. [Online]. Available: <https://www.researchgate.net/publication/357047761>
- [7] S. Vijay Gaikwad, "Text Mining Methods and Techniques," *International Journal of Computer Applications*, vol. 85, no. 17, 2014. doi: 10.5120/14937-3507
- [8] R. Prabowo dan M. Thelwall, "Sentiment Analysis: A Combined Approach" *Journal of Informetrics*, vol. 3, pp 143-157, Apr. 2009
- [9] P. Gonçalves, M. Araújo, F. Benevenuto, and M. Cha, "Comparing and combining sentiment analysis methods" *COSN 2013 - Proceedings of the 2013 Conference on Online Social Networks*, Association for Computing Machinery, pp. 27–37, 2013 doi: 10.1145/2512938.2512951.
- [10] A. R. Lubis, S. Prayudani, M. Lubis, and O. Nugroho, "Sentiment Analysis on Online Learning during the Covid-19 Pandemic Based on Opinions on Twitter using KNN Method" *International Conference on Information System and Information Technology*, pp. 106–111, 2022, doi: 10.1109/ICISIT54091.2022.9872926.
- [11] F. M. J. M. Shamrat dkk., "Sentiment analysis on Twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 1, pp. 463–470, Jul. 2021, doi: 10.11591/ijeecs.v23.i1.pp463-470.
- [12] S. Rani and N. Singh Gill, "Hybrid Model For Twitter Data Sentiment Analysis Based On Ensemble Of Dictionary Based Classifier And Stacked Machine Learning Classifiers-SVM, KNN and C5.0" *Journal of Theoretical and Applied Information Technology*, vol. 29, no. 4, pp. 4, 2020, [Online]. Available: [www.jatit.org](http://www.jatit.org)

- [13] O. F.Y, A. J.E.T, A. O, H. J.O, O. O, and A. J, “*Supervised Machine Learning Algorithms: Classification and Comparison*” *International Journal of Computer Trends and Technology*, vol. 48, no. 3, pp. 128–138, Jun 2017, doi: 10.14445/22312803/IJCTT-V48P126.
- [14] B. Mahesh, “*Machine Learning Algorithms-A Review*” *International Journal of Science and Research*, 2018, doi: 10.21275/ART20203995.
- [15] H. Wisnu, M. Afif, and Y. Ruldevyani, “*Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naïve Bayes,*” *Journal of Physics: Conference Series*, Institute of Physics Publishing, Feb. 2020. doi: 10.1088/1742-6596/1444/1/012034.
- [16] V. I. Santoso, G. Virginia, and Y. Lukito, “*Penerapan Sentiment Analysis Pada Hasil Evaluasi Dosen Dengan Metode Support Vector Machine*” 2017.
- [17] C. Kaur dan A. Sharma, “*Social issues sentiment analysis using python*” *International Conference on Computing, Communication and Security*, pp.1-6, Oct. 2020. doi: 10.1109/ICCCS49678.2020.9277251.
- [18] M. Guia, R. R. Silva, and J. Bernardino, “*Comparison of Naive Bayes, support vector machine, decision trees and random forest on sentiment analysis,*” *International Joint Conference on Knowledge Discovery*, pp. 525–531, 2019, doi: 10.5220/0008364105250531.