

Fenomena Publik Figur: Transformasi Opini dan Persepsi *Netizen*

Abid Nujaba
Fakultas Matematika dan Ilmu
Pengetahuan Alam
Universitas Gadjah Mada
Yogyakarta, Indonesia
abid.nujaba@mail.ugm.ac.id

Gelora Damayanti Manalu
Fakultas Matematika dan Ilmu
Pengetahuan Alam
Universitas Gadjah Mada
Yogyakarta, Indonesia
gelora.dam2003@mail.ugm.ac.id

Faris Zaidan Nafis
Fakultas Matematika dan Ilmu
Pengetahuan Alam
Universitas Gadjah Mada
Yogyakarta, Indonesia
faris.zaidan.nafis@mail.ugm.ac.id

Abstrak—Penelitian ini bertujuan untuk mengembangkan sebuah sistem yang dapat membandingkan pandangan netizen terhadap seorang public figur sebelum dan sesudah terjadinya kejadian kontroversial. Dengan menggunakan teknik pengolahan bahasa alami dan pembelajaran mesin, sistem ini dapat mengklasifikasikan pernyataan atau komentar netizen sebagai positif atau negatif. Melalui pengumpulan dan analisis komentar-komentar netizen pada platform Twitter, penelitian ini menyajikan wawasan tentang perubahan pandangan publik terhadap public figur setelah terjadi peristiwa kontroversial. Dengan kontribusi pada pemahaman opini publik dan pengembangan sistem analisis sentimen, penelitian ini memiliki potensi aplikasi dalam analisis opini publik, manajemen reputasi, dan pengambilan keputusan berbasis data.

Kata Kunci—*netizen, public figur, kejadian kontroversial, NLP, pembelajaran mesin, analisis sentimen, Twitter.*

I. PENDAHULUAN

A. Latar Belakang

Dalam era digital yang semakin berkembang, penggunaan media sosial telah menjadi salah satu aspek yang signifikan dalam kehidupan sehari-hari. Media sosial memberikan platform bagi masyarakat untuk berbagi pendapat, memperoleh informasi, dan mengungkapkan pandangan mereka terhadap berbagai topik, termasuk mengenai figur publik. Akan tetapi, pandangan netizen terhadap public figur seringkali mengalami perubahan yang signifikan setelah terjadinya kejadian kontroversial yang melibatkan figur publik tersebut, seperti kasus kontroversial atau perselingkuhan.

Dalam konteks ini, penelitian ini bertujuan untuk mengembangkan suatu sistem yang mampu membandingkan pandangan netizen terhadap seorang public figur sebelum dan sesudah terjadinya kejadian yang kontroversial. Sistem ini akan berfokus pada analisis sentimen komentar-komentar netizen terhadap public figur di media sosial.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini akan berfokus pada penyelesaian beberapa masalah berikut:

1. Bagaimana membangun sebuah sistem yang dapat membandingkan pandangan netizen terhadap seorang public figur sebelum dan sesudah terjadinya kejadian kontroversial?
2. Bagaimana mengembangkan model klasifikasi yang dapat mengidentifikasi apakah pernyataan atau komentar netizen bersifat positif atau negatif terkait dengan public figur?

3. Bagaimana mengumpulkan dan menganalisis komentar-komentar netizen yang relevan dari media sosial untuk mendapatkan pemahaman tentang perubahan pandangan publik terhadap public figur setelah terjadi peristiwa kontroversial?
4. Bagaimana menggunakan sistem yang dikembangkan untuk menyajikan wawasan yang berharga tentang dinamika pandangan masyarakat terhadap public figur dalam konteks media sosial?

C. Batasan

Data yang akan diamati pada penelitian ini terbatas pada fenomena yang dapat ditemui pada komentar-komentar netizen terhadap sebuah postingan di sosial media Twitter.

D. Tujuan

Penelitian ini memiliki beberapa tujuan yang ingin dicapai:

1. Mengembangkan sebuah sistem yang dapat membandingkan pandangan netizen terhadap seorang public figur sebelum dan sesudah terjadinya kejadian kontroversial. Sistem ini akan menggunakan teknik pengolahan bahasa alami dan pembelajaran mesin untuk menganalisis komentar-komentar netizen di media sosial.
2. Membangun sebuah model klasifikasi yang mampu mengidentifikasi apakah pernyataan atau komentar netizen bersifat positif atau negatif terkait dengan public figur. Model ini akan membantu dalam menganalisis sentimen dari komentar-komentar yang dikumpulkan melalui web scraping.
3. Mengumpulkan dan menganalisis komentar-komentar netizen yang relevan dari media sosial, khususnya pada platform Twitter, untuk memperoleh pemahaman yang lebih baik tentang perubahan pandangan publik terhadap public figur setelah terjadi kejadian kontroversial. Data yang dikumpulkan akan digunakan sebagai basis untuk analisis dan perbandingan.
4. Menyajikan wawasan yang berharga tentang dinamika pandangan masyarakat terhadap public figur dalam konteks media sosial. Dengan menggunakan sistem yang dikembangkan, penelitian ini bertujuan untuk menggali informasi yang dapat memberikan pemahaman yang lebih mendalam tentang bagaimana opini publik terbentuk dan berubah seiring dengan peristiwa yang melibatkan public figur.

E. Manfaat

Penelitian ini memiliki manfaat yang signifikan dalam memperkaya pemahaman tentang dinamika opini publik dan pengembangan sistem analisis sentimen. Dengan membandingkan pandangan netizen terhadap public figur

sebelum dan sesudah kejadian kontroversial, penelitian ini memberikan wawasan yang berharga tentang perubahan pandangan masyarakat. Sistem yang dikembangkan menggunakan teknik pengolahan bahasa alami dan pembelajaran mesin, sehingga memiliki potensi aplikasi dalam analisis opini publik dan pengambilan keputusan berbasis data. Penelitian ini juga memberikan manfaat bagi public figur, manajemen reputasi, dan stakeholder terkait dalam mengadaptasi strategi komunikasi. Dengan manfaat-manfaat ini, penelitian ini diharapkan dapat memberikan kontribusi yang positif dalam pemahaman opini publik dan sistem analisis sentimen di era media sosial.

II. SOLUSI USULAN

A. Gambaran Umum

Solusi yang kami ajukan untuk mengamati Fenomena Publik Figur: Transformasi Opini dan Persepsi Netizen melibatkan penggunaan model SVM, preprocessing data, dan analisis sentimen. Kami melakukan *scrapping* beberapa *tweet* terkait Ardhito Pramono di akun pribadinya dan di akun lain sebelum dan sesudah kasus kontroversial, membersihkan data, dan melatih model SVM untuk mengklasifikasikan sentimen *tweet*. Dengan menerapkan model SVM pada data tersebut, kami akan menganalisis perubahan sentimen netizen terhadap Ardhito Pramono setelah kasus tersebut. Tujuan kami adalah memberikan gambaran tentang bagaimana opini publik dan persepsi terhadap seorang publik figur dapat berubah setelah terlibat dalam kasus yang kontroversial.

B. Dataset yang Digunakan

Kami menggunakan beberapa dataset untuk menunjang keberlangsungan mini riset kami:

1) Dataset training & testing

Dataset yang digunakan untuk training dan testing diakuisisi dari dataset pihak lain yang berupa twitter dataset dengan label dan tanpa label yang dapat digunakan untuk training dan testing [1].

2) Dataset prediksi

Data yang akan diprediksi pra kasus kontroversial sentimennya adalah *tweet* dari akun Twitter Ardhito Pramono dengan username @ardhitoprnmn tentang kembali ke panggung [2] serta tentang promosi Bilibili [3].

Sedangkan, untuk data yang diprediksi sentimennya pada pasca kasus kontroversial berasal dari akun @AREAJULID [4] serta seorang user bernama MazziniGSP yang memberikan opininya terhadap kasus ini dalam suatu Twitter post [5].

Akuisisi dataset dari akun Twitter Ardhito Pramono dilakukan dengan metode *scraping* otomatis. Proses *scraping* ini kami lakukan dengan membuat program yang dapat membaca input yang diberikan berupa alamat *tweet* yang hendak diambil data komentarnya. Selanjutnya program akan melakukan *parsing* html untuk mengambil

data komentar pada elemen html tertentu dan menyimpannya dalam output dengan format csv.

3) Corpus bahasa gaul Indonesia

Twitter terkenal dengan penggunaanya yang gaul dan kekinian, sehingga besar kemungkinan bahasa yang digunakan tidak baku dan non-formal. Untuk membantu menerjemahkan bahasa slang ke bahasa formal agar mempermudah pemrosesan, diperlukan sebuah corpus berupa dictionary yang kami dapatkan dari user louisowen6 di github [6] serta dari user Nasalsabila di github juga [7].

C. Metode yang Akan Digunakan

1) Akuisisi Data

Pertama, kami akan mengakuisisi data pelatihan dan pengujian dari sumber yang berbeda, yaitu ResearchGate untuk data umum dan melakukan web *scraping* untuk mengumpulkan data *tweet* Ardhito Pramono sebelum dan sesudah kasus kontroversial. Data ini akan menjadi dasar untuk melatih dan menguji model kami.

2) Preprocessing Data

Selanjutnya, kami akan melakukan *preprocessing* data dengan menggunakan library python seperti Sastrawi, metode *stemming*, *lemmatization*, penghapusan karakter non-alphanumeric, *stopword removal* dan tokenisasi. Preprocessing data ini bertujuan untuk membersihkan dan mengorganisir data agar siap untuk analisis selanjutnya.

- a. *Stemming*: *Stemming* adalah proses mengubah kata-kata menjadi bentuk dasarnya dengan menghilangkan awalan atau akhiran yang tidak diperlukan. Contohnya, kata "berlari" akan diubah menjadi "lari". *Stemming* lebih fokus pada pemotongan kata tanpa memperhatikan konteks atau arti kata.

Contoh:

Kata asli: berlari

Stemming: lari

- b. *Lemmatization*: *Lemmatization* juga mengubah kata menjadi bentuk dasarnya, tetapi dengan mempertimbangkan konteks dan arti kata. *Lemmatization* menghasilkan kata yang valid secara gramatikal. Contohnya, kata "berlari" akan diubah menjadi "lari", tetapi kata "berita" tidak akan berubah.

Contoh:

Kata asli: berlari

Lemmatization: lari

- Menghilangkan character non-alphanumeric: Ini melibatkan penghapusan karakter yang bukan huruf atau angka dari teks. Ini dapat membantu membersihkan data dari tanda baca,

Secara keseluruhan, solusi kami melibatkan akuisisi data dari ResearchGate dan *web scraping* tweet Ardhito

Pramono, diikuti dengan preprocessing data menggunakan berbagai teknik dan representasi fitur dengan TF-IDF. Kemudian, model SVM akan digunakan untuk melakukan analisis sentimen dan memprediksi sentimen pada tweet. Analisis ini akan memberikan gambaran tentang transformasi opini dan persepsi netizen terhadap Ardhito Pramono.

D. Metrik Evaluasi yang Digunakan

Dalam konteks analisis sentimen dengan SVM dan representasi fitur TF-IDF, kami menggunakan empat metrik evaluasi yang umum digunakan untuk mengukur kinerja model kami, yaitu [8]:

1. Akurasi (*Accuracy*): Akurasi mengukur sejauh mana model SVM dapat mengklasifikasikan dengan benar sentimen tweet pada data pengujian. Akurasi dihitung dengan membagi jumlah prediksi yang benar dengan total jumlah prediksi. Sebagai contoh, akurasi sebesar 0.9 menunjukkan bahwa model SVM mampu mengklasifikasikan dengan benar sekitar 90% tweet pada data pengujian.
2. Presisi (*Precision*): Presisi mengukur sejauh mana model SVM memberikan prediksi yang benar untuk sentimen positif. Presisi dihitung dengan membagi jumlah prediksi benar positif dengan total jumlah prediksi positif yang dilakukan oleh model. Sebagai contoh, presisi sebesar 0.8 menunjukkan bahwa sekitar 80% dari tweet yang diprediksi sebagai sentimen positif oleh model SVM memang benar-benar memiliki sentimen positif.
3. Recall: Recall mengukur sejauh mana model SVM dapat menemukan kembali (*mengingat*) sentimen positif yang sebenarnya dalam data pengujian. Recall dihitung dengan membagi jumlah prediksi benar positif dengan total jumlah sentimen positif yang sebenarnya dalam data. Sebagai contoh, recall sebesar 0.6 menunjukkan bahwa model SVM dapat mengidentifikasi sekitar 60% dari keseluruhan sentimen positif dalam data pengujian.
4. F1-Score: F1-Score merupakan penggabungan antara presisi dan recall. F1-Score memberikan ukuran keseluruhan kinerja model dengan mempertimbangkan baik presisi maupun recall. F1-Score dihitung dengan mengambil harmonic mean dari presisi dan recall. Sebagai contoh, F1-Score sebesar 0.83 menunjukkan harmonisasi antara presisi dan recall dalam memprediksi sentimen tweet.

III. HASIL DAN PEMBAHASAN

Sebelum melakukan training data, dilakukan pengecekan ulang terhadap data apakah masih terdapat nilai *null*. Setelah itu, dilakukan *data splitting* dengan pembagian 80% data train sebanyak 4377 *tweet* dan 20% data test sebanyak 1095 *tweet*.

A. Evaluasi Model

	model	cross validation accuracy	standard deviations	test accuracy
3	MultinomialNB	0.730867	0.014592	0.743379
4	RandomForestClassifier	0.722641	0.008886	0.735160
5	VotingClassifier	0.730413	0.013698	0.732420
0	SVC	0.736351	0.015643	0.731507
1	KNeighborsClassifier	0.672379	0.016519	0.693151
2	DecisionTreeClassifier	0.658213	0.013843	0.667580

Fig. 2. Hasil evaluasi 6 model dengan akurasi pengujian tertinggi

Setelah *training* model, dibuat evaluasi yang berisi akurasi *cross-validation*, standar deviasi, dan akurasi pengujian. Akurasi *cross-validation* membantu memastikan bahwa model tidak hanya memiliki akurasi yang tinggi pada satu set data uji tertentu, tetapi juga mampu mengklasifikasikan data dengan baik secara umum. Didapat hasil bahwa model yang memiliki akurasi *cross-validation* tertinggi adalah SVM, diikuti dengan Naive-Bayes, Voting, *Random Forest*, KNN, dan *Decision Tree*.

Deviasi standar mengukur variasi atau ketidakpastian dalam performa model antara lipatan *cross-validation*. Nilai deviasi standar yang rendah menunjukkan bahwa model memiliki konsistensi dalam kinerjanya dan tidak terlalu dipengaruhi oleh fluktuasi dalam subset data uji yang berbeda. Model dengan standar deviasi yang terendah adalah *Random Forest*, diikuti dengan Voting, *Decision Tree*, Naive-Bayes, SVM, dan KNN.

Untuk menguji generalisasi model pada data yang belum pernah dilihat sebelumnya, semua model dievaluasi pada dataset uji terpisah. Hasil menunjukkan bahwa model dengan pengujian tertinggi adalah Naive-Bayes, diikuti dengan *Random Forest*, diikuti dengan *Voting Classifier*, SVM, KNN, dan *Decision Tree*.

Berdasarkan hasil evaluasi Fig 2, model *Random Forest* merupakan performa yang terbaik dibanding 5 model lainnya. Walaupun *Random Forest* memiliki akurasi pengujian lebih rendah dari Naive-Bayes, tetapi selisih dari akurasi tersebut sangat sedikit dibandingkan dengan selisih standar deviasi dengan tertinggi kedua yaitu Voting yang sangat jauh.

B. Testing

Pada program ini, dilakukan prediksi pada 4 data: 2 *tweet* kejadian kontroversial sebelum (pra-kejadian) [2][3] dan 2 *tweet* sesudah (pasca-kejadian) [4][5] dari seorang publik figur bernama Ardhito Pramono. Model yang digunakan adalah model terbaik pada evaluasi yaitu *Random Forest*.

1 pra_kejadian['sentiment'].value_counts()	
✓	0.0s
Positif	74
Negatif	21

Fig. 3. Jumlah opini positif dan negatif sebelum kejadian 1

	tweet	sentiment
0	Akhirnya bisa nonton bang ardhito manggung lagi	Positif
1	Wooww aku blm pernah nonton ardhito	Negatif
2	that's my baby	Positif
3	Akhirnya bisa ngrasain suasana gigs lagi setel...	Positif
4	Akhirnya bisa nonton ardhito manggung	Positif
...
90	Keren banget sumpah. Sepanjang konser ikutan n...	Positif
91	Apa kabar omnibuslaw bung? Merdeka!!	Negatif
92	mantapppp	Positif
93	Kapan dito	Positif
94	Aww sayang aku	Positif

Fig. 4. Hasil prediksi *tweet* sebelum kejadian 1

1	<code>pra_kejadian2['sentiment'].value_counts()</code>
✓	0.0s
Positif	59
Negatif	28

Fig. 5. Jumlah opini positif dan negatif sebelum kejadian 2

	tweet	sentiment
0	komen	Positif
1	Nah gini dari pada lo cuma balesin tweet gue, ...	Negatif
2	iya iya	Positif
3	This is awesome :)	Negatif
4	Sini bagi bahagiannya	Positif
...
82	~~ iklan yaa kak	Positif
83	izin numpang jualan kak	Positif
84	halooo, aku jual netflix viu dll yuk dm langsu...	Negatif
85	"	Positif
86	songs that make your day feel better"	Positif

Fig. 6. Hasil prediksi *tweet* sebelum kejadian 2

1	<code>pasca_kejadian['sentiment'].value_counts()</code>
✓	0.0s
Negatif	95
Positif	60

Fig. 7. Jumlah opini positif dan negatif setelah kejadian 1

	tweet	sentiment
0	"Ini hasil dari pengembangan kasus dan laporan...	Negatif
1	Bisa jadi sih, atau cepu tadi ya.	Positif
2	Klo nggak di list mungkin FF juga udh ditangka...	Negatif
3	polisi nyambi jd temen artis kan juga banyak bang	Positif
4	Di bandar nya ada cepu bang.	Positif
...
188	sbner nya data nya udah ada.. trus masa pengi...	Positif
189	biasa nya ini ulah bedede nya jga..	Negatif
190	udh rahasia umum...polisi mah udh pegang nama ...	Positif
191	percaya sama polisi? hahahahahaha	Positif
192	Cepu dan bandar adalah masyarakat	Positif

Fig. 8. Hasil prediksi *tweet* setelah kejadian 1

1	<code>pasca_kejadian2['sentiment'].value_counts()</code>
✓	0.0s
Negatif	125
Positif	68

Fig. 9. Jumlah opini positif dan negatif setelah kejadian 2

	tweet	sentiment
0	"Ini hasil dari pengembangan kasus dan laporan...	Negatif
1	Bisa jadi sih, atau cepu tadi ya.	Positif
2	Klo nggak di list mungkin FF juga udh ditangka...	Negatif
3	polisi nyambi jd temen artis kan juga banyak bang	Positif
4	Di bandar nya ada cepu bang.	Negatif
...
188	sbner nya data nya udah ada.. trus masa pengi...	Positif
189	biasa nya ini ulah bedede nya jga..	Negatif
190	udh rahasia umum...polisi mah udh pegang nama ...	Positif
191	percaya sama polisi? hahahahahaha	Positif
192	Cepu dan bandar adalah masyarakat	Negatif

Fig. 10. Hasil prediksi *tweet* setelah kejadian 2

Dari hasil pengujian yang ditampilkan pada Fig. 3 hingga Fig. 10, dapat dilihat hasil sentimen untuk setiap tweet beserta jumlahnya. Pra-kejadian 1 memiliki 74 opini positif dan 21 opini negatif, pra-kejadian 2 memiliki 59 opini positif dan 28 opini negatif, pasca-kejadian 1 memiliki 60 opini positif dan 95 opini negatif, dan pasca-kejadian 2 memiliki 68 opini positif dan 125 opini negatif.

Namun, perlu dicatat bahwa terdapat beberapa kesalahan dalam prediksi sentimen pada beberapa tweet. Salah satu faktor yang dapat menyebabkan kesalahan ini adalah jumlah data yang terbatas dan variasi data yang tidak mencukupi. Dengan dataset yang lebih besar dan lebih beragam, model dapat memiliki pemahaman yang lebih baik tentang berbagai sentimen yang ada dalam teks.

C. Visualisasi

Kami menggunakan visualisasi data untuk menganalisis dan menyajikan hasil dari pandangan netizen terhadap

seorang publik figur sebelum dan sesudah terjadinya peristiwa kontroversial.

Kami menggunakan berbagai jenis visualisasi seperti pie chart, bar chart, dan line chart untuk mewakili perbandingan antara pandangan sebelum dan sesudah peristiwa. Dengan menggunakan pie chart, kami dapat dengan jelas melihat proporsi komentar positif, negatif, dan netral dalam setiap periode waktu. Bar chart membantu kami membandingkan jumlah komentar positif dan negatif secara visual. Sementara itu, line chart membantu kami melihat tren perubahan sentimen dari waktu ke waktu.

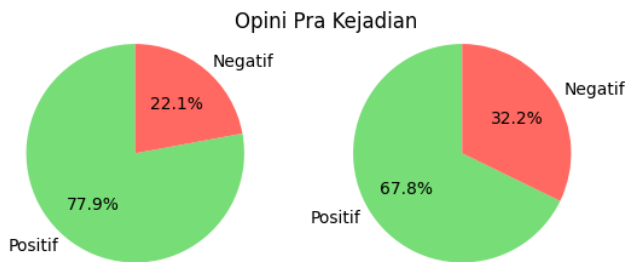


Fig. 7. Pie chart opini sebelum kejadian.

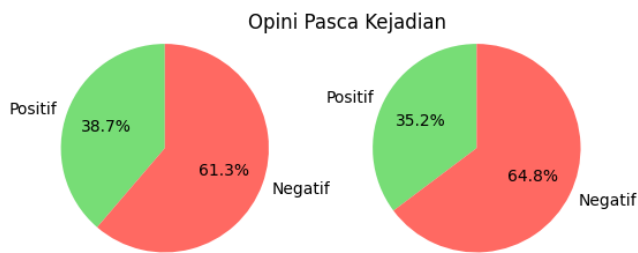


Fig. 8. Pie chart opini setelah kejadian

Pada pie chart ini, opini dibagi menjadi dua kategori, yaitu positif dan negatif. Didapat informasi bahwa terdapat 77,9% opini positif dan 22,1% opini negatif untuk pra-kejadian 1, 67,8% opini positif dan 32,2% opini negatif untuk pra-kejadian 2. Setelah kasus, terdapat 61,3% opini negatif dan 38,7% opini positif untuk pasca-kejadian 1 dan terdapat 35,2% opini positif dan 64,8% opini negatif. Terdapat perubahan dominan juga bahwa yang sebelumnya opini terbanyak adalah positif, setelah kejadian opini terbanyak adalah negatif.

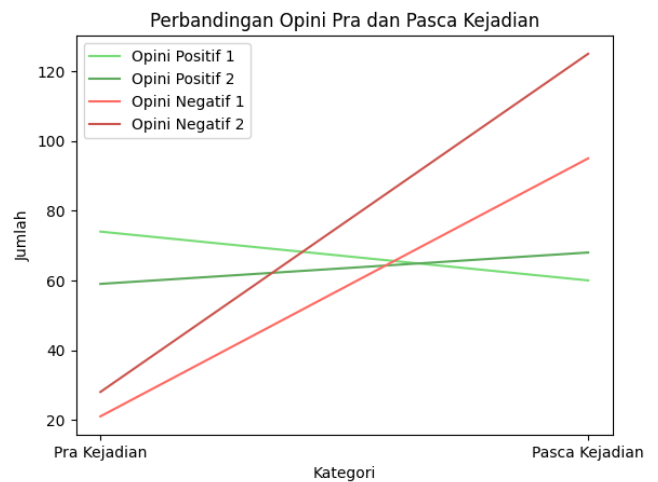


Fig. 9. Line chart opini sebelum dan sesudah kejadian

Berdasarkan Fig. 9, didapat bahwa terjadi penurunan sedikit jumlah opini positif dan peningkatan negatif yang cenderung sangat besar sampai membuat titik potong sehingga jumlah opini negatif melebihi jumlah opini positif pada pasca kejadian.

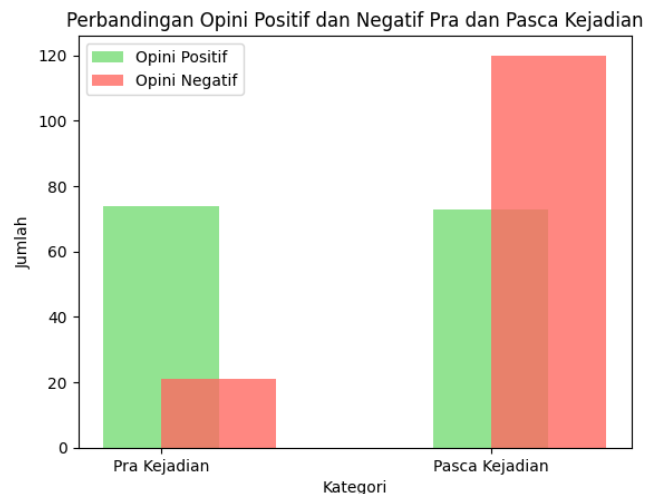


Fig. 10. Bar chart opini sebelum dan sesudah kejadian 1

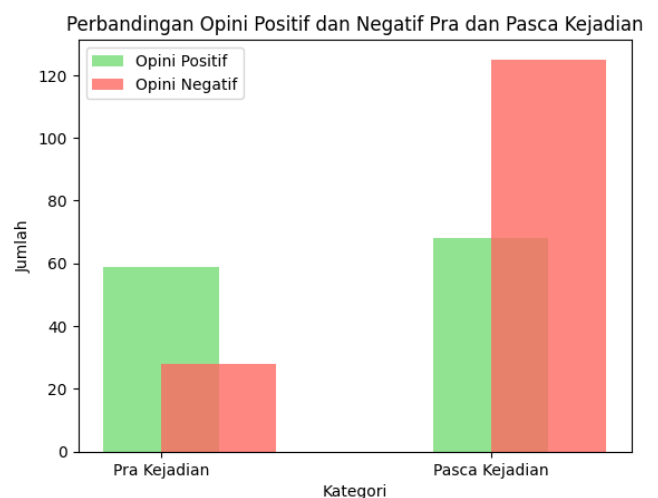


Fig. 11. Bar chart opini sebelum dan sesudah kejadian 2

IV. KESIMPULAN

Dari laporan yang dijelaskan, dapat disimpulkan bahwa kami berhasil mengembangkan sebuah sistem yang dapat membandingkan pandangan netizen terhadap seorang public figur sebelum dan sesudah terjadinya kejadian kontroversial. Melalui analisis sentimen dan model klasifikasi, kami mampu mengidentifikasi apakah pernyataan atau komentar netizen bersifat positif atau negatif terkait dengan public figur. Pengumpulan dan analisis komentar-komentar netizen dari media sosial juga memberikan pemahaman yang berharga tentang perubahan pandangan publik terhadap public figur setelah terjadi peristiwa kontroversial. Walaupun belum memiliki akurasi yang tinggi, sistem yang dikembangkan ini diharapkan memberikan wawasan yang penting dalam melihat dinamika pandangan masyarakat terhadap public figur dalam konteks media sosial.

Secara keseluruhan, penelitian ini memberikan kontribusi penting dalam memahami transformasi opini dan persepsi netizen terhadap public figur. Dengan memanfaatkan teknologi dan metode analisis yang tepat, laporan ini menyoroti pentingnya mengumpulkan dan menganalisis komentar-komentar netizen yang relevan dari media sosial. Informasi yang diperoleh dari analisis ini dapat membantu public figur dan pihak terkait dalam menghadapi peristiwa kontroversial dengan lebih baik, serta mengambil langkah-langkah yang sesuai untuk memperbaiki citra dan hubungan dengan masyarakat.

V. INFORMASI TAMBAHAN

Implementasi kode pada laporan ini dapat diakses di link Google Drive berikut:
<https://drive.google.com/drive/folders/1K6RnuptpVK6ML5Rqf5X1mtpcTR5a6HmY?usp=sharing>

REFERENSI

- [1] Ferdiana, Ridi & Fajar, Wiliam & Purwanti, Desi & Djatmiko, Fahim & Ayu, Artmita. (2018). Indonesian Sentiment Twitter Dataset. 10.13140/RG.2.2.16433.12646.
- [2] Pramono, A (2020, August 16). "Akhirnya bisa manggung lagi" [Twitter post]. Available: <https://twitter.com/ardhitoprnm/status/1295000749831684096?t=wwrmFL9K9J76jGum3nzIDA&s=19>.
- [3] Pramono, A (2021, February 11). "Karena aku nomor 1, bahagia mulainya ya dari aku, biar aku bisa membagi kebahagiaan itu ke semua orang!" [Twitter post]. Available: https://twitter.com/ardhitoprnm/status/1359836368424214528?t=OA5Pmmd9c9fp89kQ_1CX8Vg&s=19.
- [4] AREAJULID. (2022, January 13). "Wdyt? kaget ardhito nark0b4 X kaget ardhito udh punya istri ✓" [Twitter post]. Available: <https://twitter.com/AREAJULID/status/1481583097036816390?t=5jPSjPehO5vEo5abiSCp5g&s=19>.
- [5] Mazzini GSP. (2022, January 15). "Pertanyaannya. Gimana masyarakat tau artis pakai narkoba lalu lapor polisi. Informasi Ardhito udah kawin aja masyarakat gak tau, baru tau pas dia tersangkut kasus. Tapi ya udah lah ya, kita percaya aja sama Polisi." [Twitter post]. Available: https://twitter.com/mazzini_gsp/status/1482352824763949056?t=gh3jW7diN9FIaamvpm2CMw&s=19.
- [6] Louis Owen. (2021). "NLP Bahasa Resources" [GitHub repository]. Available: https://github.com/louisowen6/NLP_bahasa_resources/blob/master/combined_slang_words.txt.
- [7] Nasalsabila. (n.d.). "Colloquial Indonesian Lexicon" [CSV file]. Available: <https://raw.githubusercontent.com/nasalsabila/kamus-alay/master/colloquial-indonesian-lexicon.csv>.
- [8] Liang, S. (2020). The Most Common Evaluation Metrics in NLP. Towards Data Science. Retrieved from: <https://towardsdatascience.com/the-most-common-evaluation-metrics-in-nlp-ced6a763ac8b>.