



Analisis Sentimen *Tweet* KRI Nanggala 402 di Twitter menggunakan Metode *Naïve Bayes Classifier*

Muhammad Ariel Djamaludin ^{1*}, Agung Triayudi ², Eri Mardiani ³

^{1,2,3,4} Program Studi Informasi, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional.

article info

Article history:

Received 29 May 2021

Received in revised form

23 August 2021

Accepted 1 September 2021

Available online April 2022

DOI:

<https://doi.org/10.35870/jtik.v6i2.398>

Keywords:

Social Media; Sentiment analysis; KRI Nanggala 402; Naïve Bayes Classifier; Twitter.

Kata Kunci:

Analisis Sentiment; KRI Nanggala 402; Media Sosial; Naïve Bayes Classifier; Twitter.

abstract

Social media is one of the technological developments that has contributed greatly in making it easy for us to communicate and socialize, one of which is using Twitter social media. Twitter in this study is used as a data source to analyze *tweets* discussing KRI Nanggala 402. Analysis of KRI Nanggala 402 twitter sentiment is used to see the tendency of public responses to the sinking of the KRI Nanggala 402 submarine whether to give positive or negative opinions. This *Sentiment analysis* uses the *Naïve Bayes Classifier* method, which is a classification method. The first research stage is *crawling*, processing, classification, and evaluation. The classification stage is carried out after the processing phase, where the classification results tend to be positive or negative, using the *Naïve Bayes Classifier* method. The *accuracy* of the system in the *Sentiment analysis* of the KRI Nanggala 402 *tweet* is 73.00%.

abstrak

Media sosial adalah salah satu perkembangan teknologi yang memiliki andil besar dalam memberikan kemudahan bagi kita untuk berkomunikasi dan bersosialisasi, salah satunya menggunakan media sosial Twitter. Twitter dalam penelitian ini digunakan sebagai sumber data untuk menganalisa *tweet* yang membahas KRI Nanggala 402. Analisis sentiment twitter KRI Nanggala 402 digunakan untuk melihat kecenderungan respon masyarakat pada tenggelamnya kapal selam KRI Nanggala 402 apakah memberikan opini positif atau negatif. Analisa sentiment ini menggunakan metode *Naïve Bayes Classifier* yang merupakan metode pengklasifikasian. Tahap penelitian pertama yaitu *crawling*, processing, klasifikasi, dan evaluasi. Tahap klasifikasi dilakukan setelah fase processing, dimana hasil klasifikasi berkecenderungan positif atau negatif, menggunakan metode *Naïve Bayes Classifier*. Akurasi system pada analisis sentiment terhadap *tweet* KRI Nanggala 402 adalah 73.00%.

Author. Email: m.arieldjamiludin@gmail.com^{1}, agungtriyudi@civitas.unas.ac.id², erimardiani1@gmail.com³.

1. Latar Belakang

KRI Nanggala (402) merupakan kapal selam kedua dalam jenis kapal selam kelas Cakra. KRI Nanggala berada di bawah kendali Satuan Kapal Selam Komando Armada RI Kawasan Timur. Kapal ini merupakan kapal kedua yang menyandang nama Nanggala dalam jajaran TNI AL dan termasuk dalam armada pemukul TNI Angkatan Laut. Nama Nanggala berasal dari nama senjata pewayangan Nanggala. KRI Nanggala-402 hilang kontak pada Rabu, 21 April 2021 saat melakukan latihan penembakan torpedo di Laut Bali. Saat itu, kapal selam ini membawa 53 awak. Nanggala kemudian dinyatakan tenggelam pada Sabtu, 24 April 2021 oleh TNI AL setelah ditemukannya puing-puing yang diduga berasal dari kapal selam tersebut.

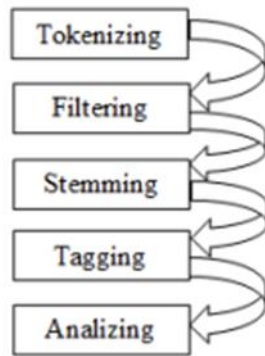
Penggunaan media sosial telah menyebar luas dan sangat pesat dalam berbagai macam lapisan masyarakat. Tidak hanya sebagai media untuk bersosialisasi, dan berkomunikasi namun juga untuk menyampaikan aspirasi dan mempresentasikan tentang apa saja yang terjadi dan dirasakan masyarakat [1]. Banyak masyarakat yang mengutarakan pendapatnya melalui media sosial Twitter dalam mengungkapkan kesedihan dan turut berduka cita atas tenggelamnya kapal selam KRI Nanggala 402. Hal ini dapat dijadikan sebagai sebuah acuan untuk mengetahui sentimen masyarakat terhadap tenggelamnya kapal selam KRI Nanggala 402.

Berdasarkan latar belakang permasalahan yang dihadapi yaitu; Bagaimana cara mengklasifikasikan Analisis Sentiment *Tweet* KRI Nanggala 402 menggunakan metode *Naïve Bayes Classifier*?. Penelitian ini membatasi permasalahan, yaitu; 1) Penelitian ini hanya untuk menilai kepuasan seseorang melalui opini berdasarkan dokumen teks, dan 2) Pengerjaan penelitian ini terbatas pada beberapa data yang dapat diambil dari Twitter.

Di dalam penelitian ini, akan dibahas tahapan yang dilalui untuk melakukan proses analisis sentimen. Dimulai dari tahap *pre-processing* (pelabelan, tokenisasi dan pembobotan), kemudian tahap analisis sentimen bagaimana mengukur kualitas hasil analisis menggunakan beberapa parameter seperti *accuracy*, *precision* dan *recall*.

Beberapa penelitian sebelumnya yang digunakan sebagai referensi. Penulis mengacu pada jurnal penelitian pertama Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode *Naïve Bayes Classifier* dapat disimpulkan bahwa tingkat keakuratan dalam metode ini sangat stabil dan memiliki nilai akurasi yang sangat baik [2]. Pada penelitian penelitian kedua Analisis Sentimen Opini Publik Mengenai Covid-19 pada Twitter Menggunakan Metode *Naïve Bayes* Dan KNN, dapat disimpulkan bahwa opini public pada Covid-19 cenderung positif dan diketahui Metode *Naïve Bayes* lebih tinggi akurasinya dibandingkan K-NN [3]. Pada penelitian ketiga Analisis Sentimen dan Klasifikasi *Tweets* Berbahasa Indonesia Terhadap Transportasi Umum MRT Jakarta Menggunakan *Naïve Bayes Classifier* ini, dapat disimpulkan bahwa penelitian ini dapat meningkatkan sistem pada MRT Jakarta, baik dalam meningkatkan layanan maupun fasilitas agar menarik masyarakat menggunakan MRT Jakarta sebagai alat transportasi [4]. Pada penelitian keempat ini tentang Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen di Twitter dapat disimpulkan bahwa ketersediaan sumber dokumen tekstual menjadi sangat besar dan mudah diakses. Salah satu contohnya yaitu Twitter, kita dapat membandingkan beberapa metode dengan sebuah cuitan atau *tweets* dengan cara menganalisa polaritas sentimennya [5]. Pada penelitian kelima *Naïve Bayes Algorithm for Sentiment analysis Windows PhoneStore Application Reviews* dapat disimpulkan bahwa, menganalisa pengguna aplikasi Windows Phone Store dengan mengklasifikasikan review ke dalam kategori opini positif atau negative [6].

Text Mining adalah salah satu bidang khusus dalam data mining yang memiliki definisi menambang data berupa teks dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen. Disisi lain teks mining dapat bekerja dalam komputer dengan tujuan mengolah informasi lama secara eksplisit sehingga menghasilkan temuan informasi baru [3]. Berikut tahapan pra-pemrosesan dalam Text Mining antara lain:



Gambar 1. Tahapan Text Mining

- 1) Tokenisasi
Tahap pemotongan string input berdasarkan tiap kata yang menyusunnya.
- 2) Filtering
Tahap pengambilan kata-kata yang penting dari hasil token. Bisa menggunakan algoritma stop list (membuang kata yang kurang penting) atau word list (menyimpan kata yang penting).
- 3) *Stemming*
Tahap mencari akar kata hasil filtering.
- 4) Tagging
Tahap mencari bentuk awal dari kata lampau hasil *Stemming*
- 5) Analisa
Tahap penentuan seberapa jauh keterhubungan antar kata-kata antar dokumen yang ada.

Sentiment analysis atau analisis sentimen dalam Bahasa Indonesia adalah sebuah teknik atau cara yang digunakan untuk mengidentifikasi bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa dikategorikan sebagai sentimen positif maupun sentimen negative [7,8]. *Sentiment analysis* atau *opinion mining* mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistic dan teks mining yang memiliki tujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu [9,10]. Dari pendapat tersebut, bisa diambil kesimpulan bahwa analisis sentiment adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen positif atau negatif.

Twitter adalah layanan jejaring sosial dan mikroblog daring yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter [11,12], akan tetapi pada tanggal 07 November 2017 bertambah hingga 280 karakter yang dikenal dengan sebutan kicauan (*tweet*) [13]. Twitter didirikan pada bulan Maret 2006 oleh Jack Dorsey, dan situs jejaring sosialnya diluncurkan pada bulan Juli. Sejak diluncurkan, Twitter telah menjadi salah satu dari sepuluh situs yang paling sering dikunjungi di *Internet* [14]), dan dijuluki dengan pesan singkat dari *Internet*. Di Twitter, pengguna tak terdaftar hanya bisa membaca kicauan, sedangkan pengguna terdaftar bisa menulis kicauan melalui antarmuka situs *web*, pesan singkat (SMS), atau melalui berbagai aplikasi untuk perangkat seluler. Berikut ini adalah beberapa istilah yang dikenal dalam Twitter:

- 1) *Mention*. *Mention* adalah menyebut atau memanggil pengguna Twitter lain dalam sebuah *tweet*. Mention dilakukan dengan menuliskan '@' diikuti dengan nama pengguna lain.
- 2) *Hashtag*. *Hashtag* digunakan untuk menandai sebuah topik pembicaraan di Twitter. Penulisan hashtag dimulai dengan tanda '#' diikuti dengan topik yang sedang dibahas. Hashtag biasa digunakan untuk meningkatkan visibilitas *tweet* pengguna.
- 3) *Emoticon*. *Emoticon* adalah ekspresi wajah yang direpresentasikan dengan kombinasi antara huruf, tanda baca dan angka. Pengguna biasa menggunakan *emoticon* untuk mengekspresikan mood yang sedang mereka rasakan.
- 4) *Trending topics*. Jika hashtag adalah cara untuk menandai sebuah topik pembicaraan di Twitter, maka *trending topics* adalah kumpulan dari topik pembicaraan yang sangat populer di Twitter.

Naïve Bayes merupakan algoritma pembelajaran yang sering digunakan untuk mengatasi masalah klasifikasi teks, merupakan salah satu metode pembelajaran mesin yang menggunakan perhitungan probabilitas [6]. Ciri utama dari *Naïve Bayes Classifier* ini adalah asumsi yang sangat kuat (naif) terhadap tingkat independensi dari masing-masing kondisi atau kejadian. Terdapat dua tahap klasifikasi dokumen *tweet* pada penelitian ini. Tahap pertama adalah proses training terhadap dokumen yang sudah diketahui kategorinya. Sedangkan tahap kedua adalah proses testing yaitu mengklasifikasikan dokumen yang belum

diketahui kategorinya. Terlepas dari kesederhanaannya, pengklasifikasi Naive Bayesian sering kali bekerja dengan sangat baik dan banyak digunakan karena sering mengungguli metode klasifikasi yang lebih canggih [15]. Dengan metode naïve bayes ini pengguna dapat mencapai probabilitas Bayesian dan bisa membuat pernyataan probabilitas. Naive bayes dapat di formulasikan menjadi persamaan 1.

$$P(H|X) = \frac{P(H|X)P(H)}{P(X)} \quad (1)$$

Naive Bayes Classifier ini menggunakan model *classifier probabilistic* yang juga termasuk salah satu algoritma yang dapat mengklarifikasi sebuah data

2. Metode Penelitian

Penelitian ini dilakukan dengan menggunakan acuan dari beberapa modul besar, yaitu pengumpulan data, processing, klasifikasi dan evaluasi.

Pengumpulan Data

Pengumpulan data merupakan proses untuk memperoleh dokumen yang akan digunakan sebagai input ke sistem. Pada penelitian ini menggunakan dokumen *tweet* yang membahas mengenai kapal selam KRI Nanggala 402. Hal ini diketahui dari penggunaan tagar yang berkaitan dengan KRI Nanggala 402 seperti #PrayForKRINanggala402, #KRINanggala402. Data yang terkumpul kemudian diberikan label kelas yaitu positif atau negatif oleh para ahli.

Pre-processing

Pada tahap ini data akan di klasifikasikan untuk tingkat akurasi ulasan perangkat lunak. Beberapa teknik yang digunakan untuk processing adalah normalisasi, case folding, tokenisasi, stopwords removal, *Stemming*.

Klasifikasi

Klasifikasi menggunakan *Naive Bayes Classifier* dan dapat di klasifikasikan ke dalam kategori tertentu berdasarkan kata-kata yang terkandung dalam dokumen tersebut.

Evaluasi

Evaluasi performansi dilakukan untuk menguji hasil klasifikasi dengan mengukur nilai kebenaran dari sistem. Parameter yang digunakan untuk mengukur nilai kebenaran yaitu *accuracy*. *Accuracy* adalah persentase dokumen yang berhasil diklasifikasikan dengan tepat oleh sistem. Perhitungan *accuracy* memanfaatkan metode confusion matrix. Perhitungan *accuracy* menggunakan persamaan 2.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (2)$$

3. Hasil dan Pembahasan

Hasil penelitian ini ditampilkan sesuai dengan urutan penelitian. Urutan proses penelitian ini adalah pengumpulan data, *Pre-processing* data, klasifikasi, dan evaluasi.

Pengumpulan Data

Data yang berhasil diperoleh adalah 53 data. Hasil dari data tersebut menghasilkan 7 *tweet* dengan sentiment positif, 7 *tweet* dengan sentiment negative dan 39 *tweet* dengan *sentiment neutral*. Contoh dokumen *tweet* yang berhasil dikumpulkan pada gambar 2.

Row No.	sentiment	text	id	Score
14	negative	dgfbl_god U-Boat KRI Nanggala 402' becau..	1405729755..	-0.193
15	negative	dgfbl_god U-Boat KRI Nanggala 402' was cu..	1405720202..	0.940
16	negative	dgfbl_god U-Boat KRI Nanggala 402' was des..	1405729268..	-1.128
17	negative	dgfbl_god Kriben Jen Uda knows who U-Bo..	1405729674..	-0.282
18	neutral	@Kemhan_RI Meman RI Prabowo Ucapan Te..	1405718070..	0
19	positive	@Kemhan_RI Penghargaan yang sebesar2nya..	1405503627..	0.205
20	positive	@Kemhan_RI Al. R. mengimban 2 Kapal Beac..	1405481504..	0.590
21	neutral	@Kemhan_RI Ucapan terkakasih tersebut bel..	1405502802..	0
22	neutral	I just saw Memhan RI Ucapan Terimakasih Men..	1405505492..	0
23	positive	Penghargaan yang sebesar2nya dan mandala..	1405501586..	0.205
24	positive	Al. R. mengimban 2 Kapal Beacawank mem..	1405481504..	0.590
25	neutral	Ucapan terkakasih tersebut betau sampikan ..	1405521508..	0

Gambar 2. Hasil Data *Crawling* Twitter

Pre-processing Data

Dari hasil *crawling* dilanjutkan ke tahap *pre-processing*, terlihat pada gambar 4.

Row No.	Text	Sentiment
1	Seeh: banget????? jadi inget tet nanggapi 402????? Click to sort bayangin sekeh apa f...	neutral
2	Tak campai di situ, gerakan ini juga memberikan bantuan pendidkan bag anak anak dari k...	neutral
3	RT @angishkasa... Mereka tidak hilang	neutral
4	RT @DamaAja... https://t.co/9k1x4f1td	neutral
5	RT @2020yflr: I got goosebumps while watched this list video. Ga kebayang sedalem apa f...	negative
6	RT @SubmarRuslan: KRI NANGGALA 402 https://t.co/0H44uVykH	neutral
7	https://t.co/9k1x4f1td	neutral
8	RT @Leonia_lesan: ADAM KAPAL SELAM KRI NANGGALA 402	positive
9	digital_god U-Boat_KR: Nanggala 402" because of those who ignore me	negative
10	digital_god U-Boat_KR: Nanggala 402" was quicker destroy to the deadly Altan on massa...	negative
11	digital_god U-Boat_KR: Nanggala 402" was destroyed and to bein consider to be raised in	negative
12	digital_god Krystan Jan Duda knows why U-Boat_KR: Nanggala 402" was destroyed	negative

ExampleSet (53 examples, 0 special attributes, 2 regular attributes)

Gambar 3. Tahap processing data

Klasifikasi

Tahap akhir yang dilakukan setelah *pre-processing* adalah klasifikasi untuk menentukan data yang diuji itu sentimen positif atau negative. Pada tahap ini digunakan algoritma Naïve Bayes. Hasil klasifikasi dikategorikan berkecenderungan positif atau negatif berdasarkan perbandingan nilai posterior probabilitas, apabila perhitungan nilai posterior probabilitas positif lebih besar dari nilai posterior probabilitas negatif maka dihasilkan nilai klasifikasi dengan kecenderungan positif, begitu juga sebaliknya.[4]

Row No.	sentiment	prediction(s...	confidence(...	confidence(...	confidence(...	Text
1	neutral	neutral	0.365	0.317	0.317	RT @Kemha...
2	positive	neutral	0.365	0.317	0.317	Penghargaan...
3	neutral	neutral	0.365	0.317	0.317	RT @bem_ta...
4	negative	neutral	0.365	0.317	0.317	RT @Domini...
5	neutral	neutral	0.365	0.317	0.317	RT @tweetm...
6	neutral	neutral	0.363	0.315	0.322	RT @Submar...
7	negative	neutral	0.363	0.315	0.322	After the previ...
8	neutral	neutral	0.363	0.315	0.322	Setelah gug...
9	neutral	neutral	0.363	0.315	0.322	Tragedi KRI...
10	neutral	neutral	0.363	0.315	0.322	RT @prabow...
11	negative	neutral	0.366	0.317	0.317	digital_god U...
12	positive	neutral	0.366	0.317	0.317	RT @Kemha...

ExampleSet (53 examples, 5 special attributes, 1 regular attribute)

Gambar 4. Tahap klasifikasi program

accuracy: 73.00% +/- 16.66% (micro average: 73.58%)

	true neutral	true negative	true positive	class precision
pred. neutral	39	7	7	73.58%
pred. negative	0	0	0	0.00%
pred. positive	0	0	0	0.00%
class recall	100.00%	0.00%	0.00%	

Gambar 5. Hasil pengujian klasifikasi.

4. Kesimpulan

Penelitian ini dibuat untuk melihat bagaimana reaksi masyarakat atas tenggelamnya kapal selam KRI Nanggala 402. Bisa dilihat dari data, masyarakat

cenderung bereaksi netral dan sentiment positif dan negative sama. Proses analisa dilakukan dengan metode Naïve Bayes untuk klasifikasi dokumen *tweet*. Metode Naïve Bayes merupakan model yang akurat dengan nilai *accuracy* sebesar 73.00%. Sehingga, penulis berharap penelitian ini dapat menambah wawasan masyarakat umum.

5. Daftar Pustaka

- [1] Ansari, M.Z., Aziz, M.B., Siddiqui, M.O., Mehra, H. and Singh, K.P., 2020. Analysis of political sentiment orientations on twitter. *Procedia Computer Science*, 167, pp.1821-1828.
- [2] Astari, N.M.A.J., Divayana, D.G.H. and Indrawan, G., 2020. Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Dan Informatika (JSI)*, 15(1), pp.27-29.
- [3] Syarifuddin, M., 2020. Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes dan Knn. *Inti Nusa Mandiri*, 15(1), pp.23-28.
- [4] Sari, D.I., Wati, Y.F. and Widiastuti, W.W., 2020. Analisis Sentimen Dan Klasifikasi Tweets Berbahasa Indonesia Terhadap Transportasi Umum Mrt Jakarta Menggunakan Naïve Bayes Classifier. *Jurnal Ilmiah Informatika Komputer*, 25(1), pp.64-75.
- [5] Hadna, N.M.S., Santosa, P.I. and Winarno, W.W., 2016. Studi literatur tentang perbandingan metode untuk proses analisis sentimen di Twitter. *Semin. Nas. Teknol. Inf. dan Komun*, 2016, pp.57-64.
- [6] P. Jurnal and R. T. Informatika, 2020. Algoritma Naïve Bayes Untuk Analisis Sentimen Ulasan Aplikasi Windows Phone Store, vol. 3, no. April 2019, pp. 1–13.
- [7] Buntoro, G.A., 2017. Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter. *INTEGER: Journal of Information Technology*, 2(1).

- [8] Liu, B., 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), pp.1-167.
- [9] Vinodhini, G. and Chandrasekaran, R.M., 2012. Sentiment analysis and opinion mining: a survey. *International Journal*, 2(6), pp.282-292.
- [10] Cambria, E., Schuller, B., Xia, Y. and Havasi, C., 2013. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent systems*, 28(2), pp.15-21.
- [11] Ravi, K. and Ravi, V., 2015. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-based systems*, 89, pp.14-46.
- [12] El Husna, R., 2020. *Analisis Klasifikasi Sentimen Pada Twitter Mengenai Netflix Yang Diblokir Oleh Telkom Menggunakan Naïve Bayes Classifier Dan Support Vector Machine* (Doctoral dissertation, Muhammadiyah University, Semarang).
- [13] Rohmah, N.N.M., 2020. Media Sosial sebagai Media Alternatif Manfaat dan Pemuas Kebutuhan Informasi Masa Pandemi Global Covid 19 (Kajian Analisis Teori Uses and Gratification). *Al-I'lam: Jurnal Komunikasi dan Penyiaran Islam*, 4(1), pp.1-16.
- [14] Faradhillah, N.Y., Kusumawardani, R.P. and Hafidz, I., 2016. Eksperimen Sistem Klasifikasi Analisa Sentimen Twitter pada Akun Resmi Pemerintah Kota Surabaya Berbasis Pembelajaran Mesin. *SESINDO 2016, 2016*.
- [15] Wagh, B., 2016. SJV, and WNR, “. *Sentimental Analysis on Twitter Data using Naive Bayes,*” *IJARCCCE*, 5(12), pp.316-319.