



KLASIFIKASI ANALISIS SENTIMEN TERHADAP KASUS PENUSUKAN WIRANTO MEDIA SOSIAL TWITTER DENGAN MENGGUNAKAN METODE *NAÏVE BAYES CLASSIFICATION*

A. Yudi Permana¹, Dirga Nuansa Mahardhika²

Program Studi Teknik Informatika Fakultas Teknik Universitas Pelita Bangsa

¹yudi@pelitabangsa.ac.id

Abstraksi

Keberadaan twitter telah digunakan secara luas oleh berbagai lapisan masyarakat dalam beberapa tahun terakhir. Kebiasaan masyarakat memposting *tweet* untuk menilai kasus penusukan wiranto adalah salah satu media dalam merepresentasikan tanggapan masyarakat terhadap kasus penusukan tersebut. Maka dari itu dalam penelitian ini akan dilakukan analisis sentimen masyarakat terhadap kasus penusukan wiranto di yang diungkapkan melalui jejaring sosial *twitter*. Analisis dilakukan dengan menggunakan klasifikasi *tweet* yang berisi sentimen masyarakat terhadap kasus penusukan wiranto. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Naive Bayes Classification (NBC)*. NBC digunakan untuk mendapatkan klasifikasi respon positif dan negatif masyarakat di twitter dan mendapatkan nilai *preference value* dari masyarakat terhadap kasus penusukan. Hasil pengujian metode persentase data wiranto 25%, 50%, 75%, dan 100% jumlah data dari data training menghasilkan akurasi sebesar 64,67%, 70,57%, 87,56%, 97,50. Dan untuk hasil pengujian respon positif masyarakat di twitter dengan nilai *preference value* sebesar 53%. Dengan demikian klasifikasi sentimen dengan menggunakan metode naive bayes classification dapat digunakan untuk mengukur respon masyarakat terhadap kasus penusukan wiranto.

Kata Kunci : Twitter, naive bayes, analisa sentimen.

Abstract

The existence of Twitter has been widely used by various levels of society in the last few years. The habit of the public to post tweets to assess the wiranto stabbing case is one of the media in representing the public's response to the stabbing case. Therefore, in this study, an analysis of public sentiment will be carried out on the case of Wiranto di stabbing which was revealed through the Twitter social network. The analysis was carried out by using the tweet classification which contains the public's sentiment regarding the Wiranto stabbing case. The classification method used in this study is the Naive Bayes Classification (NBC). NBC was used to obtain a classification of positive and negative responses to the public on Twitter and to obtain the preference value from the community for stabbing cases. The results of testing the percentage of wiranto data methods of 25%, 50%, 75%, and 100% of the amount of data from the training data resulted in an accuracy of 64.67%, 70.57%, 87.56%, 97.50. And for the results of testing the positive response of the community on Twitter with a preference value of 53%. Thus, sentiment classification using the Naive

Bayes classification method can be used to measure the public response to the Wiranto stabbing case.

Keywords: Twitter, Naive Bayes, sentiment analysis.

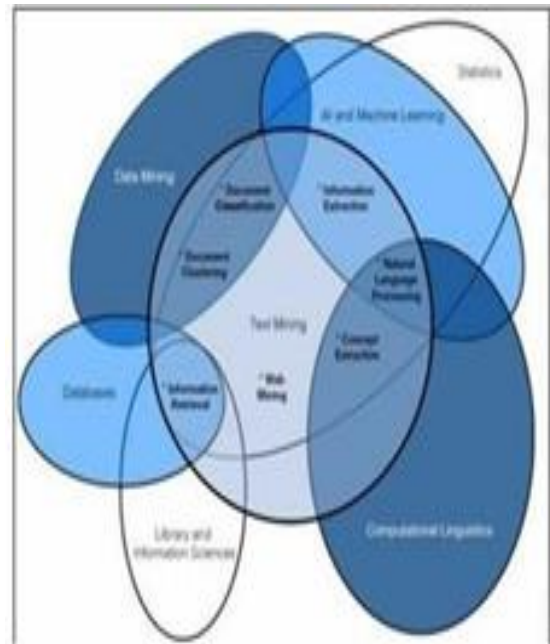
1. Pendahuluan

Di era digital ini media sosial merupakan salah satu tempat atau media untuk berekspresi dan berpendapat tentang berbagai macam topik. Sosial Media adalah satu set baru berkomunikasi dan alat kolaborasi yang memungkinkan banyak jenis interaksi yang tidak tersedia untuk orang biasa. Twitter sebagai salah satu situs microblogging mengalami banyak perkembangan, terhitung pada 21 Maret 2016 Twitter genap memasuki usianya yang ke 10. Media sosial ini secara global memiliki 332 juta pengguna bulanan dengan 500 juta kicauan dikirim setiap hari dan 200 miliar kicauan dalam setahun Indonesia adalah salah satu negara yang menganut sistem demokrasi. Hal ini ditandai dengan adanya kasus penusukan Wiranto. Menko Polhukam Wiranto ditusuk oleh seorang pria saat sedang menyapa masyarakat di Pandeglang, Banten. Wiranto disebut selama ini tidak mau dikawal ketat. Dalam peristiwa ini, anak buah Wiranto bernama Fuad juga terluka.

Sementara itu, Kapolsek Menes Kopol Daryanto juga terluka di bagian punggung akibat kejadian itu. Mudah-mudahan kepolisian segera mengungkap.

Perkembangan akses teknologi informasi yang semakin pesat pada era globalisasi saat ini dapat memberikan kemudahan untuk berkomunikasi lebih efektif dan efisien, contohnya adalah internet. Seperti halnya yang telah diungkapkan oleh Kementerian Komunikasi dan Informatika (kemenkominfo, 2017) bahwa pengguna internet di Indonesia saat ini mencapai 63 juta orang. Dari angka tersebut, terdapat 95 persennya menggunakan internet untuk mengakses media sosial. Salah satu media sosial yang diakses masyarakat adalah Twitter. Seperti akhir – akhir ini yang menjadi topic berita yaitu isu kasus penusukan Wiranto di Pandeglang, Banten. Memang semua yang berhubungan dengan kasus Wiranto menjadi banyak diperbincangkan di media sosial. Sehingga bermunculan pendapat positif maupun negatif. dari Dirga Nuansa Mahardhika 2 Jurnal Komputer dan Informatika. 23 Januari 2020 masyarakat terhadap isu tersebut. Pendapat-pendapat tersebut bisa diolah dan dianalisis agar menghasilkan data atau informasi yang sehingga informasi yang dihasilkan dapat membantu banyak pihak untuk mendukung suatu keputusan atau pilihan.

Salah satu teknik pemrosesan teks yang tepat yaitu analisa sentimen. Analisis sentimen adalah studi komputasi mengenai pendapat, perilaku dan emosi seseorang terhadap entitas. Entitas tersebut dapat menggambarkan individu, kejadian atau topik (Medhat & Korashy 2014). Media sosial yang dimaksud pada penelitian ini adalah twitter,” Twitter adalah layanan jaringan sosial dan mikroblog daring yang memungkingkan penggunaannya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter, yang dikenal dengan sebutan tweet (wikipedia)”. Tweet dapat dilihat secara publik, namun, penggunaannya dapat membatasi pengiriman status ke daftar teman yang diinginkan. Pengguna juga dapat melihat kicauan penggunaan lain, dan “#” bisa disebut hastag berfungsi sebagai topik pembahasan yang akan dibahas (Haewoon Kwak, et.al.2010). Dalam penelitian ini analisis sentiment dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek mengandung sentimen negatif atau positif menggunakan algoritma Naïve Bayes Classifier (NBC). Permasalahan yang dihadapi saat ini adalah bagaimana algoritma Naïve Bayes Classifier dapat melakukan klasifikasi sebuah tweet termasuk kedalam kelas sentimen negatif atau positif secara otomatis dan seberapa akurat algoritma Naïve Bayes Classifier dalam melakukan klasifikasi ke dalam sentimen negatif dan positif.



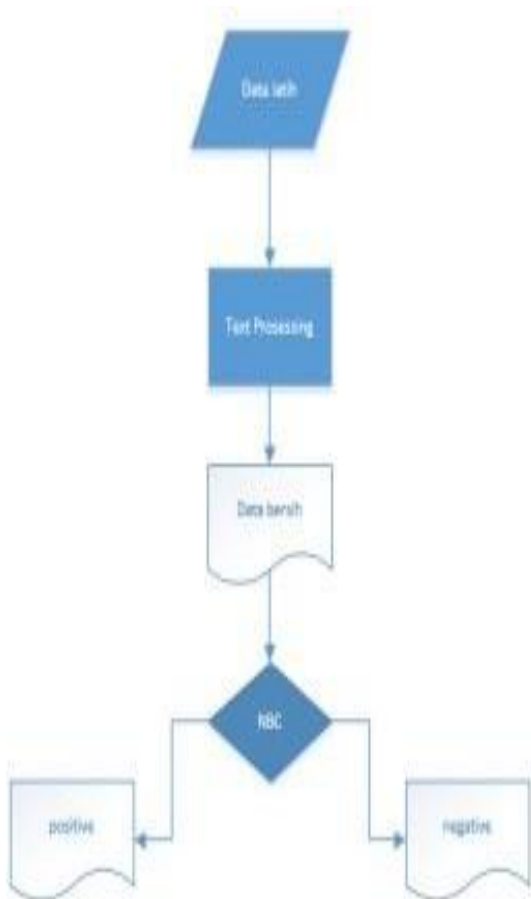
Gambar 1.1 Tahapan Text Mining

Dari uraian diatas maka dapat penulis simpulkan bahwa text mining adalah pengumpulan dokumen berupa text yang nantinya dokumen tersebut dapat diolah menjadi suatu informasi sesuai dengan tujuan pengelolaan dokumen tersebut.

2. Landasan Pemikiran

2.1 Naïve Bayes Classifier

Naïve Bayes Classifier (NBC) adalah metode classifier berdasarkan probabilitas dan teorema bayesian dengan asumsi bahwa setiap variabel X bersifat bebas (independence) (Buntoro, 2017), NBC merupakan metode klasifikasi statistik untuk memprediksi probabilitas suatu atribut dalam suatu kelas, dalam proses klasifikasinya menganggap bahwa kemunculan suatu kata pada suatu kalimat tidak dipengaruhi oleh kemunculan kata-kata yang lain pada kalimat tersebut, namun pada kenyataannya kemunculan suatu kata dipengaruhi oleh kata-kata lainnya, atau disebut dengan naïve (Susanti, 2016). NBC adalah metode dalam pembelajaran mesin, dimana pada proses klasifikasi terjadi dua tahap yaitu tahap penelitian dan klasifikasi. Tahap penelitian merupakan proses terhadap data sampel untuk menemukan term-term yang ada dalam suatu kelas dokumen, sedangkan tahapan klasifikasi yaitu tahapan penentuan kelas suatu dokumen dalam data uji berdasarkan data sampel, pengklasifikasian dokumen dalam data uji ditentukan oleh banyaknya probabilitas term-term yang muncul pada dokumen tersebut.



Gambar 2. Flowchart NBC

Evaluasi dilakukan menggunakan metode confusion matrix, confusion matrix, merupakan alat penting dalam metode visualisasi untuk menunjukkan performa akurasi, presisi dan recall, berikut adalah contoh tabel confusion matrix sesuai pada penelitian (Zy, 2017) :

Tabel 1. Confusion Matrix

		Actual class	
		Class-1	Class-2
Predicted class	Class-1	True	False
	positif	negative	
class	Class-2	False	True
	positive	negative	

Adapun dalam menghitung akurasi menggunakan rumus sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} * 100\% \quad (1)$$

$$Precision = \frac{TP}{TP + FP} * 100\% \quad (2)$$

$$Recall = \frac{TP}{TP + FN} * 100\% \quad (3)$$

Dalam menyelesaikan penelitian ini dibutuhkan sebuah kerangka pemikiran sebagai pedoman yang dilakukan secara konsisten.

Kerangka Konsep



Gambar 3. Kerangka Pemikiran

3. Metodologi Penelitian

Dalam penyusunan penelitian ini peneliti melakukan penelitian tentang Klasifikasi Analisis Sentimen Terhadap kasus penusukan Wiranto Menggunakan Naive Bayes Classifier, yang dimana pada penelitian ini dilakukan untuk mengetahui kinerja dari optimasi metode dan seleksi fitur yang digunakan serta mengetahui sentimen dari para netizen pada media sosial twitter 2019. Objek penelitian yang akan diteliti oleh peneliti ialah tentang sentimen dengan pengambilan datanya berdasarkan query (@wiranto) yang berasal dari media sosial twitter. Pemilihan objek didasari karena ingin mengetahui bagaimana pendapat masyarakat terhadap kasus penusukan wiranto, yang mana banyak muncul berbagai macam opini soal kasus penusukan ini pada media sosial twitter baik itu opini kepuasan ataupun ketidakpuasan.

Berikut adalah tahapan dalam penelitian ini :

1. Pengumpulan Data. Pada tahap ini mencari data yang tersedia, mengintegrasikan semua data kedalam data set, termasuk variabel yang diperlukan dalam proses.

2. Pengolahan Data Awal Ditahap ini dilakukan penyeleksian, pembersihan termasuk melihat data yang kosong kemudian merubah data yang diinginkan.
3. Metode Yang Diusulkan Pada tahap ini penganalisan data kemudian pengelompokan variabel yang saling berhubungan dengan yang lain, kemudian penerapan model yang sesuai data yang telah dibentuk.
4. Eksperimen dan Pengujian Metode Pada tahap ini penentuan model yang diusulkan ketika akan diuji dan melihat hasil rules yang dijadikan pengambilan keputusan.
5. Evaluasi dan Validasi Hasil Pada tahap ini melakukan hasil evaluasi yang didapat dari model yang ditetapkan sebelum dan melihat hasil akurasi dengan pengujian aplikasi terhadap metode yang digunakan.

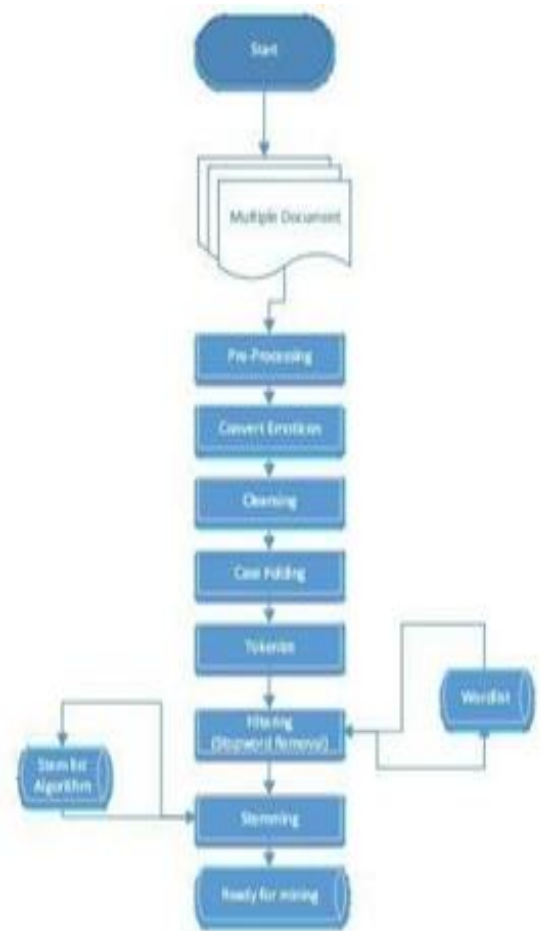
4. Pembahasan

Dari hasil data yang diperoleh kemudian dikategorikan dengan variabel yang ada dengan nilai rendah, sedang dan tinggi dan pengujian berdasarkan data training dan data testing.

Tabel 2. Data Training

id	id_tweet	text	category	sentiment	language	location	timestamp	source
1	111111
2	111112
3	111113
4	111114
5	111115
6	111116
7	111117
8	111118
9	111119
10	111120

Dari data training yang di dapatkan dari data tweet Ada enam tahapan yang penelitian gunakan dalam melakukan *preprocessing data* diantaranya akan dijelaskan pada gambar berikut:



Gambar 4. Tahapan Preprocessing

Setelah didapat dari data yang diperoleh dengan pengujian menggunakan *RapidMiner*. Berdasarkan gambar di atas, hasil yang diperoleh maka akan dilakukan fitur data seleksi terendah sampai tertinggi merupakan faktor utama yang dianggap lebih berpengaruh terhadap kasus penusukan wiranto Berikut adalah hasil *accuracy* data seleksi fitur terendah sampai tertinggi

Tabel 3. Akurasi Data Seleksi Fitur Persentase 25% Wiranto

	True positive	True negative	class precision
pred positif	110	0	72,67%
pred negatif	0	10	100,00%
class recall	100,00%	100,00%	

Dari gambar diatas dengan data persentase sebanyak 25% dari data latih maka didapatkan nilai akurasi sebesar 64,67%, nilai *precision* 56,56% dan recall 34,57%.

Tabel 4. Akurasi Data Seleksi Fitur
Persentase 50% Wiranto

accuracy: 81,67%			
	true positive	true negative	class precision
pred positive	10	0	73,17%
pred negative	0	15	100,00%
class recall	100,00%	63,33%	

Dari gambar diatas dengan data persentase sebanyak 50% dari data latih maka didapatkan nilai akurasi 81,67%, nilai *precision* 73,17% dan recall sebesar 63,33%.

Tabel 5. Akurasi Data Seleksi Fitur
Persentase 75% Wiranto

accuracy: 87,56%			
	true positive	true negative	class precision
pred positive	228	64	80,28%
pred negative	0	160	100,00%
class recall	100,00%	74,77%	

Dari gambar diatas dengan data persentase 75% dari data latih maka didapatkan nilai akurasi 87,56%, nilai *precision* 80,28%, dan nilai recall sebesar 74,77%.

Tabel 6. Akurasi Data Seleksi Fitur Persentase
100% Wiranto

accuracy: 97,50%			
	true negative	true positive	class precision
pred negative	257	1	99,61%
pred positive	14	327	95,89%
class recall	94,83%	99,70%	

Dari gambar diatas dengan data persentase 100% dari data latih maka didapatkan nilai akurasi sebesar 97,50%, nilai *precision* 95,89%, dan nilai recall sebesar 194,83%.

Tabel 7. Tabel Perbandingan Akurasi

Nilai	PERSENTASE DATA			
	25%	50%	75%	100%
Nilai akurasi	64,67	70,57	87,56	97,50
Akurasi	%	%	%	%
Recall	34,57	65,76	74,77	94,83
Recall	%	%	%	%
Precision	56,56	73,17	80,28	95,89
Precision	%	%	%	%

5. Penutup

1. Pengolahan data yang dilakukan pada pengujian ini adalah *naive bayes classifier*. Dan hasil yang diperoleh dari pengujian ialah penggunaan *naive bayes classifier* saja pada kalsifikasi mendapatkan model klasifikasi sentiment di twitter yang paling optimal sebesar 97,50% dari akun wiranto.
2. Hasil klasifikasi sentiment data *tweet* kasus penusukan wiranto yaitu dari akun wiranto dengan data testing masing-masing 600 data dengan komposisi 328 data tweet positive dan 272 data tweet negatif yang ditentukan secara manual, diperoleh jumlah data tweet dengan popularitas sentimen positif lebih banyak dibandingkan dengan data tweet dengan popularitas sentimen negatif. dan untuk menghitung nilai preference value dengan akurasi wiranto sebesar 97,50% dan tingkat respon positif masyarakat sebesar 53%

Daftar Pustaka

- [1] A. Novantirani1, M.K. Sabariah, and V. Effendy., 2015. "Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine".
- [2] Abdul, A.M. 2015. Penerapan Algoritma TF-IDF Untuk Pencarian Karya Ilmiah. Universitas Dian Nuswantoro. Semarang.
- [3] Buntoro, A. Ghulam, 2017. Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter. Universitas Muhammadiyah Ponorogo.
- [4] Charu, C. Aggarwal. 2015. Data Mining The TextBook. IBM T.J. Watson Research Center Yorktown Heights. New York USA.
- [5] Feldman, R and Sanger, J. 2007. The Text Mining Handbook: Advanced Approaches in

- Analyzing Unstructured Data. Cambridge University Press:NewYork.
- [6] Ginting, H. S., Lhaksana, K. M., & Murdiansyah, D. T. (2018). Klasifikasi Sentimen Terhadap Bakal Calon Gubernur Jawa Barat 2018 di Twitter Menggunakan Naive Bayes, 5(1), 1793–1802.
- [7] Hidayatullah, A. F., & SN, A. (2014). Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter. Seminar Nasional Informatika 2014, (August 2013), 0–8.
- [8] Kent, M. L. 2013. Public Relations Review Using social media dialogically.
- [9] Kumar, Lokesh and Bhatia, Parul Kalra. 2015. Text Mining: Concepts, Process And Application. Journal of Global Research in Computer Science. Volume 4, No. 3, March 2013.
- [10] Lesmana, P. I. (2013). Analisis Sentimen Pengguna Layanan Media Sosial Twitter Di Indonesia, Karya Akhir Ilmu Komputer. Jakarta: Universitas Indonesia.