

Analisis Sentimen Terhadap Vaksinasi *Astra Zeneca* pada Twitter Menggunakan *Metode Naïve Bayes* dan K-NN

Slamet Harry Ramadhani ^{1*}, Muhammad Iwan Wahyudin ²

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

article info

Article history:

Received 31 December 2021

Received in revised form

11 January 2022

Accepted 6 February 2022

Available online October 2022

DOI:

<https://doi.org/10.35870/jtik.v6i4.530>

Keywords:

Astra Vaccine Sentiment
Analysis; Zeneca; Naïve Bayes;
K-NN.

Kata Kunci:

Analisis Sentimen Vaksin
Astra; Zeneca; Naïve Bayes; K-
NN.

abstract

The number of positive examples of Coronavirus continues to grow every day, especially in Indonesia. In an effort to reduce the surge in the higher case, the government carried out vaccination programs provided free for all Indonesian people. One vaccine given to Masyarakat is the Astra Zeneca vaccine. In the provision of these vaccines bring benefits and losses in the community, some are supported and doubtful, some even refuse. On social media Twitter, Astra Zeneca is one of the most widely discussed in social media because of the many opinions or opinions that have sprung up from various circles. Some opinions from the community on Twitter will be used as data to examine the analysis of sentiment in the Astra Zeneca vaccine that utilizes the Naïve Bayes and K-NN methods. It is expected to produce an accurate level of accuracy. Based on the results of the study found different levels of accuracy, for the use of the Naïve Bayes method produced an accuracy rate of 90.71% +/- 4.66% (Micro Average: 90.77%) while the KNN method produced an accuracy rate of 74.78% +/- 3.74% (micro Average: 74.77%).

abstrak

Jumlah contoh positif Coronavirus terus berkembang setiap hari, khususnya di Indonesia. Dalam upaya menekan lonjakan kasus yang semakin tinggi, pemerintah melakukan program vaksinasi yang diberikan secara gratis untuk seluruh rakyat Indonesia. Salah satu vaksin yang diberikan kepada masyarakat yaitu vaksin Astra Zeneca. Dalam pemberian vaksin tersebut membawa keuntungan dan kerugian di masyarakat, beberapa ada yang mendukung dan meragukan, bahkan ada juga yang menolak. Di media sosial twitter, Astra Zeneca menjadi salah satu yang ramai diperbincangkan di media sosial karena banyaknya pendapat atau opini yang bermunculan dari berbagai kalangan masyarakat. Beberapa opini dari masyarakat di twitter akan digunakan sebagai data untuk meneliti tentang analisis sentimen pada vaksin Astra Zeneca yang memanfaatkan metode Naïve Bayes dan K-NN. Diharapkan dapat menghasilkan tingkat akurasi yang akurat. Berdasarkan dari hasil penelitian didapati tingkat keakurasian yang berbeda, untuk penggunaan metode Naïve Bayes menghasilkan tingkat akurasi sebesar 90.71% +/- 4.66% (micro average: 90.77%) sedangkan untuk metode KNN menghasilkan tingkat akurasi sebesar 74.78% +/- 3.74% (micro average: 74.77%).

Author. Email: slametharryramadhani@gmail.com ^{1}, iwan.wahyuddin@civitas.unas.ac.id ².

1. Latar Belakang

Pandemi Covid-19 merupakan wabah besar yang melanda diseluruh dunia berdasarkan sumber dari <https://smartcity.jakarta.go.id/> Awal dari kemunculan virus Covid-19 ini terjadi di Kota Wuhan, Cina, sejak akhir Desember 2019. Tanggal mulai dari kasus yang mendasarinya adalah 1 Desember 2019, akibat dari wabah besar virus ini mengakibatkan dampak yang krisis terhadap perekonomian dunia. Indonesia termasuk negara yang paling banyak terinfeksi oleh virus covid-19, pada masa pandemi ini pemerintah menyuplai berbagai jenis vaksin kepada seluruh masyarakat untuk melakukan vaksinasi dengan tujuan untuk mengurangi gejala terhadap orang yang terpapar virus covid-19 diantaranya yaitu vaksin Astra Zeneca. Awal kedatangan vaksin ini banyak dari kalangan masyarakat yang menyangkal bahwa vaksin Astra Zeneca ini merupakan vaksin dengan penggunaan dosis tinggi dan dianggap mampu untuk mengurangi terinfeksi virus covid-19, oleh karena itu menimbulkan berbagai opini yang disampaikan masyarakat dengan adanya vaksin jenis ini melalui media sosial twitter [1].

Saat ini penggunaan media sosial sangatlah umum digunakan oleh masyarakat karena perkembangan teknologi yang tumbuh begitu cepat. Media sosial tidak hanya sebagai media berkomunikasi tetapi juga digunakan untuk berbagi informasi, berita acara, media bisnis, serta berfungsi sebagai sarana bagi masyarakat untuk beropini. Oleh karena itu vaksin Astra Zeneca membuat sebagian besar masyarakat khawatir karena banyaknya berita yang beredar di media sosial khususnya twitter bahwa setelah melakukan vaksinasi dapat mengalami demam yang cukup tinggi, badan terasa sakit, sesak nafas, pembekuan darah dan lain-lain. Oleh dari itu, pada penelitian ini, penulis ingin menganalisis jumlah respon positif, negatif atau netral yang sudah disampaikan masyarakat terhadap vaksinasi Astra Zeneca yang terdapat pada dimedia sosial Twitter dengan metode *Naive Bayes* dan KNN [2].

Berdasarkan latar belakang yang ada, dalam permasalahan yang sedang dihadapi ini diperlukannya penelitian bagaimana caranya mengklasifikasikan analisis sentimen terhadap respon dari masyarakat terkait adanya vaksinasi

Astra Zeneca menggunakan metode *Naive Bayes* dan KNN [3]. Dalam penelitian, penulis membatasi masalah ini yang bertujuan untuk memperkecil cakupan pada penelitian agar penelitian menjadi lebih terarah pada satu permasalahan yang ada. Oleh karena itu diperlukannya batasan masalah penelitian, yaitu untuk menilai respon dari beberapa masyarakat melalui pendapatnya berdasarkan dokumen *tweet*, dalam tahap proses penelitian ini data yang terkumpul dari Twitter sangat terbatas. Penelitian ini bertujuan untuk mendapatkan hasil klasifikasi dan *accuracy* sentimen dari pandangan masyarakat yang terdapat pada media sosial twitter terhadap vaksin Astra Zeneca menggunakan metode *Naive Bayes* dan KNN.

Studi Literatur ini membahas beberapa tahapan yang diperlukan dalam proses analisis dan referensi yang sudah diteliti sebelumnya [4]. Tahapan yang pertama dilakukan yaitu, tahap *pre-processing*, pengumpulan data, analisis, klasifikasi dan evaluasi, selanjutnya dalam tahap analisis sentimen melibatkan pengukuran kualitas dalam penentuan hasil analisis sentimen menggunakan parameter seperti akurasi, presisi, recall [5]. Dari penelitian yang sudah dilakukan hanya sebatas untuk referensi dalam penelitian ini. Penelitian pertama dapat ditarik kesimpulan yaitu, pendapat masyarakat terhadap vaksin. Menggunakan NLP untuk memperoleh tweet dan mengklasifikasi algoritma KNN untuk data yang diproses, terlihat bahwa masyarakat umum memiliki sentimen positif yang lebih tinggi terhadap vaksin pfizer dan moderna sama dengan tingkat 47,29 dan 46,16 dibandingkan dengan tingkat 40,08 [6]. Selanjutnya penelitian kedua diambil kesimpulan mengenai pendapat masyarakat terhadap Covid-19 lebih kearah positif, sehingga dari hasil penggunaan algoritma *Naive Bayes* tingkat akurasinya tinggi jika dibandingkan dengan KNN [7]. Penelitian ketiga disimpulkan bahwa tanggapan masyarakat cenderung negatif dengan nilai akurasi 100% untuk metode *Naive Bayes* sedangkan untuk algoritma Decision Tree 50.39% [8]. Penelitian ke empat disimpulkan masyarakat sangat mendukung kebijakan PSBB karena dinilai akan lebih efektif mengurangi kasus penyebaran Covid- 19. Namun ada juga tambahan masyarakat yang bereaksi sebaliknya/negative dan merasa bahwa strategi PSBB yang dilukan oleh pemerintah gagal dengan alasan bahwa ada banyak pelanggaran [9]. Pada penelitian ke lima, disimpulkan

bahwa metode *Naive Bayes Classifier* menggunakan seleksi fitur *Chi-Squared Statistic* berbasis *forward selection* memperoleh tingkat akurasi yang lebih tinggi berbeda dengan metode seleksi fitur *Particle Swarm Optimization (PSO)* [10]. Kemudian penelitian ke enam dapat diambil kesimpulan bahwa dari ke tiga metode tersebut metode KNN yang mendapatkan nilai *accuracy* sebesar 76.21% [11]. Pada penelitian ke tujuh, ditarik kesimpulan bahwa dari ke tiga metode tersebut terdapat satu metode yang mendapatkan nilai akurasi yang tinggi, yaitu metode *Decision Tree* sebesar 61.92% [12].

Text mining adalah sebuah proses dalam bidang data mining yang dapat diartikan sebagai penambang data dalam bentuk teks ataupun data yang biasanya terdapat didalam dokumen. *Text mining* ini bertujuan untuk mencari sebuah informasi baru atau kalimat sehingga untuk melakukan proses ini diperlukan nya sebuah analisa yang saling terhubung antar dokumen. Menurut teori yang ada *text mining* mampu bekerja pada komputer dengan maksud untuk mengolah berbagai informasi lama secara efektif yang dapat menghasilkan informasi baru [13]. Berikut ini merupakan tahapan yang dilakukan dalam *text mining* sebagai berikut:

1) *Transform Case*

Transform Case adalah tahapan yang sering digunakan untuk melakukan *pre-processing* yang menggunakan parameter TF-IDF, untuk merubah teks menjadi huruf kecil semua

2) *Tokenizing*

Tokenizing merupakan tahap untuk pemilihan jumlah unit yang akan digunakan saat pemecahan data dan *tokenize* merupakan tahap pembagian atau pemecahan terhadap data pada teks.

3) *Filtering*

Tahapan *filtering* merupakan salah satu tahap yang digunakan untuk pengambilan kalimat atau kata-kata yang penting dari hasil crawling data. Kalimat umum yang biasa muncul dan tidak mempunyai arti maka sering disebut dengan istilah *stopword*. *Stopword* membantu mengurangi ukuran index pada waktu pemrosesan.

4) *Stemming*

Tahap *Stemming* adalah tahap yang dimana dapat mengurangi jumlah berbagai daftar dari satu informasi sehingga kata-kata yang memiliki

awalan akan kembali ke bentuk dasarnya.

5) *Filter Stopwords*

Merupakan proses untuk menentukan sebuah kalimat yang mengandung kata – kata yang sering keluar.

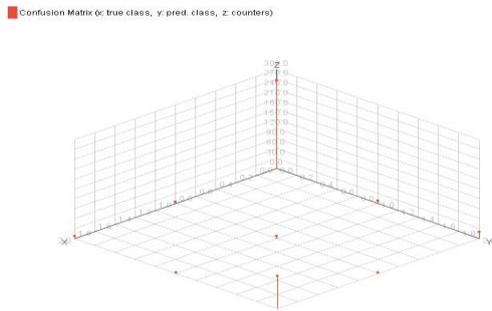
Naive Bayes merupakan algoritma yang digunakan dengan teknik data mining dan menerapkan teori Bayes. Ciri-Ciri dari *Naive Bayes* ini mempunyai perkiraan atau prediksi yang kuat berdasarkan setiap kondisi/kejadian [14]. *Naive Bayes* adalah strategi classifier probabilitas langsung yang dapat ditentukan dengan menambahkan suatu nilai iterasi dari sebuah indeks informasi yang diberikan. dan mengasumsikan seluruh atribut sehingga mereka tidak saling bergantung dari nilai yang memiliki variabel kelas. Dalam metode *Naive Bayes* terdapat dua tahapan klasifikasi [15]. Untuk tahap yang pertama dilakukan yaitu, tahap proses pembelajaran pada sebuah dokumen yang kategorinya diketahui. Pada tahap kedua proses pengujian yaitu, klasifikasi dokumen yang kategorinya tidak diketahui. Klasifikasi *Naive Bayes* dinilai dapat bekerja secara baik dibandingkan dengan model classifier yang lainnya [16]. Penggunaan metode *naive bayes* dapat dirumuskan sebagai berikut:

$$P(H \setminus X) = \frac{P(H \setminus X) + P(H)}{P(X)}$$

Analisis sentimen ialah sebuah proses pengklasifikasian emosi (positive, negative dan neutral) yang terkandung dalam tulisan menggunakan teknik analisis. Analisis sentimen juga sering disebut sebagai *Opinion Mining*. *Opinion mining* ini mengarah kepada bidang yang sangat luas, tujuannya hanya untuk melakukan sebuah analisa pendapat, sentiment evaluasi, sikap, penilaian dan emosi seseorang [17]. Dari pernyataan tersebut diketahui bahwa analisis sentiment merupakan proses penentuan perasaan atau tanggapan seseorang yang dapat digambarkan baik dalam bentuk teks maupun bahasa yang dapat dikategorikan sebagai sentiment netral, positif ataupun negative [18].

Confusion matrix

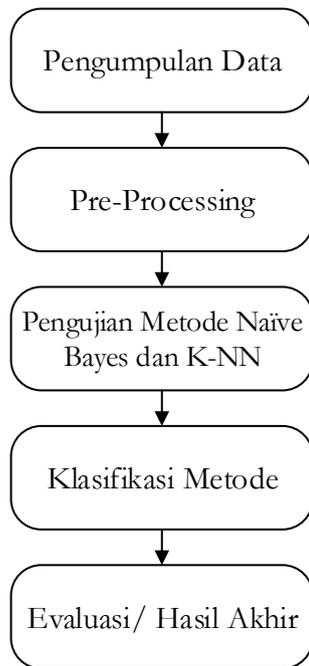
Confusion matrix disajikan dalam bentuk tabel yang terdiri dari jumlah data yang telah diuji dan yang seharusnya benar atau salah tergantung dari model klasifikasi yang digunakan.



Gambar 1. Confusion Matrix

K-Nearest Neighbor adalah teknik yang menggunakan perhitungan super-vised di mana hasil yang baru dicoba dapat diurutkan berdasarkan klasifikasi pada algoritma K-NN. Algoritma K-NN digunakan untuk mengklasifikasikan sebuah objek menurut data latih yang mendekati objek tertentu. Teknik tersebut sangat sederhana dan mudah untuk diterapkan. Menurut teori yang sudah ada metode ini sama dengan metode clustering yaitu, dengan mengelompokkan sebuah data baru berdasarkan jarak data baru tersebut kebeberapa jarak data terdekatnya [19]. Tahap algoritma ini bertujuan untuk mengklasifikasi objek berdasarkan atribut.

2. Metode Penelitian

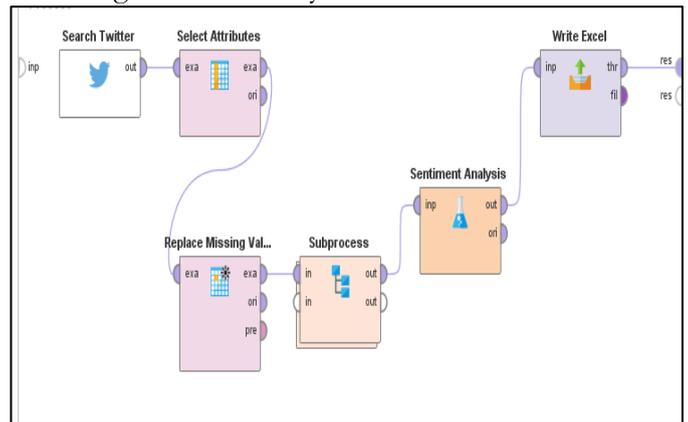


Gambar 2. Tahap Penelitian

Pada metode penelitian, terdapat beberapa tahap yang perlu dilakukan untuk analisis sentimen supaya mendapatkan hasil yang maksimal [20]. Tahapan ini terdiri dari:

Pengumpulan Data

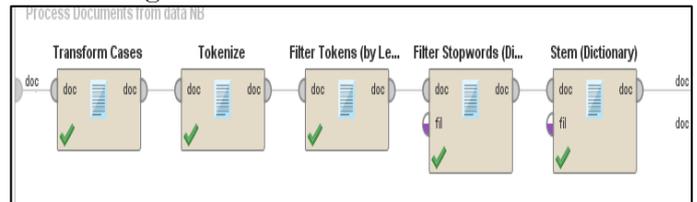
Tahap pertama yaitu pengumpulan data yang dilakukan menggunakan *Tools Rapidminer*. Twitter menyediakan layanan yang dapat diakses bebas bagi penggunaanya untuk mencari berbagai tweet dengan cara mengaktifkan Twitter *Developer* atau Twitter Api untuk memperoleh *Acces token* [21]. Setelah mendapatkan *acces token* tersebut dihubungkan pada tools rapidminer untuk melakukan proses pencarian data dengan kata kuncinya “Vaksin Astra Zeneca”.



Gambar 3. Proses Pengumpulan Data

Gambar 3 merupakan proses untuk mencari sebuah data tweet yang ada di twitter.

Pre-Processing



Gambar 4. Processing Data.

Tahap ini data yang dikumpulkan akan diklasifikasikan sebagai penentu tingkat keakurasian. Teknik yang digunakan dalam *processing* ini adalah normalisasi, *Transform case*, *Tokenize*, *Filtering Token*, *Filtering Stopword*, dan *Steam* [22].

Pengujian Metode

Tahap pengujian metode ini menggunakan algoritma Naïve Bayes dan KNN dengan parameter *CrosValidation* untuk menentukan besaran nilai akurasi yang didapatkan.

Klasifikasi

Tahap klasifikasi ini bertujuan untuk mengklasifikasikan seluruh data yang ada sebagai

penentu apakah data tersebut dikategorikan sebagai sentiment negatife, positif ataupun netral.

| Ro... | Sentime... | predicti... | confi... | confid... | confid... | text |
|-------|------------|-------------|----------|-----------|-----------|--|
| 16 | Netral | Netral | 1 | 0 | 0 | btanding sorry info vaksin astra zeneca tahap temen aing dosga dikasih vaksin thanks |
| 17 | Netral | Netral | 1 | 0 | 0 | jadwal vaksinasi massal jenis vaksin astra zeneca puskesmas madurejo Sabtu aula kantor pa |
| 18 | Netral | Netral | 1 | 0 | 0 | antoniraja rangka mendukung program pemerintah mewujudkan kekebalan komunal herd imm |
| 19 | Netral | Netral | 1 | 0 | 0 | antoniraja rangka mendukung program pemerintah mewujudkan kekebalan komunal herd imm |
| 20 | Netral | Netral | 1 | 0 | 0 | bertanyari bapaku neder vaksin astra zeneca kali gapapa vaksin gara pelayanan |
| 21 | Negatif | Negatif | 0 | 1 | 0 | kesempatan vaksinasi hidup berdampingan virus jenis vaksin tersebut Pfizer Astra Zeneca Sino |
| 22 | Positif | Positif | 0 | 0 | 1 | adikku memutuskan konsultasi dokter minum obat jadwal vaksin dokter hasilnya arman adik da |
| 23 | Positif | Positif | 0 | 0 | 1 | tubrifess vaksin astra zeneca ortu sinovac ortu gaada efek samping samsek efek sampingnya c |
| 24 | Netral | Netral | 1 | 0 | 0 | fatwa vaksin astra zeneca Pfizersehdih haram kecuali penjelasan babeh aldo |
| 25 | Positif | Positif | 0 | 0 | 1 | segenap warga ngestharjo saudaraku warga domisili ngestharjo vaksinasi covid mengikuti va |
| 26 | Positif | Negatif | 0 | 1 | 0 | nemu vaksin astrasenang |
| 27 | Netral | Negatif | 0 | 1 | 0 | fromisaya pake vaksin sino astra |

Gambar 5. Klasifikasi Naïve Bayes

| Row... | Sentim... | predicti... | confide... | confid... | confid... | text |
|--------|-----------|-------------|------------|-----------|-----------|---|
| 23 | Netral | Netral | 1 | 0 | 0 | codim enrengk koramil maiva vaksinator puskesmas maiva melaksanakan serbuan va |
| 24 | Netral | Netral | 0.591 | 0.174 | 0.236 | caca maticareca sigfess mlilh vaksin sinovac: sakit mundur (jadwalnya) vaksin kebagiannya |
| 25 | Netral | Netral | 0.590 | 0.171 | 0.239 | tedhibertfatwa vaksin astra zeneca Pfizersehdih haram kecuali penjelasan babeh aldo |
| 26 | Netral | Positif | 0.355 | 0 | 0.645 | bhabinkamtombas kelurahan kerobokan Kaja Bripta Gede pratama instansi terkait atensi k |
| 27 | Netral | Netral | 0.776 | 0 | 0.224 | hafsaaahr bchsr sentra vaksin menyediakan astra sinovac: kalo kasih astra kalo sentra vak |
| 28 | Netral | Positif | 0.137 | 0.237 | 0.626 | vaksin vaksin inova suntik astra sigra |
| 29 | Netral | Netral | 0.844 | 0 | 0.156 | baymax bull pagi sobat memiliki info vaksin dosis astra zeneca terimakasih |
| 30 | Netral | Netral | 1 | 0 | 0 | bertanyari bapaku umur udah vaksin astra neder |
| 31 | Netral | Netral | 1 | 0 | 0 | tedhibert teori konspirasi vaksin covid masuk kesuburan akumulasi vaksin ovarium |
| 32 | Positif | Netral | 1 | 0 | 0 | azilova shedding vaksin astra zeneca ovarium wanita lihat halaman ditfikan kemenkesri |
| 33 | Netral | Netral | 1 | 0 | 0 | tedhibert teori konspirasi vaksin covid masuk kesuburan akumulasi vaksin ovarium |
| 34 | Positif | Positif | 0.400 | 0 | 0.600 | azilova shedding vaksin astra zeneca ovarium wanita lihat halaman ditfikan kemenkesri |

Gambar 6. Klasifikasi KNN.

Evaluasi

Penilaian/ Eksekusi sebatas untuk menguji hasil persiapan dengan memperkirakan nilai pada sebuah system. Batas yang digunakan untuk mengukur apresiasi realitas adalah akurasi. Akurasi sendiri adalah tingkat atau hasil laporan yang telah disiapkan secara efektif oleh system. Untuk melakukan perhitungan accuracy, presisi dan recall, dapat memanfaatkan metode confusion matrix sebagai perhitungan nilai menggunakan persamaan 2.

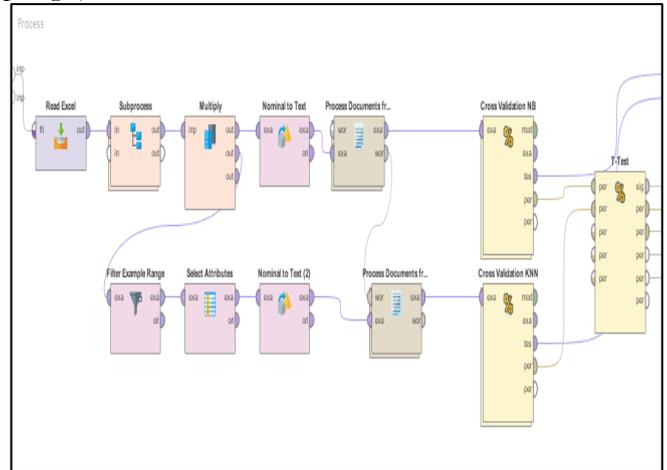
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%$$

$$Presisi = \frac{TP}{FP + TP} \times 100\%$$

$$Recall = \frac{TP}{FN + TP} \times 100\%$$

3. Hasil dan Pembahasan

Berdasarkan hasil pembahasan ini terdapat lima tahapan yang akan dibahas diantaranya, pengumpulan data, processing, pengujian metode, klasifikasi algoritma dan yang terakhir hasil akhir pengujian. Berikut ini akan dibahas secara berurutan:



Gambar 7. Proses utama analisis.

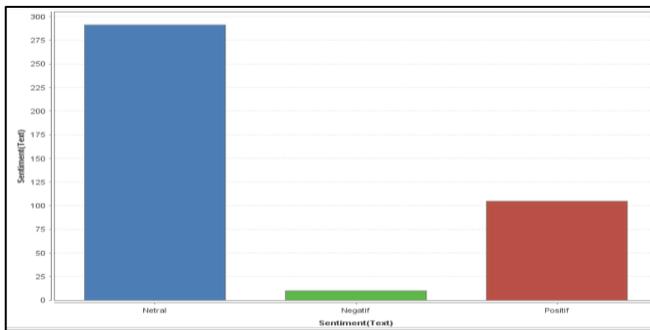
Pada gambar 7 ini proses utama dalam melakukan analisis sentiment pada penggunaan rapidminer. Didalam rapidminer ini terdapat beberapa operator yang peneliti gunakan antara lain: Read Excel, Subprocess, Multiply, Filtering, Select Atribut, Proses Dokumen, Cross-Validation, Naïve Bayes dan KNN.

Pengumpulan Data

Tahap pertama yaitu teknik pengumpulan data ini menggunakan Tools Rapidminer yang dimana proses pengumpulan data nya dengan cara pencarian data yang terdapat pada media sosial twitter dengan API-Twitter yang menghubungkan rapidminer dengan twitter developer yang sudah mendapatkan izin akses nya. Data yang berhasil dikumpulkan berjumlah 407 data. Data tersebut kemudian disimpan kedalam excell dengan format csv.

| A | B | C |
|-----|--|----------------------|
| 210 | Kondisi setelah vaksin astra dosis 1 Pegel di area suntik Pusing dikit tapi kenapa tiap hari a | 144804521798516739 |
| 211 | Gue tetep vaksin ke 10 hari gapapa gasih?? lupa soalnya pas vaksin pertama breng teme | 14460304624613978113 |
| 212 | Baru sadar weh udh 3 bulan lebih bim vaksin kedua. astra emang lama bgt apa gimana si | 1447949753145180163 |
| 213 | beres vaksin astra dosis 2, masih ga berasa efek samping nya. Fix ini mah vaksin vaksin | 1447944273488351283 |
| 214 | @sauruslicious Mitil bang, bagusan mana vaksin astra sama vaksin omien? | 1447925132893486501 |
| 215 | Log info vaksin astra daerah Kulon Progo tuur | 1447907168179539970 |
| 216 | RT @katolikg: ????? Vaksin ke berapa, pak????? ke 2, dok????? Yg pertama, vaksin apa? | 1447904765816422408 |
| 217 | gilirannya dpt jatah vaksin eh malah ke smd ?? yasadialah di lepas aja dulu vaksin, next aja | 1447903297977948627 |
| 218 | Sevual jangkui mau sharing vaksinasi di Indonesia. Vaksin di Indo ada beberapa macam | 1447899155716795074 |
| 219 | vaksin astra dosis 2, rodok kemeng to gpp?? | 1447882703219552266 |
| 220 | Btw kalo misal nya vaksin pertama astra terus vaksin kedua sinovac boleh gasih? | 1447847270881019137 |
| 221 | @dounong17 Astra emang 3 bulan kak jaraknya dari vaksin pertama sampe kedua | 1447837424667699057 |
| 222 | @boding sorry oot min. Info vaksin astra zeneca yang tahap 2 dong. Temen aing par dosi | 1447818529051348994 |
| 223 | @paktantono Putih: Sinovac, Merah, Astra Zeneca, Ijo. Pfizer, abud: Vaksin Nusantara | 144781312029442049 |
| 224 | tp w ny aja bim nemu2 vaksin ke 2 yg astra | 1447810821052977159 |
| 225 | @realina: @insidelombok_ Tadi abis vaksin astra di masbaki, lotim. | 1447809414417694720 |
| 226 | @hafsaahr @bchr_ Waktu itu di sentra vaksin yang kebetulan menyediakan astra dan | 1447807236806305747 |
| 227 | @fromisaya Pake vaksin apa sino ato astra | 144779974989223818 |
| 228 | bhabinkamtombas Kelurahan Kerobokan Kaja Bripta I Gede Eka Pratama SH bersama insta | 1447791015201423868 |
| 229 | Sentra Vaksinasi LRT Jakarta masih berlangsung??? Bagi kamu Sahabat LRT yang belum d | 144778268994190977 |
| 230 | di rs Bandung yg masih ada vaksin astra dosis 1 dimana ya? info dong ya tau. | 1447782664589570050 |
| 231 | Akhi bak tebagian vaksin yg astra :3 Emng katanya banyak keluhan shshs. Td malem pans | 144778066511875897 |
| 232 | ????? Vaksin ke berapa, pak????? ke 2, dok????? Yg pertama, vaksin apa????? Inova dok | 1447769953465802500 |

Gambar 8. Hasil Pengumpulan Data/Crawling Data.



Gambar 9. Grafik Pengumpulan Data

Dari hasil pengumpulan data atau dapat disebut crawling data digambarkan dengan grafik. Sebuah grafik penentu seberapa banyak dari masyarakat yang beropini negatif, positif, dan netral. Hasilnya adalah terdapat 291 dengan sentiment Netral, 10 sentiment Negatif dan 105 dengan sentiment positif.

Pre-processing

Selanjutnya akan dilanjutkan ketahap *processing*. Tahap ini dilakukan dengan menggunakan data yang sudah ada. Berikut adalah proses untuk melakukan *processing* data yang terdiri dari: *Transform case, Tokenize, Filtering, Stopword, Steaming*.

| Word ↓ | Attribute Name |
|---|---|
| yuhu, benar kalau ada yang menebak a... | yuhu, benar kalau ada yang menebak akdfitas vaksinasi masal. kelihatan kan para naskes yang be... |
| yfh. segenap warga ngestiharjo?? bagi ... | yfh. segenap warga ngestiharjo?? bagi saudaraku warga domisili ngestiharjo yang belum vaksinasi... |
| vaksin sinovac, astra zeneca, moderna, ... | vaksin sinovac, astra zeneca, moderna, pfizer, vaksin zifrax ?? |
| vaksin kedua astra ternyata rasanya sa... | vaksin kedua astra ternyata rasanya sama kek vaksin pertama alias tepar sobat ???? |
| vaksin dosis kedua astra zeneca termyat... | vaksin dosis kedua astra zeneca ternyata sangat berasa ya efeknya |
| vaksin astra zeneca pun tidak menimbu... | vaksin astra zeneca pun tidak menimbulkan efek samping di eug. alhamdulillah, ya kalo ini sih cuku... |
| vaksin astra zeneca @nuhahkel @din... | vaksin astra zeneca @nuhahkel @dinkeskotadepok @depokita @infocimanggis @infodepok_id @... |
| vaksin astra zenecalokasi :1 kantor koramii (*100 sasaran* hanya dosis 2)2. kantor lurah krukut (*10... | vaksin astra zenecalokasi :1 kantor koramii (*100 sasaran* hanya dosis 2)2. kantor lurah krukut (*10... |
| vaksin astra dosis 2, rodok kemeng tp g... | vaksin astra dosis 2, rodok kemeng tp gpp ?? |
| unsfrf ada info vaksin ke2 astra zeneca ... | unsfrf ada info vaksin ke2 astra zeneca di layoipig ga ya? |

Gambar 10. Proses *Transform Case*

Proses pertama yang dilakukan adalah *Transform case* untuk merubah semua kalimat dengan huruf besar menjadi kecil

| Word | Attribut... | Total O... | Docum... | Netral | Negatif | Positif |
|-----------|-------------|------------|----------|--------|---------|---------|
| ANGGOTA | ANGGOTA | 1 | 1 | 1 | 0 | 0 |
| AREAJULID | AREAJU... | 1 | 1 | 1 | 0 | 0 |
| ASTRA | ASTRA | 4 | 4 | 4 | 0 | 0 |
| Abang | Abang | 1 | 1 | 1 | 0 | 0 |
| AboutTNG | AboutTNG | 1 | 1 | 1 | 0 | 0 |
| Ada | Ada | 10 | 10 | 7 | 3 | 0 |
| Adanya | Adanya | 41 | 41 | 1 | 0 | 40 |
| Adik | Adik | 2 | 2 | 0 | 0 | 2 |
| Adikku | Adikku | 2 | 2 | 0 | 0 | 2 |
| Akibat | Akibat | 1 | 1 | 0 | 0 | 1 |
| Aktivitas | Aktivitas | 2 | 2 | 2 | 0 | 0 |

Gambar 11. Proses *Tokenize*

Selanjutnya pada tahap kedua yaitu *tokenize* untuk memisahkan setiap kalimat atau teks yang terdapat didalam sebuah dokumen

| Word | Attribute Name | Total O... | Docum... | Netral | Negatif | Positif |
|-------------------------|-------------------------|------------|----------|--------|---------|---------|
| @KUNHCNEYVaksin astra? | @KUNHCNEYVaksin astra? | 1 | 1 | 1 | 0 | 0 |
| Ada yg vaksin astra gk? | Ada yg vaksin astra gk? | 1 | 1 | 1 | 0 | 0 |

Gambar 12. Proses *Filtering*

Dari tahap *tokenize* ini kemudian dilanjutkan ketahap *filtering* yaitu untuk pengambilan kalimat yang penting dengan jumlah kata yang diinginkan.

| Word | Attribute Name |
|--|--|
| Bhabinkamtibas Kelurahan Kerobokan Kaja Bripla I Gede Eka Pr... | Bhabinkamtibas Kelurahan Kerobokan Kaja Bripla I Gede Eka Pratama |
| Bhabinkamtibas Kelurahan Kerobokan Kaja Bripla I Gede Eka Pr... | Bhabinkamtibas Kelurahan Kerobokan Kaja Bripla I Gede Eka Pratama |
| Btw kalo misalnya vaksin pertama astra terus vaksin kedua sinovac... | Btw kalo misalnya vaksin pertama astra terus vaksin kedua sinovac boleh |
| Btw yg vaksin nya bukan astra zeneca, moderna, or pfizer kt mau ke ... | Btw yg vaksin nya bukan astra zeneca, moderna, or pfizer kt mau ke Jepun |
| Contohnya, seperti kerja sama bilateral dose-sharing mengenal va... | Contohnya, seperti kerja sama bilateral dose-sharing mengenal vakinas |
| Daerah bpp ada yang mau vaksin kah???Vaksin astra tp harus skrg | Daerah bpp ada yang mau vaksin kah???Vaksin astra tp harus skrg |
| Dalam rangka turut mendukung program pemerintah guna meweju... | Dalam rangka turut mendukung program pemerintah guna mewujudkan k |
| Danramil 15J/Kalibawang Kapten Arm Trio Mardiyanto beserta Bab... | Danramil 15J/Kalibawang Kapten Arm Trio Mardiyanto beserta Babinsa, B |
| Danrem 052WKR Pantau kedatangan 245.520 dosis Vaksin Astra ... | Danrem 052WKR Pantau kedatangan 245.520 dosis Vaksin Astra Zenece |
| Di Balai Desa Sidomulyo Gelar Vaksinasi Astra Zeneca Warga Mas... | Di Balai Desa Sidomulyo Gelar Vaksinasi Astra Zeneca Warga Masyarakat |

Gambar 13. Proses *Stopword*

Kemudian dari tahap *filtering* dilanjutkan ketahap *stopwords* untuk menentukan sebuah kalimat yang sering keluar

| Word ↓ | Attribute Name |
|-----------------------------------|--|
| sdfif info vaksin astra zenece... | sdfif info vaksin astra zeneca dosis 2 di sidoarjo |
| schfess org indo yang punya... | schfess org indo yang punya paten vaksin astra zeneca, bos bca, bos mayora, ceo kok bisa |
| sarah gilbert penemu vaksi... | sarah gilbert penemu vaksin astra zeneca."covid 19 akan melemah hingga menjadi flue biasa"bila masyarakat... |
| sarah gilbert penemu vaksi... | sarah gilbert penemu vaksin astra zeneca."covid 19 akan melemah hingga menjadi flue biasa"bila masyarakat... |
| saat ini ada setidaknya 4 je... | saat ini ada setidaknya 4 jenis vaksin yang sudah ada di indonesia, yakni sinovac, astra zeneca, moderna dan ... |
| saat ini ada setidaknya 4 je... | saat ini ada setidaknya 4 jenis vaksin yang sudah ada di indonesia, yakni sinovac, astra zeneca, moderna dan ... |
| rt @uniusudarmono2 @mh... | rt @uniusudarmono2 @mharapanita @katolkg lanjut vaksin ke-3 astra zenia ?? |
| rt @oothaa: ada yg tau info v... | rt @oothaa: ada yg tau info vaksin dosis kedua jenis astra zeneca di sekitaran bekas kalau bisa yg di mall... m... |
| rt @wg4tr: vaksin astra zene... | rt @wg4tr: vaksin astra zeneca @nuhahkel @dinkeskotadepok @depokita @infocimanggis @infodepok_id ... |
| rt @tedhilbert: hanya teori k... | rt @tedhilbert: hanya teori konspirasi bahwa vaksin covid merusak kesuburan?mengapa ada akumulasi 'vaksi... |

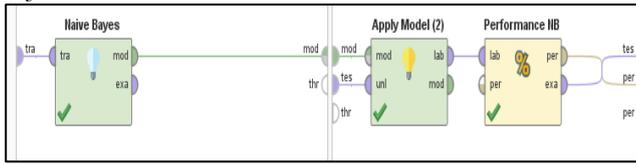
Gambar 14. Proses *Steaming*

Tahap selanjutnya pada *processing* ini yaitu *stemming* untuk mengurangi jumlah berbagai daftar dari satu informasi sehingga kata-kata yang memiliki awalan akan kembali ke bentuk dasarnya.

Pengujian Metode

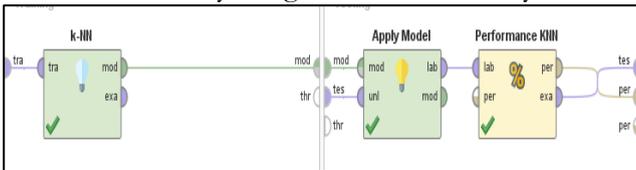
Kemudian dari tahap *processing* akan dilanjutkan ketahap pengujian metode yang pertama menggunakan metode Naïve bayes, proses pengujian dilakukan menggunakan parameter *Cross-Validation*. *Cross-Validation* merupakan parameter yang terdapat didalam Rapidminer dengan tujuan untuk

menentukan nilai akurasi pada sebuah metode. Cara kerjanya yaitu, melalui proses testing yang saling dihubungkan dengan parameter *Apply Module* dan *Performace*.



Gambar 15. Metode Bayes.

Tahap pengujian yang kedua dengan metode KNN. Cara kerjanya sama dengan pengujian yang pertama menggunakan metode Naive Bayes yang membedakan hanya tingkat nilai akurasi nya.



Gambar 16. Metode KNN.

Klasifikasi

Setelah pengujian akan keluar hasil dari sebuah data. Tujuan dari klasifikasi ini untuk menentukan besaran jumlah sentimen Netral, Positif, dan Negatif yang sudah diujikan dengan kedua metode tersebut.

Tabel 1. Tahap Klasifikasi Naive Bayes.

| Index | Nominal Value | Absolute Count | Fraction |
|-------|---------------|----------------|----------------------|
| 1 | Netral | 233 | 0.7169230769230769 |
| 2 | Positif | 84 | 0.25846153846153846 |
| 3 | Negatif | 8 | 0.024615384615384615 |

Tabel 2. Tahap Klasifikasi KNN.

| Index | Nominal Value | Absolute Count | Fraction |
|-------|---------------|----------------|----------------------|
| 1 | Netral | 233 | 0.7169230769230769 |
| 2 | Positif | 84 | 0.25846153846153846 |
| 3 | Negatif | 8 | 0.024615384615384615 |

Evaluasi/Hasil Akhir

Pada tahap terakhir yaitu, evaluasi atau hasil pengujian setelah melakukan proses pengolahan data dengan melakukan perbandingan menggunakan metode Naive Bayes dan KNN, untuk melihat nilai keakurasian pada tahap klasifikasi yang sudah dilakukan. Nilai dari akurasi yang diperoleh dengan metode Naive Bayes sebesar: 90.71% +/- 4.66% (micro average: 90.77%) sedangkan untuk metode KNN sebesar accuracy: 74.78% +/- 3.74% (micro

average: 74.77%)

Tabel 3. Hasil Pengujian Metode Naive Bayes

| | true Netral | true Negatif | true Positif | class precision |
|---------------|-------------|--------------|--------------|-----------------|
| pred. Netral | 213 | 3 | 5 | 96.38% |
| pred. Negatif | 7 | 5 | 2 | 35.71% |
| pred. Positif | 13 | 0 | 77 | 85.56% |
| class recall | 91.42% | 62.50% | 91.67% | |

Tabel 4. Hasil Pengujian Metode KNN

| | true Netral | true Negatif | true Positif | class precision |
|---------------|-------------|--------------|--------------|-----------------|
| pred. Netral | 204 | 8 | 45 | 79.38% |
| pred. Negatif | 2 | 0 | 0 | 0.00% |
| pred. Positif | 27 | 0 | 39 | 59.09% |
| class recall | 87.55% | 0.00% | 46.43% | |

4. Kesimpulan dan Saran

Berdasarkan pengujian dan pengklasifikasian yang sudah dilakukan dari tahap awal sampai akhir dapat ditentukan hasilnya. Dari hasil perbandingan algoritma pada ke dua metode tersebut yang menggunakan dataset dapat disimpulkan bahwa klasifikasi algoritma Naive Bayes merupakan sebuah algoritma klasifikasi dengan tingkat keakurasian yang paling tinggi atau paling akurat dibandingkan dengan algoritma KNN. Nilai akurasi yang di peroleh dengan metode *Naive Bayes* sebesar: 88.56% +/- 4.71% (micro average: 88.62%) sedangkan untuk metode KNN hasil yang di peroleh dari analisis sentimen yaitu, sebesar: 74.78% +/- 3.74% (micro average: 74.77%).

5. Daftar Pustaka

[1] Narulita, L.F. and Sulistyawati, D.H., 2021. Pengumpulan Data Twitter Tentang Covid-19 di Indonesia untuk Menghitung Tingkat Engagement Pengguna. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(3), pp.565-570.

[2] Kartino, A. and Anam, M.K., 2021. Analisis Akun Twitter Berpengaruh terkait Covid-19 menggunakan Social Network Analysis. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(4), pp.697-704.

- [3] [Djamaludin, M.A., Triayudi, A. and Mardiani, E., 2022. Analisis Sentimen Tweet KRI Nanggala 402 di Twitter menggunakan Metode Naïve Bayes Classifier. *Jurnal JTIIK (Jurnal Teknologi Informasi dan Komunikasi)*, 6(2), pp.161-166.
- [4] Kaparang, S., Kaparang, D.R. and Rantung, V.P., 2021. Analisis Sentimen New Normal Pada Masa Covid-19 Menggunakan Algoritma Naive Bayes Classifier. *Jointer-Journal of Informatics Engineering*, 2(01), pp.16-23.
- [5] Syarifuddin, M., 2020. Analisis sentimen opini publik terhadap efek PSBB pada twitter dengan algoritma decision tree, knn, dan naïve bayes. *INTI Nusa Mandiri*, 15(1), pp.87-94.
- [6] Shamrat, M.F.M.J., Chakraborty, S., Imran, M.M., Muna, J.N., Billah, M.M., Das, P. and Rahman, O.M., 2021. Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*, 23(1), pp.463-470.
- [7] 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.
- [8] Harun, A. and Ananda, D.P., 2021. Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve bayes dan Decision Tree: Analysis of Public Opinion Sentiment About Covid-19 Vaccination in Indonesia Using Naïve Bayes and Decision Tree. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), pp.58-64.
- [9] Lengkong, N.C., Safitri, O., Machsus, S., Putra, Y.R., Syahadati, A. and Nooraeni, R., 2021. Analisis Sentimen Penerapan Psbb Di Dki Jakarta Dan Dampaknya Terhadap Pergerakan Ihsg. *Jurnal Teknoinfo*, 15(1), pp.20-25.
- [10] Septiana, R.D., Susanto, A.B. and Tukiyat, T., 2021. Analisis Sentimen Vaksinasi Covid-19 Pada Twitter Menggunakan Naive Bayes Classifier Dengan Feature Selection Chi-Squared Statistic dan Particle Swarm Optimization. *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, 5(1), pp.49-56.
- [11] Susilawati, S., Sembiring, Z. and Muhathir, M., 2020. Motion Monitoring System Based on IoT. *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, 3(2), pp.266-271.
- [12] Putra, T.W., Triayudi, A. and Andrianingsih, A., 2022. Analisis Sentimen Pembelajaran Daring Menggunakan Metode Naïve Bayes, KNN, dan Decision Tree. *Jurnal JTIIK (Jurnal Teknologi Informasi dan Komunikasi)*, 6(1), pp.20-26.
- [13] Pamungkas, F.S. and Kharisudin, I., 2021, February. Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter. In *PRISMA, Prosiding Seminar Nasional Matematika* (Vol. 4, pp. 628-634).
- [14] Pertiwi, M.W., 2019. Analisis sentimen opini publik mengenai sarana dan transportasi mudik tahun 2019 pada twitter menggunakan algoritma naïve bayes, neural network, KNN dan SVM. *Inti Nusa Mandiri*, 14(1), pp.27-32.
- [15] Lestari, S. and Saepudin, S., 2021, September. Analisis Sentimen Vaksin Sinovac Pada Twitter Menggunakan Algoritma Naive Bayes. In *Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra* (Vol. 1, No. 01, pp. 163-170).
- [16] Suryono, S., Utami, E. and Luthfi, E.T., 2018. Analisis Sentiment Pada Twitter Dengan Menggunakan Metode Naïve Bayes Classifier. *Seminar Nasional GEOTIK 2018*.

- [17] Fairuz, A.L., Ramadhani, R.D. and Tanjung, N.A.F., 2021. Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter. *Indonesian Journal of Data Science, IoT, Machine Learning and Artificial Intelligence*, 1(1), pp.41-150.
- [18] Arsyad, Z., 2019. Text Mining Menggunakan Generate Association Rule With Weight (Garw) Algorithm Untuk Analisis Teks Web Crawler. *INTERNAL (Information System Journal)*, 2(2), pp.153-171.
- [19] Sautomo, S., Hafidz, N., Achyani, Y.E. and Gata, W., 2020. SENTIMENT ANALYSIS DUE TO "MUDIK" PROHIBITED OF COVID-19 THROUGH TWITTER. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, 6(1), pp.7-12.
- [20] Yunitasari, Y. and Putera, A.R., 2021. Analisis Sentimen Masyarakat di Twitter Terkait Pandemi Covid-19. *SMATIKA JURNAL*, 11(01), pp.22-26.
- [21] Romadhon, M.R. and Kurniawan, F., 2021, April. A Comparison of Naive Bayes Methods, Logistic Regression and KNN for Predicting Healing of Covid-19 Patients in Indonesia. In *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)* (pp. 41-44). IEEE.
- [22] Fitriana, F., Utami, E. and Al Fatta, H., 2021. Analisis Sentimen Opini Terhadap Vaksin Covid-19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes. *Jurnal Komtika (Komputasi dan Informatika)*, 5(1), pp.19-25.