

Analisis Pengaruh Meninggalnya Emmeril Terhadap Elektabilitas Ridwan Kamil Menggunakan Naïve Bayes

Risky Firmansyach¹, Ulfi Saidata Aesy²

^{1,2} Sistem Informasi, Fakultas Teknik dan Teknologi Informasi, Universitas Jendral Achmad Yani Yogyakarta.

¹riski.firmansyach@gmail.com

²ulfiaesyi@gmail.com

Abstrak— Twitter merupakan sumber informasi dan media sosial yang sangat populer dikalangan masyarakat, dimana orang-orang bisa berkomunikasi dengan pesan singkat yang disebut tweet. Dengan adanya pemberitaan tentang Eril, banyak masyarakat yang menggunakan Twitter untuk mengunggah opini dan tanggapan mereka mengenai hal ini sehingga menjadi sebuah trending topik atau perbincangan hangat di Twitter. Tanggapan masyarakat terkait Meninggalnya Almarhum Emmeril Kahn bisa memiliki peranan penting. Dengan melakukan analisis sentimen terhadap opini masyarakat kita bisa mengetahui opini-opini dari netizen mengenai Almarhum Emmeril Kahn dan apakah memiliki Pengaruh terhadap Elektabilitas Pemilu Ridwan Kamil, baik berupa tweet dan re-tweet. Dalam penelitian ini menggunakan Metode Naïve Bayes Classification dan dibantu perhitungan TF-IDF untuk membuat model analisis. Data dikumpulkan menggunakan teknik Scraping dan akan melewati tahap Preprocessing data sebelum dilakukannya penelitian. Didapatkan hasil prediksi sentimen Positif sebanyak 5701 tweet atau sebesar 67.8% dan sentimen Negatif sebanyak 2702 data tweet atau sebesar 32.2%. Dengan melakukan klasifikasi terhadap data yang telah diprediksi sehingga didapatkan nilai presisi masing-masing sentimen Positif 92% dan Negatif 80%, nilai recall masing-masing sentimen Positif 93% dan Negatif 83%. Kemudian nilai akurasi prediksi sebesar 87.9%. Sehingga dapat diketahui bahwa pemodelan yang sudah dibuat cukup bagus, dan banyak terdapat opini positif mengenai Eril dan Ridwan Kamil, sehingga masyarakat bisa mengetahui lebih jauh bagaimana sosok Ridwan Kamil di mata masyarakat sehingga bisa memberikan pengaruh positif terhadap Elektabilitasnya.

Kata Kunci— Data Mining, Analisis Sentimen, TF-IDF, Naive Bayes Classification, Ridwan Kamil.

I. PENDAHULUAN

Dr. H. Mochamad Ridwan Kamil atau biasa disapa Kang Emil. Ayah biologis Emmeril Kahn. Saat ini Ridwan Kamil menjabat sebagai Gubernur Jawa Barat sejak 2018, yang sebelumnya menjabat sebagai Walikota Bandung (Puji Lestari, 2017). Setelah masa jabatannya sebagai Gubernur berakhir pada 2023, Ridwan Kamil akan melanjutkan karir independennya di bidang politik. Dengan tetap menjadi gubernur Jawa Barat atau Calon Presiden, Ridwan Kamil akan tetap menjadi anggota partai politik. Pada saat penelitian ini dilakukan, Ridwan Kamil masih Menimbang-nimbang apakah Ridwan akan mengikuti Pilkada 2024 atau kembali mengikuti pemilihan kepala daerah Jawa Barat. ^[2]Sebelumnya, Ridwan Kamil memberikan tanggapannya atas pembahasan RUU Pilkada di DPR RI. Ada banyak poin dan sorotan, salah satunya adalah standarisasi tanggal Pilkada pada 2022 dan 2023. Jika RUU Pilkada disetujui dengan ketentuan akan ada pemilu pada 2023, ini akan mempengaruhi kemampuan Ridwan untuk mencalonkan diri sebagai gubernur lagi. Ridwan Kamil memastikan akan mematuhi segala keputusan DPR RI, termasuk saat RUU Pilkada nanti disahkan.

(Fadjar et al., 2021). Dengan adanya pemberitaan tersebut maka perlu dilakukan analisis Twitter mengenai opini publik tentang Ridwan Kamil guna melakukan analisis terhadap Elektabilitas Ridwan Kamil[1].

Emmeril Kahn Mumtaz merupakan anak pertama dari pasangan Ridwan Kamil dan Atalia Praratya. Pada tanggal 26 Mei 2022, Emmeril Kahn dikabarkan hanyut terbawa oleh arus Sungai Aare di Swiss. Selang beberapa hari kemudian, tepatnya tanggal 3 Juni 2022 pihak keluarga menyatakan bahwa Eril telah meninggal dunia. Kabar tersebut membuat Ridwan Kamil sekeluarga dan seluruh Indonesia bersedih dan berduka cita. (Erika Rizky Rachmani, 2022) Pemberitaan Eril sempat menjadi perbincangan hangat atau trending topik di Twitter sehingga banyak opini-opini masyarakat mengenai Eril dan bahkan Ridwan Kamil yang merupakan Ayah kandung Eril, sekaligus merupakan Tokoh Publik. Dari opini-opini masyarakat mengenai Ridwan Kamil bisa dijadikan bahan pertimbangan elektabilitas Ridwan Kamil[3].

Twitter merupakan sumber informasi dan media sosial yang sangat populer di masyarakat, dimana orang dapat berkomunikasi melalui pesan singkat yang dikenal dengan tweet. Dengan adanya Twitter, pengguna dapat dengan mudah memposting dan menerima informasi tentang topik diskusi terkini (Paul Gil, 2021). Dengan adanya pemberitaan mengenai Eril, banyak orang menggunakan Twitter untuk memposting pendapat dan tanggapan mereka tentang ini sehingga menjadi perbincangan hangat di Twitter. Reaksi dan opini publik tentang kematian mendiang Emmeril Kahn dapat memiliki peran penting. Dengan melakukan analisis sentimen terhadap opini publik, kita dapat mengetahui opini netizen tentang almarhum Emmeril Kahn dan pengaruhnya terhadap hak pilih Ridwan Kamil, baik berupa tweet dan re-tweet yang diklasifikasikan ke kategori positif, netral dan negative[4].

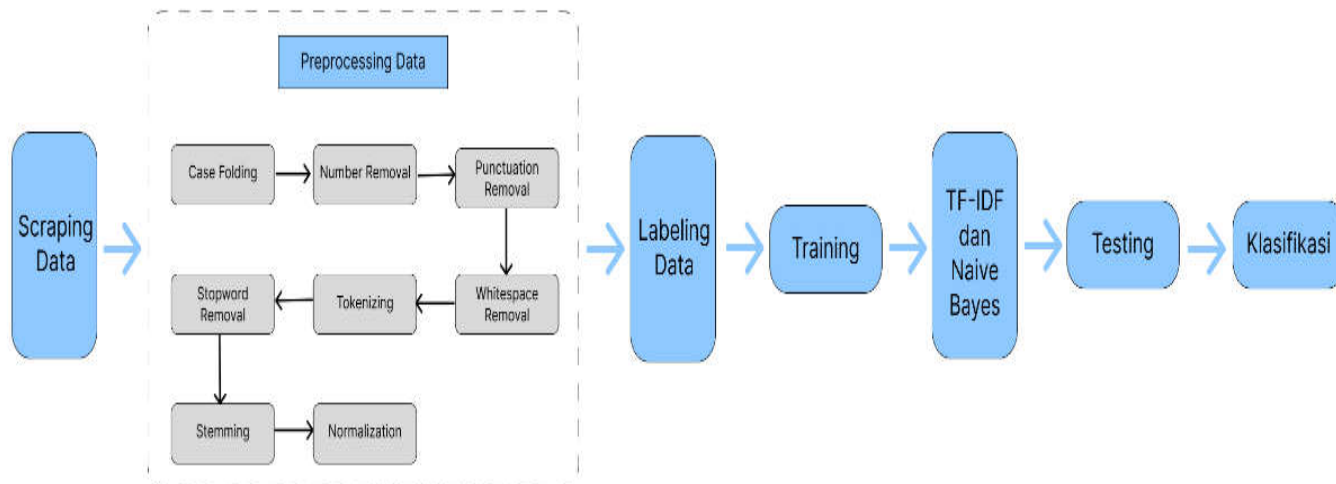
Berdasarkan dari permasalahan tersebut, maka perlu dilakukan penelitian analisis sentiment terkait Meninggalnya Almarhum Emmeril Kahn dan Pengaruhnya terhadap Elektabilitas Pemilu Ridwan Kamil di Twitter untuk mengetahui pandangan netizen mengenai Emmeril Kahn dan apakah memiliki pengaruh terhadap Elektabilitas Pemilu Ridwan Kamil yang nantinya data tersebut diambil dari Twitter, setelah itu data akan dianalisis dan dilabel ke kategori positif, netral atau negatif. ^[5]Analisis sentiment dilakukan untuk melihat kecenderungan opini terhadap suatu masalah atau objek, apakah cenderung beropini positif atau negatif (Berliana et al, 2018). Pada penelitian ini menggunakan metode Naive Bayes Classification dikarenakan memudahkan penelitian dalam mencari nilai probabilitas pada setiap kata dan dapat melakukan proses pengolahan data diskrit dan kuantitatif dengan menggunakan sampel data yang relative sedikit. ^[6]Keunggulan metode Naïve Bayes Classification dibandingkan metode lainnya yaitu lebih mudah dalam menghitung probabilitas kelas berdasarkan distribusi kata-kata yang ada dalam dokumen karena menghitung satu set probabilitas dengan menambahkan Frekuensi dan kombinasi nilai dari data training dan setiap kategori tidak saling bergantung [6].

II. METODOLOGI

Dalam penelitian ini menggunakan Metode *Naïve Bayes Classification* untuk mendapatkan nilai prediksi dari data tweet yang sudah dilabeli kemudian hasil dari klasifikasi *Naïve Bayes* dianalisis nilai presisi dari hasil model klasifikasi yang dibuat. Penelitian ini menggunakan data *tweet* dan *re-tweet* yang didapatkan dari *Twitter* dengan menggunakan keyword “Eril Ridwan Kamil” pada tanggal 1 Juni – 14 Juni 2022 dengan data tweet sebanyak 10.000, kemudian “Pemilu Ridwan Kamil”, dan “Elektabilitas Ridwan Kamil” pada tanggal 01 Januari – 14 Juni 2022 dengan jumlah data 1016 data

tweet menggunakan teknik *scraping*. Selanjutnya melakukan *preprocessing* data untuk mendapatkan hasil yang diinginkan. Nantinya data tersebut digunakan untuk memetakan informasi atau sentiment dari *netizen* di *Twitter* mengenai Meninggalnya Almarhum Emmeril Kahn dan dari sentiment tersebut apakah mempunyai pengaruh terhadap elektabilitas Pemilu Ridwan Kamil sehingga didapatkan informasi yang lebih jelas.

Penelitian ini dilakukan karena permasalahan masalah yang ada, sehingga didapatkan hasil analisis yang sesuai apa yang dibutuhkan, berikut merupakan alur diagram Metoda Penelitian dapat dilihat pada Gambar 1 berikut.



Gbr 1 Alur Diagram Metode Penelitian

Berdasarkan Gambar 1, Penelitian dimulai dengan melakukan pengambilan data dari *Twitter* menggunakan *Scraping Data* menggunakan tools *scraping* yaitu *Snsrcape*, kemudian data tersebut dilabeli menjadi 3 sentimen positif, netral, negatif secara manual dan dimuat kedalam file *csv* agar mempermudah dalam melakukan *Preprocessing Data*. Setelah melakukan *Preprocessing Data* kemudian melakukan *Labeling data* terhadap 450 data *tweet*, kemudian data tersebut diuji kedalam *Training* dan *Testing* data menggunakan metode *Naïve Bayes Classification* dan *TF-IDF* supaya mendapatkan sebuah klasifikasi sentiment untuk dijadikan prediksi. setelah itu melihat hasil dan tingkat keakuratan pemodelan yang telah dibuat.

III. HASIL DAN PEMBAHASAN

Paragraf harus Hasil dari analisis sentiment menggunakan *Naïve Bayes Classification* untuk melakukan analisis data *tweet* label sentiment positif dan netral dengan kata kunci “Eril Ridwan Kamil” menggunakan data yang diambil pada tanggal 3 Juni – 14 Juni 2022 kemudian kata kunci “Pemilu Ridwan Kamil” dan “Elektabilitas Ridwan Kamil” dengan jumlah seluruh data 11.016 data *tweet* dan *re-tweet*. Data yang digunakan dalam tahap *training data* sebanyak 450 data *tweet*. Kemudian menerapkan model klasifikasi *Naïve Bayes* yang dibuat pada data uji 11.016 data. Didapatkan hasil analisis sentiment dengan sentiment positif sebanyak 5701 dengan nilai presisi sebesar 67.8% dan Negatif 2702 data dengan nilai presisi sebesar 86% dengan nilai rata-rata akurasi sebesar 89.84%.

Dengan menggunakan program *scraping* yang sudah dibuat dengan dibantu tools *Snsrcape* untuk mengambil data dari *Twitter*. Data yang diambil merupakan data *tweet* dan *re-tweet* dengan keyword “Eril Ridwan Kamil”, “Pemilu Ridwan Kamil”, dan

“Elektabilitas Ridwan Kamil” yang menggunakan Bahasa Indonesia. Dalam data tersebut memiliki beberapa atribut seperti tanggal, username, id *tweet*, dan *tweet*. setelah 11.016 data sudah diambil kemudian melakukan *Preprocessing data* yang kemudian untuk dilakukan *Labeling data*.

Setelah melabeli data, langkah selanjutnya dalam penelitian ini adalah melatih data, menggunakan modul *scikitlearn* untuk menghitung bobot kata dari setiap *tweet*, yang kemudian akan digunakan sebagai model klasifikasi untuk menguji data. Dengan mengimpor beberapa modul dari *library sklearn* untuk melakukan data pelatihan seperti *TfidfVectorizer* untuk menghitung bobot setiap *tweet*, kemudian *MultinomialNB* untuk melakukan perhitungan *Naïve Bayes* dan terakhir *roc_auc_score* untuk menampilkan informasi kinerja informasi dari algoritma *tweet* yang diklasifikasikan.

Dengan melakukan perhitungan akurasi terhadap kelas prediksi dan kelas actual sehingga didapatkan nilai akurasi 74.4%. Dengan rincian tabel hasil klasifikasi yang dapat dilihat pada Tabel berikut.

TABEL I
CONFUSION MATRIX TRAINING DATA

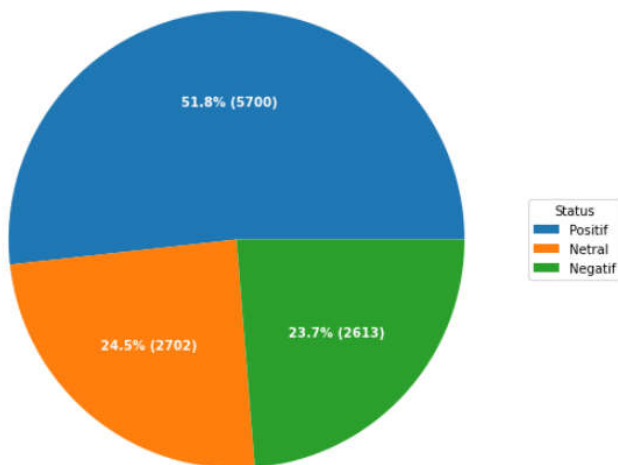
		<i>predicted</i>		
		netral (0)	positif (1)	negatif (2)
<i>actual</i>	netral (0)	15	3	5
	positif (1)	4	29	2
	negatif (2)	5	4	23

TABEL II
CLASSIFICATION REPORT TRAINING DATA

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>	<i>support</i>
0	0.62	0.65	0.64	23
1	0.81	0.83	0.82	35
2	0.77	0.72	0.74	32
<i>accuracy</i>			0.74	90
<i>macro avg</i>	0.73	0.73	0.73	90
<i>weighted avg</i>	0.75	0.74	0.74	90

Dapat dilihat pada Tabel 2 dengan hasil nilai presisi masing-masing setiment mendapatkan nilai presisi netral 62%, positif 81%, dan negatif 77%. Nilai recall netral 65%, positif 83% dan negatif 72%.

Setelah melakukan prediksi data *tweet* dengan jumlah 11.016 dan didapatkan hasil sentimen positif sebanyak 5700 atau 51.8%, Netral 2702 atau 24.5%, dan negatif sebanyak 2613 atau 23.7%, diagram hasil prediksi dapat dilihat pada Gambar 3.



Gbr 2 Hasil prediksi training data

Setelah didapatkan hasil prediksi kemudian data sentimen Netral dihapus supaya mendapatkan klasifikasi yang baik, karena analisis yang diperlukan hanya sentimen Positif dan Negatif.

Pada pengujian data *tweet* yang dilakukan terhadap 11.016 data prediksi berhasil mendapatkan nilai akurasi rata-rata yang cukup baik yaitu 87.9%, dapat diketahui bahwa hasil prediksi menggunakan Naïve Bayes telah dibuat mendapatkan nilai yang baik untuk memberikan sentimen pada data *tweet* yang disediakan. Mengenai rincian hasil klasifikasi data uji dapat dilihat pada Tabel 3 dan Tabel 4.

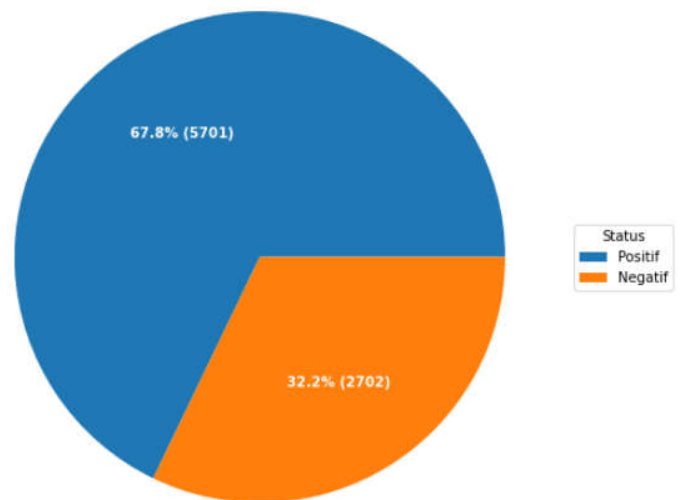
TABEL III
CONFUSION MATRIX TESTING DATA

<i>actual</i>	<i>predicted</i>	
	positif (1)	negatif (2)
positif (1)	1039	112
negatif (2)	91	439

TABEL IV
CLASSIFICATION REPORT TESTING DATA

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>	<i>support</i>
1	0.92	0.90	0.91	1151
2	0.80	0.83	0.81	530
<i>accuracy</i>			0.88	1681
<i>macro avg</i>	0.86	0.87	0.86	1681
<i>weighted avg</i>	0.88	0.88	0.88	1681

Dengan melakukan confusion matrix didapatkan hasil TP (True Positif) = 1039, FP (Fasle Positif) = 91, TN (True Negatif) = 439, FN (False Negatif) = 112. Dengan melakukan perhitungan prediksi pada *confussion matrix* rumus $TP/(TP+FP)$ kemudian menerapkan rumus tersebut pada data sentimen positif didapatkan perhitungan $1039/(1039+91)$ didapatkan hasil *precision* sentimen positif 0,92, menghitung nilai recall dengan rumus $TP/(TP+FN)$ dan jika dihitung pada sentimen positif maka $1039/(1039+112)$ maka didapatkan hasil *recall* sentimen positif 0.90 Kemudian hasil dari prediksi pada data *tweet* keseluruhan dan sudah diketahui data yang berlabel positif dan data yang berlabel negatif dengan rincian yang dapat dilihat pada Gambar 5.



Gbr 3 Histogram Sentimen Positif dan Negatif

Berdasarkan pada Gambar 5 hasil klasifikasi yang sudah dilakukan didapatkan data berlabel positif sebanyak 5701 data atau sebesar 67.8% dan data berlabel negatif 2702 data atau sebesar 32.2%. Berikut beberapa contoh *tweet* yang masuk dalam kategori positif dan negatif. Dapat dilihat pada Gambar 6 Sentimen positif pada klasifikasi ini membahas tentang beberapa hal seperti masyarakat yang turut berduka dan berbelasungkawa terhadap Eril sekeluarga, kebaikan Eril yang pernah ia lakukan selama hidupnya, dan beberapa mengenai Elektabilitas Ridwan Kamil yang tinggi.

```
0 pan jalin komunikasi dgn ridwan kamil salah ka...
1 kang emil jgn jadi dari naik elektabilitas mmg...
2 dapat amat peter gonta dorong pasang ganjarid...
3 jamiluddin amin elektabilitas ganjar ridwan ka...
4 ridwan kamil dinggap layak usung milik potensi...
...
5696 ridwan kamil jenazah eril indonesia minggu
5697 allahu akbar alhamdulillah ya allah jenazah er...
5698 antar istirahat bangga bahagia putera sayang s...
5699 kondisi jasad eril khan ridwan kamil
5700 innalillahi wainnailaihi rojiun duka cita insy...
Name: Tweet, Length: 5701, dtype: object
```

Gbr 4 Sentimen Positif

Sedangkan pada sentimen negatif data tweet beberapa membahas tentang hal-hal seperti eril hilang saja dibilang naikin elektabilitas, hal apa yang terjadi saat pemakaman Eril dan pemberitaan mengenai Ridwan Kamil elektabilitas Ridwan Kamil. Berikut beberapa contoh tweet dengan sentimen negatif dapat dilihat pada Gambar 7.

```
5701 jahat banget sm orang yang bilang kalau ridwan...
5702 eril tinggal saja netizen bdoh ngatain naikin ...
5703 eril hilang saja bilang naikin elektabilitas r...
5704 elektabilitas ridwan kamil dedi mulyadi is mou...
5705 kepala pin ridwan kamil ahaaa sy hanyut anak s...
...
8398 ridwan kamil jasad eril wangi syukur belai man...
8399 momen ridwan kamil adzankan eril telinga lahir
8400 kemendagri benar ridwan kamil cuti jemput jena...
8401 banget keluarga ridwan kamil ketemu udah tingg...
8402 momen haru video call ataliaridwan kamil mandi...
Name: Tweet, Length: 2702, dtype: object
```

Gbr 5 Sentimen Negatif

Berdasarkan hasil klasifikasi sentiment menggunakan Naïve Bayes pada data *tweet* mengenai Meninggalnya Almarhum Emmeril Kahn dan Elektabilitas Pemilu Ridwan Kamil. Banyak opini masyarakat mengenai Almarhum Eril dan Ridwan Kamil diklasifikasikan kedalam sentimen positif, dengan hasil data tersebut maka dapat digunakan untuk mengetahui *opini* dari *netizen* mengenai Emmeril Kahn dan Elektabilitas Ridwan Kamil.

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan penulis dengan judul Analisis Sentiment Twitter Terhadap Meninggalnya Almarhum Emmeril Kahn dan Pengaruhnya Terhadap Elektabilitas Pemilu Ridwan Kamil Menggunakan Naïve Bayes. Maka dapat ditarik kesimpulan bahwa : Setelah melakukan beberapa tahap

seperti Preprocessing data, Pelabelan data, perhitungan *TF-IDF* dan *Naïve Bayes*. Didapatkan hasil klasifikasi sentimen positif sebesar 67.8% dan sentimen Negatif sebesar 32.2%. Dengan melakukan klasifikasi terhadap data yang telah diprediksi menggunakan metode *Naïve Bayes* sehingga didapatkan nilai presisi masing-masing sentimen Positif sebesar 90%, nilai sentimen Negatif sebesar 80% dan Nilai recall masing-masing sentiment Positif 90% dan 83% untuk sentimen Negatif. Kemudian nilai akurasi prediksi sebesar 87.92%. Sehingga dapat diketahui bahwa pemodelan yang sudah dibuat cukup bagus. Pengaruhnya terhadap elektabilitas Pemilu Ridwan Kamil dapat diketahui dengan hasil analisis sentimen yang sudah didapatkan. Banyak terdapat opini positif mengenai Eril dan Ridwan Kamil, sehingga masyarakat bisa mengetahui lebih jauh bagaimana sosok Ridwan Kamil di mata masyarakat sehingga bisa memberikan pengaruh positif terhadap Elektabilitasnya.

Adapun saran yang dapat diberikan untuk pengembangan dari penelitian ini : Melakukan analisis terhadap beberapa kata yang diperlukan terhadap pengaruh Elektabilitas Pemilu Ridwan Kamil sehingga didapatkan hasil yang lebih baik dan sesuai. Melakukan preprocessing data dengan lebih baik sehingga didapatkan hasil preprocessing yang lebih bagus dan tidak ada kata-kata yang tidak diperlukan saat pemodelan, sehingga saat melakukan pengklasifikasian hasilnya lebih sempurna.

REFERENSI

- [1] P. Lestari. 2017. Kepemimpinan Transformatif Dalam Membangun Budaya Kewargaan : Studi Kepemimpinan Ridwan Kamil di Kota Bandung. Semarang : INTERALISTIK.
- [2] (2021) Kumparan Website [Online], <https://kumparan.com/kumparannews/ridwan-kamil-tak-ambil-pusing-kepastian-pilgub-2023-saya-percaya-takdir-tuhan-1v6BQqwaraz/full>, tanggal akses: 12 September 2022.
- [3] (2022) Zigi.id Website. [Online], <https://hits.zigi.id/profil-dan-biodata-emmeril-khan-mumtadz-agama-umur-pendidikan-ig-9690>, tanggal akses: 12 September 2022.
- [4] (2021) Livewire Website. [Online], <https://www.lifewire.com/what-exactly-is-twitter-2483331>, tanggal akses: 12 September 2022.
- [5] G. Berliana, Shaufiah, & S. Sa'adah. 2018. Klasifikasi Posting Tweet Mengenai Kebijakan Pemerintah Menggunakan Naïve Bayesian Classification. e-Proceeding of Engineering : Vol.5, No.1 Maret 2018 | Page 1562, ISSN : 2355-9365.
- [6] R. N. Devita, H. W. Herwanto, & A. P. Wibawa. 2018. Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa Indonesia. Jurnal Teknologi Informasi Dan Ilmu Komputer.