

Implementation K-Nearest Neighbor Algorithm for Sentiment Analysis on Khilafah and Radicalism Issues in Indonesia

Penerapan Algoritma K-Nearest Neighbor untuk Analisis Sentimen Terhadap Isu Khilafah dan Radikalisme di Indonesia

**Legito¹, Nindi Permata Riau², Adi Nugroho Susanto Putro³,
Eri Mardiani⁴, Nofri Yudi Arifin⁵, Sepriano⁶, Moh. Erkamim⁷**

¹Program Studi Teknik Informatika,
Sekolah Tinggi Teknologi Sinar Husni Deliserdang Sumatera Utara, Indonesia

²Program Studi Sistem Informasi, Sains dan Teknologi,
UIN Sultan Syarif Kasim Riau, Indonesia

³STAB Negeri Raden Wijaya, Wonogiri, Jawa Tengah, Indonesia

⁴Program Studi Informatika, Fakultas FTKI, Universitas Nasional, Indonesia

⁵Program Studi Teknik Informatika, Universitas Ibnu Sina, Indonesia

⁶Program Studi Sistem Informasi, Universitas Islam Negeri Sulthan Thaha Saifuddin Jambi, Indonesia

⁷Program Studi Sistem Informasi Kota Cerdas, Universitas Tunas Pembangunan Surakarta, Indonesia

E-Mail: ¹legitostt@gmail.com, ²nindipermata.riau@gmail.com, ³adinug@radenwijaya.ac.id,
⁴erimardiani1@gmail.com, ⁵nofri.yudi@uis.ac.id, ⁶sepriano@uinjambi.ac.id, ²erkamim@lecture.utp.ac.id

Received Jul 2nd 2023; Revised Sept 14th 2023; Accepted Oct 28th 2023

Corresponding Author: Legito

Abstract

As technology advances, social media is an alternative for obtaining and disseminating information quickly. One of the social media that is currently used is twitter. There are many topics discussed, one of which is the Khilafah. The Khilafah is a political institution that cannot be isolated from political activity, which emerged after the death of the Prophet to be able to continue his leadership. Khilafah is usually associated with Radicalism. Opinions about the caliphate and radicalism have never stopped being discussed among the public, therefore sentiment analysis is needed to analyze public responses in Indonesia regarding the caliphate statement. This sentiment analysis uses the K-Nearest Neighbor Algorithm or K-NN. Based on the results that have been carried out, it shows that the K-NN algorithm obtains high accuracy results of 92.11% and 88.2% on each of the Khilafah and Radicalism keywords using 5000 data contained on twitter.

Keywords: Khilafah, K-Nearest Neighbor, Radicalism, Sentiment Analysis, Twitter

Abstrak

Seiring dengan majunya teknologi, media sosial merupakan salah satu alternatif untuk mendapatkan dan menyebarkan informasi dengan cepat. Salah satu media sosial yang saat ini digunakan yaitu *Twitter*. Terdapat banyak topik yang diperbincangkan salah satunya mengenai Khilafah. Khilafah merupakan suatu institusi politik yang tidak dapat diasingkan dari aktivitas politik, dimana munculnya setelah sepeninggal Rasulullah untuk dapat meneruskan kepemimpinannya. Khilafah biasanya dikaitkan dengan yang namanya Radikalisme. Opini tentang khilafah dan radikalisme tidak pernah berhenti diperbincangkan di kalangan masyarakat, oleh karena itu dibutuhkan analisa sentimen untuk menganalisa tanggapan masyarakat di Indonesia mengenai pernyataan khilafah tersebut. Analisa sentimen ini menggunakan Algoritma *K-Nearest Neighbor* atau K-NN. Berdasarkan hasil yang telah dilakukan menunjukkan bahwa Algoritma K-NN memperoleh hasil akurasi yang tinggi yaitu 92.11% dan 88,2% pada masing-masing kata kunci Khilafah dan Radikalisme dengan menggunakan 5000 data yang terdapat pada *twitter*.

Kata Kunci: Analisis Sentimen, Khilafah, K-Nearest Neighbor, Radikalisme, Twitter.

1. PENDAHULUAN

Seiring dengan majunya teknologi pada masa modern ini, media sosial menjadi alternatif baru untuk menebarkan opini salah satunya tentang khilafah [1]. Salah satu contoh media sosial yang aktif dan banyak

digunakan oleh berbagai kalangan adalah *twitter*. Dalam sosial media saat ini dinilai banyak memberikan informasi tentang pendapat postingan seseorang yang dibuktikan dengan adanya beberapa penelitian yang menggunakan data *twitter* sebagai bahan utama untuk menganalisis sentimen opini seseorang [2].

Khilafah atau negara islam merupakan institusi politik yang tidak dapat diasingkan dari aktivitas politik [3]. Istilah khilafah dalam terminologi politik muncul setelah sepeninggal Rasulullah karena umat Islam merasa perlu melanjutkan kepemimpinannya [4]. Konsep dasar khilafah telah ditetapkan oleh Allah SWT dalam Al-quran surat Al-Baqarah: 30 [5]. Khilafah yang ideal memiliki tiga unsur yaitu pemimpin yang dapat dipercaya, adanya dominasi syariat islam dan fuqaha atau ahli agama, dan yang ketiga yaitu adanya proses yang memungkinkan setiap orang memiliki peluang menjadi khilafah [6]. Biasanya khilafah akan bersinggungan dan berkaitan erat dengan yang namanya Radikalisme.

Penelitian ini seperti halnya peneliti-penelitian terdahulu akan melakukan analisis sentimen terhadap opini seseorang dari tweet yang mereka unggah pada laman *twitter* dengan memanfaatkan teknologi *text mining*. Analisis yang dimaksud adalah untuk mengetahui keberpihakan seseorang apakah teridentifikasi dari golongan menyetujui atau tidak setuju dengan konsep khilafah. *Sentiment analysis* merupakan salah satu bidang dari *Natural Language Processing* (NLP) yang membangun sistem untuk mengenali dan meringkas opini dalam bentuk teks [7]. Analisis Sentimen disebut juga *opinion mining*, yaitu suatu proses memahami, mengekstrak dan mengolah data teks untuk mendapat informasi [2][8].

Hasil akhir dari penelitian ini yang ingin dicapai adalah suatu opini dari masyarakat secara luas yang berhubungan dengan konsep khilafah dalam bentuk persepsi negatif, positif atau netral. Analisis sentimen berdasar klasifikasinya ada 2 yaitu dari dokumen klasifikasi ke pendapat atau fakta yang disebut klasifikasi subjektifitas dan dari dokumen ke klasifikasi ke dalam positif dan negatif [9]. Difusi besar dan peran era teknologi dalam masyarakat modern merupakan salah satu hal baru yang paling menarik dalam beberapa tahun terakhir, menarik minat para peneliti, jurnalis, perusahaan, dan pemerintah [10].

Algoritma *text mining* ini digunakan karena dalam pengerjaannya dapat mengubah setiap kata dan huruf dari dokumen menjadi data numerik. *Text mining* dipercaya menjadi salah satu Algoritma untuk seleksi fitur dokumen yang baik dan banyak digunakan oleh para peneliti. Salah satu peneliti yang melakukan analisis sentimen media sosial *twitter* menggunakan teknik *text mining* yaitu Eko Yulian yang melakukan analisis arsip dengan melakukan clustering pada dokumen menggunakan menggunakan K-means clustering [11]. Selain itu, terdapat penelitian yang dilakukan oleh Ozturk dan kawan-kawan yang menganalisis sentimen *twitter* terhadap krisis yang terjadi di suriah lewat postingan yang dilakukan masyarakat. Penelitian lain mengenai *text mining* juga dilakukan oleh Zannah yang menganalisis lokasi pengaruh ISIS melalui media sosial *twitter* menggunakan *text mining*. Hal ini membuktikan bahwa teknik *text mining* menjadi *tools* yang tepat untuk suatu kasus analisis sentimen [12].

Pada penelitian ini akan digunakan Algoritma *K-Nearest Neighbor* (K-NN) untuk melakukan klasifikasi data dokumen sekaligus mengetahui adanya postingan yang berkaitan dengan khilafah dan radikalisme. Algoritma K-NN merupakan sebuah algoritma yang digunakan untuk melakukan klasifikasi terhadap objek baru berdasarkan tetangga terdekatnya dan kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi [13][14]. Kelebihan dari K-NN adalah dapat digunakan dalam memecahkan permasalahan *multiclass* yang terjadi [15]. Namun, pada Algoritma K-NN memiliki masalah untuk menemukan tetangga terdekat pada titik *query* dari dataset yang digunakan [16].

Pada penelitian sebelumnya yang dilakukan Salam dan kawan-kawan tahun 2018 menghasilkan pengujian klasifikasi analisa sentimen dokumen komentar pada jasa ekspedisi barang (J&T) di *facebook* menggunakan algoritma K-NN yang memiliki *accuracy* tertinggi yaitu 79,21% dengan jumlah dokumen *training* 101 sedangkan *accuracy* terendahnya adalah 70,3% [17]. Penelitian lainnya yang dilakukan Wisnu dan kawan-kawan tahun 2019 tentang analisis sentimen dan penambangan opini dilakukan untuk melihat kepuasan masyarakat menuju layanan pembayaran digital di Indonesia (OVO, GO-PAY dan LinkAja) menggunakan data *twitter* dan menggunakan dua algoritma Naive bayes dan K-NN. Penelitian ini menemukan bahwa K-NN memiliki akurasi yang lebih baik daripada Naive Bayes [18].

Penelitian selanjutnya yang dilakukan oleh Widagdo dan kawan-kawan tahun 2023 tentang sistem *Monitool* oleh divisi digital marketing PT Lingkar9 Titian Media berfokus pada pemantauan dan analisis merek di media sosial dan portal berita. *Monitool* menggunakan metode K-NN untuk mengukur reputasi suatu merek. Setelah diproses, sistem akan mengeluarkan resume sentimen yang terdiri dari sentimen negatif, positif, dan netral. Hasil akurasi prediksi data pengujian saat ini adalah 81% [19].

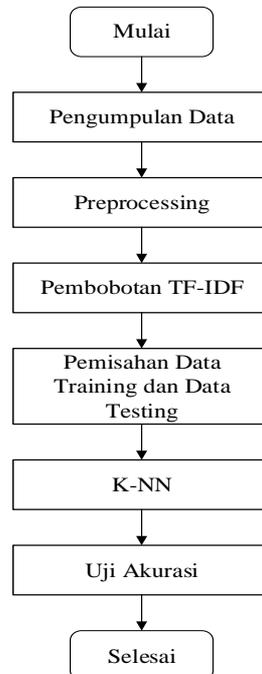
Berdasarkan penelitian yang membuktikan bahwa K-NN berhasil dalam melakukan klasifikasi dengan baik pada kasus *text mining*, maka pada penelitian ini akan dilakukan percobaan dengan menggunakan algoritma K-NN untuk memperoleh model dan akurasi klasifikasi yang baik dan optimum. Penelitian ini melakukan analisis sentimen terhadap tweet seseorang di *twitter* untuk mengetahui apakah opini yang seseorang unggah teridentifikasi dari golongan masyarakat yang menyetujui atau tidak mengenai sistem pemerintahan khilafah.

2. METODOLOGI PENELITIAN

Adapun metodologi yang diterapkan pada penelitian ini ditunjukkan pada gambar 1.

2.1 Pengumpulan Data

Tahap awal dalam melakukan proses analisis sentimen yaitu pengumpulan data. Data yang digunakan pada penelitian ini yaitu data tweet pada *twitter* dengan cara melakukan *crawling* data dengan *keyword* khilafah dengan rentang waktu dari 1 Februari 2022 - 30 Desember 2022 dengan menghasilkan 5000 data. Setelah data dikumpulkan tahap selanjutnya yaitu melakukan *preprocessing* data, pada tahapan ini terdiri dari beberapa proses yaitu *tokenization*, *cleaning*, dan *filtering*.



Gambar 1. Metodologi Penelitian

2.2 Preprocessing Data

Pada tahap *pre-processing* data, data tweet mentah terlebih dahulu dilakukan proses *case folding*, *tokenizing*, *stemming*, serta *filtering*. Hasil dari tahapan ini menghasilkan fitur yang digunakan sebagai data pembelajaran mesin oleh K-NN. Dalam tahapan *preprocessing*, terdiri dari proses *case folding*, *tokenizing*, *stemming* dan *filtering*. Pada tahap *preprocessing* data merupakan suatu proses untuk melakukan persiapan pada data mentah sebelum dilakukannya analisis data. pada tahapan ini *preprocessing* menggunakan *tools jupyter notebook* dengan bahasa pemrograman *python* untuk memperoleh data yang akan diproses. Pada tahapan *preprocessing* terdiri dari beberapa proses *cleaning*, *filtering* dan *tokenizing*.

2.3 Pembobotan Kata

Pembobotan kata adalah sebuah mekanisme yang memberikan skor terhadap frekuensi kemunculan sebuah kata dalam dokumen teks. Salah satu Algoritma dalam pembobotan kata yaitu *Term Frequency-inverse document frequency* (TF-IDF). Secara lengkap dapat ditunjukkan pada tabel 1.

Tabel 1. Hasil TF-IDF

No	Adab	Aqidah	Alam	Budaya	Berkah	...	Rasis
1	0,000	0,000	0,000	0,000	0,337	...	0,000
2	0,000	0,794	0,000	0,000	0,000	...	0,000
3	0,477	0,000	0,611	0,000	0,000	...	0,000
4	0,000	0,000	0,000	0,000	0,000	...	0,538
...
5000	0,000	0,000	0,000	0,000	0,000	...	0,000

Dataset ini dihimpun dengan menggunakan bahasa pemrograman *python* pada *google colab*, pengambilan data dengan proses *crawling* dengan memasukkan kata kunci *#khilafah* dan *#radikalisme* yang trending sejak tahun 2014 lalu hingga sekarang.

2.4 Text Mining

Text mining dapat didefinisikan sebagai suatu proses menggali informasi dimana seorang user berinteraksi dengan sekumpulan dokumen menggunakan *tools*. Tujuan dari *text mining* adalah agar mendapatkan informasi yang berguna dari beberapa dokumen [20]. Teknik yang digunakan dalam penyelesaian *Text mining* yaitu klasifikasi, *clustering*, *information extraction* dan *information retrieval* [20].

2.5 Analisa Sentimen

Sentiment analysis merupakan salah satu bidang dari *Natural Language Processing* (NLP) yang membangun sistem untuk mengenali dan mengekstraksi opini dalam bentuk teks [21]. Dengan bantuan *sentiment analysis*, informasi yang tadinya tidak terstruktur dapat diubah menjadi data yang lebih terstruktur [22]. Analisis sentimen berdasar klasifikasinya ada 2 yaitu dari dokumen klasifikasi ke pendapat atau fakta yang disebut klasifikasi subjektifitas dan dari dokumen ke klasifikasi ke dalam positif dan negatif [8]. Saat ini, analisis sentimen telah mendapatkan nilai lebih dengan munculnya jejaring sosial [7].

2.6 K-Nearest Neighbor

Algoritma K-NN merupakan sebuah algoritma untuk melakukan klasifikasi terhadap objek baru berdasarkan tetangga terdekatnya dan kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi [23][14]. Kelebihan dari K-NN adalah dapat digunakan untuk memecahkan permasalahan *multiclass*[15]. Namun, K-NN memiliki masalah untuk menemukan tetangga terdekat pada titik *query* dari dataset yang digunakan [16]. Menurut Gorunescu, Algoritma *Nearest Neighbor* (kadang disebut K-nearest Neighbor atau K-NN) merupakan Algoritma yang melakukan klasifikasi berdasarkan kedekatan lokasi (jarak) suatu data dengan data yang lain [24]. Dekat atau jauhnya lokasi (jarak) biasanya dihitung berdasarkan jarak Euclidean dengan rumus sebagai berikut:

$$d(x_i, x_j) = \sqrt{\sum_i^N + 1 (\text{diff}(X_{i1}X_{j1}))^2} \quad (1)$$

dimana:

- x_il = data testing ke-i pada variabel ke-l
- x_jl = data training ke-i pada variabel ke-l
- d(x_i,x_j) = jarak N = dimensi data variabel bebas
- diff(x_il,x_jl) = difference atau ketidaksamaan.

K-NN merupakan salah satu Algoritma yang populer untuk menghasilkan klasifikasi teks [25] yaitu dengan melakukan proses pembelajaran dari data latih untuk menentukan kelompok k objek. Sehingga dalam menentukan hasil klasifikasi K-NN melihat jarak terdekat dari objek dengan masing-masing kelompok. Jarak tersebut diperoleh dari hasil kedekatan antara data masukan dengan data yang berada dalam kelompok berdasarkan nilai sejumlah fitur yang ada [26].

2.7 Khilafah dan Radikalisme

Wakil Presiden RI, Ma'ruf Amin menyatakan bahwa khilafah adalah isu yang berpotensi merusak tatanan bernegara dan menimbulkan radikalisme. Eksistensi khilafah merupakan realita dalam sejarah peradaban Islam dan wacana khilafah yang awalnya merupakan ekspresi keagamaan sebagian kelompok umat Islam, kemudian diidentikkan sebagai ancaman eksistensial yang membahayakan negara [27]. Perkembangan wacana mengenai khilafah digambarkan sebagai perang mindset dari kelompok Radikalisme menjadi permasalahan yang serius di Indonesia [28].

3. HASIL DAN PEMBAHASAN

Hasil dari penelitian ini dapat dibagi menjadi tiga bagian. Yaitu hasil pengujian akurasi Algoritma Naïve Bayes Classifier, hasil persebaran tweet, hasil analisis sentimen terhadap khilafah, dan terakhir adalah perbandingan dengan data aktual.

3.1 Pelabelan

Pelabelan merupakan salah satu proses *text mining*, dimana pelabelan ini digunakan untuk mengkategorikan kalimat termasuk kategori positif, negatif, atau netral terhadap khilafah. Proses ini dilakukan setelah proses crawling data melalui aplikasi anaconda dengan menentukan 3 kategori yaitu:

1. Positif: merupakan tweet yang berisi penolakan pembentukan ideologi khilafah dan tetap mendukung Pancasila.
2. Negatif: merupakan tweet yang berisi tentang pendudukan antas ideologi baru khilafah da menolak pancasila
3. Netral: merupakan tweet yang berisi tentang tidak berada dikeduanya.

Tabel 2 merupakan pelabelan data dari proses *preprocessing* dengan menggunakan google colab, sebagai bahan acuan penelitian ini menggunakan pakar Bahasa Indonesia untuk memvalidasi dan menentukan sentimen positif, negatif dan netral.

Tabel 2. Preprocessing Data

No	Stemming	Count	Sentimen
1	Kebanyakan orang menganggap bahawa ini semua jahat kan kwwkkw padahal coba belajar agama lagi deh	2	Positif
2	Banyak bacot nampaknya kirain penghamba siapa itu ulu ulu makanya bos jangan belajar agama di amerika	-8	Negatif
3	Salang arah dan keblinger, masak pancasila dibandingkan sama quran, padahal orang-orang yang satu kolam itu	-5	Negatif
4	Kami menolak khilafah dan radikalisme di indonesia	0	Netral
...
5000	Islam tidak mengajarkan kejahatan bang, kalau pun jahat itu oknum jangan dicampur aduk ye	4	Positif

Setelah melakukan pelabelan data maka diperoleh data positif sebanyak 1.354 *tweets*, negatif 1.546 *tweets* dan netral 2.106 *tweets*.

3.2 Pengujian K-Nearest Neighbor

Setelah melakukan pelabelan, maka dilakukan pengujian menggunakan K-NN dengan google colab. Pengujian dilakukan dengan cara membagi *data training* menjadi 70% dan *data testing* sebanyak 30% demikian juga dengan beberapa percobaan yang lain yaitu 80% : 20% dan 90% : 10%. Dari hasil pengujian didapatkan akurasi sebesar 92.11% untuk 70% : 30%, selanjutnya dapat dilihat pada tabel 3.

Tabel 3. Akurasi Percobaan dengan Skema 1 (Khilafah)

Percobaan	Akurasi	Error	Waktu Proses (s)
90-10	0,8702	0,0321	877
80-20	0,8324	0,0472	893
70-30	0,9211	0,0564	903

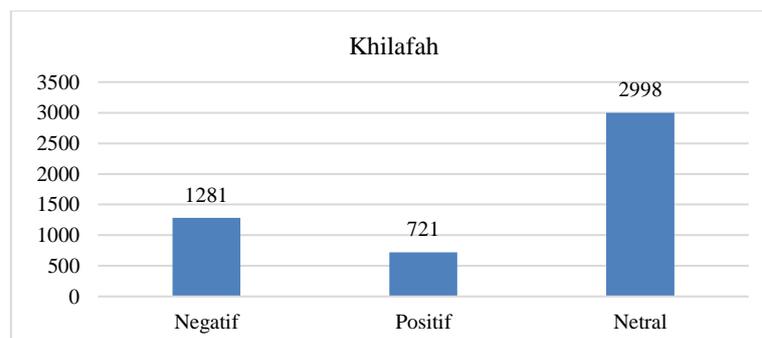
Tabel 4. Akurasi Percobaan dengan Skema 2 (Radikalisme)

Percobaan	Akurasi	Error	Waktu Proses (s)
90-10	0,8802	0,0725	886
80-20	0,8451	0,0639	921
70-30	0,8823	0,0681	897

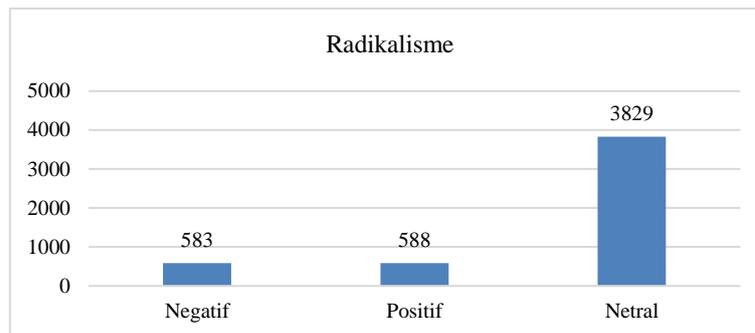
Dari tabel 3 dan 4 dapat disimpulkan bahwa akurasi KNN berada pada percobaan 70% dan 30% untuk data training dan testing. Hal ini tidak membuktikan bahwa semakin sedikit data training akan semakin lebih baik dalam proses percobaan. Tabel tersebut juga menyatakan bahwa waktu proses dan error dalam setiap percobaan tidak selalu konstan juga. Dalam penelitian lanjutan diharapkan untuk dapat mencobakan dengan beberapa algoritma klasifikasi yang sejenis untuk dapat membuktikan tingkat konstan jika menggunakan data text.

3.3 Analisis Sentimen terhadap Khilafah dan Radikalisme

Dari percobaan menggunakan keyword Khilafah dengan Nilai K terbaik dan dengan data testing 30% atau data training 70%, diperoleh nilai sentimen masyarakat yang ditunjukkan pada gambar 2.



Gambar 2. Hasil Sentimen Masyarakat terhadap Keyword Khilafah dengan K Terbaik



Gambar 3. Hasil Sentimen Masyarakat terhadap Keyword Radikalisme dengan K Terbaik

Gambar 2 dan 3 menunjukkan bahwa respon dari masyarakat secara nasional menyatakan bahwa kedua kata kunci yang digunakan dalam riset berbasis sentimen memiliki tingkat keragu-raguan yang tinggi atau sering dikatakan netral ataupun tidak memiliki pengaruh apapun terhadap objek yang dibahas. Dari gambar tersebut netral hampir 70% dalam menanggapi topik yang berkaitan dengan Kilafah dan Radikalisme, selebihnya adalah 18% dan 13% adalah Negatif dan Positif.

4. KESIMPULAN

Penelitian ini menggunakan dataset sejumlah 5000 data hasil *crawling twitter* dengan menggunakan 2 kata kunci khilafah dan radikalisme. Dari serangkaian percobaan sentimen analisis menunjukkan bahwa hampir 70% masyarakat Indonesia menanggapi hanya dengan netral terkait persoalan khilafah dan radikalisme di Indonesia. Secara umum, hanya sekitar 18 dan 13 yang menanggapi selain netral. Percobaan dilakukan dengan menggunakan algoritma K-NN dibagi dalam 3 bagian percobaan pada teknik pembagian data dan dengan jumlah K sebanyak 10 percobaan. Hasil penelitian tersebut memperoleh akurasi yaitu 92,11% dan 88,23% untuk kedua kata kunci yang digunakan. Dalam penelitian tersebut juga membandingkan error dan waktu proses algoritma, dengan hasil terbaik minimum *error* 3,2% dengan waktu proses tercepat adalah 877 detik dalam setiap proses.

REFESENSI

- [1] A. B. Puandina, S. A. Aryani, and I. I. Studies, "Konseling Lintas Agama dan Budaya dalam Fenomena Hijrah dan Perilaku Beragama Milenial di Media Sosial," vol. 9, pp. 376–385, 2023.
- [2] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowl. Inf. Syst.*, vol. 60, no. 2, pp. 617–663, 2019, doi: 10.1007/s10115-018-1236-4.
- [3] J. Subando and M. K. B. Wibowo, "Construction of the Khilafah State According to Taqiyuddin An-Nabhani: Foundation for Development of Perception Measurement Instruments about the Khilafah," *Ishlah J. Ilmu Ushuluddin, Adab dan Dakwah*, vol. 5, no. 1, pp. 55–72, 2023, doi: 10.32939/ishlah.v5i1.240.
- [4] S. Sahri, "Political thought of Al-Ghazali on Imamah: Debate between theocracy and democracy," *HTS Teol. Stud. / Theol. Stud.*, vol. 77, no. 3, pp. 1–8, 2021, doi: 10.4102/hts.v77i3.6338.
- [5] Moh. Shofan, "Konsep Khalifah Fi Al-Ardh dalam Surat Al-Baqarah Ayat 30 dan Implikasinya pada Tujuan Pendidikan Islam," vol. 2, no. 1, pp. 141–163, 2020.
- [6] E. R. Dewi, C. Hidayatullah, D. Oktaviantari, and M. Y. Raini, "Konsep Kepemimpinan Profetik," *Al-Muaddib J. Ilmu-Ilmu Sos. dan Keislaman*, vol. 5, no. 1, pp. 147–159, 2020, [Online]. Available: <http://dx.doi.org/10.31604/muaddib.v5i1.147-159>.
- [7] A. Rajput, *Natural language processing, sentiment analysis, and clinical analytics*. Elsevier Inc., 2019.
- [8] Mustakim and R. Novita, "The Implementation of Probabilistic Neural Networks to Sentiment Analysis of National Principle and Religion Issues in Indonesia," *J. Syst. Manag. Sci.*, vol. 13, no. 5, pp. 311–321, 2023, doi: 10.33168/JSMS.2023.0520.
- [9] S. M. Nagarajan and U. D. Gandhi, "Classifying streaming of Twitter data based on sentiment analysis using hybridization," *Neural Comput. Appl.*, vol. 31, no. 5, pp. 1425–1433, 2019, doi: 10.1007/s00521-018-3476-3.
- [10] S. Sun, Y. Zhai, B. Shen, and Y. Chen, "Newspaper coverage of artificial intelligence: A perspective of emerging technologies," *Telemat. Informatics*, vol. 53, no. May, 2020, doi: 10.1016/j.tele.2020.101433.
- [11] E. Yulian, "Text Mining dengan K-Means Clustering pada Tema LGBT dalam Arsip Tweet Masyarakat Kota Bandung," *J. Mat. "MANTIK"*, vol. 4, no. 1, pp. 53–58, 2018, doi: 10.15642/mantik.2018.4.1.53-58.
- [12] R. Zannah, "Analisis Sentimen Pada Media Sosial Twitter Untuk Klasifikasi Opini Islam Radikal Menggunakan Jaringan Saraf Tiruan," no. September 2017, pp. 46–54, 2019, [Online]. Available:

- <http://digilib.uinsby.ac.id/32982/>.
- [13] R. Andrian, M. A. Naufal, B. Hermanto, A. Junaidi, and F. R. Lumbanraja, "K-Nearest Neighbor (k-NN) Classification for Recognition of the Batik Lampung Motifs," *J. Phys. Conf. Ser.*, vol. 1338, no. 1, 2019, doi: 10.1088/1742-6596/1338/1/012061.
- [14] S. G. Setyorini and Mustakim, "Application of the nearest neighbor algorithm for classification of online taxibike sentiments in indonesia in the google playstore application," *J. Phys. Conf. Ser.*, vol. 2049, no. 1, pp. 0–7, 2021, doi: 10.1088/1742-6596/2049/1/012026.
- [15] M. Tanveer, A. Sharma, and P. N. Suganthan, "Least squares KNN-based weighted multiclass twin SVM," *Neurocomputing*, vol. 459, no. xxxx, pp. 454–464, 2021, doi: 10.1016/j.neucom.2020.02.132.
- [16] I. Triguero, D. García-Gil, J. Mailló, J. Luengo, S. García, and F. Herrera, "Transforming big data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 9, no. 2, pp. 1–24, 2019, doi: 10.1002/widm.1289.
- [17] A. Salam, J. Zeniarja, and R. S. U. Khasanah, "Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekpress Indonesia)," *Pros. SINTAK*, pp. 480–486, 2018.
- [18] 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," *J. Phys. Conf. Ser.*, vol. 1444, no. 1, 2020, doi: 10.1088/1742-6596/1444/1/012034.
- [19] J. Penelitian Ilmu Komputer, B. Wisnu Widagdo, and M. Adhari Adiguna, "Brand Reputation Monitoring System Based on Sentiment Analysis Using the K-Nearest Neighbor Method," vol. 1, no. 1, pp. 22–27, 2023, [Online]. Available: <https://doi.org/10.5281/zenodo.7720440https://mypublikasi.com/>.
- [20] S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, "Using text mining techniques for extracting information from research articles," *Stud. Comput. Intell.*, vol. 740, pp. 373–397, 2018, doi: 10.1007/978-3-319-67056-0_18.
- [21] G. Widarma, R. Novita, Mustakim, and N. E. Rozanda, "Comparing Classification Algorithms to Analyze Twitter Sentiment on Public Opinion on Fuel Oil," *2023 Int. Semin. Intell. Technol. Its Appl. Leveraging Intell. Syst. to Achieve Sustain. Dev. Goals, ISITIA 2023 - Proceeding*, pp. 648–653, 2023, doi: 10.1109/ISITIA59021.2023.10221118.
- [22] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Soc. Netw. Anal. Min.*, vol. 11, no. 1, 2021, doi: 10.1007/s13278-021-00776-6.
- [23] F. Hemmatian and M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," *Artif. Intell. Rev.*, vol. 52, no. 3, pp. 1495–1545, 2019, doi: 10.1007/s10462-017-9599-6.
- [24] D. Kurniawan and A. Saputra, "Penerapan K-Nearest Neighbour dalam Penerimaan Peserta Didik dengan Sistem Zonasi," *J. Sist. Inf. Bisnis*, vol. 9, no. 2, p. 212, 2019, doi: 10.21456/vol9iss2pp212-219.
- [25] K. Shah, H. Patel, D. Sanghvi, and M. Shah, "A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification," *Augment. Hum. Res.*, vol. 5, no. 1, 2020, doi: 10.1007/s41133-020-00032-0.
- [26] U. R. Gurning, Mustakim, S. T. Rizaldi, and H. Syukron, "Comparison of Naïve Bayes, C4.5 and K-Nearest Neighbor for Covid-19 Data Classification," *ISITDI*, pp. 16–21, 2022.
- [27] P. Wibowo, "Sekuritisasi Wacana Khilafah Di Indonesia," *J. Islam. Stud. Int. ...*, vol. Vol. 5, no. 1, pp. 25–49, 2020, doi: 10.5281/zenodo.4052347.
- [28] S. Rohman and S. Nurhasanah, "Paham Radikalisme Berdasarkan Perspektif Agama," *J. Terror. Stud.*, vol. 1, no. 1, pp. 18–32, 2019.