

Analisis Pilkada Medan pada Sosial Media Menggunakan Analisis Sentimen dan Social Network Analysis

M. Khairul Anam¹, Muhammad Bambang Firdaus², Triyani Arita Fitri³, Lusiana⁴, Wirta Agustin⁵, Agustin⁶

¹khairulanam@sar.ac.id, ²bambangf@fkti.unmul.ac.id, ³triyani@sar.ac.id, ⁴lusiana@sar.ac.id,

⁵wirtaagustin@stmik-amik-riau.ac.id, ⁶agustin@sar.ac.id

^{1,6}Teknologi Informasi, STMIK Amik Riau

²Informatika, Universitas Mulawarman

^{3,4}Teknik Informatika, STMIK Amik Riau

⁵Sistem Informasi, STMIK Amik Riau

Informasi Artikel

Diterima : 19 Nov 2021

Direview : 15 Dec 2021

Disetujui : 1 Mar 2022

Kata Kunci

Pilkada Medan, Support Vector machine (SVM), Naïve Bayes Classifier (NBC), Social Network Analysis, BOT

Abstrak

Pemilihan kepala daerah yang dilakukan secara serentak telah usai, namun selama kampanye sampai diputuskan menjadi kepala daerah meninggal banyak komentar baik pro maupun kontra. Kota Medan merupakan salah satu daerah yang menyelenggarakan PILKADA tahun 2020 dimasa pandemi. PILKADA Kota Medan telah memutuskan bahwa pasangan Bobby Nasution dan Aulia Rachman telah memenangkan. Kemenangan ini tentunya mendapatkan ranggam komentar pada media sosial khususnya Twitter. Penelitian ini melakukan analisis sentimen untuk melihat sentimen yang terjadi yaitu melihat komentar negatif, positif, maupun netral. Analisis sentimen ini menggunakan dua metode untuk melihat akurasi yang dihasilkan, yaitu Support Vector Machine (SVM) dan Naïve Bayes Classifier (NBC). Pada penelitian ini juga melihat interaksi yang terjadi menggunakan Social Network Analysis (SNA). Selain analisis sentimen dan SNA, penelitian ini juga melihat adanya akun BOT yang digunakan pada #PilkadaMedan. Hasil yang didapatkan pada sentimen analisis bahwa NBC merupakan akurasi tertinggi yakni 81, 72% dengan proporsi data 90:10. Kemudian pada SNA, akun @YanHarahap mendapatkan nodes tertinggi yakni 911 nodes. Kemudian dari 10326 tweet terdapat 11% yang dicurigai BOT oleh sistem DroneEmprit Academic.

Keywords

Medan Regional Head Election, Support Vector Machine (SVM), Naïve Bayes Classifier (NBC), Social Network Analysis, BOT

Abstrak

The simultaneous regional head elections were over, but during the campaign until it was decided to become regional head there were many comments, both pro and contra. The city of Medan is one of the regions that will hold the 2020 ELECTION during the pandemic. The Medan City Election has decided that the pair Bobby Nasution and Aulia Rachman have won. This victory certainly gets a variety of comments on social media, especially Twitter. This study conducts sentiment analysis to see the sentiment that occurs, namely seeing negative, positive, or neutral comments. This sentiment analysis uses two methods to see the resulting accuracy, namely Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC). This study also looks at the interactions that occur using Social Network Analysis (SNA). In addition to sentiment analysis and SNA, this study also looks at the existence of BOT accounts used in the #PilkadaMedan. The results obtained from the sentiment analysis show that NBC has the highest accuracy, which is 81, 72% with a data proportion of 90:10. Then on SNA, the @YanHarahap account got the highest nodes, namely 911 nodes. Then from 10326 tweets, 11% were suspected of being BOT by the DroneEmprit Academic system.

A. Pendahuluan

Pilkada tahun 2020 telah dilakukan secara serentak di Indonesia di beberapa daerah. Salah satu daerah yang mengadakan pilkada tahun 2020 adalah Kota Medan. Kota Medan merupakan salah satu Kota terbesar di Indonesia yang berada di provinsi Sumatra Utara. Kota Medan memiliki penduduk kurang lebih 2.210.624 orang [1]. Pilkada tahun 2020 di Medan tahun ini sedikit berbeda dengan pilkada sebelumnya, karena diselenggarakan pada saat pandemi covid 19.

Kota Medan telah memiliki pemenang dalam pilkada 2020, yakni Bobby Nasution dan Aulia Rachman. Namun pemilih kota Medan hanya sekitar 46% dari seluruh total pemilih di kota Medan [2]. Padahal saat ini Indonesia telah berupaya meningkatkan jumlah partisipasi masyarakat melalui UU nomor 7 tahun 2017 dan melakukan penyuluhan mengenai pilkada kepada masyarakat [3][4]. Penyuluhan yang dilakukan saat ini tidak hanya dilakukan secara langsung tetapi juga dilakukan melalui internet dengan memanfaatkan sosial media seperti facebook, instagram, twitter dan lain sebagainya [5]. Selain sebagai media penyuluhan media sosial juga digunakan sebagai media kampanye politik maupun pemilihan umum [6][7].

Sosial media merupakan salah satu media yang dianggap paling ideal untuk melakukan kegiatan politik [8][9]. Media sosial media Twitter, Facebook, dan Twitter mampu menciptakan masyarakat atau netizen terlibat dalam politik online [10]. Berbagai postingan dan komentar di sosial media saat ini menjadi perhatian khusus bagi peneliti. Banyak peneliti menjadikan media sosial sebagai bahan penelitiannya seperti sentimen analisis [11], Social Network Analysis [12], clustering [13], dan lain sebagainya.

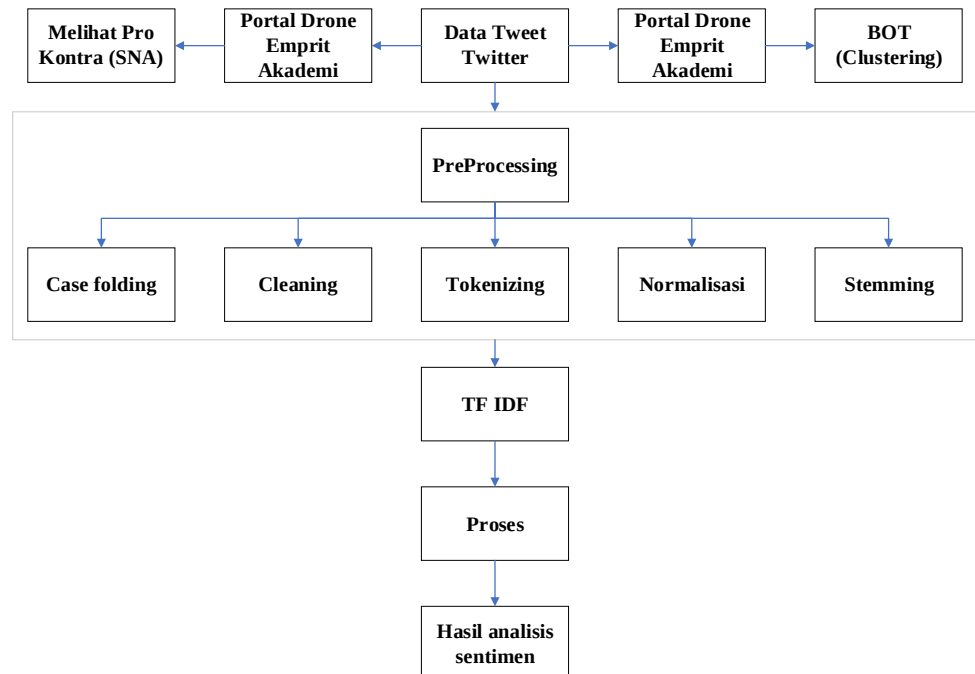
Pada penelitian yang dilakukan akan melakukan sentimen analisis terhadap pilkada Medan tahun 2020. Data pilkada yang diambil merupakan tweet atau retweet dari sosial media twitter. Media sosial twitter digunakan karena sebuah peristiwa dapat menjadi sebuah fenomena yang viral apabila mendapat banyak perhatian dari pengguna media sosial twitter [14]. Kemudian setelah berita viral di twitter barulah berita tersebut ke media sosial lain dan portal berita. Selain itu data twitter juga menjadi favorit para peneliti dalam melakukan analisa sentimen [15][16][17][18][19][20][21].

Analisis sentimen merupakan salah satu cara untuk mengukur akurasi dalam menentukan sentimen. Ada banyak metode yang digunakan untuk mengukur akurasi pada sentimen analisis, diantaranya *Naïve Bayes Classifier* [22], *Support Vector Machine (SVM)* [23], *Decision Tree* [24], *Fined-Grained* [25], *K-nearest neighbors algorithm* [26], dan lain sebagainya. Pada penelitian ini, membandingkan 2 metode yaitu *Support Vector Machine (SVM)* dan *Naïve Bayes Classifier (NBC)*. Kedua metode ini sering digunakan dan hasil akurasinya cukup tinggi. Penelitian terkait dengan SVM sering kali akurasi yang didapatkan termasuk tinggi yaitu akurasi 90% keatas [23][27][28][29]. Kemudian NBC juga memiliki akurasi yang cukup tinggi juga yaitu diatas 90% [30][31][32]. Alasan digunakan kedua metode ini adalah akurasi yang cukup tinggi dari penelitian sebelumnya.

Selain melakukan analisis sentimen penelitian ini juga melihat pro kontra yang terjadi pada pilkada Medan menggunakan *Social Network Analysis* melalui portal drone empirit akademi. Penelitian ini juga melihat penggunaan bot pada pilkada Medan di media sosial twitter menggunakan portal yang sama.

B. Metode Penelitian

Dalam melakukan penelitian dibutuhkan alur penelitian yang digunakan untuk mempermudah dalam pengerjaannya. Gambar 1 merupakan alur penelitian yang digunakan pada penelitian ini.



Gambar 1. Alur Penelitian

Berikut ini merupakan penjelasan dari gambar 1.

1. Analisis Sentimen

a. Data Tweet Twitter

Data pada penelitian ini diambil dari proses pembentukan hastag yaitu #pilkdamedan pada drone emprit akademi. Data yang diperoleh yaitu 10271, diambil dari 1 januari 2020 sampai 31 desember.

b. PreProcessing

Preprocessing data pada penelitian dilakukan menggunakan 5 tahap, yaitu:

1. Case folding merupakan proses dalam text preprocessing yang dilakukan untuk menyeragamkan karakter pada data [33].
2. Cleaning yaitu membersihkan text dari tanda baca atau HTML tag jika hasil crawling dari website [34].
3. Tokenisasi merupakan proses pemotongan string input berdasarkan kata yang menyusunnya serta membedakan karakter-karakter tertentu yang dapat diperlakukan sebagai pemisah kata atau bukan [35].
4. Normalisasi merupakan proses dimana mengubah kata yang tidak sesuai dengan EYD menjadi sesuai dengan EYD [36].
5. Stemming merupakan proses mengubah kata menjadi bentuk dasarnya [37].

c. TF-IDF

Term Frequency – Inverse Document Frequency (TF-IDF) merupakan perhitungan atau pembobotan kata melalui teknik tokenisasi, stopwords, dan stemming, dan frekuensi munculnya kata dalam dokumen yang diberikan

menunjukkan pentingnya kata itu di dalam sebuah dokumen [38]. Pada algoritma TF-IDF digunakan rumus untuk menghitung bobot (W) masing-masing dokumen terhadap kata kunci dengan rumus yaitu [39]:

$$W_{dt} = tf_{dt} * id_{ft} \quad (1)$$

Dimana:

W_{dt} = bobot dokumen ke-d terhadap kata ke-t

tf_{dt} = banyaknya kata yang dicari pada sebuah dokumen

id_{ft} = Inversed Document Frequency ($\log(N/df)$)

N = total dokumen

df = banyak dokumen yang mengandung kata yang dicari.

- d. Pada proses pembentukan analisis sentimen menggunakan 2 metode. Metode yang digunakan adalah *Naïve Bayes Classifier* dan *Support Vector Machine (SVM)*.
- e. Setelah dilakukan proses analisis sentimen, maka menghasilkan akurasi dari masing-masing metode.

2. SNA

Proses SNA pada penelitian ini digunakan untuk melihat pro-kontra yang terjadi sepanjang tahun 2020. Proses SNA dilihat dari portal drone empirit academic (dea.uui.ac.id). Hastag #pilkadamedan dibuat pada akhir tahun 2019 agar data yang diperoleh dimulai dari awal tahun 2020 sampai akhir 2020. Hal ini dilakukan untuk melihat mulai dari proses kampanye masing-masing paslon walikota Medan sampai terpilihnya salah satu paslon menjadi walikota dan wakil walikota Medan.

3. Clustering

Untuk melihat BOT yang ada pada penelitian ini, dilakukan dengan cara clustering yang menggunakan drone empirit academic. Proses ini dilakukan karena Menkominfo mengatakan bahwa pada pilkada 2020 banyak terdapat BOT yang membuat berita hoax [40].

C. Hasil dan Pembahasan

Pada tahap ini, akan membahas hasil yang diperoleh dari pengujian yang dilakukan menggunakan beberapa aplikasi.

1. Analisis Sentimen

Pada pembahasan analisis sentimen, menggunakan dua metode pengujian, masih metode akan dibagi menjadi 3 proporsi data, yakni 90:10, 80:20, 70:30.

a. Support Vector Machine (SVM)

Berikut ini merupakan hasil dataset dengan perbandingan 90:10.

Tabel 1. Hasil 90:10 SVM

Data Aktual	Data Prediksi			<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
	Neg	Net	Pos			
Neg	79	2	12	0.87	0.85	0.86
Net	6	10	11	0.77	0.37	0.50

Pos	6	1	70	0.75	0.91	0.82
-----	---	---	----	------	------	------

Pada Tabel 2 dijelaskan pada percobaan pertama digunakan *dataset* secara random diperoleh data pelatihan 90% dan data pengujian 10% yang akan diklasifikasikan menggunakan model pembelajaran SVM dengan menggunakan *confusion matrik*.

Tabel 2. Confusion matrix

Data Aktual	Data Prediksi		
	Negatif	Netral	Positif
Negatif	TNg	NgN	FN
Netral	NNg	TN	NP
Positif	FP	PN	TP

Pada Tabel 2 dijelaskan *Confusion matrix* merupakan hasil prediksi menggunakan mesin klasifikasi SVM yang diukur performa dari tiap-tiap kelas dengan cara menghitung *precision*, *recall*, dan *F1-score*, keterangan:

TNg = Kelas kata negatif terprediksi negatif

NNg = Kelas kata negatif terprediksi netral

FN = Kelas kata negatif terprediksi kata positif

NgN = Kelas kata netral terprediksi kata negatif

TN = Kelas kata terprediksi netral

NP = Kelas kata netral terprediksi sebagai kata positif

FP = Kelas kata positif terprediksi negatif

PN = Kelas kata positif terprediksi netral

TP = Kelas kata prediksi benar bernilai positif

Dari perhitungan manual tersebut sesuai dengan hasil yang diberikat oleh bahasa pemrograman python pada percobaan pertama ini dengan proporsi data 90:10.

	precision	recall	f1-score	support
0	0.87	0.85	0.86	93
1	0.77	0.37	0.50	27
2	0.75	0.91	0.82	77
accuracy			0.81	197
macro avg	0.80	0.71	0.73	197
weighted avg	0.81	0.81	0.80	197

SVM Accuracy --> 80.71065989847716

SVM Recall --> 80.71065989847716

SVM Precision --> 84.59966732440942

Gambar 2. Hasil SVM dengan proporsi data 90:10

Kemudian adalah tabel 2 merupakan hasil dari proporsi data 80:20 dengan *confusion matrix*.

Tabel 3. Hasil 80:20 SVM

Data Aktual	Data Prediksi			Precision	Recall	F1-Score
	Neg	Net	Pos			
Neg	150	3	20	0.84	0.87	0.85
Net	11	19	26	0.73	0.34	0.46
Pos	17	4	144	0.87	0.87	0.81

Dari hasil tersebut, sesuai dengan yang diuji cobakan pada bahasa pemrograman python.

	precision	recall	f1-score	support
0	0.84	0.87	0.85	173
1	0.73	0.34	0.46	56
2	0.76	0.87	0.81	165
accuracy			0.79	394
macro avg	0.78	0.69	0.71	394
weighted avg	0.79	0.79	0.78	394

SVM Accuracy --> 79.44162436548223
 SVM Recall --> 79.44162436548223
 SVM Precision --> 83.4961600267049

Gambar 3. Hasil SVM dengan proporsi data 80:20

Dari hasil kedua percobaan diatas, didapatkan hasil diatas 78%. Dibawah ini merupakan hasil percobaan ke tiga dengan proporsi data 70:30.

Tabel 4. Hasil 70:30 SVM

Data Aktual	Data Prediksi			Precision	Recall	F1-Score
	Neg	Net	Pos			
Neg	229	7	25	0.85	0.88	0.86
Net	15	28	36	0.72	0.35	0.47
Pos	27	4	220	0.78	0.88	0.83

Gambar 4, merupakan hasil menggunakan bahasa pemrograman python dengan proporsi data 70:30.

	precision	recall	f1-score	support
0	0.85	0.88	0.86	261
1	0.72	0.35	0.47	79
2	0.78	0.88	0.83	251
accuracy			0.81	591
macro avg	0.78	0.70	0.72	591
weighted avg	0.80	0.81	0.79	591

SVM Accuracy --> 80.71065989847716
 SVM Recall --> 80.71065989847716
 SVM Precision --> 84.24560945641477

Gambar 3. Hasil SVM dengan proporsi data 70:30

Dari haril ketiga uji coba didapatkan bahwa data dengan proporsi data 90:10 dan 70:30 mendapatkan hasil 80,71% sedangkan 80:20 mendapatkan akurasi sebesar 79,44%.

b. Naïve Bayes Clasification (NBC)

Pada percobaan menggunakan Naïve bayes Clasification hanya menampilkan satu hasil akurasi tertinggi yaitu proporsi data 90:10. Gambar 2 merupakan hasi perhitungan menggunakan *Confusion matrix*.

Tabel 5. Hasil 90:10 NBC

Data Aktual	Data Prediksi			Precision	Recall	F1-Score
	Neg	Net	Pos			
Neg	81	44	52	81	44	52
Net	84	48	49	84	48	49

Pos	84	45	52	84	45	52
-----	----	----	----	----	----	----

Dari perhitungan manual, langkah selanjutnya adalah melihat hasil akurasi menggunakan bahasa pemrograman python.

```

precision    recall  f1-score   support

0           0.84      0.87      0.86         93
1           0.67      0.44      0.53         27
2           0.82      0.88      0.85         77

accuracy          0.82         197
macro avg         0.78      0.73      0.75         197
weighted avg      0.81      0.82      0.81         197

NBC Accuracy --> 81.7258883248731
NBC Recall --> 81.7258883248731
NBC Precision --> 80.99141489817136

```

Gambar 4. Hasil NBC dengan proporsi data 90:10

Gambar 4 merupakan *output* yang menunjukkan besaran akurasi model dalam memprediksi data *training*.

c. Kesimpulan analisis sentiment

Tabel 6 merupakan data yang digunakan pada setiap percobaan.

Tabel 6. Data percobaan

No	Data Asli	Data Pelatihan	Training	Testing
1.	1968	70:30	1377	591
2.	1968	80:20	1574	394
3.	1986	90:10	1771	197

Dari proporsi data percobaan yang dilakukan, didapatkan hasil akurasi yang berbeda-beda baik itu menggunakan Support Vector Machine (SVM) maupun Naïve Bayes Clasification (NBC). Tabel 7 merupakan rekap dari hasil percobaan menggunakan Support Vector Machine (SVM).

Tabel 7. Rekap hasil SVM

Support Vector Machine			
Accuracy	Recall	Preccision	F1-Score
80.71 %	80.71 %	84.24 %	0.81 %
79.44 %	79.44 %	83.49 %	0.79 %
80.71 %	80.71 %	84.59 %	0.81 %

Kemudian, pada tabel 8 merupakan hasil rekap NBC dari ketiga percobaan yang telah dilakukan.

Tabel 8. Rekap hasil NBC

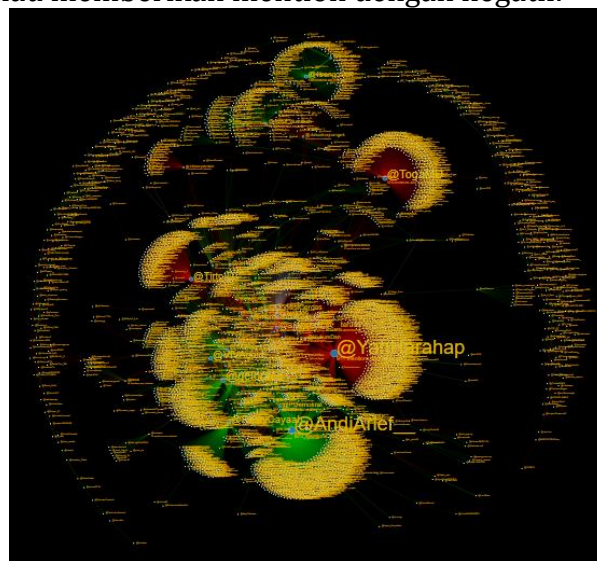
Naïve Bayes Clasification			
Accuracy	Recall	Preccision	F1-Score
80.54 %	80.54 %	79.73 %	0.81 %
80.20 %	80.20 %	79.36 %	0.80 %
81.72 %	81.72 %	80.99 %	0.82 %

Dari keenam percobaan yang dilakukan, hasil tertinggi didapatkan oleh akurasi dengan menggunakan NBC dengan proporsi data 90:10, yaitu akurasi sebesar 81.72%.

2. Social Network Analysis

Pada SNA yang terlihat pada gambar 5, terdapat 2 cluster yang berbeda. Pada cluster pertama (bagian bawah) terlihat ada 2 influencer yang begitu dominan, yaitu @YanHarahap dan @AndiArief. Kedua influencer ini merupakan kader partai demokrat yang merupakan pengusung salah satu calon walikota dan wakil walikota di Kota Medan. Namun pada kedua influencer ini sedikit berbeda pada tweet @AndiArief mendapat mention dari para followernya dengan positif, namun berbeda dengan @YanHarahap mendapatkan mention dari para followernya dengan negatif.

Kemudian pada cluster kedua (bagian atas) terlihat bahwa @togaMD mendominasi, tetapi tidak terlalu berpengaruh seperti kedua influencer di cluster pertama. Salah satu tweet yang menjadi populer saat itu adalah "Sudah banyak kali kutengok yg bahas Pilkada Medan ini. Sebagai orang Medan, ngomonglah dulu awak sikit, ya. Selain politik dinasti, mengapa tak kita bahas juga sikit "dinasti korupsi" di Medan? Bukan kaleng-kaleng itu, ah". Namun banyak para follower yang mengikuti beliau memberikan mention dengan negatif.



Gambar 5. Hasil SNA

Social Network Analysis (SNA) merupakan pemodelan terhadap user yang dilambangkan dengan titik (nodes) dan interaksi antar user tersebut yang dilambangkan dengan garis (edges) [41]. SNA membantu untuk memahami hubungan sosial yang melambangkan user dengan titik (nodes) dan hubungan antar user dilambangkan dengan garis (edges) pada Online Social Network (OSN). Pada tabel 8 merupakan 30 tertinggi dari nodes SNA influencer.

Tabel 8. Nodes SNA Influencer

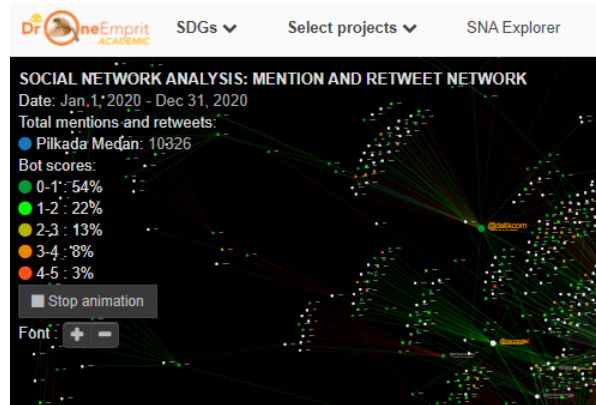
name	type	size
@YanHarahap	person	911
@AndiArief_	person	873
@TogaMD	person	315

@geloraco	person	307
@TirtoID	person	197
@VIVAcoid	person	185
@RestyCayaah	person	174
@twitpos	person	145
@SirianaGde	person	144
@1trenggalek	person	139
@MardaniAliSera	person	133
@OposisiCerdas	person	111
@_teruss	person	99
@merdekadotcom	person	93
@PDemokrat	person	88
@GeiszChalifah	person	79
@yusuf_dumdum	person	74
@HukumDan	person	70
@detikcom	person	60
@fadlizon	person	49
@SuruhMitra	person	49
@AgusYudhoyono	person	48
@musniumar	person	45
@datuakrajoangek	person	40
@CNNIndonesia	person	35
@oncomdei	person	36
@zoyamelani1	person	35
@demokrasicoid	person	34
@Pencerah__	person	32

3. BOT

Build-operate-transfer (bot) adalah sebuah program perangkat lunak yang bertindak secara otomatis yang menghasilkan pesan memungkinkan berinteraksi dengan pengguna manusia di platform media sosial [42]. Akun dapat dikatakan sebagai bot jika memiliki ciri-ciri yang dapat dideteksi secara manual [43] (Yang, Harkreader, dan Gu 2011) yaitu: jika sebuah kiriman berupa link menuju ke sebuah website tertentu, jika sebuah kiriman dengan ditambahkan sebuah foto dalam website agar lebih menarik pembaca untuk diaksesnya link tersebut, jika sebuah kiriman satu dengan kiriman yang lain tidak jauh berbeda, jika sebuah kiriman yang sedikit mendapatkan like, komentar, atau retweet karena kurang menarik pengguna lain dan serta jika kiriman yang banyak menggunakan tagar paling populer untuk mendapatkan perhatian pengguna lain. Tak hanya itu menurut [44] ciri-ciri akun bot juga dapat dikenali dari tanggal pembuatan akun yang relatif baru, banyak nama pengguna bot berisi angka, yang dapat menunjukkan pembuatan nama otomatis dan seringkali tidak ada biografi, atau memasang foto profil.

Pada gambar 6 terlihat bahwa ada sekitar 3% yang sangat dicurigai pada #pilkadamedan dari total 10326 tweet. Kemudian ada 8% masuk kategori dicurigai pada hastag ini. Hal ini membuktikan masih adanya kejahatan seperti mengirim hoak pada #pilkadamedan ini.



Gambar 6. Hasil Deteksi BOT

D. Simpulan

Setelah menyelesaikan serangkaian tahapan terhadap analisa data sentimen pada twitter dengan objek Pilkada medan, maka dapat diambil kesimpulan yaitu Hasil Naive Bayes Classifier mendapatkan hasil akurasi lebih tinggi dari Support Vector Machine. Hasil akurasi tertinggi yaitu 81,72% milik NBC dengan proporsi data 90:10. Hal ini karena dalam tahapan klasifikasiannya Naive Bayes Classifier memproses satu persatu data atribut. Berbeda dengan Support Vector Machine yang melakukan klasifikasi secara general sehingga cakupan Support Vector Machine lebih luas. Kesalahan klasifikasi pada SVM dapat terjadi dikarenakan dataset yang digunakan tidak dilakukan normalisasi terlebih dahulu sebelum dilakukan proses klasifikasi SVM.

Kemudian pada tahap SNA terdapat 3 akun paling berpengaruh yang terdapat pada 2 cluster. Cluster pertama terdapat akun influencer @YanHarahap dan AndiArief. Kemudian pada cluster ke 2 terdapat TogaMD yang paling berpengaruh. Pada #pilkadamedan juga masih terdapat 3% Tweet yang sangat dicurigai adalah bot, dan dicurigai sekitar 8% dari data tweet yang diolah menggunakan portal drone empirit.

Hasil pada penelitian ini belum sepenuhnya sempurna, masih bisa dilakukan pengembangan terhadap penelitian ini. Peneliti yang akan melanjutkan penelitian ini bisa menambahkan feature selection pada analisis sentimen sehingga bisa mendapatkan hasil akurasi yang lebih tinggi. Feature selection yang bisa digunakan bisa menggunakan PSO, Genetic Algorithm, Adaboost, dan lain sebagainya

E. Referensi

- [1] Admin, "Kependudukan," <https://pemkomedan.go.id/>, 2013. <https://pemkomedan.go.id/hal-kependudukan.html#> (accessed Aug. 26, 2021).
- [2] Z. Alfarabi, "ingkat Partisipasi Pemilih di Pilkada Medan 2020 46%, KPU: Di 2015 25%," *news.detik.com*, 2020. <https://news.detik.com/berita/d-5296958/tingkat-partisipasi-pemilih-di-pilkada-medan-2020-46-kpu-di-2015-25> (accessed Aug. 26, 2021).
- [3] M. I. Putra Alamsyah, M. A. Nasution, and R. H. Harahap, "Analisis Sosialisasi Politik Relawan Demokrasi Basis Warga Internet Kota Medan Pemilihan

- Umum Tahun 2019,” *Perspektif*, vol. 10, no. 1, pp. 65–75, 2021, doi: 10.31289/perspektif.v10i1.3923.
- [4] L. M. A. Sa’ban and A. A. M. Wijaya, “Kurangnya Partisipasi Politik Masyarakat Dalam Pemilihan Kepala Daerah Di Kabupaten Wakatobi,” *MEDIALOG J. Ilmu ...*, vol. I, no. II, pp. 35–43, 2018, [Online]. Available: <http://www.jurnal-umbuton.ac.id/index.php/Medialog/article/view/275>.
- [5] E. Nur, “Tanggapan Generasi Z Terhadap Pemanfaatan Media Sosial Dalam Mendukung Keterlibatan Dalam Pemilu Legislatif 2019,” *J. Penelit. Pers dan Komun. Pembang.*, vol. 24, no. 2, pp. 117–131, 2020, doi: 10.46426/jp2kp.v24i2.122.
- [6] Sartikasari, R. Hardi, and N. Hartaman, “Efektivitas Kampanye Politik Di Media Sosial Pada Pemilihan Kepala Daerah Tahun 2018 Di Kecamatan Sajoanging Kabupaten Wajo,” *KYBERNOLOGY J. Gov. Stud.*, vol. 1, no. 1, pp. 16–33, 2021, [Online]. Available: <https://journal.unismuh.ac.id/index.php/kybernology>.
- [7] B. Adinugroho, G. F. Prisanto, I. Irwansyah, and N. F. Ernungtyas, “Media Sosial Dan Internet Dalam Ketelibatan Informasi Politik Dan Pemilihan Umum,” *Representamen*, vol. 5, no. 02, 2019, doi: 10.30996/representamen.v5i02.2943.
- [8] Muzahid Akbar Hayat, Sjaiful Jayadiningrat, Gunawan Wibisono, and Muhammad Iwu Iyansyah, “Peran Media Sosial Dalam Komunikasi Politik,” *J. Indones. Sos. Teknol.*, vol. 2, no. 1, pp. 104–114, 2021, doi: 10.36418/jist.v2i1.61.
- [9] W. Y. Sembada and P. Sadjijo, “Strategi Pesan Aktor Politik dalam Media Sosial: Studi Kasus Akun Wali Kota Tangerang Selatan Banten,” *J. Komunikasi, Masy. dan Keamanan*, vol. 2, no. 1, pp. 48–65, 2020.
- [10] M. K. Anam, “Analisis Respons Netizen Terhadap Berita Politik Di Media Online,” *J. Ilm. Ilmu Komput.*, vol. 3, no. 1, pp. 14–21, 2017, doi: 10.35329/jiik.v3i1.62.
- [11] A. N. Ulfah and M. K. Anam, “Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM),” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 1, pp. 1–10, 2020, doi: 10.35957/jatisi.v7i1.196.
- [12] M. K. Anam, T. P. Lestari, L. B. Firdaus, and S. Fadli, “Analisis Kesiapan Masyarakat Pada Penerapan Smart City di Sosial Media Menggunakan SNA,” *J. RESTI (Rekayasa*, vol. 5, no. 1, pp. 69–81, 2021, doi: <https://doi.org/10.29207/resti.v5i1.2742>.
- [13] R. A. raffaidy Wiguna and A. I. Rifai, “Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining,” *J. Inf. Syst. Informatics*, vol. 3, no. 1, pp. 1–12, 2021, doi: 10.33557/journalisi.v3i1.78.
- [14] N. S. Fammy, A. Bambang, and D. T. Hapsari, “Fenomena Caption ‘Twitter, Please Do Your Magic!’ Membangkitkan Netizen Untuk Menolong,” *J. Digit. Educ. Commun. Arts*, vol. 3, no. 01, pp. 30–39, 2020, doi: 10.30871/deca.v3i01.1961.
- [15] K. Sailunaz and R. Alhaji, “Emotion and sentiment analysis from Twitter text,” *J. Comput. Sci.*, vol. 36, p. 101003, 2019, doi:

- 10.1016/j.jocs.2019.05.009.
- [16] M. Ghiassi and S. Lee, "A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach," *Expert Syst. Appl.*, vol. 106, pp. 197–216, 2018, doi: 10.1016/j.eswa.2018.04.006.
- [17] P. Sharma and A. K. Sharma, "Experimental investigation of automated system for twitter sentiment analysis to predict the public emotions using machine learning algorithms," in *Materials Today: Proceedings*, 2020, no. xxxx, doi: 10.1016/j.matpr.2020.09.351.
- [18] A. S. Neogi, K. A. Garg, R. K. Mishra, and Y. K. Dwivedi, "Sentiment analysis and classification of Indian farmers' protest using twitter data," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 2, p. 100019, 2021, doi: 10.1016/j.jjime.2021.100019.
- [19] S. Liu and J. Liu, "Public Attitudes toward COVID-19 Vaccines on English-language Twitter: A Sentiment Analysis," *Vaccine*, no. xxxx, 2021, doi: 10.1016/j.vaccine.2021.08.058.
- [20] M. Z. Ansari, M. B. Aziz, M. O. Siddiqui, H. Mehra, and K. P. Singh, "Analysis of Political Sentiment Orientations on Twitter," *Procedia Comput. Sci.*, vol. 167, pp. 1821–1828, 2020, doi: 10.1016/j.procs.2020.03.201.
- [21] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm," *Procedia Comput. Sci.*, vol. 161, pp. 765–772, 2019, doi: 10.1016/j.procs.2019.11.181.
- [22] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd. Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 2, no. 2, pp. 681–686, 2019, doi: <https://doi.org/10.24176/simet.v10i2.3487>.
- [23] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.
- [24] A. Harun and D. P. Ananda, "Analysis of Public Opinion Sentiment About Covid-19 Vaccination in Indonesia Using Naïve Bayes and Decission Tree Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve Bayes dan Decission Tree," *Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. April, pp. 58–63, 2021.
- [25] S. F. Pratama, R. Andrean, and A. Nugroho, "Analisis Sentimen Twitter Debat Calon Presiden Indonesia Menggunakan Metode Fined-Grained Sentiment Analysis," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 4, no. 2, p. 39, 2019, doi: 10.31328/jointecs.v4i2.1004.
- [26] N. T. Romadloni, I. Santoso, and S. Budilaksono, "Perbandingan Metode Naive Bayes , Knn Dan Decision Tree Terhadap Analisis Sentimen Transportasi Krl," *J. IKRA-ITH Inform.*, vol. 3, no. 2, pp. 1–9, 2019.
- [27] F. Romadoni, Y. Umaidah, and B. N. Sari, "Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 9, no. 2, pp. 247–253, 2020, doi: 10.32736/sisfokom.v9i2.903.
- [28] E. Indrayuni, "Analisa Sentimen Review Hotel Menggunakan Algoritma

- Support Vector Machine Berbasis Particle Swarm Optimization,” *J. Evolusi Vol. 4 Nomor 2 - 2016*, vol. 4, no. 2, pp. 20–27, 2016.
- [29] A. Z. Rasyida, I. D. Wijaya, and Y. Yunhasnawa, “Analisis Sentimen Kualitas Layanan Online Marketplace Di Indonesia Menggunakan Metode Support Vector Machine,” in *Seminar Informatika Aplikatif Polinema (SIAP) 2020*, 2020, pp. 70–75.
- [30] Muljono, D. P. Artanti, A. Syukur, A. Prihandono, and D. R. I. M. Setiadi, “Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes,” in *Konferensi Nasional Sistem Informasi 2018*, 2018, pp. 8–9.
- [31] Y. H. Kristian, K. R. Prilianti, and P. L. Tirma Irawan, “Implementasi Text Mining Untuk Analisis Preferensi Masyarakat Terhadap Tempat Wisata Di Indonesia,” *J. Simantec*, vol. 7, no. 2, pp. 72–81, 2019, doi: 10.21107/simantec.v7i2.6710.
- [32] P. Arsi, B. A. Kusuma, and A. Nurhakim, “Analisis Sentimen Pindah Ibu Kota Berbasis Naive Bayes Classifier,” *J. Inform. Upgris*, vol. 7, no. 1, pp. 1–6, 2021, doi: 10.26877/jiu.v7i1.7636.
- [33] F. S. Jumeilah, “Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 1, pp. 19–25, 2017, doi: 10.29207/resti.v1i1.11.
- [34] I. Subagyo, L. D. Yulianto, W. Permadi, and A. W. Dewantara, “Sentiment Analisis Review Film Di IMDB Menggunakan Algoritma SVM,” *J. Sist. Inf. DAN Teknol. Inf.*, vol. 8, no. 1, pp. 47–56, 2019.
- [35] D. I. Sari, Y. F. Wati, and Widiastuti, “Analisis Sentimen Dan Klasifikasi Tweets Berbahasa Indonesia Terhadap Transportasi Umum Mrt Jakarta Menggunakan Naïve Bayes Classifier,” *J. Ilm. Inform. Komput.*, vol. 25, no. 1, pp. 64–75, 2020, doi: 10.35760/ik.2020.v25i1.2427.
- [36] I. Najiyah and I. Hariyanti, “Sentimen Analisis Covid-19 Dengan Metode Probabilistic Neural Network Dan Tf-Idf,” *J. Responsif Ris. Sains dan Inform.*, vol. 3, no. 1, pp. 100–111, 2021, doi: 10.51977/jti.v3i1.488.
- [37] L. Ardiani, H. Sujaini, and T. Tursina, “Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak,” *J. Sist. dan Teknol. Inf.*, vol. 8, no. 2, pp. 44–51, 2020, doi: 10.26418/justin.v8i2.36776.
- [38] A. Ryansyah and S. Andayani, “Implementasi Algoritma TF-IDF Pada Pengukuran Kesamaan Dokumen,” *JuSiTik J. Sist. dan Teknol. Inf. Komun.*, vol. 1, no. 1, p. 53, 2017, doi: 10.32524/jusitik.v1i1.218.
- [39] R. Melita, V. Amrizal, H. B. Suseno, and T. Dirjam, “Penerapan Metode Term Frequency Inverse Document Frequency (Tf-Idf) Dan Cosine Similarity Pada Sistem Temu Kembali Informasi Untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Hadits Shahih Bukhari-Muslim),” *J. Tek. Inform.*, vol. 11, no. 2, pp. 149–164, 2018, doi: 10.15408/jti.v11i2.8623.
- [40] D. Muliya, “Menkominfo Temukan 47 Isu Hoax dalam Pilkada 2020, Polisi Tindaklanjuti Pelanggaran Tindak Pidananya,” *www.kompas.tv*, 2020. <https://www.kompas.tv/article/129659/menkominfo-temukan-47-isu-hoax-dalam-pilkada-2020-polisi-tindaklanjuti-pelanggaran-tindak-pidananya?page=all> (accessed Aug. 27, 2021).

-
- [41] M. K. Bratawisnu and A. Alamsyah, "Social Network Analysis Untuk Analisa Interaksi User Di Media Sosial Mengenai Bisnis E-Commerce (STUDI KASUS: LAZADA, TOKOPEDIA DAN ELEVENIA)," *J. Manaj. DAN BISNIS*, vol. 2, no. 2, pp. 107–115, 2018, [Online]. Available: <http://journal.unla.ac.id/index.php/almana/article/view/143>.
- [42] E. Kušen and M. Strembeck, "Something draws near, I can feel it: An analysis of human and bot emotion-exchange motifs on Twitter," *Online Soc. Networks Media*, vol. 10–11, pp. 1–17, 2019, doi: 10.1016/j.osnem.2019.04.001.
- [43] C. Yang, R. Harkreader, and G. Gu, "Die Free or Live Hard? Empirical Evaluation and New Design for Fighting Evolving Twitter Spammers," in *Sommer R., Balzarotti D., Maier G. (eds) Recent Advances in Intrusion Detection. RAID 2011. Lecture Notes in Computer Science*, 2011, vol. 6961, pp. 1280–1293, doi: 10.1007/978-3-642-23644-0_17.
- [44] Symantec Security Response, "How to Spot a Twitter Bot," <https://symantec-enterprise-blogs.security.com/>, 2018. <https://symantec-enterprise-blogs.security.com/blogs/election-security/spot-twitter-bot> (accessed Aug. 30, 2021).