



Analisis Sentimen Masyarakat Kota Medan Terhadap Persiapan PON XXI Sumut-Aceh 2024 Menggunakan Metode Naive Bayes

Muhammad Suhery^{1*}, Gema Ramadhan², Abdul Halim Hasugian³

^{1,2} Universitas Islam Negeri Sumatera Utara Medan, Indonesia

suhery2001@gmail.com^{1*}, gema.ramadhan02411@gmail.com², abdulhalimhasugian@uinsu.ac.id³

Alamat: Jl. William Iskandar Ps. V, Medan Estate, Kec. Percut Sei Tuan, Kabupaten Deli Serdang, Sumatera Utara 20371

Korespodensi email: suhery2001@gmail.com

Abstract. *The North Sumatra Aceh National Sports Week (PON) which will be held in September 2024 in Papu, North Sumatra has drawn many pros and cons from the public. This topic allows the public to provide criticism, suggestions and opinions regarding the 2024 North Sumatra Aceh PON. Instagram is a popular social media for conveying public opinion. The sentiment analysis process can find and resolve problems based on public opinion on social media such as Instagram. The classification method used in this research is the Naïve Bayes Classifier. Datasets can be obtained from the data crawling process using the Google Chrome extension: IGCommentExport. The data is labeled positive, neutral, or negative. The results of the labeling process showed 770 negative data, 256 neutral data and 920 positive data. Then pre-processing is carried out on the data that has been previously labeled, and a word weighting process is also carried out using TF-IDF. After that, modeling was carried out using the Naïve Bayes Classifier and the final process was evaluation-testing. The high accuracy results from the fourth experiment which compared 90% of the training data with 10% of the testing data resulted in an accuracy of 75%. Meanwhile, the sentiment test results show that positive sentiment is more numerous than negative sentiment and neutral sentiment.*

Keywords: PON 2024, Sentiment analysis, Naïve Bayes

Abstrak. Pekan Olahraga Nasional (PON) Aceh Sumut yang digelar pada September 2024 di PapuSumatera Utara menuai banyak pro dan kontra dari masyarakat. Topik ini memungkinkan masyarakat untuk memberikan kritik, saran, dan pendapat mengenai PON Aceh Sumut 2024. Instagram merupakan salah satu media sosial yang populer untuk menyampaikan opini masyarakat. Proses analisis sentimen dapat menemukan dan menyelesaikan permasalahan berdasarkan opini masyarakat di media sosial seperti Instagram. Metode klasifikasi yang digunakan dalam penelitian ini adalah Naïve Bayes Classifier. Dataset dapat diperoleh dari proses crawling data menggunakan ekstensi google chrome: IGCommentExport. Data tersebut diberi label positif, netral, atau negatif. Hasil proses pelabelan menunjukkan 770 data negatif, 256 data netral, dan 920 data positif. Kemudian dilakukan pre-processing terhadap data yang telah diberi label sebelumnya, juga dilakukan proses pembobotan kata menggunakan TF-IDF. Setelah itu dilakukan pemodelan dengan menggunakan Naïve Bayes Classifier dan proses terakhir adalah evaluasi-pengujian. Hasil akurasi yang tinggi dari percobaan keempat yang membandingkan 90% data pelatihan dengan 10% data pengujian menghasilkan akurasi 75%. Sedangkan hasil uji sentimen menunjukkan bahwa sentimen positif lebih banyak dibandingkan sentimen negatif dan sentimen netral.

Kata kunci: PON 2024 , Analisis sentiment , Naïve bayes

1. LATAR BELAKANG

Naskah Naskah Menurut Pasal 1 ayat (4) Undang-Undang Nomor 3 Tahun 2005, “Olahraga adalah kegiatan yang melatih tubuh manusia tidak hanya jasmani tetapi juga mental.” Menurut ayat 4 Undang-Undang Nomor 3 Tahun 2005, tujuan olahraga nasional adalah “Tujuan olahraga nasional adalah untuk memelihara dan meningkatkan kesehatan dan kebugaran, prestasi, mutu hidup manusia, nilai moral dan akhlak mulia, sportivitas, mendisiplinkan, memantapkan dan memupuk persatuan dan kesatuan bangsa, memperkuat ketahanan nasional serta mengangkat harkat dan martabat bangsa(Nurfaizin

& Anam, 2022).

Sistem Undang-Undang Keolahragaan Nasional tahun 2005 Nomor. 3 menyebutkan bahwa olahraga dibagi atas tiga bagian : (1). Olahraga Prestasi, (2). Olahraga Pendidikan, (3). Olahraga Rekreasi. Ketiga jenis olahraga tersebut tidak dapat dipisahkan satu dengan yang lainnya, dan saling ketergantungan mempunyai hubungan yang sinergi jika ingin mencapai prestasi olahraga. Namun sekedar untuk menjaga dan memelihara kebugaran memiliki program yang tidak seberat seperti atlet yang ingin mencapai prestasi puncak. Kenyataannya (Makorohim & Apriani, 2017). Olahraga merupakan bagian dari aktivitas sehari-hari manusia sehingga membentuk jasmani dan rohani yang sehat dan kuat (Pitajaya et al., 2023), olahraga tidak lagi hanya diperuntukkan bagi sebagian ahli di bidangnya saja, namun untuk semua orang, dan sudah menjadi bagian dari keseharian dan gaya hidup masyarakat. Indonesia perlu menyadari bahwa olahraga itu penting, memperkuat upayanya, dan fokus pada peningkatan pendidikan secara keseluruhan (Putra & Kafrawi, 2023). Pada hakekatnya kesegaran jasmani berhubungan dengan komponen-komponen tubuh manusia dan merupakan dasar kemampuan seseorang untuk melakukan kegiatan sehari-hari dengan baik (Apriani et al., 2021). *pointers*.

Pekan Olahraga Nasional 2024 (biasa disingkat PON XXI/PON ACEH-SUMUT 2024) adalah penyelenggaraan ke-21 dari Pekan Olahraga Nasional (PON) yang merupakan ajang multi-olahraga nasional utama yang dijadwalkan berlangsung dari 8 hingga 20 September 2024 dengan Aceh dan Sumatera Utara sebagai tuan rumah. [1] Edisi ini menandai pertama kalinya PON diselenggarakan oleh dua provinsi sekaligus, juga menandai kedua kalinya diadakan di Sumatera Utara (sebelumnya di tahun 1953), dan pertama kalinya diadakan di Aceh. PON ini merupakan keempat kalinya diadakan di Sumatera (sebelumnya Medan 1953, Palembang 2004, Pekanbaru 2012) dan pertama kali sejak 2012. Provinsi termuda, Papua Barat Daya, Papua Pegunungan, Papua Tengah dan Papua Selatan akan memulai debutnya di ajang PON XXI ini.

Dengan adanya hal tersebut, masyarakat Indonesia mengungkapkan pendapatnya melalui media sosial, dan popularitas media sosial tidak dapat disangkal karena sebagai titik fokus penyebaran opini dan informasi. Namun seiring berjalannya waktu, masyarakat telah memanfaatkan media sosial untuk menyebarkan opini dan informasi. Meskipun demikian, penggunaan media sosial cenderung mengganggu lingkungan sekitarnya dan kebebasan berekspresi dilakukan tanpa memperhatikan moralitas dan peraturan yang berlaku di masyarakat. Instagram merupakan salah satu platform media

sosial yang banyak digunakan masyarakat saat ini. Instagram adalah aplikasi berbagi foto yang memungkinkan pengguna mengunggah foto secara langsung dari ponsel mereka. Instagram juga memiliki fitur komentar postingan yang menjadi kelebihan instagram. Karena komentar instagram dibatasi hingga 2.200 karakter, sedangkan twitter hanya 280 karakter. Sebanyak 36.4% pengguna instagram di Indonesia berusia antara 18 sampai dengan 24 tahun. Angka-angka tersebut menunjukkan bahwa instagram memiliki peran yang penting di Indonesia, sehingga menyebabkan meningkatnya jumlah instansi pemerintahan yang tertarik memanfaatkan instagram untuk berkomunikasi dengan publik. Hal ini karena daya tarik media sosial yang meluas, penggunaan yang relatif sederhana, kemampuan untuk menjangkau khalayak luas, biaya rendah. Lantaran semakin banyaknya pengguna media sosial yang memanfaatkan instagram untuk mengungkapkan pendapatnya, diperlukan tindakan berupa analisis sentimen untuk memecahkan masalah berdasarkan opini publik saat ini. Analisis sentimen dapat digunakan sebagai alat untuk menilai pandangan publik terhadap peristiwa tertentu, sehingga dapat digunakan sebagai bahan untuk mengevaluasi peristiwa di masa mendatang. Analisis sentimen memeriksa polaritas teks dalam komentar untuk menentukan bahwa pendapat dalam komentar tersebut positif, netral, atau negatif. Metode klasifikasi yang digunakan yaitu Naïve Bayes Classifier. Metode Naïve Bayes Classifier merupakan algoritma yang populer untuk menentukan nilai probabilitas terbaik untuk mengkategorikan data uji dalam suatu kelompok kelas. Metode klasifikasi menggunakan metode probabilistik dan statistic untuk memperkirakan probabilitas berdasarkan peristiwa sebelumnya (Teorema Bayes), dengan asumsi yang kuat (naif) antara setiap keadaan atau peristiwa. Penelitian ini berfokus pada sentimen komentar masyarakat penyelenggaraan Pekan Olahraga Nasional 2024 Aceh-Sumut pada instagram yang mengandung maksud positif, netral atau negatif. Uji evaluasi juga dilaksanakan untuk menguji model klasifikasi yang telah dihasilkan. Penelitian ini diharapkan dapat memberikan kontribusi bagi masyarakat untuk mengetahui komentar tentang Pekan Olahraga Nasional 2024 yang bersifat netral, negatif atau positif, serta membantu pemerintah sebagai bahan evaluasi dan untuk menentukan langkah-langkah atau kebijakan Pekan Olahraga Nasional selanjutnya.

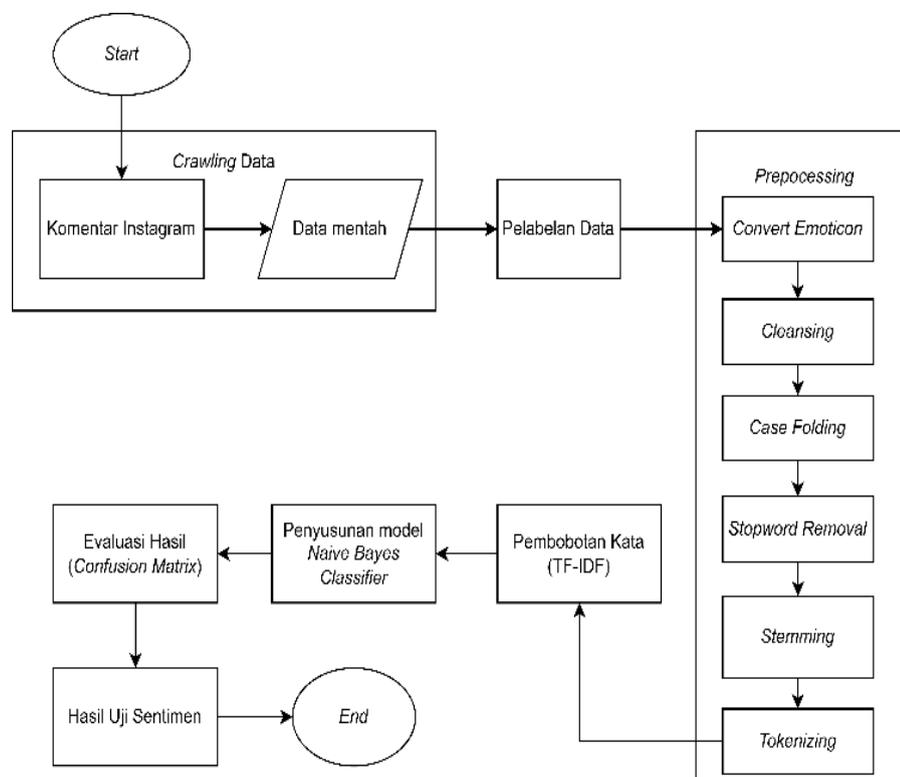
2. METODE PENELITIAN

Bagian Metode Pengumpulan Data

Metode pengumpulan data pada penelitian ini dilakukan dengan mengambil data komentar dari instagram menggunakan IGCommentExport yang merupakan *extensions google chrome*. Pengumpulan data dikumpulkan dari postingan instagram yang dibuat oleh @ponxxAceh-sumut2024 antara tanggal 8 hingga 20 September 2024. Jumlah data yang diambil sebanyak 2.000 data komentar.

Tahapan Penelitian

Tahapan penelitian merupakan urutan langkah-langkah penelitian. Urutan tersebut dilakukan mulai dari pengumpulan data, pengolahan data, hingga hasil akhir penelitian. Alur tahapan penelitian dapat dilihat pada gambar 1:



Gambar 1. Tahapan Penelitian

Dari flowchart gambar 1 menunjukkan bahwa penelitian ini diawali dengan proses *crawling* data di instagram menggunakan IGCommentExport yang merupakan *extensions google chrome*. Dataset tersebut selanjutnya diberi label positif, negatif atau netral oleh ahli bahasa. Selanjutnya data diolah pada tahap *text preprocessing* yang terdiri dari *convert emoticon, cleansing, case folding, stopwords removal, stemming, dan tokenizing*.

Fungsi *text preprocessing* untuk mengurangi *noise* pada dataset. Tahap selanjutnya adalah pembobotan kata dengan memanfaatkan algoritma *Term Frequency – Inverse Document Frequency* (TF-IDF).

3. HASIL DAN PEMBAHASAN

Bagian Hasil Dataset

Dataset untuk penelitian ini berasal dari proses *crawling* data di *instagram*. Data komentar *instagram* dapat diambil menggunakan *extension Google Chrome IGCommentExport*. Akun *google chrome* yang telah terpasang *extension* tersebut dapat melakukan *crawling* dengan menginputkan link postingan *instagram* yang hendak di *crawling*. Dataset yang diperoleh diberi label secara manual oleh Sarjana Pendidikan Bahasa dan Sastra Indonesia. Berdasarkan hasil pelabelan didapatkan 770 data negatif, 255 data netral dan 920 data positif. Sehingga jumlah keseluruhan dataset pada penelitian ini sebanyak 2000 data. Dataset tersebut berisi data *training* dan data *testing*. Pada tahap klasifikasi, data *training* digunakan untuk membuat bahan pembelajaran sehingga hasil klasifikasi dapat diterapkan pada data *testing*. Dataset untuk penelitian ini dibagi menjadi data latih dan data uji dengan perbandingan sebagai berikut: 60%:40%, 70%:30%, 80%:20%, dan 90%:10%. Ahli bahasa telah melabeli contoh data pada tabel 1 sebagai berikut:

Tabel 1. Contoh Pelabelan Dataset

Komentar	Sentimen
Kalau udah berurusan sama pegawai pemerintahan gawatlah,sampai hari ini aja honor pelatih green volouner belum dibayar.	Negatif
@disiniodhe pon Aceh-Sumut kan boleh dihadiri penonton dengan kapasitas 2000 orang	Netral
Keren banget PON ACEH-SUMUT !!!	Positif

a. Hasil Pre-Processing

Pada tahap *preprocessing* dataset diawali dengan tahap *convert emoticon* yaitu tahap mengubah menjadi *string* yang bersesuaian. Hasil proses *convert emoticon* dapat dilihat pada tabel 2:

Tabel 2. Hasil *Convert Emoticon*

Komentar	Hasil <i>Convert Emoticon</i>
Tolong ya kalau udah berurusan sama pegawai pemerintahan gawatlah,sampai hari ini aja honor pelatih green volounter belum dibayar.	Tolong ya kalau udah berurusan sama pegawai pemerintahan gawatlah,sampai hari ini aja honor pelatih green volounter belum dibayar.
@disiniodhe pon Aceh-Sumut kan boleh dihadiri penonton dengan kapasitas 2000 orang	@disiniodhe pon Aceh-Sumut kan boleh dihadiri penonton dengan kapasitas 2000 orang
Keren banget PON ACEH-SUMUT !!!	Keren banget PON ACEH-SUMUT !!!

Proses mengonversi data dari huruf besar ke huruf kecil (*lowercase*) dikenal sebagai

b. *case folding*.

Hasil proses *case folding* dapat dilihat pada tabel 3:

Tabel 3. Hasil *Case Folding*

Komentar	Hasil <i>Case Folding</i>
Kalau udah berurusan sama pegawai pemerintahan gawatlah,sampai hari ini aja honor pelatih green volounter belum dibayar.	Kalau udah berurusan sama pegawai pemerintahan gawatlah,sampai hari ini aja honor pelatih green volounter belum dibayar.
Pon Aceh-Sumut kan bolehdihadiri penonton dengan kapasitas 2000 orang	pon Aceh-Sumut kan boleh dihadiri penonton dengan kapasitas 2000 orang
Keren banget PON ACEH-SUMUT !!!	Keren banget PON ACEH-SUMUT !!!

Proses *stopword removal* kemudian digunakan untuk menghapus kata dari klasifikasi yang tidak memiliki arti yang penting. Contoh list kata *stopword removal* adalah ‘haha’, ‘kah’, ‘wah’, ‘wkwk’, dan seterusnya. Hasil proses *stopword removal* dapat dilihat pada tabel 4:

Tabel 4. Hasil *Stopword Removal*

Komentar	Hasil <i>Stopword Removal</i>
tolong ya panitia evaluasi buat kedepannya wasitnya berlisensi lah kayak apa ya bukan gondok sama pemainnya malah jadi gondok ke wasitnya	tolong panitia evaluasi buat kedepannya wasitnya berlisensi kayak apa bukan gondok sama pemainnya malah jadi gondok wasitnya
Keren banget PON ACEH-SUMUT !!!	Keren banget PON ACEH-SUMUT !!!

Stemming adalah penghapusan imbuhan, seperti awalan dan akhiran untuk pengambilan kata dasar atau *root*. Pada proses ini dibantu oleh *library sastrawi*. Hasil proses *stemming* dapat dilihat pada tabel 5:

Tabel 5. Hasil *Stemming*

	Komentar	Hasil <i>Stopword Removal</i>
	tolong panitia evaluasi buat kedepannya wasitnya berlisensi kayak apa bukan gondok sama pemainnya malah jadi gondok wasitnya	tolong panitia evaluasi buat depan wasit lisensi kayak apa bukan gondok sama main malah jadi gondok wasit
	Keren banget PON ACEH-SUMUT	Keren banget PON ACEH-SUMUT

kata dalam penelitian ini dengan menggunakan fungsi *feature_extraction* pada *library scikit learn*. Sebagai contoh perhitungan proses *TF-IDF* dalam penelitian ini menggunakan sampel data *testing* dan *training*.

Hasil Uji Evaluasi

Proses uji evaluasi menggunakan *confusion matrix* yang berguna untuk menentukan keakuratan model yang sebenarnya. Proses perhitungan *confusion matrix* meliputi akurasi, presisi, *recall* dan *f-measure*. Pada penelitian ini, penulis melakukan empat percobaan dengan membandingkan data *training* dengan data *testing* dalam proporsi berikut: 60%:40%, 70%:30%, 80%:20%, dan 90%:10%. Hasil uji evaluasi dapat dilihat pada tabel 6:

Tabel 6. Hasil Uji Evaluasi

Keterangan	Percobaan			
	1	2	3	4
Data Training	60%	70%	80%	90%
Data Testing	40%	30%	20%	10%
Akurasi	70%	70%	74%	75%
Presisi	82%	82%	84%	84%
Recall	53%	53%	57%	59%
F-Measure	50%	51%	56%	58%

Hasil Uji Sentimen

Uji sentimen dilakukan agar dapat mengetahui sentimen dari komentar yang ditelaah diuji mengandung sentimen positif, netral atau negatif. Data komentar yang digunakan dalam tahap uji sentimen berbeda dari dataset. Proses uji sentimen dilakukan dalam empat percobaan dengan membandingkan data *training* dan data *testing* sebesar 60%:40%, 70%:30%, 80%:20%, dan 90%:10%. Contoh hasil percobaan uji sentimen dapat dilihat pada tabel 7 :

Tabel 7. Contoh Hasil Uji Sentimen Perbandingan 60%:40%

Komentar	Hasil Uji Sentimen
@jovimaspaitella pernah tuh PON 2016 bahkan DUTA nya bukan dari kelahiran jabar tapi gak dipermasalahkan ini cuma jadi IKON kok rame sejagad raya ya?	Negatif
Betul tuh min, kearifan lokal jgn sampe di lupakan dan wajib di lestarikan.	Netral
@wennyprd kami siswa SMAN 2 Padang Sidempuan siap mendukung dan mensukseskan PON XXI Aceh SUMUT 2024	Positif

4. KESIMPULAN

Disimpulkan bahwa 2000 data opini publik pada instagram terhadap kegiatan Pekan Olahraga Nasional ACEH SUMUT 2024 menunjukkan hasil positif yang lebih tinggi dibandingkan data negative atau netral. Hasil uji sentimen menunjukkan sentimen

negatif ketika komentar yang diuji pada tahap uji sentimen mengandung konotasi negatif, penolakan, dan ketidaksetujuan. Sedangkan komentar yang diuji pada tahap uji sentimen tidak berisi pernyataan negatif maupun positif dan kalimat tersebut merupakan kalimat pernyataan yang tidak memihak maka hasil uji sentimen menunjukkan netral. Dan hasil uji sentimen akan menunjukkan hasil sentimen positif apabila komentar tersebut bermakna positif, setuju dan mendukung pernyataan tersebut.

DAFTAR REFERENSI

- Arifin, H. S., Widyowati, W., & Hernawaty, T. (2017). Freedom of Expression di Media Sosial bagi Remaja Secara Kreatif dan Bertanggung Jawab: Bagi Siswa SMA Al-Ma'Soem Rancaek dan SMA Muhammadiyah Pangandaran. *Jurnal Pengabdian Kepada Masyarakat*, 1(5), 332–337.
- CNN Indonesia. (2021). *Menpora: Ketakutan Gangguan Keamanan dan Lonjakan Covid-19 di PON Papua Tidak Terjadi*. Indonesia: www.youtube.com.
- CNN Indonesia. (2021). *Kapolda Papua Janji Bayar Sengketa Lahan Adat Usai PON XX*. <https://www.cnnindonesia.com/nasional/20210831093436-12-687670/kapolda-papua-janji-bayar-sengketa-lahan-adat-usai-pon-xx> (accessed Feb. 25, 2022).
- Dyah, E. (2021). PON XX Papua Tingkatkan Pendapatan bagi Pelaku UKM hingga 4x Lipat. *news.detik.com*. <https://news.detik.com/berita/d-5756185/pon-xx-papua-tingkatkan-pendapatan-bagi-pelaku-ukm-hingga-4x-lipat> (accessed Feb. 25, 2022).
- F. Janati. (2021). Nagita Slavina Jadi Duta PON XX Papua, Arie Kriting: Seharusnya Sosok Perempuan Papua. *Kompas.com*. <https://www.kompas.com/hype/read/2021/06/02/163654966/nagita-slavina-jadi-duta-pon-xx-papua-arie-kriting-seharusnya-sosok?page=all> (accessed Feb. 25, 2022).
- Humas. (2021). Presiden Jokowi: Penyelenggaraan PON XX Gambarkan Kemajuan Papua. *setkab.go.id*. <https://setkab.go.id/presiden-jokowi-penyelenggaraan-pon-xx-gambarkan-kemajuan-papua/> (accessed Feb. 25, 2022).
- Ponne, L. (2017). Berlomba di Daerah Bergolak: Penyelenggaraan Pekan Olahraga Nasional (PON) IV di Kota Makassar. *Lensa Budaya Jurnal Ilmu-Ilmu Budaya*, 12(1), 70–81. <https://doi.org/10.34050/jlb.v12i1.3114>
- Rialdi, I. F. (2021). Viral Permainan ‘Bar-bar’ Jatim vs Jabar, Kepemimpinan Wasit Dipertanyakan. *bolatimes.com*. <https://www.bolatimes.com/arena/2021/10/01/191431/viral-permainan-bar-bar-jatim-vs-jabar-kepemimpinan-wasit-dipertanyakan> (accessed Feb. 25, 2022).
- Widyasari, K. L., Nuranna, F. S., & Sulasmi, A. L. (2018). *Memaksimalkan Penggunaan Media Sosial dalam Lembaga Pemerintahan* (1st ed.). Jakarta: Direktorat Jenderal Informasi dan Komunikasi Publik, Kementerian Komunikasi dan Informatika.