Analisis Sentimen Masyarakat Terhadap Fenomena Sound Horeg Pada Platform X Dengan Menggunakan Metode K-Nearest Neighbour (K-NN)

Bilaldi Alim Ghunadi[1], Paramitha Nerisafitra[2]

[1,3] Jurusan Teknik Informatika Fakultas Teknik Universitas Negeri Surabaya
[1]bilaldi.19079@mhs.unesa.ac.id
[2]paramithanerisafitra@unesa.ac.id

***Abstrak*— tujuan dari penelitian ini yaitu untuk mengkaji pendapat publik tentang fenomena suara horeg di platform X dengan menerapkan algoritma K-nearest neighbor (K-NN). Publik melihat fenomena ini sebagai hiburan atau sebagai sumber polusi suara yang mengganggu ketertiban umum. Data penelitian terdiri dari 762 tweet yang didapatkan melalui platform X, yang kemudian diproses dan diekstraksi fiturnya menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). Kemudian, data yang sudah didapat displit menjadi data latih dan uji menggunakan proporsi 70:30 sebelum algoritma K-NN diterapkan untuk proses pengklasifikasian sentimen menjadi positif, netral, ataupun negatif. Hasil dari penelitian ini menyatakan metode K-NN dapat mengkategorikan sentimen dengan akurasi 63,76%, precision 63,13%, recall 63,76%, dan untuk F1-Score 62,01%. Temuan ini memperlihatkan bahwa algoritma K-NN dapat digunakan untuk menganalisis pendapat publik mengenai fenomena Sound Horeg di Platform X**

***Kata Kunci*— Analisis Sentimen, Sound Horeg, K-Nearest Neighbor, TF-IDF, Platform X.**

I. PENDAHULUAN

Dalam beberapa tahun terakhir, khususnya di Jawa Timur dan Jawa Tengah, fenomena sound horeg telah menarik banyak perhatian publik. Istilah Jawa "horeg" berarti "menggerakkan" atau "bergetar". Dalam praktiknya, istilah sound horeg mengacu pada penggunaan sistem suara yang menghasilkan suara yang sangat keras, terutama di rentang bass, sehingga menghasilkan getaran di sekitarnya. Praktik ini umum di berbagai pesta, festival, atau acara hiburan publik. Meningkatnya penggunaan sound horeg di berbagai daerah juga disebabkan oleh fakta bahwa sound system dapat disewa dengan mudah dan murah.

Penduduk setempat telah menunjukkan berbagai reaksi terhadap Sound Horeg. Sebagian orang menganggapnya sebagai hiburan yang dapat dinikmati oleh berbagai kelompok dan merupakan bagian dari budaya lokal. Sebaliknya, banyak orang menganggap Sound Horeg sebagai sumber polusi suara yang mengganggu masyarakat, terutama karena volume suara yang tinggi dan seringkali berlangsung hingga larut malam. Komunitas memiliki banyak pendapat berdasarkan perspektif yang berbeda ini.

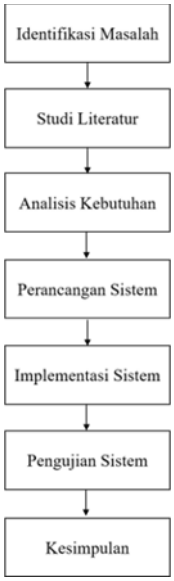
Dengan munculnya media sosial, masyarakat dapat mengekspresikan pendapat mereka tentang berbagai topik, seperti fenomena sound horeg. Platform X merupakan media sosial yang banyak dipergunakan oleh masyarakat umum untuk mengeluarkan kritik, dukungan, atau berbagi pengalaman terkait fenomena tersebut. Analisis sentimen mampu dimanfaatkan sebagai sumber data untuk dapat mengenali tren pendapat publik dari informasi yang dipublikasikan oleh pengguna. Teknik ini memungkinkan data teks diklasifikasikan menjadi sentimen positif, sentimen netral, atau sentimen negatif, sehingga menggambarkan persepsi publik terhadap topik tertentu.

Algoritma K-nearest neighbor merupakan algoritma yang kerap dipilih untuk proses pengklasifikasian teks dalam industri penambangan teks. Algoritma ini menempatkan dataset ke dalam kategori berdasarkan berapa banyak tetangga terdekatnya. Menurut Jabal Tursina (2019), metode K-NN memiliki prosedur yang sederhana dan mudah diterapkan, dan ketika nilai K dipilih dengan benar, memberikan akurasi klasifikasi yang tinggi. Untuk memungkinkan pemrosesan numerik sebelum klasifikasi menggunakan algoritma K-NN, penerapan metode TF-IDF dipakai untuk memperlihatkan data teks pada awal penelitian.

Mempertimbangkan hal-hal tersebut , kajian ini berfokus untuk mengevaluasi pendapat masyarakat mengenai fenomena Sound Horeg yang terjadi di platform X dengan menerapkan algoritma KNN. Hasilnya dimaksudkan untuk memberikan gambaran umum tentang tren opini publik terkait fenomena tersebut dan berfungsi sebagai referensi untuk penelitian mendatang serta sebagai dasar bagi pihak berwenang yang relevan untuk lebih memahami persepsi publik terhadap Sound Horeg.

II. METODOLOGI PENELITIAN

Kajian ini dilaksanakan melalui serangkaian tahapan yang sistematis dan bertujuan untuk mengidentifikasi pendapat masyarakat mengenai isu sound horeg pada platform X. Proses penelitian diawali dari pengumpulan data, dilanjutkan dengan preprocessing data, lalu ekstraksi fitur dengan mempergunakan metode TF-IDF, proses klasifikasi mempergunakan algoritma K-NN, dan diakhiri evaluasi terhadap kinerja model. Urutan tahapan penelitian tersebut ditampilkan pada Gambar 1.



Gbr. 1 Alur Penelitian

A. Pengumpulan Data

Data penelitian didapat melalui proses scraping terhadap platform X dengan menggunakan kata kunci "sound horeg". Pengambilan data dilakukan secara otomatis melalui pemanfaatan X API yang dijalankan menggunakan Google Colab. Keseluruhan data yang telah dikumpulkan selanjutnya disimpan dengan format CSV dan menghasilkan dataset yang terdiri atas 762 tweet. Sebelum memasuki tahap pengolahan, dataset terlebih dahulu diperiksa untuk menghilangkan data yang duplikat serta dilakukan pelabelan sentimen yang kemudian dimasukkan ke dalam 3 kelas, yaitu kelas netral, kelas positif, dan kelas negatif. Pada Gambar 2 ini merupakan alur dari scraping data

Gbr. 2 Alur Scraping Data

B. Preprocessing Data

Prosedur preprocessing ini dimaksudkan untuk menyiapkan data teks guna memiliki kualitas yang lebih baik sebelum dilakukannya tahap klasifikasi. Pada preprocessing ini meliputi proes cleaning, tokenizing, case folding, normalisasi, filtering, dan juga stemming.

1. Pada tahap cleaning, berbagai macam unsur yang tidak berpengaruh bagi analisis sentimen, seperti halnya link URL, angka, hashtag, tanda baca, mention dan karakter non-alfanumerik dihapus dari data teks.
2. Selanjutnya, case folding diterapkan dengan mengkoversi keseluruhan karakter kedalam lowercase atau huruf kecil.
3. Lalu terdapat tahapan tokenizing, yang bertugas membagi kalimat kedalam sejumlah token berupa kata.
4. Setelah itu dilakukan normalisasi yaitu proses mengkoversi kata yang awalnya tidak baku kedalam bentuk baku,
5. Lanjut pada proses filtering, yang mana digunakan untuk menghilangkan stopword yang tidak memiliki pengaruh signifikan terhadap proses analisis.
6. Lalu untuk tahap final dari preprocessing yaitu stemming, yang bertujuan mengubah kata yang awalnya berimbuhan menjadi kata dasar agar diperoleh data yang semakin konsisten sehingga dapat diproses untuk tahap selanjutnya.

C. Ekstraksi Fitur Menggunakan TF-IDF

Data yang telah diolah pada tahap preprocessing selanjutnya dikonversi menjadi bentuk representasi numerik dengan menerapkan metode Term Frequency-Inverse Document Frequency (TF-IDF). Penggunaan metode ini dimaksudkan untuk memberi bobot kedalam masing-masing kata yang mengacu pada frekuensi kemunculannya pada suatu dokumen dan tingkat penyebarannya pada seluruh dokumen. Hasil pembobotan tersebut selanjutnya dipergunakan untuk representasi fitur yang akan dijadikan masukan oleh algoritma K-Nearest Neighbor dalam proses klasifikasi. Rumus yang digunakan dalam perhitungan TF-IDF disajikan berikut ini :

$$idf = log \frac{N}{df} \quad (1)$$

$$w(k,d) = tf(k,d) \times idf \quad (2)$$

$$w(k,d) = tf(k,d) \times log \frac{N}{df}$$

D. Klasifikasi Menggunakan Algoritma K-Nearest Neighbor

Algoritma K-Nearest Neighbor (K-NN) diterapkan kedalam proses klasifikasi terhadap data yang telah direpresentasikan menggunakan TF-IDF. Sebelum klasifikasi dilakukan, dataset displit menjadi data latih dan uji menggunakan proporsi 70:30. Algoritma K-NN menentukan kategori sentimen suatu data berdasarkan kedekatan jaraknya dengan sejumlah data latih yang menjadi tetangga terdekat sesuai nilai K yang digunakan. Melalui proses tersebut, setiap data uji dikelompokkan menjadi salah satu kategori sentimen, berupa sentimen positif, sentimen negatif, atau sentimen netral. Perhitungan jarak antardata menggunakan metode Euclidean Distance dilakukan berdasarkan Persamaan (3) sebagai berikut :

$$d(x,y) = \sqrt{\sum_{i=1}^{p}(x_i - y_i)^2} \quad (3)$$

E. Evaluasi Model

Pada tahap evaluasi ini dijalankan untuk menilai tingkat performa model klasifikasi yang sudah dirancang. Pengukuran dilakukan menggunakan Confusion Matrix dengan cara melakukan perbandingkan hasil prediksi yang dihasilkan model dari label sebenarnya berdasarkan data uji. Selanjutnya, kinerja model diidentifikasi berdasarkan metrik evaluasi yang meliputi Akurasi, Precision, Recall, dan F1-Score. Nilai dari masing-masing metrik digunakan sebagai dasar dalam menilai kemampuan algoritma K-Nearest Neighbor dalam mengklasifikasikan opini masyarakat mengenai fenomena sound horeg pada platform X.

III. Hasil Dan Pembahasan

A. Pengumpulan Data

Proses pengumpulan data dijalankan melalui proses scraping terhadap platform X dengan memanfaatkan X API mempergunakan kata kunci "sound horeg". Data yang dikumpulkan berupa tweet berbahasa Indonesia yang dipublikasikan selama periode Januari 2025 hingga Januari 2026. Proses pengambilan data menghasilkan 762 tweet yang selanjutnya dimuat kedalam format CSV sebagai dataset penelitian. Sebelum digunakan pada tahapan analisis, dataset terlebih dahulu melalui proses pemeriksaan untuk menghilangkan data yang terduplikasi, kemudian setiap tweet diberikan label sentimen yang dibagi menjadi beberapa kategori yaitu positif, netral, dan negatif.

B. Distribusi Sentimen

Hasil dari distribusi sentimen yang telah dipisahkan berdasarkan masing-masing jenis sentimennya seperti yang tertera pada Tabel 1 berikut :

TABEL 1
JUMLAH DATA TWEET

Hasil pelabelan menunjukkan bahwa dataset terdiri atas 172 tweet berkategori positif, 212 tweet berkategori netral, dan 378 tweet berkategori negatif. Distribusi tersebut memperlihatkan bahwa sentimen negatif menempati jumlah terbesar daripada kategori sentimen positih dan netral. Temuan tersebut mengindikasikan bahwa sebagian besar pengguna platform X memberikan tanggapan yang cenderung menyoroti dampak negatif dari fenomena sound horeg, khususnya yang berkaitan dengan kebisingan dan gangguan terhadap kenyamanan lingkungan. Sementara itu, sentimen positif dan netral tetap ditemukan, meskipun jumlahnya lebih sedikit dibandingkan sentimen negatif.

C. Hasil Preprocessing Data

Pada Tahapan preprocessing ini diterapkan agar kualitas data teks dapat ditingkatkan sebelum masuk ke tahap klasifikasi. Tahapan yang dijalankan terdiri atas cleaning, case folding, tokenizing, normalisasi, filtering, dan terakhir stemming. Melalui tahapan tersebut, berbagai karakter yang tidak memiliki kontribusi terhadap sentimen, seperti halnya link URL, angka, hashtag, tanda baca, mention dan karakter non-alfanumerik berhasil dihilangkan. Selain itu, penulisan kata diseragamkan, kata tidak baku diubah menjadi bentuk baku, stopword dihapus, dan kata berimbuhan dikembalikan ke bentuk dasarnya. Proses tersebut membentuk data yang semakin konsisten sehingga selanjutnya dapat diterapkan proses ekstraksi fitur dan kemudian klasifikasi.

Setelah seluruh tahapan preprocessing selesai dilaksanakan, diperoleh hasil akhir pengolahan data yang contohnya disajikan Tabel 2 berikut.

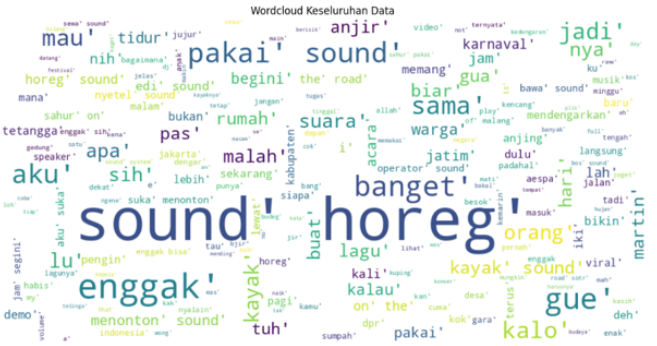
TABEL 2
CONTOH HASIL PREPROCESSING DATA

Tahapan	Hasil
Data Awal Tweet	RT @user Sound horeg di kampung saya bikin resah banget, suaranya keras sampai tengah malam!!! #soundhoreg
Cleaning	Sound horeg di kampung saya bikin resah banget suaranya keras sampai tengah malam
Case Folding	sound horeg di kampung saya bikin resah banget suaranya keras sampai tengah malam
Tokenizing	[sound, horeg, di, kampung, saya, bikin, resah, banget, suaranya, keras, sampai, tengah, malam]
Normalisasi	[sound, horeg, di, kampung, saya, bikin, resah, banget, suara, keras, sampai, tengah, malam]
Filtering	[sound, horeg, kampung, bikin, resah, banget, suara, keras, tengah, malam]
Stemming	[sound, horeg, kampung, bikin, resah, banget, suara, keras, tengah, malam]
Data Akhir Tweet	sound horeg kampung bikin resah banget suara keras tengah malam

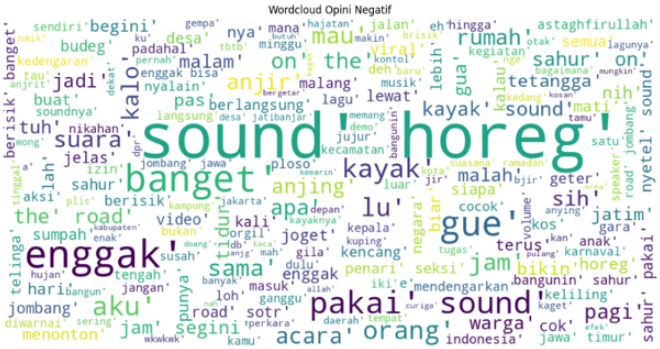
D. Visualisasi Data

Opini	Jumlah Data
Positif	172
Negatif	378
Netral	212
Total	**762**

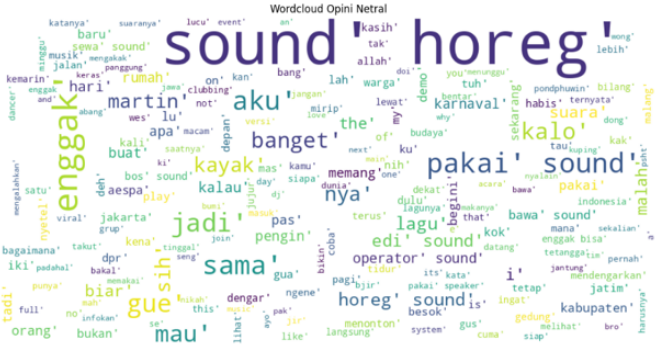
Visualisasi data dilakukan dengan memanfaatkan wordcloud untuk memberikan gambaran mengenai kata apa saja yang frekuensi kemunculannya tertinggi pada keseluruhan dataset maupun pada masing-masing kategori sentimen ditunjukkan dalam Gambar 3, 4, 5, dan 6.



Gbr. 3 Wordcloud Keseluruhan Data



Gbr. 4 Wordcloud Sentimen Negatif



Gbr. 5 Wordcloud Sentimen Netral

Gbr. 6 Wordcloud Sentimen Positif

Hasil visualisasi menunjukkan bahwa kata-kata seperti sound, horeg, ganggu, keras, dan bising memiliki frekuensi kemunculan yang relatif tinggi. Dominasi kata-kata tersebut mengindikasikan bahwa pembahasan mengenai kebisingan menjadi topik yang paling banyak disampaikan oleh pengguna platform X. Selain itu, perbedaan karakteristik kosakata pada setiap kategori sentimen memberikan gambaran awal mengenai kecenderungan opini masyarakat sebelum dilakukan proses klasifikasi dengan menerapkan algoritma K-Nearest Neighbor.

E. Hasil Pengujian Parameter K

Pada Tabel 3 dibawah ini ditunjukkan pengujian terhadap beberapa split data dan nilai K untuk mencari kombinasi terbaik.

TABEL 3
HASIL PENGUJIAN DATA SPLIT DAN NILAI K

Split	K	Accuracy	Precision	Recall	F1-Score
70:30	3	0.5982	0.6033	0.5982	0.5965
70:30	5	0.6113	0.6197	0.6113	0.6043
70:30	7	0.6026	0.6018	0.6026	0.5962
70:30	9	0.6113	0.6130	0.6113	0.6022
70:30	**11**	**0.6375**	**0.6313**	**0.6375**	**0.6200**
80:20	3	0.5294	0.5347	0.5294	0.5244
80:20	5	0.5816	0.5918	0.5816	0.5774
80:20	7	0.5816	0.5875	0.5816	0.5800
80:20	9	0.5816	0.5934	0.5816	0.5783
80:20	11	0.5816	0.5961	0.5816	0.5720
90:10	3	0.5194	0.5101	0.5194	0.5087
90:10	5	0.5584	0.5320	0.5584	0.5380
90:10	7	0.5584	0.5516	0.5584	0.5422
90:10	9	0.5454	0.5540	0.5454	0.5203
90:10	11	0.5324	0.5164	0.5324	0.5097

Pengujian model dilakukan menggunakan beberapa ragam nilai K, yaitu 3, 5, 7, 9, dan 11, dengan tiga konfigurasi distribusi data latih dan uji, yaitu yang pertama 70:30, kedua 80:20, dan terakhir 90:10. Tujuan dari pengujian tersebut yaitu untuk memperoleh kombinasi parameter yang menghasilkan performa klasifikasi terbaik berdasarkan nilai Akurasi, Precision, Recall, dan F1-Score. Dan menurut dari hasil pengujian, konfigurasi paling optimal dihasilkan pada penggunaan K = 11 dengan pembagian data 70:30, yang menghasilkan nilai Akurasi mencapai 63,76%, Precision mencapai 63,13%, Recall mencapai 63,76%, dan F1-Score mencapai 62,01%. Hasil tersebut menunjukkan bahwa pemilihan nilai K memberikan pengaruh terhadap kemampuan algoritma K-Nearest Neighbor pada proses klasifikasi sentimen.

F. Evaluasi Model

Evaluasi dijalankan untuk mengidentifikasi tingkat performa dari model klasifikasi yang sebelumnya telah dirancang. Pengukuran diuji menggunakan Confusion Matrix dengan metrik evaluasi yang terdiri dari Akurasi, Precision, Recall, dan F1-Score.

TABEL 4
EVALUASI HASIL

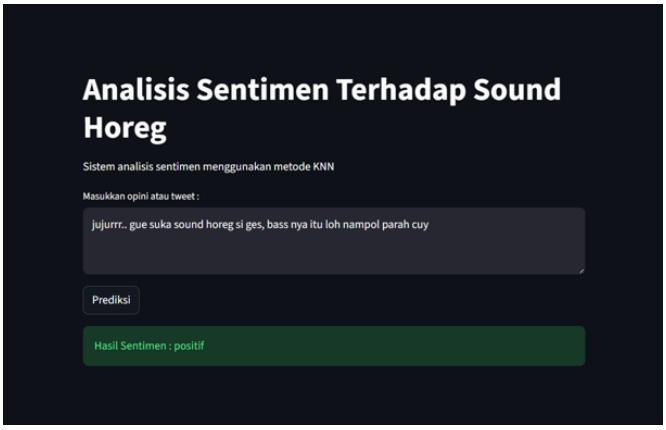
Metrik Evaluasi	Nilai (%)
Akurasi	63.76
Precision	63.13
Recall	63.76
F1-Score	62.01

Hasil evaluasi mengindikasikan bahwa model mendapatkan nilai Akurasi yakni 63,76%, Precision 63,13%, Recall 63,76%, serta F1-Score 62,01%. Nilai tersebut menunjukkan bahwa algoritma K-Nearest Neighbor mampu mengklasifikasikan sentimen publik mengenai fenomena sound horeg ke dalam klasifikasi positif, netral, dan juga negatif dan menunjukkan performa yang cukup baik. Walaupun demikian, masih ditemukan sejumlah kesalahan dalam klasifikasi yang menyebabkan nilai F1-Score lebih rendah dibandingkan nilai Akurasi dan Recall.

G. Implementasi Sistem

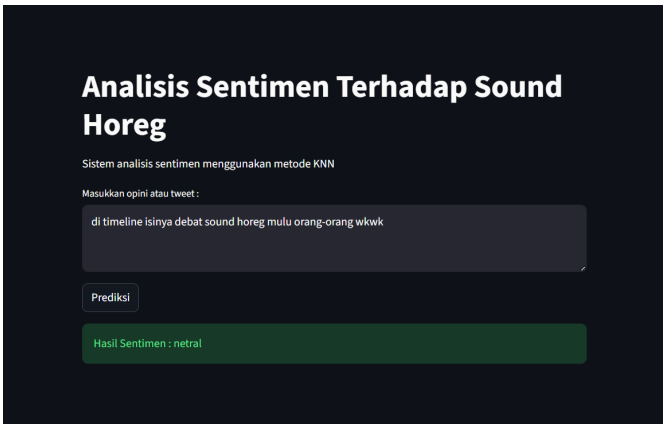
Model klasifikasi yang sudah melewati proses pengujian serta evaluasi kemudian dikembangkan dalam bentuk web menggunakan Streamlit. Aplikasi tersebut memungkinkan pengguna memasukkan teks opini mengenai fenomena sound horeg, kemudian sistem secara otomatis melakukan proses prediksi menggunakan model K-Nearest Neighbor yang telah dibangun. Hasil prediksi ditampilkan dalam bentuk klasifikasi opini positif, negatif, ataupun netral sehingga aplikasi dapat digunakan sebagai media untuk menganalisis sentimen secara sederhana atas opini publik pada platform X.

Gbr. 7 Tampilan Utama Aplikasi Web



Gbr. 8 Tampilan Hasil Prediksi Opini Positif

Gbr. 9 Tampilan Hasil Prediksi Opini Negatif



Gbr. 10 Tampilan Hasil Prediksi Opini Netral

IV. KESIMPULAN

Berlandaskan pada temuan penelitian atas analisis sentimen masyarakat mengenai fenomena sound horeg pada platform X dengan menerapkan algoritma K-Nearest Neighbor (K-NN), didapat kesimpulan meliputi.

1. Analisis sentimen mengenai fenomena sound horeg pada platform X berhasil dilaksanakan melalui serangkaian tahapan yang terdiri dari preprocessing data, pembobotan fitur dengan menerapkan metode TF-IDF, dan juga proses pengklasifikasian dengan menerapkan algoritma K-Nearest Neighbor.
2. Hasil pengujian terhadap beberapa variasi nilai K dan konfigurasi pembagian data latih serta uji memperlihatkan jika konfigurasi paling optimal diperoleh pada nilai K = 11 dengan proporsi pembagian data 70:30. Kombinasi parameter tersebut menghasilkan performa klasifikasi yang lebih baik dibandingkan konfigurasi lainnya.
3. Model klasifikasi yang dibangun mampu mengelompokkan sentimen masyarakat terhadap fenomena sound horeg dengan nilai Akurasi yakni 63,76%, Precision 63,13%, Recall 63,76%, dan F1-Score 62,01%.
4. Berdasarkan hasil analisis sentimen, kategori negatif menjadi sentimen yang paling dominan dengan jumlah 378 data (49,61%), diikuti oleh sentimen netral sebanyak 212 data (27,82%), dan sentimen positif sebanyak 172 data (22,57%). Dominasi sentimen negatif mengindikasikan bahwa sebagian besar masyarakat memandang fenomena sound horeg memberikan dampak yang kurang baik, terutama berkaitan dengan tingkat kebisingan serta gangguan terhadap lingkungan sekitar. Meskipun demikian, sebagian masyarakat lainnya tetap memberikan tanggapan positif dengan menganggap sound horeg sebagai salah satu bentuk hiburan sekaligus bagian dari aktivitas sosial masyarakat.
5. Hasil penelitian menunjukkan bahwa algoritma K-Nearest Neighbor (K-NN) mampu diimplementasikan untuk melakukan klasifikasi sentimen dalam platform X. Namun, tingkat performa model masih dipengaruhi oleh kualitas dataset, ketidakseimbangan distribusi kelas sentimen, serta karakteristik bahasa informal yang banyak digunakan dalam tweet.

V. SARAN

Berlandaskan pada temuan penelitian yang sudah diperoleh, disarankan beberapa hal yang diharapkan dapat membantu untuk penelitian mendang.

1. Peningkatan kualitas pelabelan data

   Penelitian berikutnya disarankan menerapkan proses pelabelan data yang lebih sistematis dan konsisten, misalnya dengan melibatkan lebih dari satu annotator atau menggunakan pedoman pelabelan yang lebih terstruktur. Langkah tersebut diharapkan mampu mengurangi subjektivitas dalam proses pemberian label sehingga kualitas dataset menjadi lebih baik.
2. Penyeimbangan distribusi dataset

   Dalam penelitian berikutnya, sebaiknya lebih memperhatikan keproporsionalan dataset dalam distribusi kelas sentimen. Pada penelitian ini, jumlah data dengan sentimen positif dan juga netral lebih sedikit jika dikomparasikan dengan jumlah sentimen negatif dan hal tersebut dapat berpotensi memengaruhi kemampuan model dalam mengenali seluruh kategori sentimen secara seimbang.
3. Pengembangan metode klasifikasi

   Penelitian selanjutnya dapat dilakukan komparasi antara satu algoritma dengan algoritma lain, seperti

contoh algoritma K-NN dengan algoritma Naive Bayes, SVM, atau Random Forest, untuk mencari mana performa paling optimal dari beberapa klasifikasi tersebut.

4. Penyempurnaan tahapan preprocessing

   Penelitian berikutnya dapat mengembangkan proses preprocessing dengan menambahkan tahapan yang lebih komprehensif, seperti normalisasi bahasa gaul, penanganan kata negasi, koreksi kesalahan penulisan (typo), serta penggunaan kamus sentimen yang disesuaikan dengan konteks fenomena sound horeg. Pengembangan tersebut diharapkan dapat meningkatkan akurasi hasil klasifikasi sentimen.

5. Pengembangan sistem analisis sentimen

   Sistem analisis sentimen yang telah dikembangkan dapat disempurnakan dengan menambahkan fitur visualisasi data, dashboard interaktif, serta kemampuan analisis secara real-time. Pengembangan tersebut diharapkan mampu menyajikan informasi yang lebih informatif dan mudah dipahami oleh pengguna.

REFERENSI

[1] Damarta, R., Hidayat, A., & Abdullah, A. S. (2021). The application of k-nearest neighbors classifier for sentiment analysis of PT PLN (Persero) twitter account service quality. Journal of Physics: Conference Series, 1722(1). https://doi.org/10.1088/1742-6596/1722/1/012002

[2] Dharmawan, L. R., Arwani, I., & Ratnawati, D. E. (2020). Analisis Sentimen pada Sosial Media Twitter Terhadap Layanan Sistem Informasi Akademik Mahasiswa Universitas Brawijaya dengan Metode K- Nearest Neighbor. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 4(3), 959–965. http://j-ptiik.ub.ac.id/index.php/j- ptiik/article/view/7099

Elcom. (2010). Twitter untuk pemula. Jakarta: Elex Media Komputindo.

[3] Furqan, M., Sriani, S., & Sari, S. M. (2022). Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia. Techno.Com, 21(1), 51–60. https://doi.org/10.33633/tc.v21i1.5446

[4] Han, J., Kambe, M., & Pe, J. (2011). Data Mining: Concepts and Techniques. In Data Mining: Concepts and Techniques. https://doi.org/10.1016/C2009-0-61819-5

[5] Hanafi, A., Adiwijaya, A., & Astuti, W. (2020). Klasifikasi Multi Label pada Hadis Bukhari Terjemahan Bahasa Indonesia Menggunakan Mutual Information dan k-Nearest Neighbor. Jurnal Sisfokom (Sistem Informasi Dan Komputer), 9(3), 357–364. https://doi.org/10.32736/sisfokom.v9i3.980

[6] Harahap, E. P., Purnomo, H. D., Iriani, A., Sembiring, I., & Nurtino, T. (2020). Trends in sentiment of Twitter users towards Indonesian tourism using KNN. International Journal of Advanced Computer Science and Applications, 11(3), 456–462. https://doi.org/10.14569/IJACSA.2020.0110358

[7] Harlian, M. (2006). Text mining: Analisis teks dalam data mining.

[8] Isnain, A. R., Supriyanto, J., & Kharisma, M. P. (2021). Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 15(2), 121. https://doi.org/10.22146/ijccs.65176

[9] Jabal Tursina, M. (2019). Sentimen Analisis Sistem Zonasi Sekolah Pada Media Sosial Youtube Menggunakan Metode K- Nearest Neighbor Dengan Algoritma Levenshtein Distance. Universitas Islam Negeri Syarif Hidayatullah Jakarta. https://eprints.utdi.ac.id/9135/

[10] Liu, B. (2012). Sentiment analysis and opinion mining. Morgan & Claypool Publishers.

[11] Malik, A. Z., Utami, E., & Raharjo, S. (2019). Analisis Sentiment Twitter Terhadap Capres Indonesia 2019 dengan Metode K- NN. Jurnal INFORMA Politeknik Indonesia Surakarta, 5(2), 1–7.

[12] Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. Proceedings of the 2nd International Conference on Knowledge Capture. https://doi.org/10.1145/945645.945658

[13] Neal, J. W., & Neal, Z. P. (2021). Prevalence and characteristics of childfree adults in Michigan (USA). PLoS ONE, 16(6 June), 1–18. https://doi.org/10.1371/journal.pone.0252528

[14] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1–2), 1–135. https://doi.org/10.1561/1500000011

[15] Ramadhani, K. W., & Tsabitah, D. (2022). Fenomena Childfree dan Prinsip Idealisme Keluarga Indonesia dalam Perspektif Mahasiswa. LoroNG: Media Pengkajian Sosial Budaya, 11(1), 17–29. https://doi.org/10.18860/lorong.v11i1.2107

[16] Rindu Fajar Islamy, M., Siti Komariah, K., Mayadiana Suwarma, D., & Hafidzani Nur Fitria, A. (2022). Fenomena Childfree di Era Modern: Studi Fenomenologis Generasi Z serta Pandangan Islam terhadap Childfree di Indonesia. Sosial Budaya, 19(2), 81–89. http://dx.doi.org/10.24014/sb.v19i2.16602

[17] Sari, R. (2020). Analisis Sentimen Pada Review Objek Wisata Dunia Fantasi Menggunakan Algoritma K-Nearest Neighbor (K- Nn). EVOLUSI : Jurnal Sains Dan Manajemen, 8(1), 10–17. https://doi.org/10.31294/evolusi.v8i1.7371

[18] Satrio, R. H., & Fauzi, M. A. (2019). Klasifikasi Tweets Pada Twitter Menggunakan Metode K-Nearest Neighbour (K-NN) Dengan Pembobotan TF-IDF. 3(8), 2548–2964. http://j-ptiik.ub.ac.id

[19] Qurotul, A. & Muhammad, R. (2023). Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Di Sma Negeri 15 Pekanbaru. https://doi.org/10.57152/ijirse.v3i1.484.

[20] Utomo, B. (2013). Stemming bahasa Indonesia untuk text mining.

[21] Wenando, F. A., Septiadi, R., Gunawan, R., Mukhtar, H., & Syahril. (2022). Analisis sentimen komentar merdeka belajar menggunakan KNN. Jurnal Informatika, 10(2), 55–63. https://doi.org/10.55681/sentri. v5i1.5425