

**ANALISIS SENTIMEN TERHADAP PENGEMBANGAN VAKSIN COVID-19 DI
INDONESIA PADA PLATFORM *TWITTER* MENGGUNAKAN METODE *NAÏVE
BAYES* DAN *SUPPORT VECTOR MACHINE***

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata-1
Pada Jurusan Teknik Industri Fakultas Teknologi Industri**



Nama: Nikmah Rizki Ramadhani

No. Mahasiswa: 15 522 322

**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA**

2022

SURAT PERNYATAAN KEASLIAN

Demi Allah saya akui bahwa karya ini adalah karya saya sendiri kecuali kutipan dan ringkasan yang setiap salah satunya telah saya jelaskan sumbernya. Jika ditemukan dikemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah dalam karya tulis dan hak kekayaan intelektual maka saya bersedia ijazah yang saya terima untuk ditarik oleh Universitas Islam Indonesia.

Yogyakarta, 21 April 2022



Nikmah Rizki Ramadhani



FAKULTAS
TEKNIK INDUSTRI

Gedung IX, Mei Heris
Jl. Sekeloa No.14.5 Yogyakarta
Telp. (0274) 895287, 898444 ext 2513,
Fax. (0274) 899007

SURAT KETERANGAN PENELITIAN
Nomor : 229/A/Ka.Lab DATMIN/FTI-UII/IV/2022

Assalamu'alaikum Warahmatullahi Wabarakatuh

Kami yang bertanda tangan dibawah ini, menerangkan bahwa mahasiswa dengan keterangan sebagai berikut :

Nama : Nikmah Rizki Ramadhani
No. Mhs : 15522322
Dosen Pembimbing : Annisa Uswatun Khasanah, S.T., M.Sc.

Telah selesai melaksanakan penelitian yang berjudul "ANALISIS SENTIMEN TERHADAP PENGEMBANGAN VAKSIN COVID-19 DI INDONESIA PADA PLATFORM TWITTER MENGGUNAKAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE" di Laboratorium Data Mining, Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia tercatat mulai tanggal 13 Januari 2021 sampai dengan tanggal 23 Januari 2021

Demikian surat keterangan kami keluarkan, agar dapat dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Yogyakarta, 22 April 2021 M

Kepala Laboratorium
Data Mining

Annisa Uswatun Khasanah, S.T., M.Sc.

LEMBAR PENGESAHAN PEMBIMBING

**ANALISIS SENTIMEN TERHADAP PENGEMBANGAN VAKSIN COVID-19 DI
INDONESIA PADA PLATFORM TWITTER MENGGUNAKAN METODE NAÏVE BAYES
DAN SUPPORT VECTOR MACHINE**

TUGAS AKHIR

Disusun Oleh

Nama : Nikmah Rizki Ramadhani

No. Mahasiswa 15 522 322

Yogyakarta, 22 April 2022

Dosen Pembimbing



Annisa Uswatun Khasanah, S.T., MBA, M.Sc.

LEMBAR PENGESAHAN PENGUJI

**ANALISIS SENTIMEN TERHADAP PENGEMBANGAN VAKSIN COVID-19 DI
INDONESIA PADA PLATFORM TWITTER MENGGUNAKAN METODE NAÏVE BAYES
DAN SUPPORT VECTOR MACHINE**

TUGAS AKHIR

Oleh:

Nama : Nikmah Rizki Ramadhani
No. Mahasiswa : 15 522 322
Fak/Jurusan : FTI/Teknik Industri

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat
untuk memperoleh gelar Sarjana Satrata-1 Teknik Industri

Yogyakarta, April 2022

Tim Penguji

Annisa Uswatun Khasanah, S.T., MBA, M.Sc.

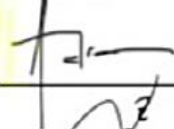
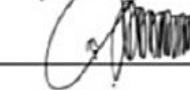
Ketua

Sri Indrawati, S.T., M.Eng.

Anggota I

Abdullah 'Azzam, S.T., M.T.

Anggota II

Mengetahui,

Ketua Program Teknik Industri

Fakultas Teknologi Industri

Universitas Islam Indonesia



Immawan, S.T., M.M


HALAMAN PERSEMBAHAN

*Dengan Nama Allah yang Maha Pengasih dan Maha Penyayang.
Alhamdulillahirobbil'alamin. Puji syukur tiada henti atas kehadiran Allah SWT atas nikmat yang telah diberikan. Skripsi ini saya persembahkan kepada:*

Kedua orang tua saya, Bapak Imroni Salam dan Ibu Catur Prasetyawati. Terima kasih telah mendoakan, mendidik, membina, dan selalu memberikan dukungan dari kecil hingga saat ini. Beliau adalah sosok utama dan segalanya dalam hidup saya yang tidak pernah akan terganti. Betapa segala pengorbanan tidak akan pernah bisa membalas segala bulir keringat dan air mata.

Teruntuk kakak, dan keluarga besar saya, terima kasih atas semangat, dukungan, dan motivasi yang selalu diberikan hingga pada tahap ini. Serta sahabat dan teman-teman saya yang senantiasa hadir mendengarkan, membantu dan menemani hari-hari saya selama masa perkuliahan.

HALAMAN MOTTO


 فَبِأَيِّ آلَاءِ رَبِّكُمَا تُكَذِّبَانِ

“Maka nikmat Tuhanmu yang manakah yang kamu dustakan?.” (QS. Ar Rahman: 13)



KATA PENGANTAR

Assalamu'allaikum Wr.Wb.

Segala puji dan syukur atas kehadiran Allah SWT atas segala rahmat dan hidayah-Nya serta shalawat dan salam semoga selalu tercurah kepada Nabi Muhammad SAW sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul penelitian “ANALISIS SENTIMEN TERHADAP PENGEMBANGAN VAKSIN COVID-19 DI INDONESIA PADA PLATFORM *TWITTER* MENGGUNAKAN METODE *NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINE*”.

Penulis menyadari bahwa penulisan Tugas Akhir ini tidak akan lancar tanpa adanya bimbingan dan dorongan dari semua pihak. Dengan segala kerendahan hati penulis menyampaikan terima kasih dan penghargaan yang setinggi-tingginya kepada semua pihak yang telah berjasa memberikan motivasi dalam rangka menyelesaikan Tugas Akhir ini. Penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Hari Purnomo, M.T. selaku Dekan Fakultas Teknologi Industri, Universitas Islam Indonesia.
2. Bapak Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D. selaku Ketua Jurusan Teknik Industri Fakultas Teknologi Industri, Universitas Islam Indonesia.
3. Bapak Dr. Taufiq Immawan, S.T., M.M. selaku Ketua Program Studi Teknik Industri Fakultas Teknologi Industri, Universitas Islam Indonesia.
4. Ibu Annisa Uswatun Khasanah, S.T., MBA., M.Sc. selaku dosen pembimbing I dan Ibu Harwati, S.T., M.T. selaku dosen pembimbing II Tugas Akhir program studi Teknik Industri Universitas Islam Indonesia.
5. Bapak Imroni Salam dan Ibu Catur Prasetyawati selaku orang tua saya, kakak tersayang Muhammad Andhika Pratama dan Mahma Matahari, serta keponakan saya tercinta Aubrienne Rahmandhika dan Arrumi Rahmandhika yang selalu memberikan kasih sayang dan support kepada saya hingga saat ini.
6. Seluruh keluarga besar Teknik Industri angkatan 2015 yang telah menemani perjuangan untuk mencapai kesuksesan masa depan.

7. Sahabat saya Bella Azis Dewanti Putri, yang sentiantiasa dengan senang hati berbagi ilmu dan mendengarkan segala keluh kesah, terimakasih atas dukungan dan semangat yang diberikan kepada penulis.

Penulis menyadari dalam menyusun dan menulis laporan Tugas Akhir ini masih terdapat kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran bersifat membangun untuk lebih baiknya Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat bagi semua pihak.

Wassalamu'alaikum Wr.Wb

Yogyakarta, 21 April 2022



Nikmah Rizki Ramadhani

ABSTRAK

Twitter merupakan salah satu media sosial yang digunakan untuk menyalurkan opini dan membahas berbagai masalah yang ada. Masalah yang banyak dibahas pada masa pandemic ini adalah salah satunya adalah Vaksin Covid-19. Opini dan masalah tersebut dapat diolah dengan cara yaitu analisis sentimen. *Naive Bayes* dan *Support Vector Machine* (SVM) merupakan teknik *machine learning* yang populer untuk pengklasifikasian. Kedua algoritma digunakan, untuk kemudian dibandingkan hasil akurasi. Tujuan dari penelitian ini adalah mengoptimasi algoritma dengan akurasi tertinggi antara *Naive Bayes* dan *Support Vector Machine* (SVM). Dalam penelitian ini *Naive Bayes Classifier* menghasilkan akurasi yang lebih tinggi dibandingkan dengan *Support Vector Machine* (SVM). *Naive Bayes* merupakan algoritma yang digunakan dalam pengkategorian teks, dengan menggabungkan probabilitas kata-kata dan kategori untuk memprediksi probabilitas kategori dari sebuah dokumen. Penggunaan metode naive bayes menghasilkan akurasi paling optimal sebesar 74,74%. Dapat disimpulkan bahwa hasil klasifikasi sentimen opini masyarakat terhadap pengembangan Vaksin Covid-19, data sentimen positif sejumlah 422 dan sentimen negatif sejumlah 346. Kemudian sentimen negatif dianalisa dan akar masalah divisualisasikan menggunakan diagram *fishbone*.

Kata kunci : Analisis Sentimen, *Twitter*, *Naive Bayes*, *Support Vector Machine* (SVM), *Fishbone Diagram*

DAFTAR ISI

COVER	i
SURAT PERNYATAAN KEASLIAN.....	ii
SURAT KETERANGAN PENELITIAN	iii
LEMBAR PENGESAHAN PEMBIMBING	iv
LEMBAR PENGESAHAN PENGUJI	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
KATA PENGANTAR	viii
ABSTRAK	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Sistematika Penulisan	5
BAB II.....	7
TINJAUAN PUSTAKA.....	7
2.1 Kajian Induktif	7
2.2 Kajian Deduktif	13
2.2.1 Analisis Sentimen	13
2.2.2 Twitter	13
2.2.3 Vaksin COVID-19	13
2.2.4 <i>Support Vector Machine</i>	14
2.2.5 <i>Naïve Bayes</i>	15
BAB III	17
METODE PENELITIAN	17
3.1 Populasi dan Sampel Penelitian.....	17

	xii
3.2	Jenis dan Sumber Data..... 17
3.3	Variabel Penelitian 17
3.4	Metode Pengumpulan Data 18
3.5	Analisa Data..... 19
3.6	Diagram Alur Kerangka Penelitian 19
BAB IV 23
HASIL DAN PEMBAHASAN 23
4.1	Proses Pengolahan Dataset..... 23
4.1.1	Proses Pengumpulan Data <i>Crawling</i> 23
4.1.2	Proses Seleksi Data..... 24
4.1.3	Proses <i>Cleansing</i> 24
4.2	Proses <i>pre-prossesing</i> 29
4.3	Pengujian Algoritma (<i>Cross Validation</i>) 35
4.3.1	Naive Bayes..... 35
4.3.2	<i>Support Vector Machine</i> 35
4.4	Hasil Analisa Akurasi 36
4.4.1	Tabel Perbandingan akurasi..... 37
4.4.2	Hasil Klasifikasi Sentimen 37
4.5	Diagram <i>Fishbone</i> untuk Analisa Sentimen Negatif 38
BAB IV 41
PENUTUP 41
5.1	Kesimpulan 41
5.2	Saran 41
DAFTAR PUSTAKA	
LAMPIRAN	



DAFTAR TABEL

Tabel 2 1 Posisi Penelitian dan Penelitian Sebelumnya	12
Tabel 4 1 Tabel Perbandingan Akurasi	37



DAFTAR GAMBAR

Gambar 1.1 Grafik Tingkat Penggunaan Media Sosial di Indonesia	2
Gambar 2.1 Pencarian Hyperplane <i>Support Vector Machine</i>	15
Gambar 3.1 Diagram Alur Penelitian	20
Gambar 4. 1 Hasil Proses Pengumpulan Data <i>Crawling file CSV</i>	23
Gambar 4. 2 Hasil data dalam bentuk <i>file Microsoft Excel</i>	24
Gambar 4. 3 Hasil Proses Seleksi Data.....	24
Gambar 4. 4 Proses <i>Select Attributes Data</i>	25
Gambar 4. 5 Hasil Proses <i>Select Attributes Data</i>	26
Gambar 4. 6 Tampilan Desain dari Proses <i>Replace Link</i>	26
Gambar 4. 7 Hasil Proses <i>Replace Link</i>	27
Gambar 4. 8 Tampilan Desain dari Proses <i>Replace @</i>	27
Gambar 4. 9 Hasil <i>Replace @</i>	27
Gambar 4. 10 Tampilan Desain dari Proses <i>Replace #</i>	28
Gambar 4. 11 Hasil Proses <i>Replace #</i>	28
Gambar 4. 12 Hasil Cleansing dengan <i>Microsoft Excel</i>	29
Gambar 4. 13 Tampilan Desain dari <i>Pre-Pocessing</i>	29
Gambar 4. 14 Tampilan Desain dari Operasi <i>Tokenize</i>	30
Gambar 4. 15 Tampilan Desain dari Operasi <i>Tokenize (2)</i>	30
Gambar 4. 16 Tampilan Hasil Proses <i>Tokenize</i>	30
Gambar 4. 17 Tampilan Desain Proses <i>Transform Cases</i>	31
Gambar 4. 18 Tampilan Hasil Proses <i>Transform Cases</i>	31
Gambar 4. 19 Tampilan Desain Proses Filter <i>Tokens (By Length)</i>	32
Gambar 4. 20 Tampilan Hasil Proses Filter <i>Tokens (By Length)</i>	32
Gambar 4. 21 Tampilan Desain Proses Filter <i>Stopwords (Dictionary)</i>	32
Gambar 4. 22 Tampilan Hasil Proses Filter <i>Stopwords (Dictionary)</i>	33
Gambar 4. 23 Hasil Stemming dengan <i>Microsoft Excel</i>	33
Gambar 4. 24 Tampilan Desain Proses Stem (<i>Dictionary</i>)	34
Gambar 4. 25 Tampilan Hasil Proses Stem (<i>Dictionary</i>).....	34
Gambar 4. 26 Tampilan Desain Proses <i>Cross Validation</i> dengan <i>Algoritma Naïve Bayes</i>	35
Gambar 4. 27 <i>Confussion Matrix Naïve Bayes</i>	35
Gambar 4. 28 Tampilan Desain Proses <i>Cross Validation</i> dengan <i>Algoritma Support Vector Machine</i>	36
Gambar 4. 29 <i>Confussion Matrix Support Vector Machine</i>	36
Gambar 4. 30 Hasil Polaritas Sentimen dan Prioritas (Sentimen).....	37
Gambar 4. 31 Diagram Hasil Analisa Klasifikasi Sentimen	38
Gambar 4. 32 Diagram <i>Fishbone</i> Sentimen Negatif Vaksin Covid-19.....	39

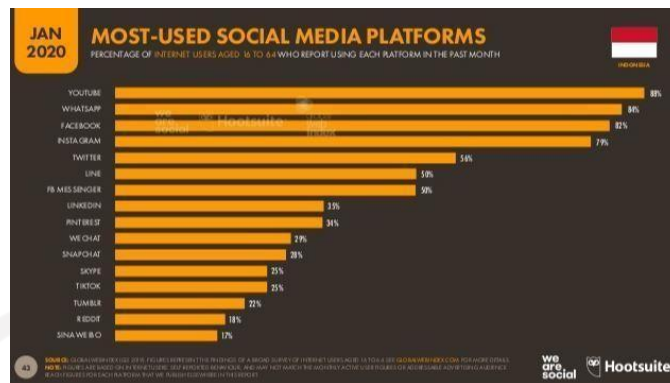
BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi di Indonesia saat ini bertumbuh sangat pesat, salah satunya perkembangan teknologi pada jaringan internet. Menurut riset yang dilakukan *We Are Social* bertajuk "*Global Digital Reports 2020*" pada akhir Januari 2020 ini menunjukkan hampir 64% penduduk Indonesia sudah terkoneksi dengan jaringan internet. Selain itu jumlah pengguna internet di Indonesia sudah mencapai 175,4 juta orang dari total jumlah penduduk Indonesia sekitar 272,1 juta. Riset ini juga menunjukkan bahwa 95% orang menggunakan internet untuk mengakses media sosial (Kementerian Komunikasi dan Informatika, 2017). Media sosial merupakan media di internet yang memungkinkan pengguna untuk mewakili dirinya maupun berinteraksi, bekerja sama, berbagi, berkomunikasi dengan pengguna lain, dan membentuk ikatan sosial secara virtual (Nasrullah, 2015). Beberapa media sosial yang biasa digunakan masyarakat yaitu *Facebook*, *Instagram* dan *Twitter*.

Twitter adalah salah satu media sosial yang diminati di Indonesia. *Twitter* merupakan jaringan informasi yang terdiri atas pesan-pesan sepanjang 140 karakter yang disebut *Tweet* (*Twitter*, 2013). *Twitter* dapat digunakan untuk membaca pikiran atau sentimen pemilik akun, menyebarkan berita terkini dan pengalaman, menyebarkan informasi dan komentar, serta menjaga kedekatan dengan sekitar (O'Reilly, dkk., 2009). Pada awal tahun 2020 *Twitter* berada di peringkat 5 sebagai platform yang paling sering digunakan di Indonesia. Berikut ini grafik tingkat penggunaan media sosial di Indonesia



Gambar 1.1 Grafik Tingkat Penggunaan Media Sosial di Indonesia

Salah satu fitur yang dimiliki *Twitter* adalah percakapan antar user yang dapat dilakukan dalam bentuk kolom komentar atau “*tweet*”. Percakapan antar-user di *Twitter* seringkali berisi informasi acak, namun data yang tidak berstruktur ini dapat disaring berdasarkan kata kunci tertentu.

Topik hangat di Indonesia yang masih diperbincangkan terutama pada media sosial *Twitter* adalah *Coronavirus Disease 2019 (COVID-19)*. Virus ini menjadi sebuah pandemi yang telah menyebar ke seluruh dunia dengan sangat cepat, termasuk Indonesia. Pandemi yang sudah berjalan lebih dari lima bulan di Indonesia hingga saat ini belum ada tanda-tanda penurunan penularan. Pemerintah mengungkapkan kasus COVID-19 pada pertengahan Agustus ini mencapai 141.370 orang, terhitung sejak diumumkannya pasien pertama pada 2 Maret 2020 (Kompas, 2020).

Beberapa cara telah dilakukan oleh pemerintah, institusi, hingga kalangan masyarakat untuk memutus rantai penyebaran virus COVID-19. Upaya tersebut seperti gerakan physical distancing, membuat protokol kesehatan, Pembatasan Sosial Berskala Besar (PSBB) dan juga pengembangan vaksin COVID-19. Vaksin COVID-19 ini dikembangkan oleh Lembaga Biologi Molekuler Eijkman bekerja sama dengan beberapa universitas negeri di Indonesia, serta dua perusahaan kimia Indonesia yang telah bekerja sama dengan lembaga bioteknologi asing menuai berbagai opini di masyarakat.

Aspirasi, opini dan kritik masyarakat sangat sulit disampaikan secara langsung karena beberapa kendala seperti keterbatasan waktu dan ruang. Namun, media sosial mampu membantu menyalurkan aspirasi tersebut, salah satunya melalui *Twitter*. Beragam opini

menjadi perbincangan terkait pengembangan vaksin COVID-19 di Indonesia menjadi topik yang sering diperbincangkan di *Twitter*. Namun setiap *tweet* dari pengguna seringkali tidak mengandung makna yang selaras, sehingga perlu dilakukan analisis terhadap opini mereka (Syarifuddin, 2020).

Analisis opini dalam *tweet* dilakukan untuk menyelaraskan dan memberi pandangan baru mengenai suatu isu tentang pengembangan vaksin COVID-19. Penggunaan analisis sentimen sebagai tolak ukur dalam menilai suatu hal, kejadian atau produk telah sering dilakukan. Dalam beberapa tahun terakhir, teknik ini telah digunakan untuk menganalisis ulasan produk online, berita online, pemilihan umum, bencana, pasar saham, dan media sosial, terutama *Twitter* (Gormantara, 2020). Adapun jenis analisis sentimen yang sering digunakan dalam kalangan peneliti yaitu, analisis sentimen dalam bentuk dokumen atau kalimat (Fink, dkk., 2011).

Analisis sentimen termasuk dalam cabang ilmu dari Text Mining. Ilmu ini merupakan bidang penelitian berkelanjutan yang berada di antara berbagai bidang seperti Data Mining, Natural Language Processing (NLP) dan Machine Learning yang berfokus pada ekstraksi sentimen dalam sebuah kalimat berdasarkan isinya (Al-Ayyoub, dkk., 2019). Dalam melakukan analisis sentimen diperlukan metode yang menunjang klasifikasi. *Support Vector Machine* (SVM) merupakan salah satu metode klasifikasi dengan menggunakan metode *machine learning* (*supervised learning*) yang memprediksi kelas berdasarkan pola dari hasil proses training. Sedangkan *Naïve Bayes* merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dg menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dr *Naïve Bayes Classifier* ini adalah asumsi yg sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian (Olson Delen, 2008)

Oleh karena itu, penelitian ini bertujuan untuk mengetahui gambaran umum pengembangan vaksin COVID-19 berdasarkan opini pada *Twitter*. Hasil klasifikasi dengan akurasi tertinggi yang diperoleh dengan perbandingan hasil metode *Support Vector Machine* (SVM) dan *Naïve Bayes* dapat digunakan sebagai acuan dalam kebijakan yang perlu diambil

dalam proses pengembangan vaksin COVID-19 di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah tersebut, maka rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana performa algoritma *Support Vector Machine* dan *Naïve Bayes* dalam mengklasifikasikan opini menjadi kelas optimis dan pesimis?
2. Bagaimana opini masyarakat di *Twitter* mengenai gambaran umum pengembangan vaksin COVID-19 dengan menggunakan metode *Support Vector Machine* dan *Naïve Bayes*?
3. Apakah informasi yang didapatkan setelah proses klasifikasi dilakukan?

1.3 Batasan Masalah

Penetapan batasan-batasan masalah dalam penelitian dilakukan dengan tujuan agar penelitian yang dilakukan lebih terarah. Adapun batasan dalam penelitian ini adalah opini masyarakat berupa *tweet* di *Twitter* mengenai pengembangan vaksin COVID-19 mulai tanggal 13 sampai 23 Januari 2021.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

1. Mengetahui performa algoritma *Support Vector Machine* dan *Naïve Bayes* dalam mengklasifikasikan opini menjadi kelas optimis dan pesimis.
2. Mengetahui opini masyarakat di *Twitter* mengenai gambaran umum pengembangan vaksin COVID-19 dengan menggunakan metode *Support Vector Machine* dan *Naïve Bayes*
3. Mengetahui informasi yang didapatkan setelah proses klasifikasi dilakukan.

1.5 Manfaat Penelitian

1. Bagi Pemerintah

Hasil penelitian ini diharapkan dapat menjadi bahan masukan dan evaluasi dalam pengembangan kegiatan vaksinasi COVID-19

2. Bagi Peneliti

Menambah pengetahuan dan memperluas wawasan peneliti mengenai pengukuran opini masyarakat di *Twitter* dengan menggunakan performa algoritma *Support Machine Vector*.

1.6 Sistematika Penulisan

Untuk lebih terstrukturanya penelitian ini, maka selanjutnya sistematika penulisan dari penelitian disusun sebagai berikut.

BAB I PENDAHULUAN

BAB I merupakan pembukaan isi keseluruhan Tugas Akhir yang menguraikan latar belakang permasalahan, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penelitian.

BAB II KAJIAN LITERATUR

BAB II menyajikan dan menjelaskan dasar-dasar teori yang berkaitan dengan permasalahan dalam penelitian ini. Selain itu juga memuat uraian tentang hasil penelitian yang pernah dilakukan sebelumnya oleh peneliti lain yang memiliki hubungan dengan penelitian yang dilakukan.

BAB III METODE PENELITIAN

BAB III berisi bagan alur penelitian beserta penjelasannya, teknik yang digunakan dalam penelitian serta data yang akan digunakan sebagai data kajian.

BAB IV PENGUMPULAN DAN PENGOLAHAN DATA

BAB IV menyajikan pengumpulan data yang diperoleh selama penelitian serta hasil pengolahan data menggunakan metode yang diterapkan agar tercapainya tujuan dari penelitian. Pengolahan data juga termasuk analisis

yang dilakukan terhadap hasil yang diperoleh. Pada sub bab ini merupakan acuan untuk pembahasan yang akan ditulis pada BAB V yaitu pembahasan.

BAB V PEMBAHASAN

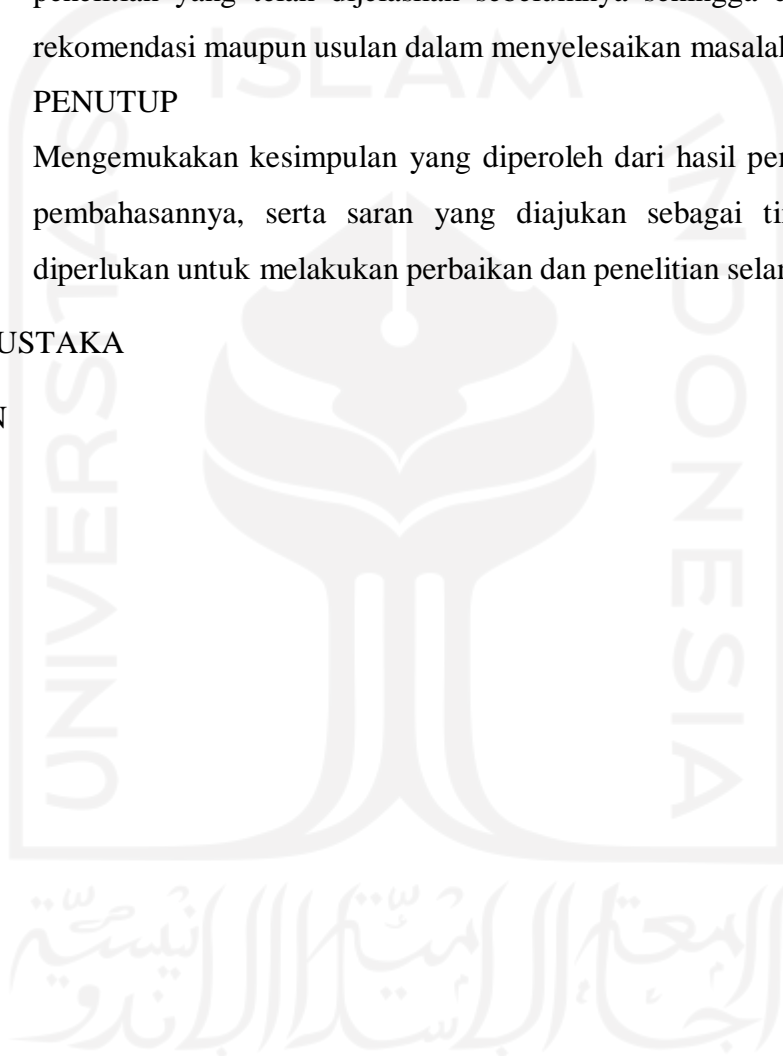
BAB V memaparkan pembahasan mengenai hasil pengolahan data serta batasannya yang telah dilakukan dengan mengacu pada teori dan alur penelitian yang telah dijelaskan sebelumnya sehingga dapat memberikan rekomendasi maupun usulan dalam menyelesaikan masalah.

BAB VI PENUTUP

Mengemukakan kesimpulan yang diperoleh dari hasil pengolahan data dan pembahasannya, serta saran yang diajukan sebagai tindak lanjut yang diperlukan untuk melakukan perbaikan dan penelitian selanjutnya.

DAFTAR PUSTAKA

LAMPIRAN



BAB II

TINJAUAN PUSTAKA

2.1 Kajian Induktif

Dari penelitian dengan menggunakan metode analisis sentimen telah banyak dilakukan pada berbagai objek. Hennie Tuhuteru (2020) melakukan penelitian dengan menggunakan metode SVM untuk mengetahui sentimen masyarakat terhadap rencana penerapan PSBB di Kota Ambon melalui *tweet* dan komentar di *Twitter*. Dari 1075 *tweet* menunjukkan sentimen positif sebesar 28%, sentimen negatif sebesar 27%, dan sentimen netral sebesar 45% yang artinya sentimen masyarakat terhadap rencana penerapan PSBB di Kota Ambon berdasarkan komentar di media sosial cukup berimbang antara sentimen positif dan negatif serta didominasi komentar dengan sentimen netral.

Rony Budianto (2019) melakukan penelitian untuk mengetahui jumlah sentimen positif dan sentimen negatif masyarakat pada Institusi Komisi Pemilihan Umum Indonesia (KPU) di Pemilu 2019 melalui *Twitter*, penelitian ini memiliki komposisi data latih 80% dan data uji 20% menghasilkan hasil sentimen negatif sebesar 589 (58.9%) dan sentimen positif sebesar 411 (41.1%). Penggunaan metode SVM dirasa cukup baik karena memiliki tingkat akurasi sebesar 80.55%, presisi sebesar 74%, *recall* sebesar 77% dan *f1-score* sebesar 75.4%.

Penggunaan metode SVM juga bisa dikaitkan dengan metode lain seperti *Maximum Entropy* untuk melihat *review* konsumen pada *e-commerce*. Pratiwi (2018) dengan menggunakan metode SVM dan *Maximum Entropy* untuk mengklasifikasikan data *review* dari *e-commerce* Bukalapak. Jika dilihat dari tingkat akurasinya, metode SVM memiliki tingkat akurasi sebesar 91,95%. Sedangkan dengan metode *Maxent* memiliki tingkat akurasi yang lebih tinggi yaitu sebesar 92,98%. Selanjutnya, diketahui juga bahwa sentimen positif diantaranya terkait barang, transaksi, fitur, pelayanan, pesanan, pengiriman, respon, berbelanja, akulaku, kebutuhan dan cicilan. Sedangkan pada sentimen negatif yang sering dikeluhkan diantaranya barang, update, server, chat, email, transaksi, upload, promo, voucher, bukadompet dan upgrade.

Selain melakukan analisis sentimen dengan metode SVM, tentunya ada metode lain

yang bisa dibandingkan juga akurasiya yaitu *Naïve Bayes*. Evi (2018) membandingkan kedua metode tersebut untuk mengetahui tanggapan mengenai Traveloka yang diperoleh melalui media sosial *Twitter* pada bulan Maret dan bulan April 2018. SVM bekerja dengan cara mencari hyperplane terbaik yang memisahkan dua buah kelas, sedangkan *Naive Bayes* bekerja dengan cara mencari nilai probabilitas bersyarat terbesar. Berdasarkan tingkat akurasi terbesar dengan menggunakan 10 kombinasi dan pada masing- masing kombinasi dilakukan 10 kali pengacakan, diperoleh kesimpulan bahwa algoritma SVM memiliki tingkat akurasi lebih tinggi dibandingkan algoritma *Naive Bayes* dengan rata-rata besarnya tingkat akurasi sebesar 79,14% pada SVM dan 75,61% pada *Naive Bayes*.

Penggunaan metode SVM juga dilakukan Erfina, dkk. (2020) dalam penelitiannya mengenai aplikasi pembelajaran online di play store pada masa pandemi Covid-19 yang menyatakan bahwa SVM memiliki tingkat keakuratan yang dinilai cukup tinggi untuk menentukan aplikasi *E-learning* mana yang terbaik. Dari 1500 ulasan yang didapat pada masing-masing aplikasi menunjukkan bahwa tingkat keakuratan untuk aplikasi terbaik pertama yaitu Ruang Guru sebesar 83,33%, kedua Zenius sebesar 82,67%, ketiga Kelas Pintar sebesar 82,00%, keempat Edmodo sebesar 74,67%, dan kelima Google Classroom sebesar 64,33%.

Nooraeni, dkk., (2020) melakukan penelitian dengan menggunakan analisis sentimen dan menggunakan metode *Support Vector Machine* (SVM) untuk menganalisis respon masyarakat mengenai ketidakterbukaan DPR atas keputusannya untuk melakukan revisi UU KPK bahwa terdapat 60,9% merupakan sentimen negatif menunjukkan sikap tidak setuju terhadap revisi Undang-Undang KPK. Sementara sisanya, 39,1% merupakan sentimen positif menunjukkan sikap setuju terhadap revisi Undang-Undang KPK. Penggunaan metode SVM juga cukup baik karena memiliki nilai akurasi sebesar 81,32%, nilai sensitivitas sebesar 71,47% dan nilai spesifisitas sebesar 87,64%.

Umumnya penggunaan metode SVM dilakukan untuk menganalisis sentimen melalui berbagai platform seperti media sosial. Maulana, dkk. (2018) melakukan penelitian untuk mengetahui kecepatan penyebaran informasi melalui media sosial. Terlihat bahwa hasil feedback dari pemasaran atau penyebaran informasi yang didapatkan dari konsumen,

masyarakat bahkan instansi tertentu. Adapun hasil penelitian adanya perbedaan data yang didapat dari *Facebook* dan *Twitter* tidak mempengaruhi kecepatan hasil akhir dari pengelompokan nilai positif, negatif dan netral.

Namun, analisis sentimen tidak selalu menggunakan metode SVM pada pengukurannya. Syarifuddin (2020) melakukan penelitian mengenai untuk mengetahui sentimen opini publik mengenai Covid-19 pada *Twitter* menggunakan metode *Naïve Bayes* Dan Knn. Dari 1098 opini dengan kata kunci “COVID-19” di *Twitter*, metode *Naïve Bayes* memiliki nilai lebih tinggi, salah satunya adalah tingkat akurasi sebesar 63.21%, sedangkan metode KNN sebesar 58.10%, dan didapatkan pula kecenderungan opini masyarakat di *Twitter* condong positif, hal tersebut dapat dilihat dari jumlah opini positif sebesar 610 sedangkan negatif 488, ditunjang dengan hasil pengujian precision di metode *Naïve Bayes* dengan nilai positif lebih tinggi daripada negatif yaitu 66.40% : 58.94%.

Begitu juga dengan Feby, dkk., (2018) melakukan penelitian dengan metode klasifikasi K-NN. Bahwa terdapat sentimen netral mengenai program imunisasi measles rubella di Indonesia pada *Twitter* lebih banyak muncul dengan pelabelan manual dan sentimen positif lebih banyak muncul dengan *RapidMiner*. *Twitter* dapat digunakan memprediksi opini masyarakat tentang imunisasi MR di Indonesia. Namun untuk prediksi manual diperlukan waktu banyak sehingga diperlukan tool yang dapat membantu mempercepat prediksi tetapi *tool* tersebut perlu memperhatikan keseimbangan data latih dan penggunaan dataset yang di dalamnya terdapat kata yang sama.

Dwiraswati (2019) melakukan penelitian dengan menggunakan metode pengklasifikasian berbasis machine learning, yaitu *Naive Bayes Classifier*. Metode ini dapat diterapkan untuk melakukan analisis sentimen terhadap penggunaan antibiotik di Indonesia. Hasil analisis digunakan sebagai pertimbangan dalam menentukan strategi kampanye penggunaan antibiotik yang tepat.

Penggunaan metode SVM juga terdapat pada penelitian yang dilakukan oleh Tineges (2020) untuk opini masyarakat mengenai layanan indihome melalui *Twitter* dan berdasarkan 290 data *tweet* yang ada memiliki hasil nilai positif sebesar 18,4% dan hasil nilai negatif sebesar 81,6% yang artinya tingkat kepuasan pengguna layanan *Indihome* cukup rendah.

Adapun dengan nilai akurasi sebesar 87%, ketepatan antara hasil prediksi dengan data sebenarnya (*precision*) sebesar 86%, tingkat keberhasilan sistem dalam memprediksi sebuah data (*recall*) sebesar 95%, tingkat kesalahan semua data yang diprediksi (*error rate*) sebesar 13%, sedangkan untuk nilai perbandingan rata-rata *precision* dan *recall* (*f1score*) adalah sebesar 90%.

Metode SVM juga dapat digunakan untuk mengetahui opini masyarakat terhadap seseorang. Taufik, dkk., (2018) melakukan penelitian untuk mengetahui analisis sentimen terhadap tokoh publik. Kinerja algoritma SVM ini menggunakan presisi, akurasi, dan *recall*. Hasil menunjukkan bahwa Kernel linear memiliki tingkat presisi yang paling baik sekitar 80%. Kernel *sigmoid* memiliki tingkat *recall* yang paling baik sekitar 85%. Kernel *sigmoid* memiliki tingkat akurasi yang paling baik sekitar 81%.

Aplikasi *Twitter* juga menjadi platform yang memudahkan peneliti untuk mendapatkan data yang dibutuhkan. Seperti yang dilakukan dalam penelitian Rahutomo (2018) untuk mendapatkan data opini konsumen pada *review* film. Dengan menggunakan data SVM dari 1.027 *tweet* yang didapatkan mengenai film populer tahun 2016 memiliki hasil akurasi klasifikasi nilai *precision* positif memiliki rata-rata sebesar 79,97%, 78,71%, 84,02%, dan 85,54%. Nilai *precision* negatif memiliki rata-rata sebesar 81,73 %, 87,41%, 87,37%, dan 93,61% dan nilai *precision* netral memiliki rata-rata sebesar 67,13%, 69,47%, 74,08%, dan 74,14%.

Dalam penelitian Amar, P., (2020) *Support Vector Machine* (SVM) digunakan untuk mengetahui analisis sentimen keberlanjutan pembelajaran daring dengan hasil menunjukkan bahwa evaluasi tertinggi pada 8-Fold Cross Validation dengan *accuracy* sebesar 86,00%, *precision* sebesar 87,38%, dan *recall* sebesar 85,02%.

Ratino, dkk., (2020) dalam melakukan penelitian menggunakan *Support Vector Machine* dan *Naïve Bayes*. Hasil menunjukkan bahwa algoritma klasifikasi yang digunakan yaitu *Naïve Bayes* dengan hasil akurasi 78,02% dan AUC 0,714, sedangkan *Support Vector Machine* menghasilkan akurasi sebesar 80,23% dan AUC 0,904. Memiliki selisih akurasi 2,21%. Setelah di optimasi dengan operator Particle Swarm Optimization, algoritma *Naïve Bayes* (PSO) menghasilkan akurasi sebesar 79,07% dan AUC 0,729, sedangkan algoritma

Support Vector Machine (PSO) menghasilkan akurasi sebesar 81,16% dan AUC 0,903. Memiliki selisih akurasi sebesar 2,09%. Hasil pengujian algoritma, *Support Vector Machine* berbasis PSO maupun tidak, selalu dapat menghasilkan akurasi yang lebih tinggi.

Dalam penelitian ini peneliti menjadikan isu vaksin Covid-19 sebagai pembahasan utama yang akan diteliti menggunakan dua metode *Support Vector Machine* (SVM) dan *Naïve Bayes* yang kemudian dibandingkan nilai akurasinya, sumber data dalam penelitian ini didapatkan dari *Twitter*. Berikut adalah tabel yang menjelaskan posisi penelitian jika dibandingkan dengan penelitian sebelumnya:



Tabel 2 1 Posisi Penelitian dan Penelitian Sebelumnya

Penulis	Tahun	Aplikasi	Objek			Metode		
			Peraturan Pemerintah	Isu Covid-19	Tokoh Publik	SVM	Naive	KNN
Hennie Tuhuteru	2020			√		√		
Rony Budianto	2019		√	√		√		
Pratiwi	2018	√				√		
Evi	2018	√					√	
Erfina, dkk.	2020	√				√		
Nooraeni, dkk.	2020		√			√		
Maulana, dkk.,	2018	√				√		
Syarifuddin	2020			√			√	√
Feby, dkk.,	2018		√					√
Dwiraswati	2019	√					√	
Tineges	2020	√				√		
Taufik, dkk.,	2018				√	√		
Rahutomo	2018	√				√		
Amar,P.	2020	√				√		
Ratino, dkk.,	2020					√	√	
Usulan Penelitian	2021			√		√		

2.2 Kajian Deduktif

Kajian deduktif merupakan kumpulan dari dasar-dasar teori yang berkaitan dengan permasalahan serta kajian-kajian teoritis dari para ahli terkait dengan metode yang akan digunakan. Kajian deduktif pada penelitian ini adalah sebagai berikut.

2.2.1 Analisis Sentimen

Analisis sentimen merupakan salah satu bidang dari ilmu komputer yang bertujuan untuk menganalisa emosi, penilaian, sikap, pendapat, sentimen, evaluasi seseorang terhadap seorang pembicara atau penulis berkenaan dengan suatu produk, layanan, organisasi, individu, tokoh publik, topik, acara, ataupun kegiatan tertentu (Liu, 2012). Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah 13 masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif (Liu, 2010)

Analisis sentimen biasanya dilakukan untuk mengumpulkan dan mengetahui opini masyarakat dalam postingan Blog, *Twitter*, Facebook, dan yang lainnya untuk mengetahui opini publik terhadap suatu objek (Pintoko, dkk., 2018). Analisis sentimen sangatlah berguna sebagai pemroses penganalisis komentar seperti pendapat kemudian diproses menjadi sesuatu yang lebih bermakna.

2.2.2 Twitter

Twitter adalah situs web dimiliki dan dioperasikan oleh *Twitter, Inc.*, yang menawarkan jaringan sosial berupa microblog. Disebut microblog karena situs ini memungkinkan penggunaannya mengirim dan membaca pesan blog seperti pada umumnya namun terbatas hanya sejumlah 140 karakter yang ditampilkan pada halaman profil pengguna. Pesan-pesan ini disebut *Tweet* (*Twitter*, 2013). Menurut O'Reilly dan Milstein (2009) *Twitter* adalah layanan perpesanan yang membagikan berbagai karakteristik dengan alat komunikasi yang sudah digunakan. *Twitter* memiliki unsur-unsur yang mirip dengan surat elektronik (email), *instant messenger*, pesan singkat (SMS), *blogging*, RSS, jejaring sosial, dan sebagainya.

2.2.3 Vaksin COVID-19

Vaksin adalah bahan atau senyawa yang mengandung virus atau bakteri yang dilemahkan atau dimatikan (*World Health Organization*, 2019). Vaksin menstimulasi

sistem kekebalan tubuh untuk menghasilkan antibodi, persis seperti terkena penyakit tersebut dan setelah divaksinasi, tubuh akan mengembangkan kekebalan terhadap penyakit tersebut, tanpa harus terkena penyakit tersebut terlebih dahulu (*Centers for Disease Control and Prevention*, 2012). Vaksin dibutuhkan untuk menangani pandemi yang sedang terjadi. COVID-19 adalah penyakit menular yang disebabkan oleh virus korona yang paling baru ditemukan dimana virus dan penyakit baru ini tidak diketahui sebelum wabah dimulai di Wuhan, China, pada Desember 2019 dan COVID-19 sekarang menjadi pandemi yang menyerang banyak negara secara global (*World Health Organization*, 2020).

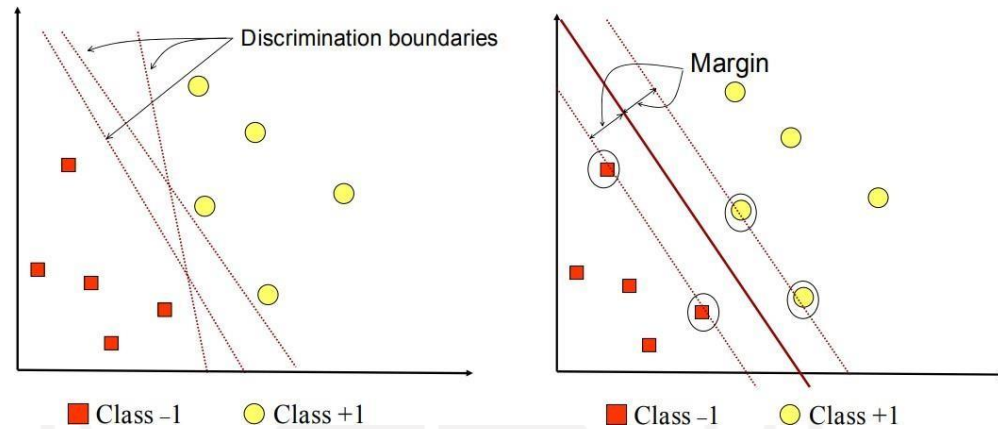
Saat ini terdapat lebih dari 100 kandidat vaksin COVID-19 yang sedang dikembangkan, dan beberapa di antaranya masih dalam tahap uji coba pada manusia. Tidak terkecuali untuk Indonesia, saat ini Indonesia memiliki 2 pilihan vaksin untuk COVID-19, pertama vaksin Merah Putih yang dikembangkan Kementerian Riset dan Teknologi, Badan Riset dan Inovasi Nasional serta Lembaga Biologi Molekuler Eijkman (LBME) dan pilihan kedua adalah mengembangkan kerjasama internasional (*Tim Komunikasi Komite Penanganan COVID-19*, 2020).

Menurut Menteri Riset dan Teknologi Prof Bambang Brodjonegoro, vaksin Corona Merah Putih yang dikembangkan Eijkman ditargetkan akhir 2020 akan memasuki uji klinis pada hewan (Astuti, 2020). Lembaga Eijkman sedang mengembangkan vaksin Merah Putih berbasis virus yang bersirkulasi di Indonesia dengan menggunakan platform protein rekombinan yang menargetkan Protein S dan Protein N (Sagita, 2020). Selain Kementerian Riset dan Teknologi, Badan Riset dan Inovasi Nasional serta Lembaga Biologi Molekuler Eijkman, beberapa institusi lain yang juga mengembangkan Vaksin Corona Merah Putih adalah Universitas Indonesia, Institut Teknologi Bandung, Universitas Airlangga, serta Lembaga Ilmu Pengetahuan Indonesia (LIPI) (Sagita, 2020).

2.2.4 *Support Vector Machine*

Support Vector Machine (SVM) merupakan salah satu metode klasifikasi dengan menggunakan machine learning (*supervised learning*) yang memprediksi kelas berdasarkan model atau pola dari hasil proses training (Novantirani, dkk., 2015).

Klasifikasi dilakukan dengan mencari hyperplane atau garis pembatas (decision boundary) yang memisahkan antara suatu kelas dengan kelas lain. Penggambaran proses pencarian hyperplane pada SVM ditunjukkan pada Gambar 2.1.



Gambar 2.1 Pencarian Hyperplane *Support Vector Machine*

Hyperplane pemisah terbaik antara kedua kelas diperoleh dengan cara mengukur margin dari *hyperplane* dan mencari margin terbesar. Margin adalah jarak antara *hyperplane* dengan data terdekat dari masing-masing kelas. Data yang paling dekat dengan *hyperplane* disebut sebagai support vector. Garis solid pada Gambar 2.1 menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua kelas, sedangkan titik merah dan kuning yang berada dalam lingkaran hitam adalah *support vector* (Nugroho, dkk., 2003). Metode ini dipilih karena memiliki akurasi yang lebih baik dari metode yang lain seperti *Naïve Bayes Classifier* (NBC) dan K-Nearest Neighbor (K-NN) (Tuhuteru, dkk., 2018).

2.2.6 *Naïve Bayes*

Naive Bayes Classifier merupakan teknik klasifikasi berdasarkan Teorema Bayes dengan asumsi independensi di antara para prediktor. *Naive Bayes Classifier* memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. *Naïve bayes* merupakan metode paling sederhana dari pengklasifikasian probabilitas, memiliki tingkat akurasi yang sangat tinggi ketika diaplikasikan pada *database* dengan *big data* (Wati, 2016). Selain itu model *naïve bayes* juga memiliki tingkat kesalahan yang sangat minimum

dibandingkan dengan algoritma klasifikasi lainnya (Liu dkk., 2016).

Algoritma naïve bayes hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Karena yang diasumsikan sebagai variabel *independent*, maka hanya variansi dari suatu variabel dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarians.

2.2.7 Diagram *Fishbone*

Fishbone diagram sering disebut *Cause and Effect* diagram adalah sebuah diagram yang menyerupai tulang ikan yang dapat menunjukkan sebab akibat dari suatu permasalahan (John Bank, 1992). *Fishbone* diagram juga merupakan salah satu tool dari *7 basic quality tools*. *Fishbone* diagram digunakan ketika kita ingin mengidentifikasi kemungkinan penyebab masalah dan terutama ketika sebuah team cenderung jatuh berpikir pada rutinitas. Bagian tulang dari pada diagram tulang ikan dan permasalahan yang ingin diketahui penyebabnya terletak pada bagian kepala ikan. Setiap faktor pada tulang memiliki akar permasalahannya masing-masing, melalui *Fishbone* diagram maka akar-akar permasalahan dengan mudah untuk diketahui. Berikut ini langkah-langkah untuk membuat *Fishbone*:

- a). Tulis permasalahan utama pada bagian kanan (kepala ikan). Gambar garis panah dari kiri ke kanan mengarah ke permasalahan.
- b). Identifikasi semua kategori utama penyebab masalah.
- c). Gunakan panah yang lebih kecil untuk menjelaskan akar permasalahan sehingga menjadi lebih detail.
- d). Ulangi langkah (c) berulang-ulang sehingga menemukan akar permasalahan yang paling mendasar.

BAB III

METODE PENELITIAN

3.1 Populasi dan Sampel Penelitian

Populasi dalam penelitian ini adalah opini berupa *tweet* mengenai pengembangan vaksin COVID-19. Sedangkan untuk sampel yang digunakan adalah opini berupa *tweet* mengenai pengembangan vaksin COVID-19 mulai tanggal 13 Januari 2021 hingga 23 Januari 2021.

3.2 Jenis dan Sumber Data

Jenis data yang digunakan pada penelitian ini adalah data primer. Data primer (Sekaran, 2011) merupakan data yang mengacu pada suatu informasi yang diperoleh dari tangan pertama oleh peneliti yang dapat berkaitan dengan variabel minat untuk tujuan spesifik studi. Data yang diperoleh berupa opini berupa *tweet* mengenai pengembangan vaksin COVID-19 mulai tanggal 13 Januari 2021 hingga 23 Januari 2021 sebanyak 1.294 *tweet*.

3.3 Variabel Penelitian

Variabel penelitian ini digunakan sebagai batasan yang diperhatikan dalam memilah opini atau *tweet* di *Twitter*. Adapun variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

1. *Date* merupakan tanggal opini berbentuk *tweet* mengenai opini pengembangan vaksinasi COVID-19
2. Opini merupakan pendapat, ide atau pikiran untuk menerangkan preferensi atau kecenderungan tertentu terhadap ideologi dan perspektif yang memiliki sifat tidak objektif.

3.4 Metode Pengumpulan Data

Metode yang digunakan untuk proses pengumpulan data dari penelitian ini adalah sebagai berikut:

1. *Naïve Bayes Classifier*

Penelitian ini dilakukan dengan membagi data menjadi dua kelompok data yaitu data training dan data testing. Data dalam penelitian ini akan melalui beberapa proses sehingga menghasilkan data yang siap digunakan untuk melakukan klasifikasi. Pengumpulan data dilakukan dengan cara mengunduh data *tweet* dari *Twitter* dan menghasilkan data mentah. Data mentah yang telah didapatkan akan melalui proses *preprocessing* data. Dalam *preprocessing* data terdapat lima sub proses yang dilalui yaitu *lowercase conversion*, *URL eliminations*, *@username removing*, *hashtag removing* dan *stop words removing*. Setelah melalui proses *preprocessing* data maka akan menghasilkan dataset yang siap digunakan untuk melakukan klasifikasi dengan menggunakan *Naïve Bayes Classifier*.

Data yang telah dilabeli akan dijadikan acuan untuk menentukan klasifikasi sentimen untuk data *tweet* yang belum dilabeli dengan menggunakan *Naïve Bayes Classifier*. Setelah semua data memiliki label, selanjutnya akan diukur tingkat akurasi dari metode *Naïve Bayes Classifier* pada aplikasi *RapidMiner*. Dengan adanya email baru, algoritma tersebut kemudian akan menghasilkan prediksi apakah email baru itu spam atau tidak. Proses klasifikasi *Naïve Bayes Classifier* terhadap dokumen yaitu dengan mempresentasikan setiap dokumen dengan atribut “ $X_1, X_2, X_3, \dots, X_n$ ” yang mempunyai makna bahwa X_1 untuk kata pertama, X_2 adalah kata kedua, dan seterusnya. Untuk himpunan kategori komentar dipresentasikan dengan V . Tujuan dari proses klasifikasi adalah untuk menentukan sebuah kalimat apakah termasuk sebagai anggota kelas opini positif, negatif atau sebagai anggota kelas opini netral yang ditentukan berdasarkan nilai perhitungan probabilitas Bayes yang lebih besar. Jika hasil probabilitas Bayes kalimat tersebut untuk kelas opini positif lebih besar maka kalimat tersebut masuk kategori opini positif demikian juga sebaliknya.

RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan

yang paling baik. Solusi yang diusung antara lain: Integrasi data, Analitis ETL, Data Analisis, dan Pelaporan dalam satu suite tunggal.

2. Studi Pustaka

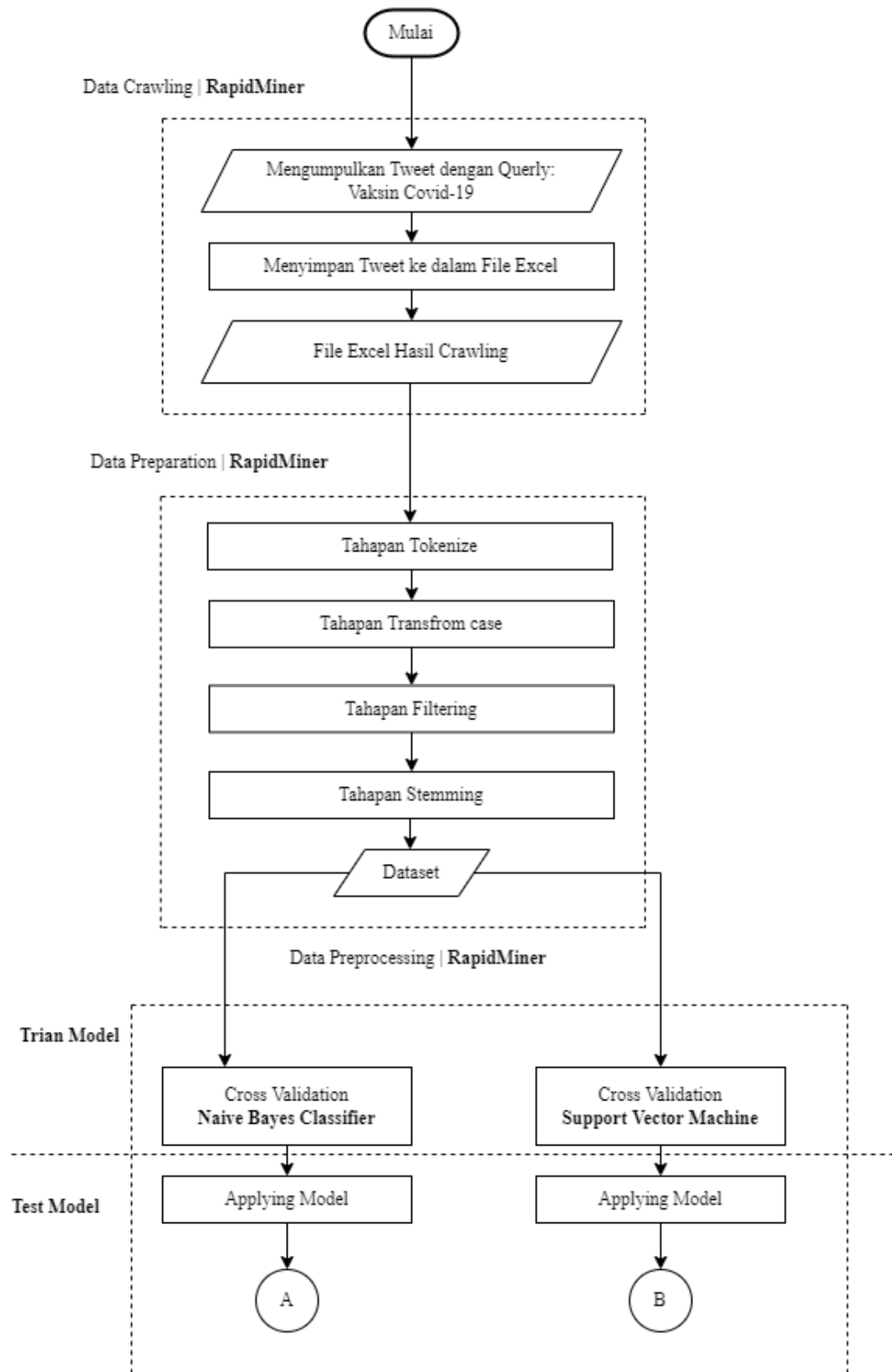
Studi pustaka dilakukan dengan menggunakan beberapa kajian literatur, buku, maupun referensi jurnal yang sekiranya berkaitan dengan tujuan, rumusan, batasan, dan metode penelitian

3.5 Analisa Data

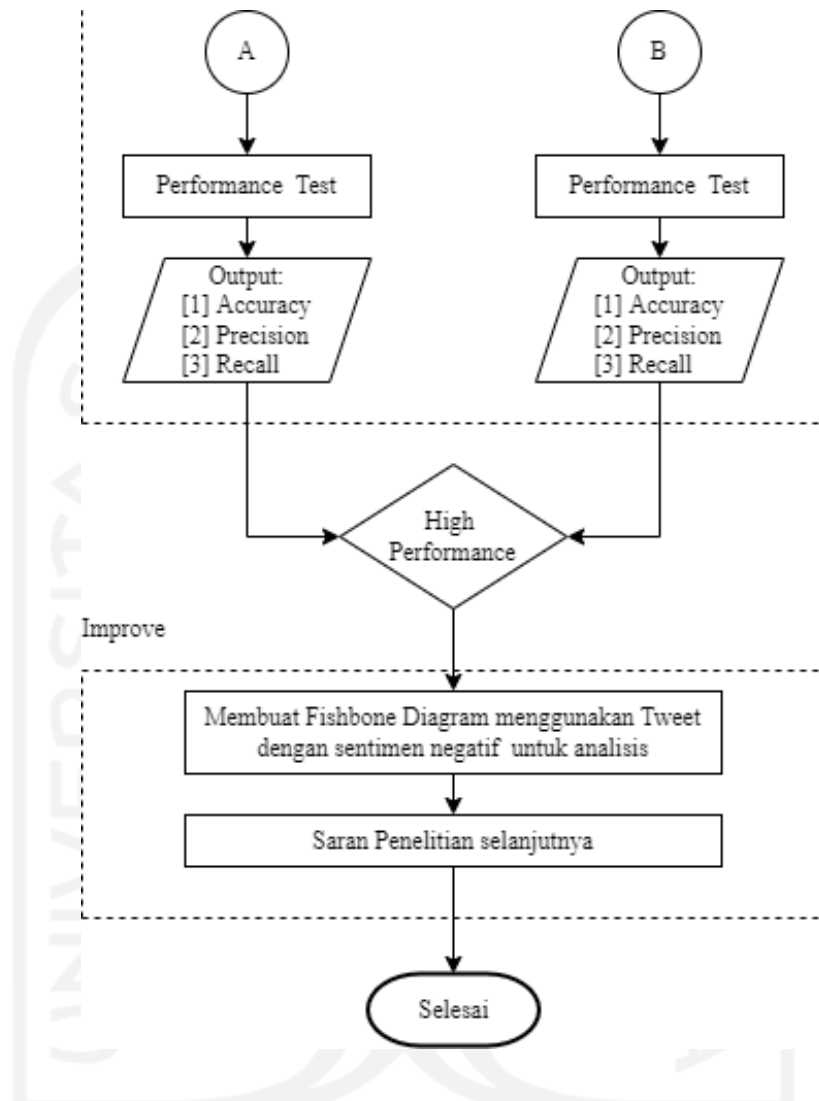
Adapun beberapa metode analisis data yang digunakan dalam penelitian ini yang dapat membantu dalam menganalisis data adalah, antara lain:

1. Metode analisis deskriptif, yang digunakan dalam menggambarkan secara umum opini pengembangan vaksin COVID 19 pada aplikasi *Twitter* .
2. Analisis sentimen berbasis kamus lexicon, digunakan untuk melakukan pelabelan data ke dalam kelas sentimen positif dan negatif.
3. Metode machine learning yaitu *Support Vector Machine* yang dalam penelitian ini digunakan untuk mengklasifikasikan opini pengembangan vaksin COVID 19 pada aplikasi *Twitter* yang dapat berbentuk kelas positif maupun kelas negatif.
4. Diagram *Fishbone* (Tulang Ikan), metode analisis yang digunakan untuk mengidentifikasi sebab dan akibat yang timbul dari hasil sentimen negatif perkembangan COVID 19 pada aplikasi *Twitter* untuk mencari tahu penyebab spesifik dan memisahkan akar penyebab masalahnya.

3.6 Diagram Alur Kerangka Penelitian



Gambar 3.1 Diagram Alur Penelitian



Gambar 3.1 Diagram Alur Penelitian (Lanjutan)

Berikut ini penjelasan mengenai tahapan dalam metode penelitian di atas, yaitu:

- a. Mengumpulkan *Tweet* Sesuai dengan *Query*: Vaksin Covid-19
- b. Menyimpan *Tweet* ke dalam File excel, Menghimpun data ke dalam file excel agar lebih mudah dalam proses analisis data
- c. Mendapatkan hasil dari proses mengindeks informasi pada halaman menggunakan URL (*Uniform Resource Locator*) dengan menyertakan API (*Application Programming Interface*)

- d. Tahap *Transform Case* Tahapan transform cases sendiri dalam penelitian ini bertujuan untuk menyeragamkan bentuk huruf menjadi *lower case* karena pada komentar *Twitter* sendiri memiliki bentuk huruf yang beragam.
- e. Tahap *Tokenize* *Tokenize* merupakan sebuah proses untuk memisahkan atau pemisahan dokumen menjadi per kata atau beberapa kata dalam suatu token/unit kecil.
- f. Tahap *Filtering* merupakan tahapan proses yang bertujuan untuk menghapus atau menyaring kata yang tidak memiliki arti dan tidak berhubungan dengan ucapan atau kata sentimen.
- g. Tahap *Stemming* adalah tahap menghilangkan imbuhan yang masih melekat sehingga dapat diperoleh kata dasar, dalam penelitian ini proses stemming akan dilakukan dengan menggunakan data stem nolimit. 6) *Generate N-Grams* Proses ini merupakan proses pemotongan string atau dokumen (n-gram) berdasarkan limit angka tertentu. Dan dalam penelitian ini sendiri n-gram yang akan digunakan adalah bigram. Tujuan penggunaan N-gram sendiri karena dalam Bahasa Indonesia banyak frase yang tidak hanya terdiri dari satu kata.
- h. Klasifikasi dalam penelitian ini algoritma yang digunakan adalah *Naiïve Bayes Classifier* dan berikut merupakan tampilan penerapan algoritma tersebut di dalam sebuah sub proses pada operator cross.
- i. Membuat *Fishbone Diagram* menggunakan *Tweet* dengan sentimen negatif untuk analisis. Selain itu untuk mengidentifikasi berbagai sebab potensial dari satu efek atau masalah, dan menganalisis masalah tersebut melalui sesi brainstorming.
- j. Setelah mendapatkan hasil penelitian, peneliti dapat memberikan saran pada pihak terkait. Seperti, Pemerintah, masyarakat, tenaga kesehatan, relawan, media.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Proses Pengolahan Dataset

Pada penelitian ini dataset yang digunakan adalah data yang berasal dari media sosial *Twitter*. Objek penelitiannya sendiri adalah vaksin Covid-19. Adapun tahapan proses pengolahannya hingga menjadi *dataset* yang nantinya akan digunakan untuk pengujian penelitian ini, berikut tahapan proses pengolahan *dataset*:

4.1.1 Proses Pengumpulan Data *Crawling*

Pengumpulan data dilakukan dengan cara *Crawling* menggunakan perangkat lunak *RapidMiner* jumlah data *tweet* yang diperoleh dari proses tersebut adalah sebanyak **4.589 tweet**. Data tersebut diambil mulai dari 13 Januari 2021 sampai 23 Januari 2021 dengan merupakan data yang sudah terseleksi dengan kata kunci terkait ‘vaksin covid’.

Data yang dikumpulkan merupakan data teks berbahasa indonesia yang diperoleh dari akun media sosial *Twitter*. Pengolahan dilakukan dengan menyeleksi *Tweet* yang mengandung kata kunci terkait ‘vaksin covid’ dan *Tweet* original. Data yang telah ditemukan kemudian disimpan dalam *Microsoft Excel*. Berikut merupakan contoh hasil dari proses pengumpulan data *Crawling* dalam bentuk file CSV, yang nantinya akan diubah ke file *Microsoft excel*.

Id;InputID;Text;Sentiment									
1	33106E+18	Vaksin Merah Putih Siap Diproduksi Pada 2021	https://t.co/cWSiydUxbl;neu						
1	33106E+18	Vaksin Merah Putih Siap Diproduksi Pada 2021	https://t.co/cWSiydUxbl;neu						
1	33106E+18;3;	Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021	https://t.co/LBw5wDphKA;neu						
1	33092E+18;4;	"https://t.co/OcUurMt1il							
		#covid19 #vaksinmerahputih							
		#bumninccom	https://t.co/yejiwcyOpH";neu						
1	33092E+18	before that please stay safe							
		Though M Biofarma	etc alreac and we already order an amount of dosage too";neg						
1	33089E+18;6;	@jasa_kuli: Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.	https://t.co/K5Qv3LD5UJ;neu						
1	33089E+18;7;	Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.	https://t.co/K5Qv3LD5UJ;neu						
1	33088E+18	tetapi juga soal kedaulatan nasional. ?????							
		#IsolasiRizieqdanFPI #YangYangBestBoy #BTSONGMA	https://t.co/3SwMcOuJ4k";pos						
1	33088E+18	apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah putih	https://t.co/OdYMuA1MwJ;neu						
1	33088E+18;10;	@itarahmadanialf Kok ngga digawe sukarelawan uji vaksin Merah Putih ae kerumunane.	wqwqsjal;neu						
1	33084E+18	klo vaksin merah-putih kira2 akurasinya berapa %? .;neg							
1	33084E+18;12;"	@RakyatMerdeka99: Siap Diproduksi Akhir 2021							
		Kemenristek: Vaksin Merah Putih Bisa Perkuat Kedaulatan Negara							

Gambar 4. 1 Hasil Proses Pengumpulan Data *Crawling* file CSV

A	B	C	D
Id	InputID	Text	Sentiment
1	1,33106E+18	1 @Jawapos: Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021 https://t.co/cWSiydUxbl	neu
2	1,33106E+18	2 Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021 https://t.co/cWSiydUxbl	neu
3	1,33106E+18	3 Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021 https://t.co/LBw5wDphKA	neu
4	1,33092E+18	4 https://t.co/OcUurMt1I#Covid19 #vaksinmerahputih #bumnincocom https://t.co/yejiwcyOpH	neu
5	1,33092E+18	5 Let's pray for our Vaksin Merah Putih with local strain virus (Eijkman and Biofarma), before that please stay safe Though MUI, Biofarma, etc already went to Sino	neg
6	1,33089E+18	6 @Jasa_kuli: Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik. https://t.co/K5Qv3LD5UI	neu
7	1,33089E+18	7 Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik. https://t.co/K5Qv3LD5UI	neu
8	1,33088E+18	8 Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulatan nasional. ?????#isolasiRizieqdp	pos
9	1,33088E+18	9 AstraZeneca menang jauh dari Moderna, apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah putih https://t.co/OdYMuAlMv	neu
10	1,33088E+18	10 @itarahmadanialf Kok ngga digawe sukarelawan uji vaksin Merah Putih ae kerumunane. wqwqsi	neu
11	1,33084E+18	11 Sudah 3 vaccine trial yg publikasi akurasi & efektivitas. Pak @Jokowi, klo vaksin merah-putih kira2 akurasi berapa %?. ?	neg
12	1,33084E+18	12 @RakyatMerdeka99: Siap Diproduksi Akhir 2021Kemenristek: Vaksin Merah Putih Bisa Perkuat Kedaulatan Negara@KemenristekBRIN #Vaksin #C...	neu
13	1,33084E+18	13 Siap Diproduksi Akhir 2021Kemenristek: Vaksin Merah Putih Bisa Perkuat Kedaulatan Negara@KemenristekBRIN #Vaksin #Corona #coronavirus #COVID19 #COV	neu
14	1,33083E+18	14 Erick Thohir Kawal Vaksin Merah Putih. https://t.co/RnvQqWnD9E	neu
15	1,33082E+18	15 Kabar baiknya bibit vaksin merah putih siap diujikan pada hewan dan kita bisa mengetahui efeknya #DukungErickPimpinPDIhttps://t.co/q8Vfu0iuj	neu
16	1,33082E+18	16 Harapan yang terbaik untuk pengembangan vaksin merah putih semoga bisa diproduksi tahun 2022 nanti dan bisa memutus rantai penyebaran covid #DukungE	neu
17	1,33082E+18	17 Tahun 2022 Indonesia tidak akan lagi mengimpor vaksin dari luar negeri karena akan menggunakan vaksin Covid buatan dalam negeri sendiri #DukungErickPim	neu
18	1,33082E+18	18 Dalam upayanya untuk mengurangi ketergantungan pada vaksin Covid impor pemerintah mengembangkan vaksin covid buatan sendiri yang diberi nama vaksin	neu
19	1,33082E+18	19 Pengembangan Vaksin Covid dalam negeri terus dilakukan agar nanti kita tak perlu lagi mengimpor vaksin dari luar negeri #DukungErickPimpinPDI https://t.co/	neu

Gambar 4. 2 Hasil data dalam bentuk file *Microsoft excel*

4.1.2 Proses Seleksi Data

Tahapan ini bertujuan untuk menyeleksi data *tweet* yaitu untuk mengambil data *tweet* yang akan digunakan nantinya. Data *tweet* yang didapat dari proses *Crawling* masih memiliki data-data yang tidak diperlukan. Pada penelitian ini hanya membutuhkan komentar *tweet* terkait Vaksin Covid-19 dengan unsur sentimen optimis dan pesimis atau sentimen positif dan negatif. Total data yang digunakan berdasarkan sentimen positif dan negatif adalah sejumlah 768 *tweet*. Contoh hasil proses seleksi data dapat dilihat pada Gambar 4.3.

InputID	Text	Sentiment
5	Let's pray for our Vaksin Merah Putih with local strain virus (Eijkman and Biofarma), before that please stay safe	neg
8	Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulatan nasional?	pos
11	Sudah 3 vaccine trial yg publikasi akurasi & efektivitas. Pak @Jokowi, kalau vaksin merah-putih kira-kira akurasi berapa persen?	neg
30	Pemerintah beruoaaya optimal untuk memutus rantai virus covid, sekarang itu membuah hasil kabarnya bibit vaksin merah putih siap diujikan pada hewan	neg
40	Sementara sebelum adanya vaksin merah putih Indonesia mamanfaatkan vaksin yang seadanya untuk mencegah covid?	pos
67	Harapan yang terbaik untuk pengembangan vaksin merah putih semoga bisa diproduksi tahun 2022 nanti dan bisa memutus rantai penyebaran covid?	pos
68	Vaksin merah putih menggunakan isolat virus yang bertransmisi di Indonesia	neg
72	Erick thohir mengatakan bahwa langkah penting untuk pengadaan vaksin merah putih. Semangat terus!	pos
73	Sebelum ada vaksin merah putih mari kita selalu menjaga protokol kesehatan yang ada?	pos
78	Udah tahu belum teman-teman kalau vaksin yang dikembangkan di Indoensia dinamai vaksin merah putih	pos
109	Pemerintah kawal akselerasi vaksin Merah Putih karya anak bangsa. LUAAAR BIASAA...??	pos
131	Vaksin Merah Putih merupakan strategi jangka panjang penanganan pandemi COVID-19 yang memanfaatkan isolat virus yang bersirkulasi di Indonesia dalam proses per	neg
141	Peneliti Rampungan Vaksin Merah Putih, #AzisSyamsuddin Ajak Kepala Daerah Redam Sebaran Wabah	neg
146	President @jokowi : Pemerintah Kawal Akselerasi Vaksin PRODUK ANAK BANGSA Merah Putih ??BRAVO Indonesia,,,!!! ????	pos
149	Sabar yah! Vaksin Merah Putih Baru Tersedia 2022	pos
152	Sabar yah! Vaksin Merah Putih Baru Tersedia 2022	pos
177	Jokowi Gaspol Lawan Pandemi.Pemerintah akan kawal akselerasi proses pembuatan Vaksin Merah Putih.Semakin cepat semakin baik!!	pos
179	Pemerintah kawal akselerasi vaksin merah putih!	pos
184	Merah Putih produk anak bangsa masih dalam tahap proses belum sampe produksi. di support penuh oleh p Jokowi.. si @aniesbaswedan dan kawan-kawan kan anti rez	pos

Gambar 4. 3 Hasil Proses Seleksi Data.

4.1.3 Proses *Cleansing*

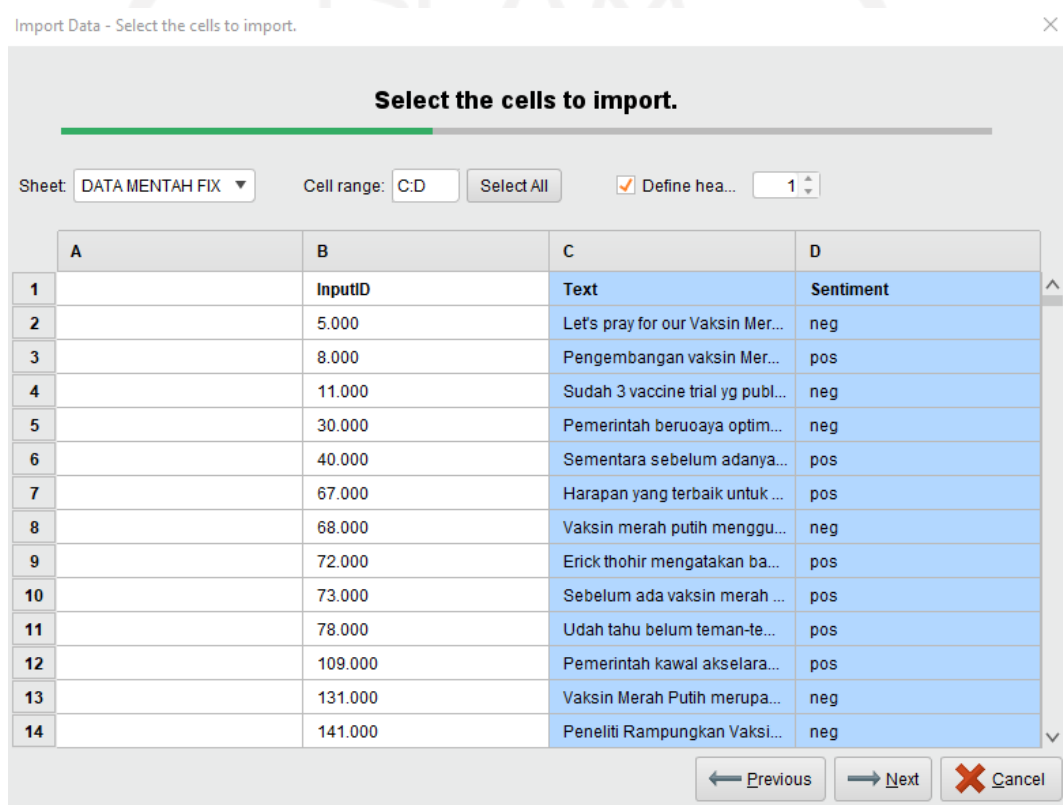
Tahapan proses *cleansing* pada penelitian ini adalah untuk menghapus kata, karakter dan simbol yang tidak diperlukan dalam penelitian, berikut ini proses *cleansing*:

Tahapan *cleansing* pada penelitian ini untuk menghapus kata, karakter dan simbol yang tidak diperlukan dalam penelitian, yaitu antara lain seperti: Karakter HTML (<, >, dll), Ikon

emosi (:0, :(,dll), *Hashtag* (#), *Username* (@username), *Url* (<http://website.com>), *Email* (nama@website.com), Tanda baca atau *punctuation* (—., —!, —?!, —!;, dll), dan *Retweet* (RT).

a. Proses *Select Attributes* Data

Berikut merupakan tampilan proses *select attributes* pada *software RapidMiner* yang ditunjukkan pada Gambar 4.4. Tahap *select attributes* bertujuan untuk memfokuskan data *tweet* yang berbentuk *text* saja agar proses *sentiment analysis* dapat dilakukan dengan berfokus pada salah satu atribut, hasil dari proses *select attributes* data ditunjukkan pada Gambar 4.5.



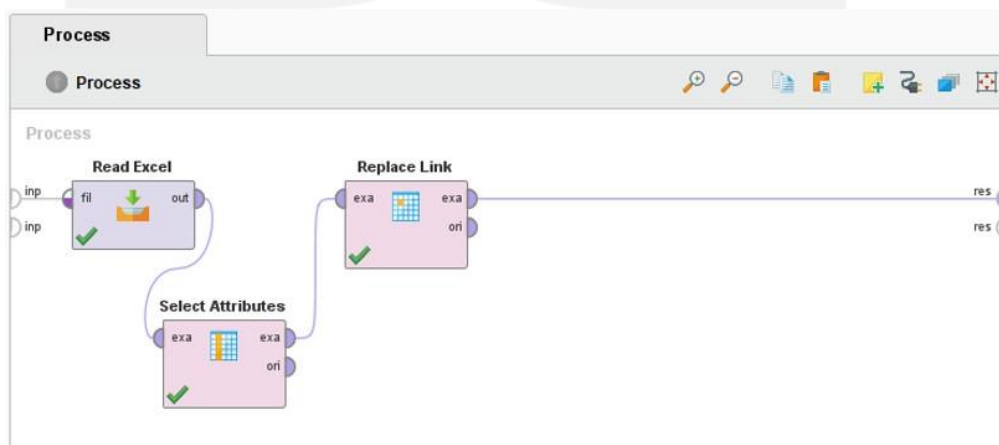
Gambar 4. 4 Proses *Select Attributes* Data.

Text
@jawapos: Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021 https://t.co/cWSiydUxbl
Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021 https://t.co/cWSiydUxbl
Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021 https://t.co/LBw5wDphKA
https://t.co/OcUurMt1il
Let's pray for our Vaksin Merah Putih with local strain virus (Eijkman and Biofarma), before that please stay safe
@jasa_kuli: Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik. https://t.co/K5Qv3LD5UI
Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik. https://t.co/K5Qv3LD5UI
Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulatan nasional. ?????
Astrazeneca menang jauh dari Moderna, apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah putih https://t.co/...
@itarahmadanialf Kok ngga digawe sukarelawan uji vaksin Merah Putih ae kerumunane. wqwqsial
Sudah 3 vaccine trial yg publikasi akurasi & efektivitas. Pak @Jokowi, klo vaksin merah-putih kira2 akurasinya berapa %?. ?
@RakyatMerdeka99: Siap Diproduksi Akhir 2021
Siap Diproduksi Akhir 2021
Erick Thohir Kawal Vaksin Merah Putih.

Gambar 4. 5 Hasil Proses *Select Attributes* Data

b. *Replace*

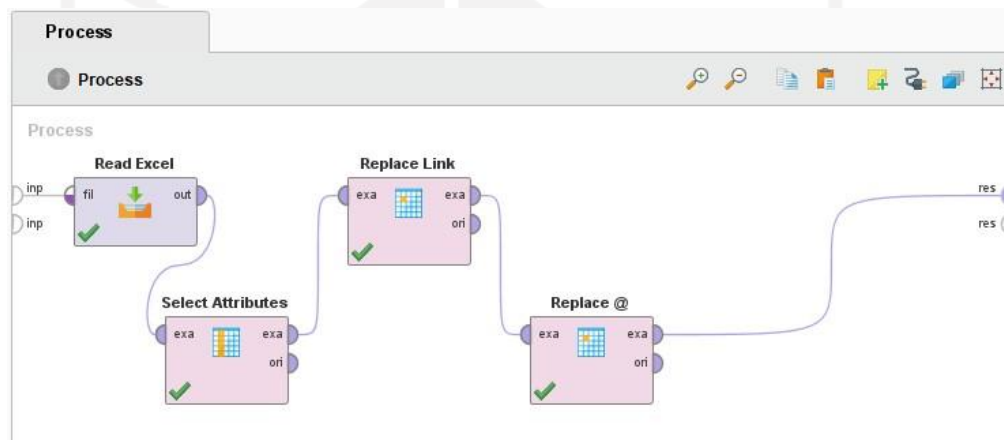
Dalam *Software RapidMiner* proses untuk menghapus karakter dan simbol yang tidak diperlukan dalam penelitian adalah dengan menggunakan operasi *replace*. Pada penelitian ini proses *replace* dilakukan beberapa kali. Gambar 4.6 menunjukkan tampilan desain proses *replace link* pada *software RapidMiner* beserta Gambar 4.7 merupakan hasil *replace link*. Gambar 4.8 menunjukkan tampilan desain proses *replace @* pada *software RapidMiner* beserta Gambar 4.9 merupakan hasil *replace @*. Gambar 4.10 menunjukkan tampilan desain proses *replace #* pada *software RapidMiner* beserta Gambar 4.11 merupakan hasil *replace #*.



Gambar 4. 6 Tampilan Desain dari Proses *Replace Link*.

Text
@jawapos: Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021
Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021
Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021
#covid19 #vaksinmerahputih #bumninccom
Let's pray for our Vaksin Merah Putih with local strain virus (Eijkman and Biofarma), before that please stay safe
@jasa_kuli: Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.
Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.
Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulata...
Astrazeneca menang jauh dari Moderna, apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah ...
@itarahmadanialf Kok ngga digawe sukarelawan uji vaksin Merah Putih ae kerumunane. wqwqsial
Sudah 3 vaccine trial yg publikasi akurasi & efektivitas. Pak @Jokowi, klo vaksin merah-putih kira2 akurasi berapa %?. ?
@RakyatMerdeka99: Siap Diproduksi Akhir 2021

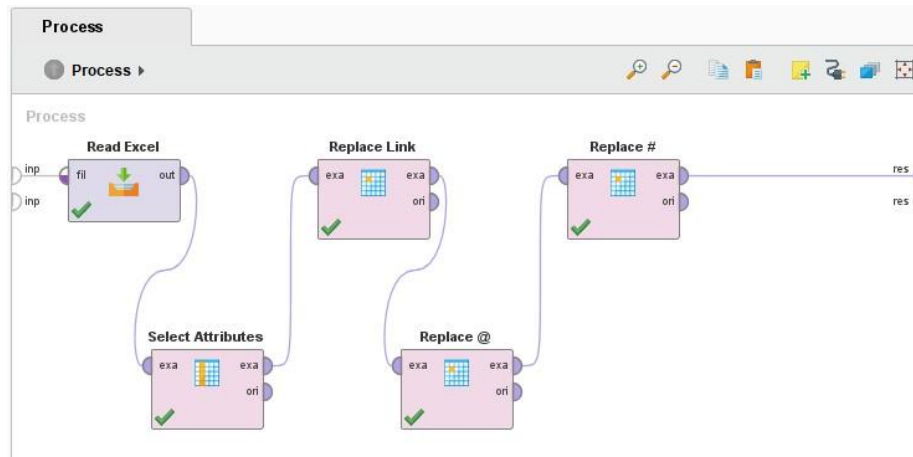
Gambar 4. 7 Hasil Proses *Replace Link*.



Gambar 4. 8 Tampilan Desain dari Proses *Replace @*

Text
Jawapos: Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021
Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021
Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021
#covid19 #vaksinmerahputih #bumninccom
Let's pray for our Vaksin Merah Putih with local strain virus (Eijkman and Biofarma), before that please stay safe
jasa_kuli: Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.
Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.
Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulatan nasion...
Astrazeneca menang jauh dari Moderna, apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah putih
itarahmadanialf Kok ngga digawe sukarelawan uji vaksin Merah Putih ae kerumunane. wqwqsial
Sudah 3 vaccine trial yg publikasi akurasi & efektivitas. Pak Jokowi, klo vaksin merah-putih kira2 akurasi berapa %?. ?
RakyatMerdeka99: Siap Diproduksi Akhir 2021
Siap Diproduksi Akhir 2021

Gambar 4. 9 Hasil *Replace @*



Gambar 4. 10 Tampilan Desain dari Proses *Replace #*

Text
jawapos: Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021
Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021
Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021
covid19 vaksinmerahputih bumninccom
Let's pray for our Vaksin Merah Putih with local strain virus (Eijkman and Biofarma), before that please stay safe
jasa_kuli: Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.
Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.
Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulatan nasional. ?????
Astrazeneca menang jauh dari Moderna, apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah putih
itarahmadanialf Kok ngga digawe sukarelawan uji vaksin Merah Putih ae kerumunane. wqwqisial
Sudah 3 vaccine trial yg publikasi akurasi & efektivitas. Pak Jokowi, klo vaksin merah-putih kira2 akurasi berapa %? ?
RakyatMerdeka99: Siap Diproduksi Akhir 2021
Siap Diproduksi Akhir 2021

Gambar 4. 11 Hasil Proses *Replace #*

Hasil dari proses *cleansing* kemudian disimpan dalam bentuk *file Microsoft Excel*, yang nantinya akan di proses lagi menggunakan bantuan *software Microsoft Excel* untuk melakukan *cleansing* data-data yang tidak perlu seperti nama user yang menulis *tweet*.

c. Proses *cleansing* dengan *software Microsoft Excel*

Hasil yang diperoleh dari tahapan *cleansing* dengan *RapidMiner* kemudian akan diolah lagi menggunakan bantuan *Microsoft Excel* untuk menghilangkan *username* dari data ini. Jumlah data yang diolah menggunakan *software RapidMiner* pada proses *cleansing* sejumlah 768 data. Data hasil *cleansing* ini akan digunakan kedalam proses *pre-processing*. Berikut Gambar 4.12 yang menunjukkan hasil *cleansing* dengan menggunakan *Microsoft Excel*.

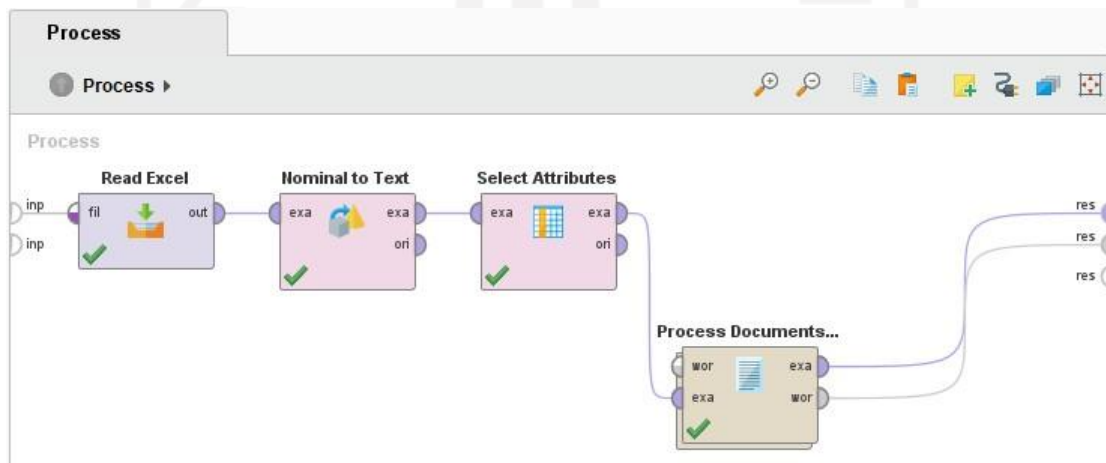
1	Text
2	Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021
3	Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021
4	covid19 vaksin merah putih bumh inccom
5	Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik.
6	Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulatan nasional.
7	Astrazeneca menang jauh dari Moderna, apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah putih
8	Ssudah 3 vaksin trial yg publikasi akurasi & efektivitas. Pak Jokowi, klo vaksin merah-putih kira2 akurasinya berapa %?
9	Vaksin Merah Putih Bisa Perkuat Kedaulatan Negara KemenristekBRIN Vaksin Covid-19
10	Vaksin Merah Putih Bisa Perkuat Kedaulatan Negara Kemenristek BRIN Vaksin Corona coronavirus COVID-19

Gambar 4. 12 Hasil *Cleansing* dengan *Microsoft Excel*

4.2 Proses *pre-processing*

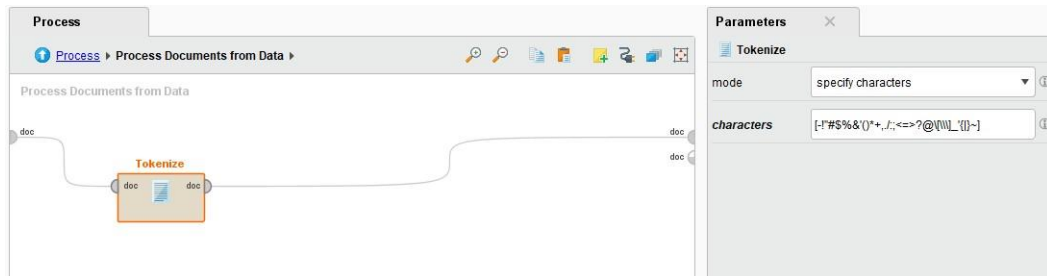
Data *tweet* yang telah mengalami pemrosesan pada tahapan *cleansing* kemudian akan dilakukan proses *pre-processing*. Proses *pre-processing* dalam *RapidMiner* menggunakan operasi *Process Document from Data*, dan memiliki sub operasi lain diantaranya, *Tokenize*, *Transform Cases*, *Filter Tokens (by length)*, *Filter Stopwords (dictionary)*. Sebelum masuk ke operasi *process document from data*, terlebih dahulu melalui proses *Nominal to Text* dan *Select Attribute* yang bertujuan untuk memfokuskan operasi *process document from data* pada atribut *text*. Berikut desain proses dan hasil dari masing-masing tahapan, Gambar 4.13 menunjukkan tampilan desain dari *pre-processing*, Gambar 4.14 menunjukkan tampilan desain dari operasi *Tokenize*.

a. Proses *Nominal to Text* dan *Select Attribute*



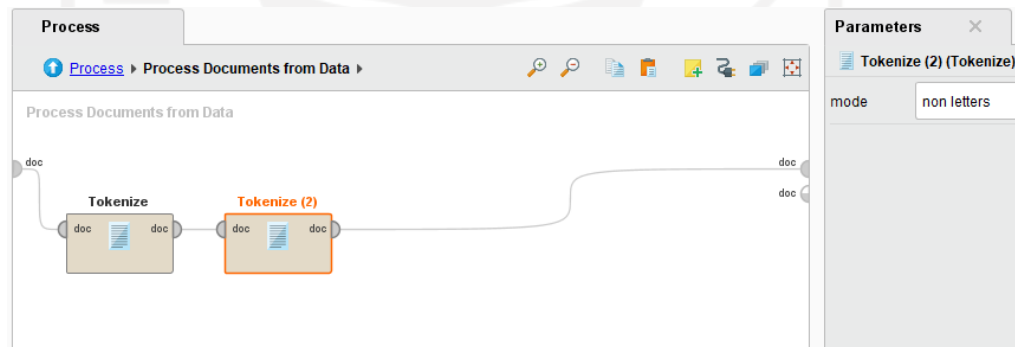
Gambar 4. 13 Tampilan Desain dari *Pre-Pocessing*

b. Proses *Tokenize*



Gambar 4. 14 Tampilan Desain dari Operasi *Tokenize*

Operasi *Tokenize* pada penelitian ini akan dilakukan dua kali, untuk operasi *Tokenize* pertama ini bertujuan untuk menghilangkan karakter-karakter yang tidak perlu dengan menggunakan mode *Specify Characters*. Untuk *Tokenize* kedua bertujuan untuk menghilangkan angka dengan menggunakan mode *Non Letters* seperti yang ditunjukkan pada Gambar 4.15. Hasil dari proses *Tokenize* mode *Non Letters* dapat dilihat pada Gambar 4.16.



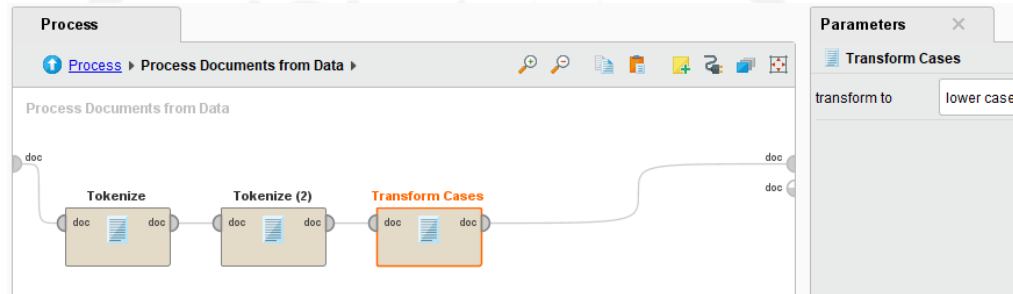
Gambar 4. 15 Tampilan Desain dari Operasi *Tokenize (2)*

Word	Attribute Name	Total Occure... ↓	Document Occurences	pos	neg
vaksin	vaksin	671	523	263	408
Covid	Covid	260	236	85	175
Vaksin	Vaksin	247	231	101	146
covid	covid	239	212	115	124
di	di	234	196	103	131
dan	dan	202	164	79	123
yang	yang	187	150	59	128
virus	virus	170	165	5	165

Gambar 4. 16 Tampilan Hasil Proses *Tokenize*

c. Proses *Transform Cases*

Hasil yang diperoleh dari proses *Tokenize* adalah sejumlah 4399 kata. Selanjutnya dilakukan proses *Transform Cases* dengan tujuan untuk menyeragamkan bentuk *text* menjadi ukuran huruf kecil semua, dengan menggunakan operasi *Transform Cases* dengan *transform to: lower case*, seperti yang ditunjukkan pada Gambar 4.17. Kemudian hasil dari proses *transform to: lower case* dapat dilihat pada Gambar 5.18.



Gambar 4. 17 Tampilan Desain Proses *Transform Cases*

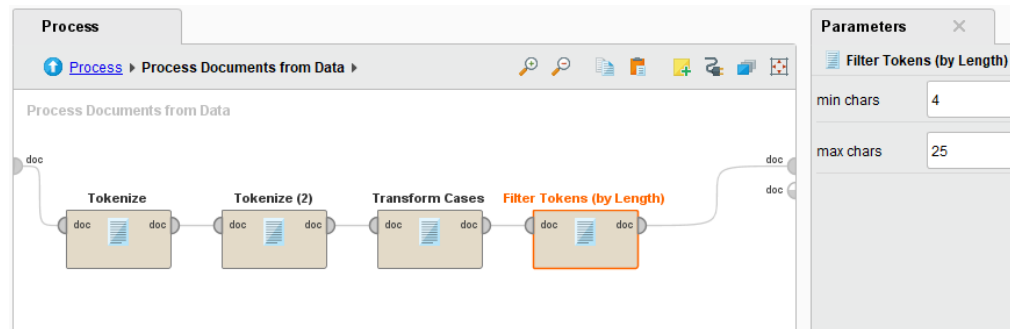
Word	Attribute Name	Total Occure... ↓	Document Occurences	pos	neg
vaksin	vaksin	925	716	368	557
covid	covid	596	518	235	361
di	di	245	202	108	137
virus	virus	225	216	5	220
dan	dan	219	179	85	134
yang	yang	202	157	66	136
corona	corona	172	163	27	145

Gambar 4. 18 Tampilan Hasil Proses *Transform Cases*

d. Proses *Filter*

Berdasarkan hasil yang diperoleh dari proses *transform cases* terdapat data sejumlah 3718 kata. Pada proses ini akan dilakukan proses *filter* dengan dua operasi diantaranya *filter tokens (by length)* dengan *minimal chars : 4* dan *maximal chars : 25*, yang bertujuan untuk membatasi minimal dan maksimal karakter/huruf dalam satu kata, dan yang kedua *filter stopwords dictionary* yang bertujuan untuk menghapus kata-kata yang tidak memiliki arti apabila kata tersebut berdiri sendiri, pada tahap ini peneliti terlebih dahulu mempersiapkan file kata-kata yang tidak memiliki arti jika berdiri sendiri yang nantinya akan di inputkan ke dalam proses *filter*

stopwords (dictionary) ini. Berikut desain dan hasil dari proses *filter* ditunjukkan pada Gambar 4.19 sampai Gambar 4.22.

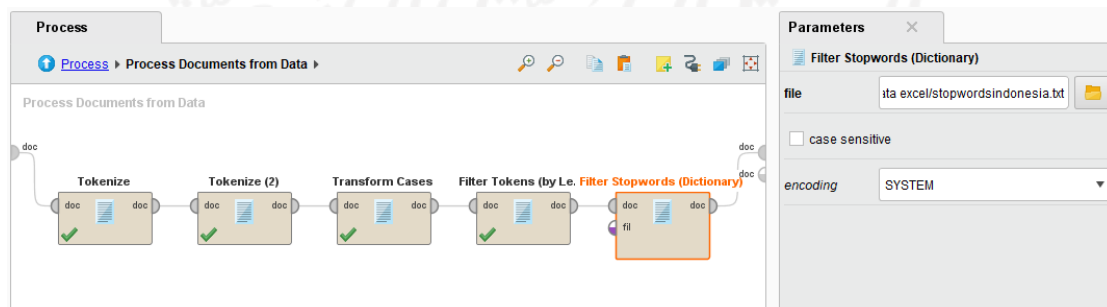


Gambar 4. 19 Tampilan Desain Proses Filter *Tokens (By Length)*

Word	Attribute Name	Total Occure... ↓	Document Occurences	pos	neg
vaksin	vaksin	925	716	368	557
covid	covid	596	518	235	361
virus	virus	225	216	5	220
yang	yang	202	157	66	136
corona	corona	172	163	27	145
untuk	untuk	152	142	45	107
merah	merah	121	115	70	51
putih	putih	118	115	71	47
indonesia	indonesia	93	83	40	53

Gambar 4. 20 Tampilan Hasil Proses Filter *Tokens (By Length)*

Pada tahap *filter tokens (by length)* ini data yang dihasilkan berjumlah 3250 kata, yang kemudian akan dilanjutkan ke tahap *filter stopwords (dictionary)*.



Gambar 4. 21 Tampilan Desain Proses Filter *Stopwords (Dictionary)*

Word	Attribute Name	Total Occurences ↓	Document Occurences	pos	neg
vaksin	vaksin	925	716	368	557
covid	covid	596	518	235	361
virus	virus	225	216	5	220
corona	corona	172	163	27	145
merah	merah	121	115	70	51
putih	putih	118	115	71	47
indonesia	indonesia	93	83	40	53
jokowi	jokowi	76	71	32	44
vaksinasi	vaksinasi	75	60	20	55

Gambar 4. 22 Tampilan Hasil Proses Filter *Stopwords (Dictionary)*

e. *Stemming*

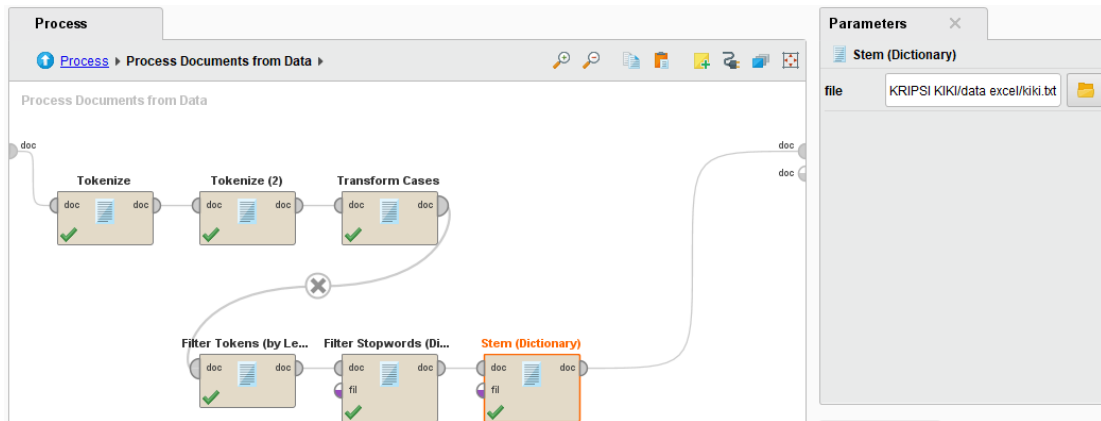
Pada tahap *stemming* merupakan suatu proses untuk mengubah kata-kata yang terdapat dalam suatu dokumen ke dalam kata-kata akarnya dengan menggunakan aturan-aturan tertentu. Proses *stemming* bahasa Indonesia dilakukan dengan menghilangkan *sufiks*, *prefiks*, dan *konfiks* pada dokumen. Pada penelitian ini proses *stemming* menggunakan bantuan *software Microsoft Excel*, selanjutnya hasil *stemming* ini di simpan dalam bentuk *file txt* yang nantinya akan di inputkan ke dalam proses *stem (dictionary)*. Berikut Gambar 4.23. menunjukan hasil dari proses *stemming*, sedangkan Gambar 4.24 menunjukan tampilan desain proses *stem (dictionary)*

```

aamiin : amin
aamin : amin
abis : abis
absurdnya : absurd
aceh : aceh
acting : act
adakah : ada
adakan : ada

```

Gambar 4. 23 Hasil Stemming dengan *Microsoft Excel*



Gambar 4. 24 Tampilan Desain Proses *Stem (Dictionary)*

Word	Attribute Name	Total Occurenc... ↓	Document Occurences	neg	pos
vaksin	vaksin	877	652	499	378
covid	covid	541	478	310	231
virus	virus	177	167	172	5
corona	corona	151	140	126	25
merah	merah	104	98	44	60
putih	putih	102	99	41	61
indonesia	indonesia	82	72	46	36
suntik	suntik	69	57	39	30
jokowi	jokowi	61	59	30	31

Gambar 4. 25 Tampilan Hasil Proses *Stem (Dictionary)*

Berdasarkan hasil *pre-processing* yang dilakukan kata yang mengandung sentimen adalah sejumlah 2452 kata. Kemudian, dari 768 *tweet* terdapat 403 sentimen negatif dan 365 sentimen positif. Dari hasil *pre processing* terdapat kata yang paling banyak muncul adalah kata Vaksin dengan total kemunculan sebanyak 877 kata dalam 652 *tweet*.

f. *Smote Upsampling*

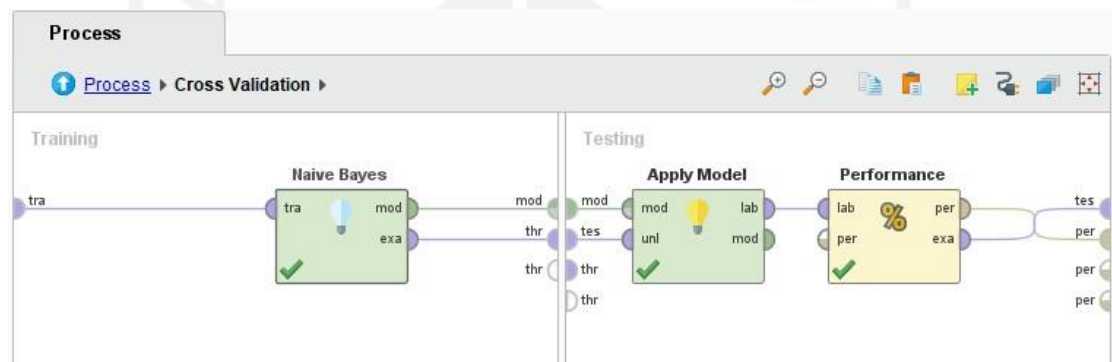
Sebelum melakukan uji *cross validation* sebelumnya operasi yang harus dilakukan adalah operasi *smote upsampling* yang berfungsi untuk menyeimbangkan data label. Dimana dalam penelitian ini sudah terdapat label yaitu positif dan negatif yang tidak seimbang, yaitu data positif sebanyak 365 dan data negatif sebanyak 403.

Sehingga dalam proses ini perlu dilakukan operasi *smote upsampling* sehingga hasil yang diperoleh lebih akurat.

4.3 Pengujian Algoritma (Cross Validation)

4.3.1 Naive Bayes

Pengujian dilakukan menggunakan *Tools RapidMiner 9.3.001*. Pada penelitian ini dilakukan pengujian dengan menggunakan algoritma *Naïve Bayes*. Akurasi dari algoritma klasifikasi *Naïve Bayes*, maka dataset secara otomatis dibagi menjadi training dan testing data. Proses pengujian pada Gambar 2.27 dilakukan dengan teknik *cross validation* dengan tujuan untuk memperoleh hasil nilai akurasi terbaik.



Gambar 4. 26 Tampilan Desain Proses *Cross Validation* dengan *Algoritma Naïve Bayes*

Hasil pengujian *cross validation* dapat dilihat pada Gambar 4.27 *Confusion Matrix Naive Bayes*. Berdasarkan hasil pengujian di bawah, dapat diperoleh bahwa hasil nilai akurasi 74,74.%

Table View Plot View

accuracy: 74.74% +/- 4.71% (micro average: 74.74%)

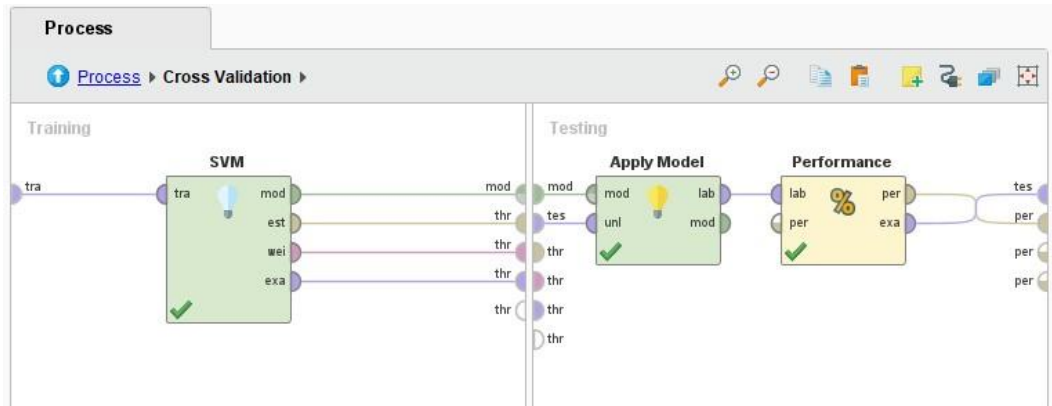
	true neg	true pos	class precision
pred. neg	259	69	78.96%
pred. pos	125	315	71.59%
class recall	67.45%	82.03%	

Gambar 4. 27 *Confusion Matrix Naive Bayes*

4.3.2 Support Vector Machine

Pengujian dilakukan menggunakan *Tools RapidMiner 9.3.001*. Pada penelitian ini sebagai pembandingan algoritma *Naïve Bayes* maka dilakukan pengujian dengan menggunakan

algoritma lain yakni algoritma *Support Vector Machine*. Akurasi dari algoritma klasifikasi *Support Vector Machine*, maka dataset secara otomatis dibagi menjadi training dan testing data. Proses pengujian SVM pada Gambar 4.28. juga dilakukan dengan teknik *cross validation* dengan tujuan untuk memperoleh hasil nilai akurasi terbaik.



Gambar 4. 28 Tampilan Desain Proses *Cross Validation* dengan Algoritma *Support Vector Machine*

Hasil pengujian *cross validation* dapat dilihat pada Gambar 4.29 *Confusion Support Vector Machine* Berdasarkan hasil pengujian di bawah, dapat diperoleh bahwa hasil nilai akurasi 74,08%

Table View Plot View

accuracy: 74.08% +/- 5.84% (micro average: 74.09%)

	true neg	true pos	class precision
pred. neg	218	33	86.85%
pred. pos	166	351	67.89%
class recall	56.77%	91.41%	

Gambar 4. 29 *Confusion Matrix Support Vector Machine*

4.4 Hasil Analisa Akurasi

Performa atau akurasi merupakan tingkat kedekatan atau seberapa dekat nilai prediksi dari nilai sebenarnya. Berdasarkan dari pengujian menggunakan *Naïve Bayes* dan *Support Vector Machine* yang telah dilakukan, dapat dilihat performa hasil dari nilai akurasi yang diperoleh, Berikut adalah tabel perbandingan dari hasil nilai akurasi masing-masing pengujian yang telah dilakukan menggunakan teknik *cross validation*.

4.4.1 Tabel Perbandingan akurasi

Performa atau akurasi merupakan tingkat kedekatan atau seberapa dekat nilai prediksi dari nilai sebenarnya. Berdasarkan dari pengujian diatas yang telah dilakukan dapat dilihat performa atau hasil dari nilai akurasi yang diperoleh menggunakan algoritma *Naive Bayes* dan *Support Vector Machine*. Berikut ini adalah Tabel 4.1 perbandingan dari hasil nilai akurasi kedua algoritma yang telah dilakukan menggunakan teknik *cross validation* :

Tabel 4 1 Tabel Perbandingan Akurasi

Algoritma	<i>Naive Bayes</i>	SVM
Akurasi	74,74%	74,08%

Berdasarkan dari tabel pengujian diatas, diperoleh hasil bahwa akurasi *Naive Bayes* yang memiliki nilai akurasi 74,74% dibandingkan dengan *Support Vector Machine* lebih tinggi dengan nilai 74,08%.

4.4.2 Hasil Klasifikasi Sentimen

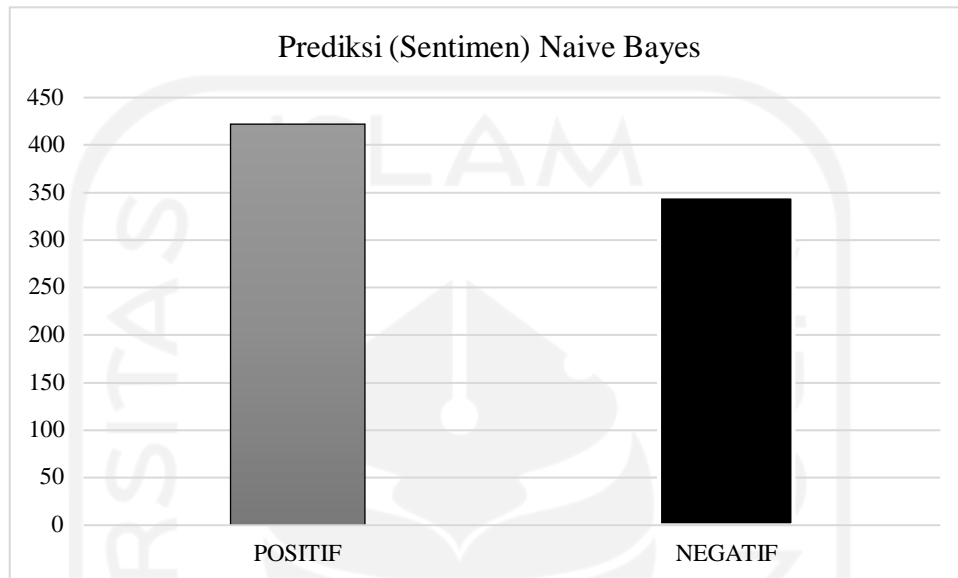
Hasil klasifikasi ini adalah sentimen yang didapatkan dari hasil klasifikasi data testing dengan menggunakan model klasifikasi yang dibuat dan dilakukan, berikut ini adalah Gambar 4.30 Hasil polaritas sentimen dan prioritas (sentimen):

Name	Type	Missing	Statistics	Values
Label Sentiment	Binominal	0	Negative neg	Positive pos neg (384), pos (384)
Prediction prediction(Sentiment)	Binominal	0	Negative neg	Positive pos pos (422), neg (346)
Confidence_pos confidence(pos)	Real	0	Min 0	Max 1 Average 0.549
Confidence_neg confidence(neg)	Real	0	Min 0	Max 1 Average 0.451

Gambar 4. 30 Hasil Polaritas Sentimen dan Prioritas (*Sentimen*)

Gambar diatas menunjukkan hasil proses klasifikasi yaitu pada baris prediction (sentimen) dengan jumlah negatif sebanyak 768 data *tweet* dan jumlah positif sebanyak 422 data *tweet*, dan jumlah negatif sebanyak 346. Sebagaimana pada proses manual jumlah negatif sebanyak 384 data *tweet* dan positif sebanyak 364 data *tweet* seperti yang terlihat pada baris sentimen, dari statistik diatas dapat diketahui bahwa akurasi

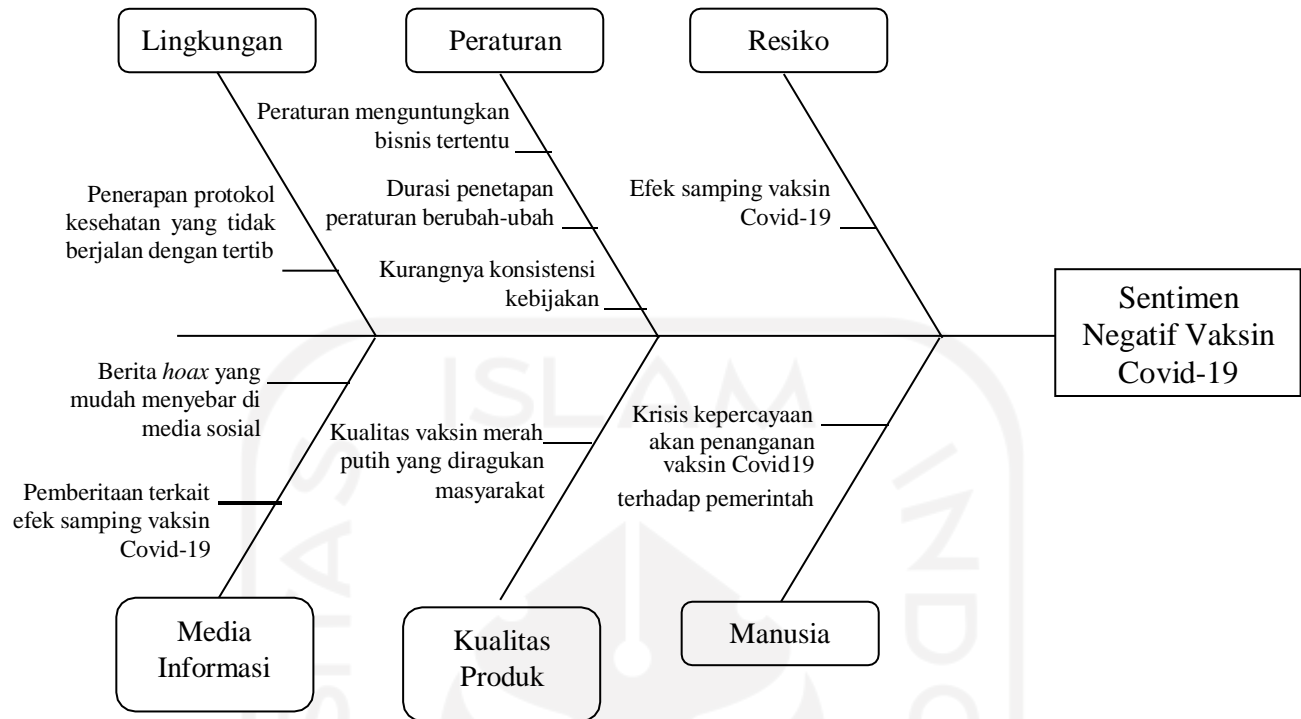
prediksi 74.73% yang merupakan hasil pengujian algoritma *Naïve Bayes*. Untuk perbandingan hasil analisa klasifikasi sentimen bisa dilihat pada Gambar 4.31 diagram berikut :



Gambar 4. 31 Diagram Hasil Analisa Klasifikasi Sentimen

4.5 Diagram *Fishbone* untuk Analisa Sentimen Negatif

Berdasarkan hasil sentimen negatif yang didapatkan dari data, maka dapat diperoleh informasi mengenai masalah yang terjadi terkait dengan komentar vaksin Covid-19 di media sosial *Twitter* dan akar penyebab masalahnya yang dipresentasikan dalam diagram *Fishbone* pada Gambar 4.32. berikut.



Gambar 4. 32 Diagram *Fishbone* Sentimen Negatif Vaksin Covid-19

Pada Gambar diatas, dapat diperoleh bahwa terdapat 6 kategori sebab utama munculnya komentar sentimen Vaksin Covid-19 pada media social *Twitter*, dan akar masalah pada setiap kategori, sebagai berikut :

a. Lingkungan Sosial

Terdapat opini negatif yang pesimis terhadap pengembangan vaksin Covid-19 dikarenakan kondisi lingkungan dan *social* yang tidak patuh terhadap penerapan protokol kesehatan yang tidak berjalan dengan tertib. Hal tersebut dikarenakan kurangnya kesadaran masyarakat akan bahaya virus corona.

b. Peraturan

Pandangan masyarakat terkait pembuatan peraturan pemerintah tentang penanganan pengembangan vaksin Covid-19 yang dinilai menguntungkan bisnis tertentu seperti regulasi frekuensi penyuntikan Vaksin dan juga tetap diwajibkannya tes PCR atau antigen walaupun sudah melakukan vaksin 2 dosis. Kemudian kurangnya konsistensi kebijakan dan juga durasi penetapan peraturan berubah-ubah

c. Resiko

Masyarakat mengeluhkan, kurangnya sosialisasi mengenai efek samping vaksin

Covid-19 sesuai dengan produk vaksin nya, seperti efek samping dari Vaksin *Sinovac, Astra Zeneca, Pfizer, Moderna*, dan terutama Vaksin Merah Putih.

d. Media Informasi

Berita *hoax* yang mudah menyebar di media *social* dapat meningkatkan pandangan negatif masyarakat terhadap vaksin Covid-19. Selain itu, pemberitaan terkait efek samping negatif vaksin Covid-19 yang bahkan sampai menyebabkan pasien meninggal dunia meningkatkan opini negatif atau pesimis.

e. Kualitas Produk

Kualitas vaksin merah putih yang diragukan masyarakat.

f. Manusia

Krisis kepercayaan akan penanganan vaksin Covid19 terhadap pemerintah.



BAB IV

PENUTUP

5.1 Kesimpulan

1. Sentimen positif yang diprediksi oleh algoritma *Support Vector Machine* berjumlah 517 dan sentimen negatif berjumlah 251. Sedangkan sentimen positif yang diprediksi oleh *Naïve Bayes* sebesar 422 dan sentimen negatif sebanyak 346. Dari uji yang telah dilakukan, dapat disimpulkan bahwa opini masyarakat yang disampaikan melalui *Twitter* mengenai pengembangan Vaksin Covid-19 di Indonesia lebih banyak masyarakat yang optimis dibandingkan dengan masyarakat yang pesimis.
2. Opini masyarakat di *Twitter* mengenai gambaran umum pengembangan vaksin Covid-19 dengan menggunakan metode *Support Vector Machine* menghasilkan nilai akurasi sebesar 74.08% lebih rendah dibandingkan dengan nilai akurasi *Naïve Bayes* sebesar 74.73%.
3. Sentimen negatif vaksin Covid-19 yang divisualisasikan melalui diagram *Fishbone*, terdapat 6 kategori sebab utama, diantaranya lingkungan social, peraturan, risiko, media informasi, kualitas produk, dan manusia.

5.2 Saran

Penelitian ini masih memiliki banyak kekurangan, oleh sebab itu penulis memberikan saran untuk penelitian selanjutnya yaitu :

1. Proses klasifikasi menggunakan algoritma yang lain, seperti algoritma linear regression, *logistic regression*, *lexical-based*, dan lain sebagainya
2. Penggunaan fitur seleksi dan penggunaan model preprocessing yang lain.
3. Penggunaan software *text mining* yang lain seperti *python*.

DAFTAR PUSTAKA

- Angreni, I. A. A., Adisasmita, S. A., dan Ramli, M. I. (2018). Pengaruh Nilai K pada Metode K-Nearest Neighbor (KNN) Terhadap Tingkat Akurasi Identifikasi Kerusakan Jalan. *Rekayasa Sipil p-ISSN 2252-769*. 7(2), 63–70.
- Biro Humas Kemenkominfo. (2020). Hasil Survei Indeks Literasi Digital Nasional 2020, Akses Internet Makin Terjangkau: https://kominfo.go.id/content/detail/30928/siaran-pers-no149hmkominfo112020-tentang-hasil-survei-indeks-literasi-digital-nasional2020-akses-internet-makin-terjangkau/0/siaran_pers (2 Maret 2020).
- Budiyanto, R. (2019). Analisis Sentimen Terhadap Institusi Komisi Pemilihan Indonesia (KPU) Pada Pemilu 2019 Di Media Sosial Twitter Menggunakan Metode Klasifikasi Support Vector Machine (Svm) dan Fitur Lexicon Based. S1 Skripsi, Universitas Gunadarma Jakarta.
- Clayton R. Fink, dkk. 2011. Coarse- and Fine-Grained Sentimen Analysis of Social Media Text. *Johns hopkins apl technical digest*, volume 30, number 1.
- Dwiraswati, O., Siregar, K. N., Pengawas, B., dan Indonesia, U. (2019). Penggunaan Antibiotik Di Indonesia Dengan Naive Bayes Classifier Sentimen Analysis on Twitter About the Use of Antibiotics in Indonesia With Naive Bayes Classifier. *15(1)*, 1–9.
- Erfina, A., Basryah, E. S., Saepulrohman, A., dan Lestari, D. (2020). Analisis Sentimen Aplikasi Pembelajaran Online Di Play Store Pada Masa Pandemi Covid-19 Menggunakan Algoritma Support Vector Machine (Svm). *Seminar Nasional Informatika (SEMNASIF)*, 1(1), 145–152.
- Gormantara (2020) Pemodelan Dan Simulasi Indoor Wireless System Menggunakan Metode Finite Volume Dengan Gpu Cuda. *S2 Thesis, Universitas Atma Jaya Yogyakarta*.
- Harafani, H., dan Maulana, A. (2019). Penerapan Algoritma Genetika pada Support Vector Machine Sebagai Pengoptimasi Parameter untuk Memprediksi Kesuburan. *Jurnal Teknik Informatika STMIK Antar Bangsa*, V(1), 51–59.
- Kompas.com. (2020, Desember 6). virus Corona di Asia: Indonesia Nomor 4 Kasus COVID-19 Terbanyak. Retrieved Desember 6, 2020, from KOMPAS.com: <https://www.kompas.com/tren/read/2020/12/06/080100965/virus-corona-di-asia--indonesianomor-4-kasus-COVID-19-terbanyak?page=all>

- Latifah, E. (2018). Perbandingan Kinerja Machine Learning Berbasis Algoritma Support Vector Machine dan Naive Bayes (Studi Kasus: Data Tanggapan Mengenai Traveloka Melalui Media Sosial Twitter). *SI Skripsi, Universitas Islam Indonesia Yogyakarta*.
- Lengkong, N. C., Safitri, O., Machsus, S., Putra, Y. R., Syahadati, A., dan Nooraeni, R. (2021). Analisis Sentimen Penerapan PSBB Di Dki Jakarta Dan Dampaknya Terhadap Pergerakan Ihsg. *Jurnal Teknoinfo*, 15(1), 20.
- Liu, J., Tian, Z., Liu, P., Jiang, J., dan Li, Z. (2016). An Approach of Semantic Web Service Classification Based on Naive Bayes. *Proceedings - 2016 IEEE International Conference on Services Computing, SCC 2016*, 356–362.
- M. Al-Ayyoub, A. A. Khamaiseh, Y. Jararweh, and M. N. Al-Kabi, "A Comprehensive Survey Of Arabic Sentimen Analysis," *Information Processing dan Management*, vol. 56, pp. 320-342, 2019.
- Nasrullah, Rulli. (2015). Media Sosial Perspektif Komunikasi, Budaya dan Sosio teknologi. Bandung. *Simbiosis Rekatama Media*
- Natasuwarna, A. P. (2020). Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring. *Techno.Com*, 19(4), 437–448.
- O'Reilly, C. A., Chatman, J., dan Caldwell, D. F. (1991). People and Organizational Culture: a Profile Comparison Approach To Assessing Person-Organization Fit. *Academy of Management Journal*, 34(3), 487–516.
- Pratiwi, R. W., H, S. F., Dairoh, D., Af'idah, D. I., A, Q. R., dan F, A. G. (2021). Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM). *INISTA: Journal of Informatics, Information System, Software Engineering and Applications*, 4(1), 40–46.
- Syarifuddin, M. (2020). Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode *Naïve Bayes* dan KNN. *Inti Nusa Mandiri*, 15(1), 23-28.
- Taufik, I dan S.A.Pamungkas. (2018). Analisis Sentimen Terhadap Tokoh Publik Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal "LOG!K@,"* 8(1), 71–79.
- Tineges, R., Triayudi, A., dan Sholihati, I. D. (2020). *Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter dengan Metode Klasifikasi Support Vector Machine*

(SVM). *Jurnal Media Informatika Budidarma*, 4(3), 650.

Tuhuteru, H., dan Kristen Indonesia Maluku Jl Ot Pattimaipauw, U. (2020). Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine. *Information System Development (Isd)*, 5(2), 7–13.

Wati, R. (2016). Penerapan Algoritma Genetika Untuk Seleksi Fitur Pada Analisis Sentimen Review Jasa Maskapai Penerbangan Menggunakan Naive Bayes Risa. *Jurnal Evolusi* Volume, 4.



LAMPIRAN

Data Hasil *Crawling Twitter*

No	Tweet	Sentimen
1	@jawapos: Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021 https://t.co/cWSiydUxbl	neu
2	Capai Kemandirian, Vaksin Merah Putih Siap Diproduksi Pada 2021 https://t.co/cWSiydUxbl	neu
3	Vaksin Merah Putih Produksi Indonesia Siap Diproduksi Akhir 2021 https://t.co/LBw5wDphKA	neu
4	https://t.co/OcUurMt1il #covid19 #vaksinmerahputih #bumninccom https://t.co/yejiwcYOpH	neu
5	Let's pray for our Vaksin Merah Putih with local strain virus (Eijkman and Biofarma), before that please stay safe Though MUI, Biofarma, etc already went to Sinovac facility in Mainland China, and we already order an amount of dosage too	neg
6	@jasa_kuli: Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik. https://t.co/K5Qv3LD5UI	neu
7	Vaksin Merah Putih lebih cepat lebih baik. Tapi tak mesti cepat yang penting baik. https://t.co/K5Qv3LD5UI	neu
8	Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19, tetapi juga soal kedaulatan nasional. ????? #IsolasiRizieqdanFPI #YangYangBestBoy #BTSONGMA https://t.co/3SwMcOuJ4k	pos
9	Astrazeneca menang jauh dari Moderna, apalagi Pfizer dari sisi transportability dan biaya. Mantap sih ini daripada vaksin merah putih https://t.co/OdYMuAIMwJ	neu
10	@itarahmadalf Kok ngga digawe sukarelawan uji vaksin Merah Putih ae kerumunane. wqwqsial	neu
11	Sudah 3 vaccine trial yg publikasi akurasi dan efektivitas. Pak @Jokowi, klo vaksin merah-putih kira2 akurasinya berapa %?. ?	neg
12	@RakyatMerdeka99: Siap Diproduksi Akhir 2021 Kemenristek: Vaksin Merah Putih Bisa Perkuat Kedaulatan Negara @KemenristekBRIN #Vaksin #C...	neu

13	<p>Siap Diproduksi Akhir 2021 Kemenristek: Vaksin Merah Putih Bisa Perkuat Kedaulatan Negara</p> <p>@KemenristekBRIN #Vaksin #Corona #coronavirus #COVID19 #COVID?19 #COVID_19 #Covid_19 #COVID #RakyatMerdeka #RMCoid</p> <p>https://t.co/J04qbToGfu</p>	neu
14	<p>Erick Thohir Kawal Vaksin Merah Putih.</p> <p>https://t.co/RnvQqWnD9E</p>	neu
15	<p>Kabar baiknya bibit vaksin merah putih siap diujikan pada hewan dan kita bisa mengetahui efeknya #DukungErickPimpinPDIP</p> <p>https://t.co/q8VFu0iuju</p>	neu
16	<p>Harapan yang terbaik untuk pengembangan vaksin merah putih semoga bisa diproduksi tahun 2022 nanti dan bisa memutus rantai penyebaran covid</p> <p>#DukungErickPimpinPDIP https://t.co/7C2FWIQ5cw</p>	neu
17	<p>Tahun 2022 Indonesia tidak akan lagi mengimpor vaksin dari luar negeri karena akan menggunakan vaksin Covid buatan dalam negeri sendiri</p> <p>#DukungErickPimpinPDIP https://t.co/Pg2goIHnmI https://t.co/aOxY9kCOL3</p>	neu
18	<p>Dalam upayanya untuk mengurangi ketergantungan pada vaksin Covid impor pemerintah mengembangkan vaksin covid buatan sendiri yang diberi nama vaksin merah putih #DukungErickPimpinPDIP https://t.co/czRb15lpdV</p>	neu
19	<p>Pengembangan Vaksin Covid dalam negeri terus dilakukan agar nanti kita tak perlu lagi mengimpor vaksin dari luar negeri #DukungErickPimpinPDIP https://t.co/K6iRtHkmix https://t.co/FZk2fD16yp</p>	neu
20	<p>Vaksin merah putih buatan dalam negeri akan ditempatkan diproduksi pada tahun 2022 #DukungErickPimpinPDIP https://t.co/VV4WcwSaz1 https://t.co/ySJ4Q1McBc</p>	neu
...
4586	<p>@GOV_BN: Kajiselidik kebangsaan mengenai penerimaan orang ramai terhadap penggunaan vaksin COVID-19.</p> <p>- National survey on the public acc...</p>	neu
4587	<p>Dr tahun kemaren sbnrnya ada rencana vaksin influenza. Bahkan sbml covid dateng. Krn setahun bisa flu tiap bulan ? miniamal 6x setahun. Tp gak lagi.. gak lagi. Dlu mikirnya cape klo mnum vitamin c tiap hri. Tktnya ngerusak hati/ginjal jg tkt lupa byk mnum air putih</p>	pos
4588	<p>China produsen sinovac sedang lockdown menghadapi serangan kedua covid yg mematikan.. Sementara ina merasa tenang karena sudah memiliki vaksin buatan china.. Hingga gembira dan teriak makan tuh woy China vaksin lu dah Ai borong.. Wkwkwk pengen sekali tag prof mahfud md..</p>	neu

Data Hasil *Cleansing* dan *Pre-processing*

No	Sentimen	Tweet
1	pos	Pengembangan vaksin Merah Putih bukan hanya untuk kemandirian dalam produksi vaksin Covid-19 tetapi juga soal kedaulatan nasional
2	neg	Sudah 3 mencoba vaksin yg publikasi akurasi dan efektivitas. Pak Jokowi kalau vaksin merah-putih kira - kira akurasinya berapa persen
3	neg	Pemerintah berupaya optimal untuk memutus rantai virus covid sekarang itu membuahkan hasil kabarnya bibit vaksin merah putih siap diujikan pada hewan
4	pos	Sementara sebelum adanya vaksin merah putih Indonesia mamfaatkan vaksin yang seadanya untuk mencegah covid
5	pos	Harapan yang terbaik untuk pengembangan vaksin merah putih semoga bisa diproduksi tahun 2022 nanti dan bisa memutus rantai penyebaran covid
6	neg	Vaksin merah putih menggunakan isolat virus yang bertransmisi di Indonesia
7	pos	Erick thohir mengatakan bahwa langkah penting untuk pengadaan vaksin merah putih. Semangat terus!
8	pos	Sebelum ada vaksin merah putih mari kita selalu menjaga protokol kesehatan yang ada
9	pos	Udah tahu belum teman-teman kalau vaksin yang dikembangkan di Indoensia dinamai vaksin merah putih
10	pos	Pemerintah kawal akselerasi vaksin Merah Putih karya anak bangsa. LUAR BIASA
11	neg	Vaksin Merah Putih merupakan strategi jangka panjang penanganan pandemi COVID-19 yang memanfaatkan isolat virus yang bersirkulasi di Indonesia dalam proses pengujiannya.
12	neg	Peneliti Rampungkan Vaksin Merah Putih AzisSyamsuddin Ajak Kepala Daerah Redam Sebaran Wabah
13	pos	President jokowi : Pemerintah Kawal Akselerasi Vaksin PRODUK ANAK BANGSA Merah Putih BRAVO Indonesia!!!
14	pos	Sabar yah! Vaksin Merah Putih Baru Tersedia 2022
16	pos	Jokowi Gaspol Lawan Pandemi. Pemerintah akan kawal akselerasi proses pembuatan Vaksin Merah Putih. Semakin cepat semakin baik!!
17	pos	Pemerintah kawal akselerasi vaksin merah putih!
18	pos	Merah Putih produk anak bangsa masih dalam tahap proses belum sampe produksi. didukung penuh oleh Pak Jokowi. Si aniesbaswedan dan kawan-kawan kan anti rezim Pingin sembuh cepet ya vaksin dari cina Yg berjubel di bandara tempo
19	pos	Kembali ke pekerjaan Presiden Jokowi mengawal pengadaan vaksin Merah Putih

20	pos	Pemerintah kawal akselerasi vaksin merah putih produksi anak bangsa Indonesia bangga dengan karya anak bangsa jokowi. Lawan Pandemi
21	pos	Vaksin Merah Putih
22	pos	Duh semoga vaksin merah putih dapat segera hadir yaa
23	pos	sudah tidak sabar menanti hadirnya vaksin merah putih
24	neg	Presiden Joko Widodo terus mengawal akselerasi Vaksin Merah Putih yang sedang dikembangkan oleh 6 perguruan tinggi dan lembaga penelitian terkemuka di Indonesia. Pak jokowi menginginkan Vaksin yang terbaik dan teruji klinis dalam rangka melawan virus Covid-19 ini.
25	pos	Belum lagi ngapain bikin vaksin merah putih kalo toh akhirnya ternyata impor juga. rakyat dikerjain terus tapi anehnya pada seneng. Gak kapok2
26	pos	Dok kabar vaksin merah putih gimana Ini eropa sudah beli vaksin dan gratis untuk yang di Belgia. Saya pingin mudik tapi kasian keluarga saya diindo vaksin merah putih apa kabar yaa dok ini update nya Saya baca baca kok belum ada kabar terba
27	pos	Sabar Ya! Eijkman Janji Pengembangan Vaksin Merah Putih Cepat dan Aman
28	neg	Sejumlah kandidat vaksin Covid-19 rencananya akan digunakan di Indonesia. Tidak hanya dari luar negeri vaksin Merah Putih buatan Indonesia juga sedang dikembangkan dalam rangka mengidentifikasi target untuk terapi dan vaksin virus corona.
29	pos	Mohon maaf terlalu lama 2022 keburu ada virus baru om. Mending beli vaksin yg udah terbukti saja
30	neg	Saya berharap Vaksin Merah Putih menjadi solusi pencegah infeksi Virus Corona di negara kita Jangan impor melulu dong Pak
31	pos	Kawan laut ku Andai kau tahu Vaksin merah putih Bukan asli merah putih Pesan dari laut selatan
32	pos	jalan-jalan ke Aceh! Ibu Susi mencari tahu apa budaya Smong dan mencicipi makanan khas Simeulue. Ibu Susi juga menggali perkembangan vaksin Merah Putih Universitas Airlangga dan Eijkman.
33	neg	mungkin vaksin merah putih ini akan jadi vaksin pamungkas yang akan sangat cocok digunakan untuk strain virus lokal semoga sebelum vaksin itu selesai pandeminya mereda
34	neg	Jadi sekali lagi apa yang sudah kami lakukan dan kita lakukan bersama dengan industri untuk memberikan vaksin secepatnya tapi aman dan manjur' tutupnya.
35	neg	Pengembangan vaksin merah putih untuk menanggulangi pandemi virus Corona merupakan simbol kemandirian suatu bangsa. Cegah Kovid Dan Radikalisme
36	pos	Masih setahun lagi baru produksinya belon distribusi. Udahlah mending nunggu yang import aja Sabar! Vaksin COVID-19 Merah Putih Baru Bisa Diproduksi Desember 2021

37	neg	Bukti yang sangat terlihat jelas bahwa pemerintah sangat sekali peduli kepada masyarakat agar tidak terhindar dari virus
38	neg	Pengembangan vaksin merah putih utk menanggulangi pandemi virus corona merupakan simbol kemandirian suatu bangsa.
39	neg	Pengembangan vaksin merah putih utk menanggulangi pandemi virus corona merupakan simbol kemandirian suatu bangsa. Teladani Rasul Atasi Pandemi
40	pos	Vaksin merah putih ditambah doa kyai dan habib insya Allah efektif 110
41	pos	Datamin mengamati respons publik melalui media sosial dan pemberitaan media online atas pengembangan Vaksin Merah Putih pada 1 Oktober-8 November. Selama itu terdapat 2.614 berita di portal daring dan 5.333 cuitan netizen di Twitter.
42	neg	Sobat GEN Pemerintah meluncurkan proyek 'Vaksin Merah Putih' apakah ini bisa jadi solusi Lengkapnya kayak gimana langsung dengerin aja Spektrum Radio Jatim hari ini jam 9 pagi di 91 stasiun radio anggota Persatuan Radio Siaran Swasta Nasional Indonesia
43	neg	Harapan Vaksin Merah Putih
...
765	neg	Sebenarnya pengen di vaksin covid tapi apalah daya masih punya baby
766	pos	Ada yg dapet broadcast ini di WA Kata WA ini 1. Vaksin presiden Jokowi salah teknik sehingga harus diulang 2. Injeksi intramuskular harus minimal 3cc baru berefek 3. Vaksin covid harus tiap 6 bulan seumur hidup 4.
767	pos	Covid dateng. Panik dong aku soalnya imun aku paling lemah diantara temen2 atau keluarga. Dari maret hampir gapernah stop vit c. Dan dr kantor juga memang dikasih utk tiap bulannya. Dari maret sudah ga pernah flu lagi jadinya kelupaan sm si vaksin influenza karena keenakan sm vitamin c
768	pos	Dari tahun kemaren sbnrnya ada rencana vaksin influenza. Bahkan sblm covid dateng. Karena setahun bisa flu tiap bulan miniamal 6x setahun. Tp gak lagi. gak lagi. Dulu mikirnya capek kalau minum vitamin c tiap hari. Takutnya merusak hati/ginjal juga takut lupa byk minum air putih