Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

Analisis Sentimen Media Sosial Twitter Terhadap Calon Presiden RI Tahun 2024 Menggunakan Klasifikasi Algoritma Naïve Bayes

Muhammad Makmun Effendi, Ahmad Turmudi Zy*, Asep Arwan

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Kabupaten Bekasi, Indonesia
Jl. Inspeksi Kalimalang No.9,Cibatu, Rumbai, Kabupaten Bekasi, Jawa Barat, Indonesia
Email: [1]effendiyan@pelitabangsa.ac.id, [2,*]turmudi@pelitabangsa.ac.id, [3]aseparwan@pelitabangsa.ac.id
Email Penulis Korespondensi: turmudi@pelitabangsa.ac.id
Submitted: 22/05/2024; Accepted: 31/05/2024; Published: 31/05/2024

Abstrak−Kemajuan media sosial saat ini sudah banyak dirasakan oleh masyarakat Indonesia saat ini, salah satu media sosial yang sering digunakan adalah Twitter yang merupakan media posting penyampaian informasi. Saat ini *posting* yang viral adalah pemilihan Calon Presiden (capres) Republik Indonesia yang akan dilaksanakan tahun 2024, seiring hal tersebut maka dari Komisi Pemilihan Umum (KPU) menyelenggarakan debat capres yang diselenggarakan diberbagai media televisi di Indonesia dan dari hasil debat tersebut masyarakat Indonesia biasanya memberikan opini ataupun komentar terhadap debat tersebut dari sisi positif dan negatif terhadap calon presiden yang tampil saat itu yaitu Anis, Prabowo dan Ganjar Pranowo. Untuk mengetahui hasil sentimen terhadap capres tersebut maka peneliti melakukan analisis dengan menggunakan klasifikasi tweet yang berisi sentimen masyarakat terhadap pencalonan presiden 2024 yaitu Anis, Prabowo dan Ganjar dengan metode klasifikasi yang digunakan dalam penelitian ini adalah *Naive Bayes Classification(NBC)*. dataset anies baswedan 61,35% pengguna *twitter* berkomentar negatif dan 39,65% pengguna twitter berkomentar positif, dataset Ganjar Pranowo 59,12% pengguna twitter berkomentar negatif dan 41,88% pengguna *twitter* berkomentar positif, dataset Ganjar Prabowo Subianto 49,25% pengguna *twitter* berkomentar negatif dan 51,75% pengguna *twitter* berkomentar positif. Perbandingan hasil dari ketiga kandidat calon presiden, nilai akurasi anies baswedan lebih kecil dari pada kedua kandidat lainnya karena komentar negatif dari anies baswedan lebih banyak daripada kedua kandidat lainnya. Anies baswedan mendapatkan nilai *akurasi* 67,23%, prabowo subianto 83,42% dan ganjar pranowo sebesar 88,15%, Banyaknya data mempengaruhi hasil dari analisa sentimen, semakin banyak data akan semakin baik nilai akurasi yang didapat.

Kata Kunci: Twitter; Naive Bayes; Analisa Sentimen; KPU; Pemilihan Umum; Capres

Abstract−The progress of social media is currently being felt by many Indonesian people, one of the social media that is often used is Twitter, which is a media for posting information. Currently the viral post is the election of Presidential Candidates (capres) of the Republic of Indonesia which will be held in 2024, in line with this, the General Election Commission (KPU) is holding a presidential candidate debate which will be held on various television media in Indonesia and from the results of this debate the Indonesian people usually give opinions or comments on the debate from the positive and negative sides of the presidential candidates who appeared at that time, namely Anis, Prabowo and Ganjar Pranowo. To find out the results of sentiment towards the presidential candidates, the researchers carried out an analysis using a classification of tweets containing public sentiment towards the 2024 presidential candidacy, namely Anis, Prabowo and Ganjar with the classification method used in this research is Naive Bayes Classification (NBC). Anies Baswedan dataset 61.35% of Twitter users have negative comments and 39.65% of Twitter users have positive comments, Ganjar Pranowo dataset 59.12% of Twitter users have negative comments and 41.88% of Twitter users have positive comments, Ganjar Prabowo Subianto dataset 49.25% Twitter users commented negatively and 51.75% of Twitter users commented positively. Comparing the results of the three presidential candidates, Anies Baswedan's accuracy value is smaller than the other two candidates because Anies Baswedan has more negative comments than the other two candidates. Anies Baswedan got an accuracy value of 67.23%, Prabowo Subianto 83.42% and Ganjar Pranowo 88.15%. The amount of data affects the results of sentiment analysis, the more data the better the accuracy value obtained.

Keywords: Twitter; Naive Bayes; Sentiment Analysis; KPU; General Election; Presidential Candidate

1. PENDAHULUAN

Indonesia merupakan salah satu negara yang menganut sistem demokrasi. Hal ini ditandai dengan diadakannya suatu pemilihan umum terhadap presiden dan wakil presiden setiap 5 tahun sekali secara periodik. Pemilu atau singkatan dari Pemilihan Umum yang merupa proses demokratis untuk memilih wakil rakyat atau pejabat pemerintahan secara langsung oleh warga negara suatu negara.Pemilihan Umum merupakan mekanisme penting dalam sistem demokrasi modern yang memungkinkan rakyat untuk berpartisipasi dalam menentukan pemimpin dan kebijakan negara. Seorang tokoh politik yang ingin maju sebagai calon presiden tentu akan melihat atau mempertimbangkan popularitas mereka berdasarkan opini dari masyarakat. Pada tahun 2024 Indonesia akan mengadakan pemilihan presiden, supaya masyarakat Indonesia dapat mempertimbangkan pilihannya sebelum diselenggarakan Pemilihan Umum(Pemilu) maka penyelenggara pemilu yaitu Komisi Pemilihan Umum (KPU) mengadakan acara Debat Presiden di televisi nasional, untuk mengetahui sentimen positif maupun negatif terhadap calon presiden dari masyarakat maka salah satu media sosial yang saat ini banyak digunakan masyarakat untuk menyampaikan opininya melalui Twitter [1].

Seorang tokoh politik yang ingin maju sebagai calon presiden tentu akan melihat atau mempertimbangkan popularitas mereka berdasarkan opini dari masyarakat. Pada tahun 2024 Indonesia akan mengadakan pemilihan presiden, supaya masyarakat Indonesia dapat mempertimbangkan pilihannya sebelum diselenggarakan

Copyright © 2024 **Author**, Page 739
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

Pemilihan Umum(Pemilu) maka penyelenggara pemilu yaitu Komisi Pemilihan Umum (KPU) mengadakan acara Debat Presiden di televisi nasional, untuk mengetahui sentimen positif maupun negatif terhadap calon presiden dari masyarakat maka salah satu media sosial yang saat ini banyak digunakan masyarakat untuk menyampaikan opininya melalui Twitter

Pemilu di Indonesia dilakukan secara serentak diseluruh provinsi yang ada di Indonesia sehingga partisipasi masyarakat harus meningkat, dan jangan lagi ada kotak suara yang kosong ataupun tidak hadir, menurut peneliti sebelumnya melakukan penelitian untuk meningkatkan partisipasi Pemilu 2024 dengan menggunakan pendekatan *Stakeholder Mapping Analysis* sebagai langkah strategi untuk meningkatkan partisipasi politik masyarakat dalam Pemilu 2024. Beberapa rekomendasi antara lain, KPU mensosialisasikan Pemilu 2024 dengan metode dan media sosial yang sering digunakan masyarakat agar muncul kesadaran berpartisipasi pemilu pada masyarakat [2].

Banyaknya opini publik warganet Indonesia terkait bakal calon presiden Indonesia 2024 di media sosial Twitter sulit dianalisis. Penentuan klasifikasi oleh manusia membutuhkan tenaga ahli dan waktu lama. Oleh karena itu, diperlukan sistem visualisasi klasifikasi opini publik secara real-time [3]. Pengguna di Indonesia masih cukup tinggi, mencapai 24 juta pengguna. Dalam peringkat dunia, Indonesia menempati peringkat kelima pengguna Twitter, Dilansir dari Twitter The World Ranking @worldranking_, peringkat pertama ditempati oleh Amerika Serikat dengan 95,4 juta. "Pengguna Twitter," katanya, dikutip Selasa (30/5/2023) terbanyak di dunia.

Calon Presiden Tahun 2024 di Indonesia adalah Anis Bawesdan dengan nomor urut 1 , Prabowo Subianto dengan nomor urut 2 sedangkan Ganjar Pranowo dengan nomor urut 3. Twitter merupakan wadah masyarakat menyampaikan opini tersebut oleh sebab itu maka salah satu metode untuk mengetahui sentimen positif maupun negatif tiap calon presiden maka salah satu metode yang digunakan adalah algoritma Naïve bayes. Naïve bayes merupakan salah satu algoritma algoritma dengan teknik klasifikasi yang mencari nilai probabilitas tertinggi menerapkan teorema bayes, algoritma ini memiliki kelebihan yaitu kesederhanaan dalam komputasi untuk menangani data dengan jumlah yang besar. penelitian ini untuk mengklasifikasikan text bahwa data tweet masuk kedalam kategori sentimen positif atau negatif sehingga dapat mengetahui nilai accuracy, precision, dan recall [4].

Analisis sentimen sangat diperlukan sekali dalam memonitoring media sosial. Terlebih lagi dalam memonitor sentimen masyarakat mengenai pilpres 2024. Di media sosial, analisis sentimen sangat berguna supaya mendapatkan gambaran umum dari suatu opini yang berasal dari masyaarakat. Pemilihan presiden 2024 akan dilaksanakan pada tanggal 14 Februari , namun sampai penelitian ini dibuat sudah banyak sekali bermacam-macam sentimen dari masyarakat baik itu sentimen positif maupun negatif untuk para calon presiden Indonesia. Tetapi bagi masyarakat yang umum terhadap politik maupun yang tidak mengikuti perkembangan di media sosial, mereka masih belum menemukan pilihan presiden yang tepat di hati untuk memimpin negara Indonesia, sehingga diperlukan analysis ini untuk membantu masyarakat dalam menentukan pilihannya.

Untuk menentukan masa depan suatu bangsa maka Pemilu merupakan salah satu alat yang digunakan untuk memilih Calon Kepala pemerintahan dan negara dan juga wakil rakyat yang duduk menjadi anggota legislatif. Diharapkan juga peran aktif dari masyarakat untuk ikut serta dalam proses pemilu ini karena hak setiap warga negara adalah memilih dan dipilih dalam Pemilu yang dilaksanakan tiap 5 tahun sekali di Indonesia.

Peneliti sebelumnya melalui proses analisis sentimen Terhadap Pilpres 2024 Berdasarkan Opini Dari Twitter Menggunakan Naïve Bayes memiliki nilai accuracy 96.63%., precision 94.30%, dan recall 98.90% sedangkan jumlah sentimen positif sebanyak 1004 text dan untuk sentimen negatif sebanyak 1102 text [5]. Peneliti selanjutnya menghasilkan Analisis sentimen calon presiden Indonesia 2019 dari media sosial twitter menggunakan metode Naive Bayes mendapatkan nilai akurasi yaitu 64,6% untuk paslon 01, dan 58% untuk paslon 02 pada pengujian 3 kelas. Sedangkan pengujian 2 kelas 77,7% untuk paslon 01 dan 84% untuk paslon 02 [6].Penelitian yang lainnya dengan menggunakan metode Naïve Bayes Classifier untuk mengklasifikasikan data ke dalam kategori yang tepat, dimana data akan dibagi menjadi tiga sentimen yaitu positif, negatif dan netral. Dan Bag of Words digunakan sebagai metode untuk ekstraksi fitur. Hasil dari proses klasifikasi lebih dominan label positif dengan data yang digunakan sebanyak 300 data dengan rasio pembagian data 70% (210 data) untuk data latih dan 30% (90 data) buat data uji. Lalu dilakukan pengujian sehingga menghasilkan accuracy 75.56%, precision 75.56% dan recall 100.00% untuk calon presiden Ganjar Pranowo [7]. Untuk Penelitian selanjutnya menghasilkan Algoritma naïve bayes classifier dapat digunakan untuk klasifikasi sentiment berbahasa Indonesia menjadi 3 kategori yaitu sentiment positif, negative, dan netral. Dalam melakukan pengujian menggunakan 2000 tweet dapat dilihat pada hasil pengujian sistem didapatkan hasil akurasi 95,71%, precision 97,99%, recall 96,29% dan nilai F-1 97,01% [8].

2. METODOLOGI PENELITIAN

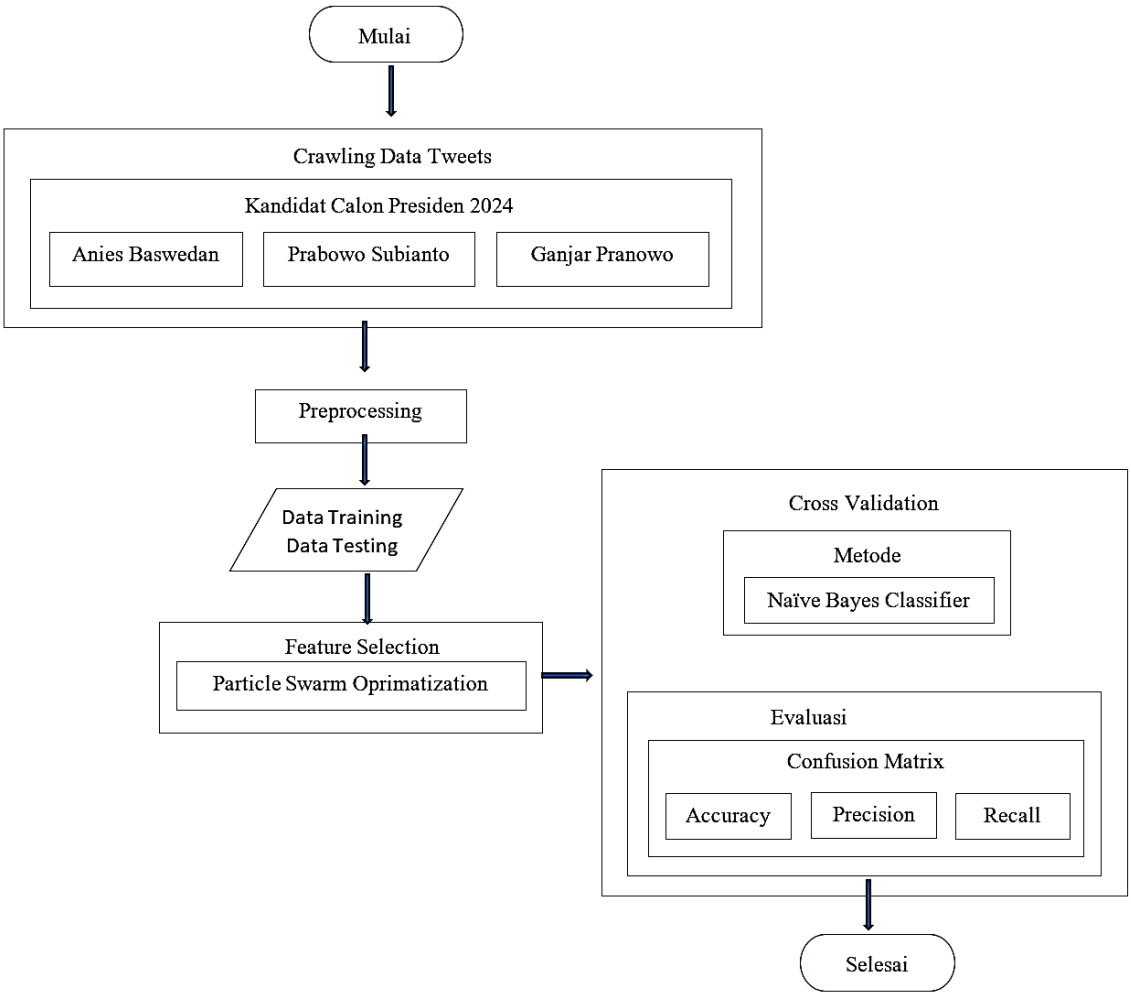
2.1 Tahapan Penelitian

Penelitian ini menggunakan metode kualitatif, metode kualitatif merupakan memiliki sifat deskriptif dan cenderung menggunakan analisis pendekatan induktif sehingga proses dan makna berdasarkan perspektif subyek

Copyright © 2024 **Author**, Page 740
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

lebih ditonjolkan dalam penelitian kualitatif ini [9]. Metode kualitatif digunakan dalam hal ini untuk mengolah data yang berasal dari twitter yang diposting masyarakat melalui media sosia Twitter mengenai Pilpres tahun 2024. Data yang digunakan dari Twitter 1200 dataset selama bulan Januari 2024. Untuk tahapan penelitian yang digunakan terlihat pada gambar 1 berikut:



Gambar 1. Tahapan Penelitian

Penjelasan pada Gambar 1.

a. Pengumpulan data yakni dengan istilah *Crawling* data Twitter adalah pemanfaatan program bot internet untuk menelusuri alamat di World Wide Web, dapat dilakukan pengindeksan dan untuk menyimpan data penting sesuai dengan tujuan bot dibuat dengan menggunakan sumber data Twitter Januari 2024[10]. Kemudian data tersebut diberikan labeling sentimen positif dan negatif.
b. *Preprocessing* merupakan kegiatan untuk melakukan pembersihan data sehingga menghasilkan data bersih yang meliputi casefolding, filtering, tokenisasi, slang replacement dan stopwordremoval.
c. *Feature Selection* method merupakan salah satu metode yang banyak digunakan dalam melakukan reduksi fitur. Metode ini banyak digunakan sebagai metode tambahan dalam melakukan klasifikasi sehingga mampu meningkatkan nilai feature sebelum dilatih (train) dengan algoritma klasifikasi seperti Naïve Bayes [11].
d. *Cross Validation* merupakan uatu metode untuk membagi data ke dalam beberapa bagian (fold) sebanyak k untuk menentukan data training, dan data testing [12].
e. Evaluasi hasil yang didapatkan dari metode *Cross Validation*

2.2 Metode Pengumpulan dan pengolahan Data

Pengumpulan data atau *Crawling* data Twitter didapatkan dari hasil penelusuruan alamat World Wide Web dengan menggunakan sumber data dari media sosial Twitter dengan *query* "Anies Baswedan, Ganjar Pranomo, Prabowo", jenis hasil "*recent or popular*", batas "19000", dan Bahasa *"id"*untuk masing masing calon presiden. proses akan selesai ketika *crawling* data yang diambil dari database Twitter mencapai limit yang telah ditetapkan [13]. Setelah itu data dilakukan cleansing. Cleasing merupakan cara untuk untuk menghilangkan karakter khusus (simbol), *retweet* (RT), *username* dan *link* yang ada pada dataset. Proses *cleansing* dilakukan dengan menggunakan operator *replace* pada *RapidMiner* dengan memasukan kata yang ingin di *replace (replace what)*

Copyright © 2024 **Author**, Page 741
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

dan mengkosongkan kata pengganti *(replace by)*[14] . Kemudian melakukan removal duplicat data yang merupakan menghapus data yang tidak penting[15], kemudian melakukan seleksi data, normalisasi dan labelisasi yang merupakan proses untuk menandakan respon *tweet* yang ada di dalam dataset yang dilakukan menggunakan *python* dengan menggunakan *library transformers*, *pipeline*, dan *pandas*, komentar yang ada didalam dataset dibagi menjadi dua yaitu koemntar positif, dan negatif [16].

Dari pengumpulan data dari Twitter setelah diberi labeling data maka didapatkan data sentimen positif dan negatif untuk ketiga kandidat calon presiden tersebut adalah

Tabel 1. Data Tweet Labeling

Nama Capres 2024	Positif	Negatif
Anies Baswedan	1783	2771
Prabowo Subianto	3775	3576
Ganjar Pranowo	2614	3540

Kemudian data pada tabel 1 tersebut dilakukan tahapan pemisahan data dengan menggunakan operator *split- data*, yang bertujuan untuk memisahkan secara total data menjadi dua bagian, yaitu data training dan data testing. Tahapan pemisahan data dilakukan karena proses klasifikasi membutuhkan data training dan data testing. Setelah melalui beberapa tahapan seleksi data, data *tweet* yang berisi opini publik dibagi menjadi dua bagian, 70% sebagai data training dan 30% sebagai data testing.

Setelah mendapatkan labeling data tweet maka masing masing calon presiden dan wakil presiden terhadap penilain positif dan negatif para netizen maka setelah itu dilakukan proses pengolahan data untuk masing masing data training dan data testing diklasifikasikan pada opini setelah diberikan labeling sentimen positif dan sentimen negatif, maka menghasilan data training sebanyak 3188 dan data testing sebanyak 1366 menghasilkan data positif dan data negatif, adapun data datanya dapat dilihat pada tabel 2 berikut:

Tabel 2. Data Training dan Data Testing

Training	Testing	Training		Testing	
		Positif	Negatif	Positif	Negatif
3188	1366	776	1940	535	831
5145	2205	1029	2960	1869	1268
4308	1846	862	2095	2286	898

Setelah itu melakukan proses *text processing*, Text preprocessing merupakan proses untuk menghilangkan kata kata yang tidak diperlukan [17]. *Preprocessing* memiliki beberapa tahapan yaitu: *tokenized*, *transform cases*, *filter tokens*, *filter stopword*, dan *stemming* [18].Setelah itu melakukan evaluasi terhadap data yang dikumpulkan dan di olah [19]. Particle Swarm Optimization merupakan salah satu teknik optimasi yang sangat sederhana untuk memodifikasi beberapa parameter serta digunakan untuk memecahkan masalah optimasi pada feature selection[20].

3. HASIL DAN PEMBAHASAN

Pada tahapan ini dengan metode algoritma Naive Bayes dan pemilihan fungsi “Particle Swarm Optimization” untuk tahap training maupun testing dilakukan pengujian sebanyak tiga kali terhadap dataset calon presiden untuk mengetahui hasil perbandingan akurasi, presisi dan recall masing-masing calon presiden.
(training).

3.1 Proses Tahapan Pengujian

Pada tahap proses pengujian merupakan serangkaian tahapan untuk menguji data data yang terkait dengan calon presiden yang telah ditetapkan sehingga tidak terjadinya kesalahan atau tercampurnya data dari masing masing calon presiden tersebut. Proses pengujian dilakukan beberapa tahapan:

1. Pada proses pengujian pertama dilakukan pengujian terhadap Anies Baswedan,
2. Pada proses pengujian kedua dilakukan pengujian terhadap Prabowo Subianto
3. Proses pengujian ketiga dilakukan pengujian terhadap Ganjar Pranowo.

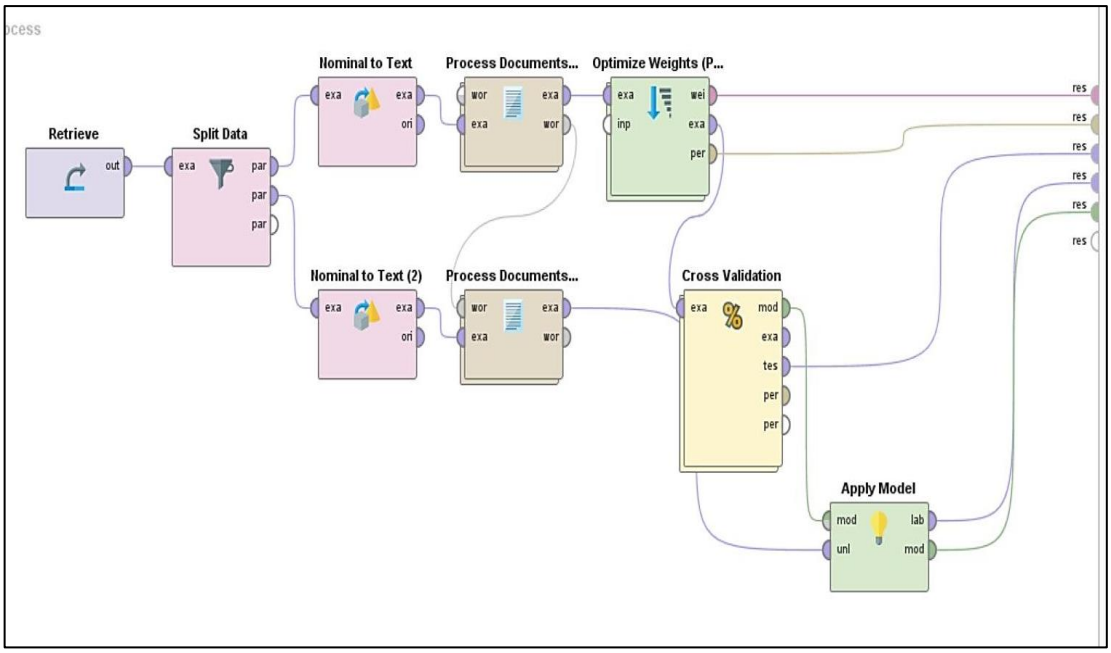
Proses pengujian pertama, kedua dan ketiga dilakukan dengan menggunakan Naive Bayes dan Particle Swarm Optimization. Untuk menghitung hasil akurasi, presisi maupun recall dengan menggunakan operator apply model dan operator performance yang menunjukan confusion matrix maka bisa dilihat pada Gambar 2 yang merupakan tahapan pengujian metode dan Gambar 3 adalah dokumen proses kemudian Gambar 4 adalah model cross validation yang terdiri dari dua bagian yang terpisah yaitu bagian pengujian *(testing)* dan bagian pelatihan *(training)* dengan menggunakan Rapidminer Studio.

RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan munggunakan bahasa java

This Journal is licensed under a Creative Commons Attribution 4.0 International License

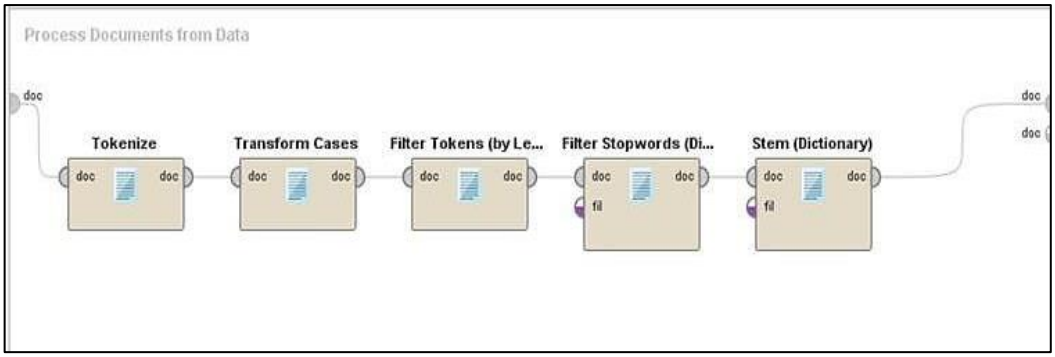
Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

sehingga dapat bekerja di semua sistem operasi. RapidMiner menyediakan UI untuk mendesain pipa analisis, di mana akan menghasilkan file XML yang dapat menjelaskan proses analisis yang ingin diterapkan oleh pengguna ke data. RapidMiner akan membaca file ini untuk menjalankan analisa secara otomatis. RapidMiner menyediakan tampilan (UI) yang ramah pengguna, sehingga memudahkan pengguna saat menggunakannya. Tampilan yang terdapat pada RapidMiner disebut Perspective. Terdapat 3 *Perspective*, yaitu *Welcome Perspective, Design Perspective dan Result Perspective*.



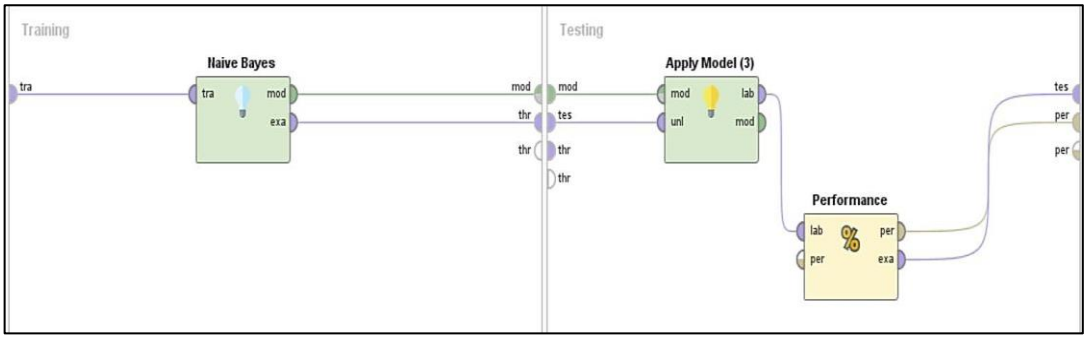
Gambar 2. Proses pengujian *Naïve Bayes* dan *PSO*

Gambar 2 ini menunjukkan langkah-langkah pengujian menggunakan algoritma Naïve Bayes dan Particle Swarm Optimization (PSO). Langkah-langkah termasuk pengumpulan data, preprocessing, pemilihan fitur, validasi silang, dan evaluasi hasil.



Gambar 3. *Proses Document from data*

Gambar 3 ini Menunjukkan proses dokumen dari data yang dihasilkan. Serta Menunjukkan alur data dari pengumpulan hingga hasil akhir.



Gambar 4. Pemodelan *Cross Validation* pada metode *Naïve Bayes* dan *PSO*

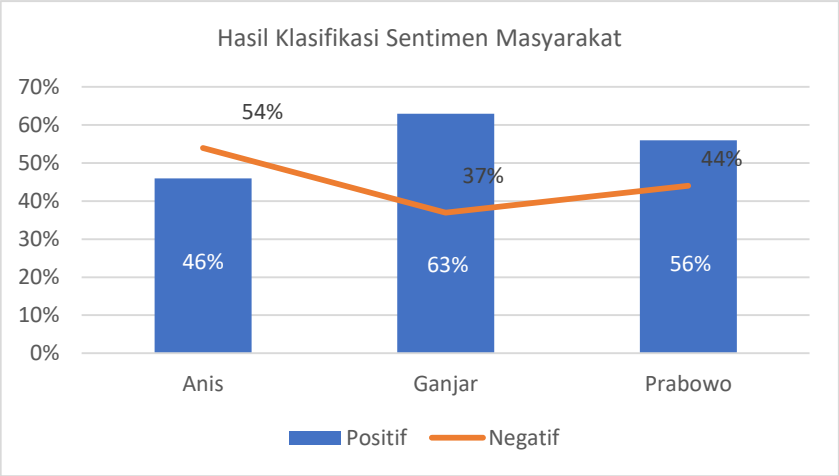
Copyright © 2024 **Author**, Page 743
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

Gambar 4 ini menggambarkan pemodelan validasi silang untuk algoritma Naïve Bayes dan PSO. Memisahkan data menjadi bagian pengujian dan pelatihan untuk mengevaluasi kinerja model.

3.1.1 Hasil Klasifikasi Opini Sentimen Positif dan Negatif

Data yang diperoleh dari data *training* yang sebelumnnya di testing dengan menggunakan algoritma Naïve Bayes dan PSO maka hasil klasifikasi opini sentimen positif maupun sentimen negatif masyarakat Indonesia terhadap calon presiden RI pada tahun 2024 terlihat pada gambar 5 berikut:

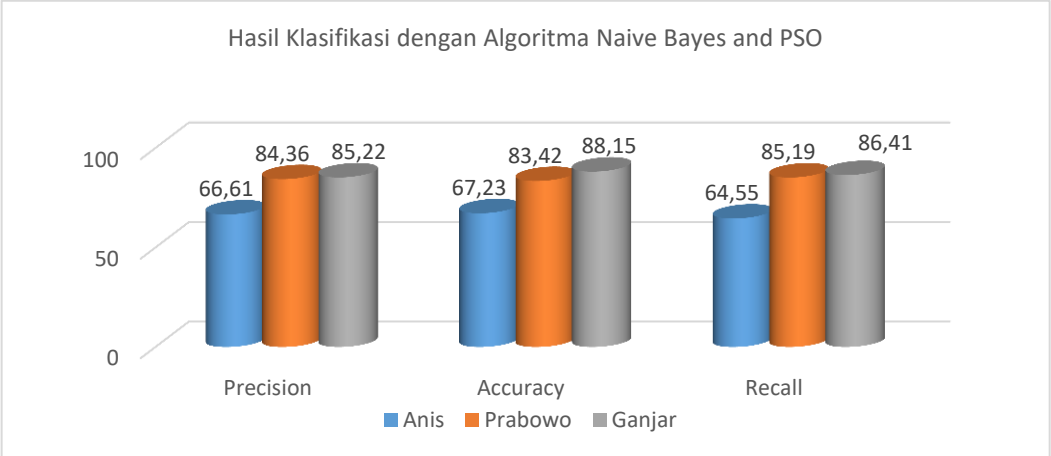


Gambar 5. Hasil klasifikasi Sentimen Masyarakat

Gambar 5 ini menunjukkan hasil klasifikasi sentimen positif dan negatif masyarakat terhadap calon presiden. Hasil klasifikasi menunjukkan proporsi sentimen positif dan negatif untuk masing-masing calon presiden berdasarkan data yang dikumpulkan dan diolah.

3.2 Hasil Analisis Opini Sentimen Masyarakat

Pengujian dilakukan dengan menggunakan teknik *cross validation* dengan nilai k=10. Di bawah ini adalah hasil dari *confusion matrix* saat menggunakan metode *Naïve Bayes* di perangkat lunak *RapidMiner Studio*. Pada penelitian ini dilakukan pengujian sebanyak 3 kali, pengujian pertama, kedua dan ketiga dilakukan dengan menggunakan metode yang sama yaitu *Naïve Bayes* dengan pemilihan fitur *Particle Swarm Optimization*. Di bawah ini adalah gambar grafik yang membandingkan hasil *akurasi*, *presisi*, dan *recall* dari 3 dataset yang berbeda dan berisikan dari rapidminer untuk metode pengujian.



Gambar 6. Hasil Analisis Sentimen Masyarakat

3.3 Pembahasan

Penelitian yang dilakukan dengan menggunakan particle swarm optimization pada Naive Bayes untuk menganalisis opini masyarakat umum tentang ketiga calon presiden 2024 pada sosial media twitter mendapatkan nilai accuracy, precision, dan recall. Untuk Anies Baswedan mendapatkan nilai *Precison* yang didapatkan 66,61%, nilai *accuracy* yang didapatkan 67,23%, dan nilai *recall* yang didapatkan 64,55%. Untuk Prabowo Subianto didapatkan nilai *Precison* 84,36%, nilai *accuracy* yang didapatkan 83,42%, dan nilai *recall*

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

yang didapatkan 85,19%. Sedangkan Ganjar Pranowo didapatkan nilai *Precison* 85,22%, nilai *accuracy* yang didapatkan 88,15%, dan nilai *recall* yang didapatkan 86,41%.

Tabel 3. Data Hasil Opini Sentimen Masyakat Umum

Hasil	Anies Baswedan	Prabowo Subianto	Ganjar Pranowo
%Precision	66,61%	84,36%	85,22%
%Accuracy	67,23%	83,42%	88,15%
%Recall	64,55%	85,19%	86,41%

Tabel 3 ini menampilkan nilai presisi, akurasi, dan recall untuk masing-masing calon presiden setelah analisis sentimen dilakukan. Anies Baswedan dengan perolehan Presisi 66.61%, Akurasi 67.23% dan Recall 64.55%. Prabowo Subianto dengan nilai Presisi 84.36%, Akurasi 83.42% dan Recall 85.19%. Sedangkan Ganjar Pranowo perolehan Presisi 85.22%, Akurasi 88.15% dan Recall 86.41%.

4. KESIMPULAN

Hasil penelitian analisis sentimen dengan menggunakan algoritma *naïve bayes* dan *feature selection particle swarm optimization* dengan *cross validation* k = 10 dan pelabelan data dengan menggunakan *method transformers* memudahkan pelabelan otomatis untuk pelabelan yang memiliki data banyak secara cepat, sehiingga dapat diketahui bahwa dataset anies baswedan 61,35% pengguna *twitter* berkomentar negatif dan 39,65% pengguna twitter berkomentar positif, dataset Ganjar Pranowo 59,12% pengguna twitter berkomentar negatif dan 41,88% pengguna *twitter* berkomentar positif, dataset Ganjar Prabowo Subianto 49,25% pengguna *twitter* berkomentar negatif dan 51,75% pengguna *twitter* berkomentar positif. Perbandingan hasil dari ketiga kandidat calon presiden, nilai akurasi anies baswedan lebih kecil dari pada kedua kandidat lainnya karena komentar negatif dari anies baswedan lebih banyak daripada kedua kandidat lainnya. Anies baswedan mendapatkan nilai *akurasi* 67,23%, prabowo subianto 83,42% dan ganjar pranowo sebesar 88,15%,
Banyaknya data mempengaruhi hasil dari analisa sentimen, semakin banyak data akan semakin baik nilai akurasi yang didapat.Peneliti berharap untuk penelitian selanjutnya bisa dilakukan oleh penelitian yang lainnya dengan metode yang berbeda dan membandingkan beberapa metode agar bisa mendapatkan data yang lebih baik daripada peneliti lakukan saat ini

REFERENCES

[1] A. E. Subiyanto, “Pemilihan Umum Serentak yang Berintegritas sebagai Pembaruan Demokrasi Indonesia,” *J. Konstitusi*, vol. 17, no. 2, p. 355, 2020, doi: 10.31078/jk1726.
[2] D. Iswanto and D. B. Pamungkas, “Meningkatkan Partisipasi Pemilih dalam Pemilu 2024: Pendekatan Stakeholders Mapping Analysis,” *J. Adhyasta Pemilu*, vol. 6, no. 1, pp. 15–27, 2023.
[3] M. Rizki, M. F. Hidayattullah, and D. Intan, “Klasifikasi Opini Publik di Twitter Terhadap Bakal Calon Presiden Indonesia Tahun 2024 Menggunakan LSTM Secara Realtime Berbasis Website,” vol. 14, no. 02, pp. 286–294, 2024, doi: 10.35970/infotekmesin.v14i2.1908.
[4] R. S. Wahono, *Data Mining Data mining*, vol. 2, no. January 2013. 2023.
[5] T. Rosyida, H. P. Putro, H. Wahyono, F. Teknik, and U. Krisnadwipayana, “OPINI DARI TWITTER MENGGUNAKAN NAÏVE BAYES DAN SVM,” vol. 26, no. 1, 2024.
[6] S. N. J. Fitriyyah, N. Safriadi, and E. E. Pratama, “Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes,” *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 3, p. 279, 2019, doi: 10.26418/jp.v5i3.34368.
[7] F. Salsabila and U. Budiyanto, “Implementasi Naïve Bayes Classifier Terkait Pencalonan Ganjar Pranowo Sebagai Calon Presiden 2024 Di Twitter,” *Semin. Nas. Mhs. Fak. Teknol. Inf.*, vol. 2, no. September 2023, pp. 1–10, 2023.
[8] D. A. Faroek, M. Yusuf, and G. Syatauw, “3261-Article Text-11304-2-10-20231130,” vol. 17, no. 2, pp. 216–226, 2024.
[9] d n Katresna and f m Dzikry, “Implementasi Algoritma Naive Bayes Pada Klasifikasi Tweet Untuk Mengetahui Tingkat Kemalasan Siswa,” *J. Siliwangi Seri Sains Dan …*, vol. 6, no. 2, pp. 66–70, 2020.
[10] L. Lathifah, E. Handoyo, and Y. A. Adi Soetrisno, “Sistem Crawling Data Instrumen Akreditasi Berbasis Selenium Dan Pandas,” *Transient J. Ilm. Tek. Elektro*, vol. 10, no. 1, pp. 84–91, 2021, doi: 10.14710/transient.v10i1.84-91.
[11] H. M. Lumbantobing, J. Sisingamangaraja, J. Sisingamangaraja, and J. Sisingamangaraja, “Penerapan Metode Feature Selection pada Algoritma Naïve Bayes dalam Kasus Keyword Extraction,” pp. 6–8, 2020.
[12] C. Journal, I. W. Saputro, B. W. Sari, P. Studi, I. Komputer, and J. Informatika, “Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa,” vol. 6, no. 1, pp. 1–11, 2019.
[13] I. Pendahuluan, “Penerapan Algoritma Naive Bayes dan PSO pada Analisis Sentimen Kandidat Calon Presiden 2024,” vol. 7, pp. 1367–1380, 2024.
[14] A. P. Nardilasari, A. L. Hananto, S. S. Hilabi, T. Tukino, and B. Priyatna, “Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM Pada Media Sosial Twitter,” *JOINTECS (Journal Inf. Technol. Comput. Sci.*, vol. 8, no. 1, p. 11, 2023, doi: 10.31328/jointecs.v8i1.4265.
[15] R. Q. Rohmansa, N. Pratiwi, and M. J. Palepa, “ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI DISCORD MENGGUNAKAN METODE K-NEAREST NEIGHBOR,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran*

This Journal is licensed under a Creative Commons Attribution 4.0 International License

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 3, May 2024, Page 739-746
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i3.5210

Inform., vol. 9, no. 1, pp. 368–378, 2024.
[16] M. H. Al-Areef and K. Saputra S, "Analisis Sentimen Pengguna Twitter Mengenai Calon Presiden Indonesia Tahun 2024 Menggunakan Algoritma LSTM," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 22, no. 2, p. 270, 2023, doi: 10.53513/jis.v22i2.8680.
[17] A. F. Firdaus and W. I. Firdaus, "Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi:(Sebuah Ulasan)," *JUPITER J. Penelit. Ilmu dan Teknol. Komput.*, vol. 13, no. 1, pp. 66–78, 2021.
[18] Asyharudin *et al.*, "Comparing Algorithm for Sentiment Analysis in Healthcare and Social Security Agency (Bpjs Kesehatan)," *Techno Nusa Mandiri J. Comput. Inf. Technol. As an Accredit. J. Rank*, vol. 19, no. 1, pp. 31–37, 2022.
[19] H. F. Husniah, T. Arifin, U. Adhirajasa, R. Sanjaya, U. Adhirajasa, and R. Sanjaya, "Implementasi Algoritma Naïve Bayes Berbasis Particle Swarm Optimization Untuk," vol. 14, no. 2, pp. 37–49, 2019.
[20] R. Aulianita and A. Rifai, "Optimasi Particle Swarm Optimization Pada Naive Bayes Untuk Sentiment Analysis Furniture," *Inf. Manag. Educ. Prof.*, vol. 3, no. 1, p. 78839513, 2018.

This Journal is licensed under a Creative Commons Attribution 4.0 International License