

ANALISIS SENTIMEN TANGGAPAN MASYARAKAT TERHADAP BANTUAN SOSIAL PEMERINTAH DI MASA PANDEMI COVID-19 PADA PLATFORM TWITTER

Melani Asta Rosari¹⁾ Wasino²⁾ Tony³⁾

¹⁾²⁾³⁾ Sistem Informasi, Fakultas Teknologi Informasi, Universitas Tarumanagara
Jl. Letjen S. Parman, No.1 Jakarta 11440 Indonesia

email : melaniasta00@gmail.com¹⁾ email : wasino@fti.untar.ac.id²⁾, email : tony@fti.untar.ac.id³⁾

ABSTRACT

The study was conducted to analyze public opinion about the government's efforts in overcoming the Covid-19 pandemic by providing social assistance (bansos) in the form of goods, money and or services. Through social media Twitter with the topic of social assistance, the classification of positive, neutral or negative sentiments will be carried out. The results of this classification will reveal what social assistance topics are often discussed on Twitter. This classification process uses the Naive Bayes method and uses the RapidMiner application. The data used in this analysis process is 747 Twitter comment text data with a data collection time span from October to November. The classification process is supported by the Term Frequency-Inverse Document Frequency feature as the word weighting stage. This classification produces 2,382 word attributes or word vectors from 747 data, with 370 sample data for model testing which produces an accuracy value of 24.32%, a true neutral recall value of 100%, and a true neutral precision value of 24.32%. The word that most often appears from the results of this sentiment analysis is the word "bantuan".

Key words

social assistance, Naive Bayes Classification, Twitter.

1. Pendahuluan

Pandemi Covid-19 ini berdampak pada kesejahteraan masyarakat Indonesia [1]. Pemerintah Indonesia menyikapi permasalahan sosial pandemi dengan memberikan bantuan sosial kepada seluruh masyarakat yang terkena dampak. Bantuan sosial tersebut disalurkan dalam banyak bentuk, dalam tahun 2021 ini terdapat bantuan sosial kartu prakerja, subsidi listrik, bantuan langsung tunai UMKM, Program Keluarga Harapan (PKH), program sembako, bantuan sosial tunai (BST).

Dengan topik bantuan sosial oleh pemerintah, dilakukan analisis untuk melihat secara umum kedalam ruang diskusi media sosial Twitter. Topik bantuan sosial

barang atau jasa yang paling sering menjadi pembicaraan. Bernilai seperti apakah topik pembicaraan tersebut, lalu seperti apakah bantuan sosial yang diinginkan masyarakat. Supaya dapat diketahui dari hasil analisis sentimen klasifikasi ini.

Klasifikasi ini menggunakan metode Naive Bayes sebagai algoritma untuk menghitung probabilitas sentimen dalam data Twitter yang didapatkan. Penggunaan metode Naive Bayes didukung oleh penelitian sebelumnya yang menguji Naive Bayes apakah baik dalam melakukan klasifikasi data [2]. Dengan hasil penelitian yang mengatakan bahwa Naive Bayes cukup berpotensi baik dalam klasifikasi dokumen dibandingkan metode lain dalam hal akurasi serta efisiensi komputasi. Dalam penelitian sebelumnya metode Naive Bayes juga dilakukan untuk menganalisis kebijakan PPKM oleh pemerintah selama masa pandemi Covid-19 dari komentar Twitter, yang menghasilkan nilai 98% opini Twitter masuk kedalam polaritas positif dan 2% polaritas negatif [3]. Penelitian menggunakan sumber data dari Twitter ini didukung oleh penelitian sebelumnya mengenai Twitter sebagai ruang publik pemerintah lokal yang parsitifatif di era modern saat ini [4].

2. Tinjauan Pustaka

2.1 Analisis Sentimen

Analisis sentimen adalah proses analisis data yang diperoleh dari berbagai *platform* media sosial maupun internet. Proses analisis ini memerlukan pemahaman dan pengolahan data teks untuk mendapatkan informasi sentimen dari suatu teks ataupun kalimat opini dalam *platform* media sosial [3]. Analisis sentimen ini dilakukan dengan tujuan untuk mengetahui, apakah teks opini tersebut mengarah ke pandangan yang positif atau negatif.

2.2 Bantuan Sosial

Bantuan sosial adalah pemberian bantuan berupa uang, barang atau jasa yang diberikan kepada seseorang, keluarga, kelompok atau masyarakat miskin, tidak mampu, dan/atau rentan terhadap resiko sosial [5]. Bantuan sosial yang diberikan pada saat pandemi covid-19 adalah upaya pemerintah pusat dan daerah untuk membantu seluruh masyarakat yang terdampak pandemi. Koordinasi kerjasama antara Kementerian sosial dan lembaga pemerintahan serta Kementerian Keuangan dalam mengatur pengelolaan dana bantuan sosial, diatur dalam PMK-43-2020 [6]. Beberapa bentuk bantuan sosial yang diberikan selama pandemi di akhir tahun 2021 dari bulan September hingga Desember adalah : kartu prakerja dan subsidi gaji, kartu sembako, bantuan listrik, bantuan kuota kemendikbudristek, PKH (Program Keluarga Harapan), BPUM (Bantuan Produktif Usaha Mikro) atau Bantuan Langsung Tunai UMKM [7].

2.3 Text Mining

Text Mining adalah proses penemuan pengetahuan menggunakan *Natural Language Processing* (NLP) dengan cara menggali informasi dari sebuah data berformat teks. *Text mining* sendiri bertujuan untuk mencari kata-kata dan mendapatkan informasi yang berguna dimana informasi tersebut dapat mewakili isi dari dokumen yang berkaitan sehingga dapat dilakukan analisa yang berhubungan antar dokumen [8]. Sumber data yang digunakan pada *text mining* adalah kumpulan teks yang memiliki format tidak terstruktur.

2.4 Naive Bayes Classifier

Naive Bayes Classifier adalah sebuah metode untuk melakukan klasifikasi sederhana berdasarkan algoritma bayesian, dengan probabilitas sederhana supaya dapat diketahui bahwa tidak ada keterkaitan antar kelas yang satu dan yang lain [9]. Model Naive Bayes ini dinilai cukup berpotensi baik dalam klasifikasi dokumen dibandingkan metode lain dalam hal akurasi serta efisiensi komputasi [2], karena memungkinkan setiap atribut berkontribusi dalam keputusan akhir atribut yang setara dan independen. Penelitian tersebut menguatkan penulis untuk menggunakan metode Naive Bayes dalam analisis sentimen ini. Selain karena prosesnya yang mudah dipahami, Naive Bayes ini sering digunakan dalam penelitian serupa dengan hasil akurasi maksimal walaupun dengan data latih yang sedikit.

2.5 Twitter

Twitter adalah layanan jejaring sosial dan mikroblog daring yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter akan tetapi pada tanggal 07 November 2017 bertambah hingga 280 karakter yang dikenal dengan sebutan kicauan (tweet) [10]. Salah satu daya tarik dari Twitter adalah hastag atau kata yang dapat menjadi

trending, baik pada kolom pencarian ataupun dalam *trending* tweet atau unggahan tweet. Dalam proses analisis sentimen Twitter ini, penulis memanfaatkan *trending* pada tweet dengan topik bantuan sosial agar dapat diambil datanya untuk proses analisis penelitian.

2.6 Preprocessing

Preprocessing adalah tahap awal pengolahan data yang belum terstruktur [3]. Tujuannya untuk mempermudah proses pencarian query dalam dokumen, mempercepat dalam pemrosesan terhadap dokumen, dan mempermudah proses pengurutan data yang diambil [11]. Preprocessing ini mempunyai tahapan-tahapan dalam mengolah suatu data teks, tahap tersebut meliputi [3]:

1. Case folding, proses merubah semua huruf di dalam dokumen menjadi huruf kecil. Selain huruf dihilangkan.
2. Cleaning, proses untuk memperbaiki atau menghapus data yang tidak akurat. Dilakukan *missing value* atau mengisi nilai yang hilang dan menghilangkan data yang bersifat *noise*.
3. Remove Stopword, kata-kata yang kurang bermakna atau tidak mempunyai arti dihilangkan atau dihapus
4. *Tokenization*, memecah kalimat dokumen menjadi potongan-potongan bagian yang disebut token.
5. *Stemming*, mengubah kata menjadi bentuk kata dasarnya, dengan cara menghilangkan kata imbuhan baik awalan atau akhiran kata.

2.7 Term Frequency Inverse Document Frequency (Tf-Idf)

Tf-Idf merupakan metode yang digunakan untuk menentukan nilai frekuensi sebuah kata di dalam sebuah dokumen atau artikel dan juga frekuensi di dalam banyak dokumen [8]. Fitur ini mempermudah untuk berjalannya proses klasifikasi.

2.8 PEM (Performance Evaluation Measure)

PEM merupakan suatu rangkaian evaluasi dari model yang digunakan. Evaluasi ini bertujuan untuk mengetahui keakuratan kinerja dan performa dari uji model data training [8].

2.9 RapidMiner

RapidMiner merupakan perangkat lunak terbuka (open source) yang digunakan sebagai salah satu solusi dalam melakukan analisis dalam data mining, teks mining, dan analisis prediksi [12].

2.10 PHP

Hypertext Preprocessor atau PHP merupakan bahasa pemrograman *server side scripting* yang didesain untuk pengembangan suatu web [13].

3. Metodologi

3.1 Data Penelitian

Data penelitian yang digunakan adalah data dokumen yang berasal dari aplikasi Twitter. Data tersebut berupa data teks komentar yang nantinya akan menjadi dokumen untuk di klasifikasi. Proses pengambilan data menggunakan aplikasi RapidMiner dengan bantuan *query* bantuan sosial covid-19 pada parameter operator *search twitter*. *Crawling* pada Twitter menghasilkan 1.385 dokumen kalimat yang masih banyak duplikasi dan terdapat kalimat bermakna sama karena hasil *buzzer* Twitter (akun yang dibayar untuk meramaikan topik dengan *broadcast* tweet), maka setelah dilakukan penghapusan duplikat dihasilkan 747 data siap proses.

Setelah data didapatkan maka dilakukan *training labelling* untuk acuan perbandingan dari *labelling* manual dengan hasil dari klasifikasi sentimen. Data yang sudah di *labelling* selanjutnya masuk ke tahap pembersihan *data mining*, atau *preprocessing*.

3.2 Preprocessing

Preprocessing ini untuk membersihkan kata yang tidak diperlukan atau kata yang tidak bermakna. Langkah-langkah dari *preprocessing* harus disesuaikan dengan kondisi data komentar Twitter yang ada. Berikut contoh langkah *preprocessing* :

“Jangan main-main dengan bansos covid19, bansos harus diterima utuh oleh penerima manfaat.”

1. Tahap Case Folding
jangan main-main dengan bansos covid19, bansos harus diterima utuh oleh penerima manfaat.
2. Tahap Cleaning
jangan main main dengan bansos covid19 bansos harus diterima utuh oleh penerima manfaat
3. Tahap Tokenization
jangan | main | main | dengan | bansos | covid19 | bansos | harus | diterima | utuh | oleh | penerima | manfaat
4. Tahap Remove Stopword
main | main | bansos | covid19 | bansos | diterima | utuh | penerima | manfaat
5. Tahap Stemming
main | main | bansos | covid19 | bansos | terima | utuh | terima | manfaat

3.3 Ekstraksi Fitur

Pada tahap ini dilakukan representasi data dengan *word vector* (matriks yang mewakili kata dokumen). Selain proses tersebut terdapat fitur kemunculan kata dengan Tf-Idf (*Term frequency - Inverse Document Frequency*) [14].

Gambar 1 Ekstraksi Fitur

No.	Twitter hasil preprocessing	Klasifikasi
1.	indonesia (1) tangan (1) covid (1) baik(1) pesta(1) warga(2) gopoh (2) bawa (1) oksigen(1) beli (1) obat(1) vitamin (1) nakes(1) jibaku(1) bansos(1) begal (1) pemerintah menteri (1)	Negatif
2.	main (2) bansos(2) covid (1) terima (2) utuh (1) manfaat(1)	Positif
3.	temu (1) terima (1) bansos (3) valid(1) sembako (1) kemarin(1) nilai(1) triliun (1) dana(1) lari (1) kemana(1)	Negatif
4.	bansos (1) tangan (1) covid (1) salur (1) hingga (1) dua (3) ribu (1) puluh (1)	Positif
5.	masyarakat (1) terima (1) bansos (1) covid (1) saran (1) punya (1) rekening (1) bank (1)	Netral

Pada tabel ekstraksi Tf-Idf menggunakan rumus menghitung nilai bobot dokumen :

$$W_{dt} = TF_{dt} * IDF_{ft} \dots\dots\dots(1)$$

Keterangan:

W_{dt} = Nilai dokumen ke-d pada kata ke-t

TF_{dt} = Jumlah kata yang dicari dalam satu dokumen

IDF_{ft} = Inverse Document Frequency $(\log(\frac{D}{df}))$

D = Jumlah dokumen

df = Jumlah dokumen yang mengandung kata

3.4 Klasifikasi Naive Bayes

Metode ini beroperasi dengan menggunakan teorema Bayesian yang akan mencari probabilitas dari data sentimen Twitter yang sering muncul. Klasifikasi Naive Bayes ini dilakukan pada data dokumen Twitter yang sudah diproses dalam tahap *preprocessing* dan sudah melalui tahap ekstraksi fitur sehingga dapat lebih mempermudah proses klasifikasi.

Berikut rumus Bayes yang digunakan dalam menghitung probabilitas [15]:

$$P(H|X) = \frac{P(H|X) P(H)}{P(X)} \dots\dots\dots(2)$$

Keterangan :

X = Data class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

$P(H|X)$ = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)

$P(H)$ = Probabilitas hipotesis H (prior prob.)

$P(X|H)$ = Probabilitas X berdasarkan kondisi tersebut

$P(X)$ = Probabilitas X

Dengan aturan sebagai berikut :

Jika $P(H1|X) < P(H2|X)$, maka X termasuk sebagai H2. Pernyataan $P(H1|X)$ menandakan probabilitas hipotesis H1 berdasarkan kondisi X, begitu pula dengan h2. Sehingga dapat klasifikasi dari x sesuai dengan probabilitas terbesar diantara probabilitas x terhadap semua kelas.

3.5 PEM (Performance Evaluation Measure) & Confusion Matrix

Berikut adalah rumus performance evaluation measure, untuk evaluasi model klasifikasi :

1. Precision adalah tingkat ketepatan antara request pengguna dengan jawaban sistem. Rumus precision dapat dilihat pada persamaan 3.

$$pre = \frac{TP}{FP + TP} \dots \dots \dots (3)$$

2. Accuration adalah perbandingan antara informasi yang dijawab oleh sistem dengan benar dengan keseluruhan informasi. Rumus accuration dapat dilihat pada persamaan 4.

$$acc = \frac{TN + TP}{FN + FP + TN + TP} \dots \dots \dots (4)$$

3. Recall adalah ukuran ketepatan antara informasi yang sama dengan informasi yang sudah pernah dipanggil sebelumnya. Rumus recall dapat dilihat pada persamaan 5.

$$rec = \frac{TP}{FN + TP} \dots \dots \dots (5)$$

Melalui persamaan rumus PEM maka mendapatkan hasil yang akan ditampilkan dalam tabel confusion matrix atau tabel akurasi model perbandingan dengan dataset [8].

True Class	Predicted Class		
	Positif	Negatif	Netral
Positif	TP	FN	FNet
Negatif	FP	TN	FNet
Netral	FN	FP	TNet

Keterangan :

- TP (true positive) : jumlah data nilai positif yang diprediksi benar sebagai positif
- TN (true negative) : jumlah data nilai negatif yang diprediksi benar sebagai negatif
- FP (false positive) : jumlah data nilai negatif yang diprediksi salah sebagai positif
- FN (false negative) : jumlah data nilai positif yang diprediksi salah sebagai negatif
- FNet (false netral) : jumlah data nilai netral yang diprediksi positif atau negatif
- TNet (true netral) : jumlah data nilai netral yang diprediksi sebagai netral

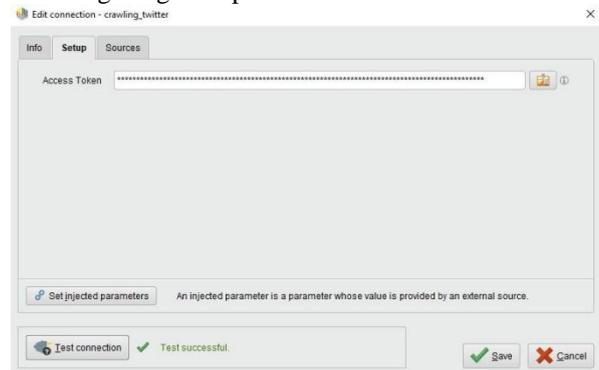
4. Hasil dan Pembahasan

4.1 Deskripsi Data

Data yang digunakan untuk proses penelitian ini merupakan data dari Twitter. Data Twitter ini berupa dokumen teks yang berasal dari komentar Twitter oleh user dengan berbagai topik. Dalam penelitian dibutuhkan data Twitter dengan topik bantuan sosial selama covid-19, dengan spesifikasi waktu data dari bulan Oktober hingga November 2021.

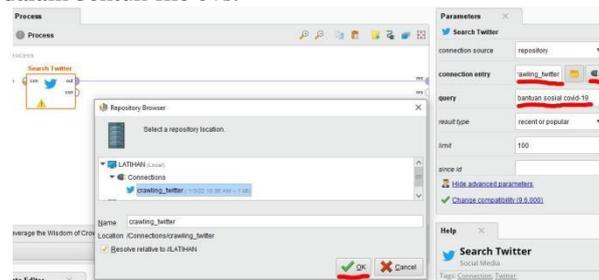
Pengambilan data membutuhkan akun Twitter karena proses perizinan membutuhkan akun agar mendapat access token. Aplikasi RapidMiner sebagai alat

pengambilan data dibantu dengan operator search twitter yang memiliki *connection* antara aplikasi Twitter dan RapidMiner. *Connection* dibuat dengan menggunakan *access token* yang didapatkan lalu dilakukan test connection untuk menguji apakah Twitter sudah terhubung dengan RapidMiner.



Gambar 2 Tampilan Test Connection

Proses selanjutnya yaitu menentukan parameter yaitu *query*, *limit data crawling* dan *connection entry* sebagai file penyimpanan otomatis setelah crawling oleh operator. Setelah proses dijalankan maka hasil akan otomatis tersimpan di penyimpanan perangkat komputer dalam bentuk file cvs.



Gambar 3 Parameter Search Twitter

4.2 Pengolahan & Penyajian Data

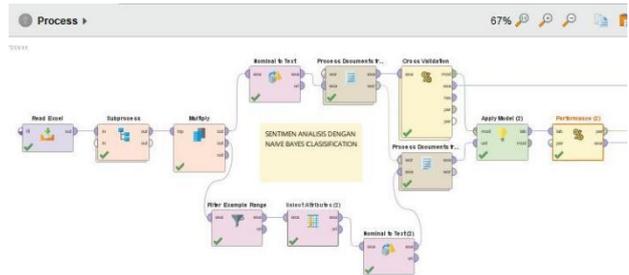
Data yang telah di-crawling dan tersimpan dalam file .cvs selanjutnya di-labelling secara manual sebagai *training label*. *Labelling* ini berguna untuk menjadi acuan pembandingan bagi label data sebelum dilakukan klasifikasi dan label data setelah di klasifikasi.

Text	Label
polynomial	polynomial label
89 RT @temponewsroom: Menteri Sosial Juliani Batubara didug...	Negatif
90 Kementerian Sosial bersama KOTICA salurkan bantuan COVI...	Positif
91 Jangan sampai program prorakyat dalam bentuk bantuan so...	Netral
92 Upaya Mengurangi Ketimpangan Ekonomi Akibat Pandemi C...	Positif
93 @TAKDIR Menemukan Penyelewengan pada Bantuan Sosial...	Negatif
94 Partai Demokrat di Sulawesi Barat melaksanakan Bakti sosia...	Positif
95 @geloraco Proyek Kereta Cepat Jkt-Bdg yg tak layak disuntik ...	Netral
96 RT @hariansederhana: Informasikan Perihal Bantuan Covid...	Netral
97 Kami yg terancam makan sejak awal pandemi yg diming-im...	Negatif
98 Bertempat di bangsal kepatihan kantor Gubernur DIY Senin (...)	Positif
99 Bhayangkara Untuk Negeri Polsek Sungai Raya bersama MP...	Positif
100 @OwHumas_Polri Semoga dengan hal baik ini akan terus be...	Positif

Gambar 4 Data Labelling

Proses selanjutnya dilakukan menggunakan rangkaian operator RapidMiner. Berikut proses dalam RapidMiner :

1. Langkah pertama adalah meng-input data dari hasil *labelling* manual ke dalam operator *read.csv*
2. Kemudian data masuk ke dalam subproses yang akan membersihkan data dari atribut-atribut asal Twitter yang tidak ingin digunakan seperti @, url, dan karakter lain yang tidak mempunyai arti.
3. Selanjutnya data masuk ke operator *multiply* yang akan membuat salinan data, sehingga dapat disebarakan menjadi 2 *output* yang sama jumlah dan spesifikasi datanya.
4. Berlanjut pada data yang di-input di jalur atas, yaitu data yang akan dilakukan perubahan angka menjadi teks.
5. Lalu masuk kedalam subproses yang memiliki fungsi *preprocessing*, dan ekstraksi fitur dengan parameter Tf-Idf. Didalam proses tersebut terdapat perubahan karakter emoji menjadi suatu label ekspresi, misal emoji :) secara otomatis akan tergantikan oleh kamus menjadi kata “senang”.
6. Selanjutnya kalimat yang telah manjadi pecahan kata berbobot masuk dalam *operator cross validation*. Operator ini memiliki subproses lagi yaitu proses operator Naive Bayes untuk klasifikasi, serta operator *apply model* dan *performance* yang berfungsi sebagai penerapan model data yang belum dilabel untuk dilakukan label model dan menilai *performance model* yang diterapkan oleh operator.
7. *Output* sampel data pada operator *cross validation* langsung disambungkan ke *port result*, sedangkan *output* model dilanjutkan ke operator *apply model (2)* untuk dilakukan perbandingan data dengan yang belum di beri label. Data yang belum di beri label adalah data yang berasal dari rangkaian bawah, yang tidak melewati proses klasifikasi Naive Bayes.
8. Operator jalur bawah terdapat operator *example range* untuk mengatur dengan parameter seberapa banyak akan mengambil data *training*.
9. Selanjutnya data *training* di-input ke operator yang mengubah angka menjadi teks.
10. Kemudian *output* dari operator *nominal to text* akan terhubung ke operator *process document from data*, yaitu didalamnya terdapat proses *preprocessing*, ekstraksi fitur menggunakan parameter tf-idf untuk data *training*.
11. Data *training* dan data hasil klasifikasi Naive Bayes di pertemukan dalam operator *apply model (2)* untuk membandingkan model dengan data yang belum di label oleh hasil klasifikasi.
12. Data hasil label tersebut akan terhubung untuk *input* ke operator *performance (2)* yang membandingkan 2 proses, yang pertama melalui klasifikasi yang kedua tidak melalui proses klasifikasi. *Output* dari operator *performance* adalah *performance* akurasi data klasifikasi yang sudah diterapkan model dengan data *training* yang tidak di klasifikasi.



Gambar 5 Rangkaian Operator RapidMiner

4.3 Hasil Penelitian

Terdapat 3 result dari proses klasifikasi menggunakan algoritma Naive Bayes ini. Klasifikasi ini menghasilkan 2.382 atribut hasil proses pembobotan dari 747 dokumen dengan hasil sentimen label yaitu nilai sentimen 295 positif, 255 negatif, dan 196 netral. Lalu menghasilkan 370 example set dengan label Netral 143, Positif 137, Negatif 90. Serta menampilkan hasil performance vector yaitu tingkat akurasi 23.42% . Data hasil di presentasikan dengan tabel *confusion matrix*.

accuracy: 24.32%

	true Negatif	true Neutral	true Positif	class precision
pred. Negatif	0	0	0	0.00%
pred. Neutral	143	90	137	24.32%
pred. Positif	0	0	0	0.00%
class recall	0.00%	100.00%	0.00%	

Gambar 6 Confusion Matrix

Untuk mencapai manfaat penulisan laporan dan penelitian analisis ini yaitu mengetahui topik mana yang paling banyak dibicarakan antara topik bantuan sosial tunai dan bantuan sosial barang/sembako. Agar para penyelenggara bantuan sosial dapat terbantu dengan hasil yang paling banyak muncul dalam analisis ini. Maka diketahui pada proses *word vector & tf-idf* terdapat yang mengandung arti sama antara bantuan sosial tunai dan bantuan sosial barang. Atribut tersebut yaitu “duit” sebagai bentuk dari topik bantuan sosial tunai, serta “beras” sebagai bentuk dari topik bantuan sosial barang/sembako.

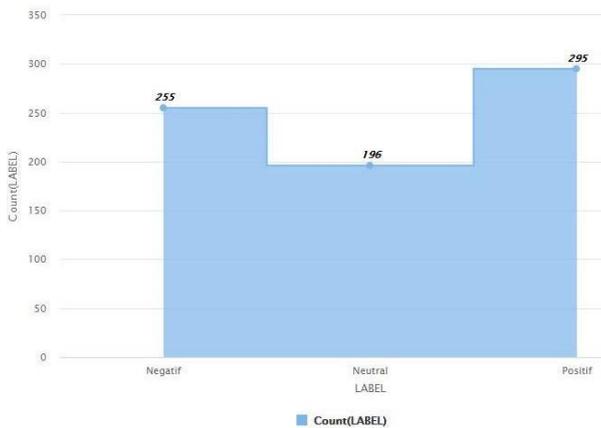
Tabel 1 Perbandingan algoritma A dan algoritma B

Atribut	Term / Banyaknya	Sentimen	Total Bobot
beras (Bantuan sosial barang)	12	2 Negatif	0.857428448
		7 Positif	1.473513773
		3 Neutral	1.090393701
duit (Bantuan sosial tunai)	16	11 Negatif	4.121794485
		1 Positif	0.320885902
		4 Neutral	0.921058625

4.4 Penyajian Hasil

Penyajian hasil dari aplikasi RapidMiner dapat di-export ke dalam bentuk file penyimpanan secara satu per satu karena RapidMiner tidak bisa melakukan sekali export dengan banyak data. Hasil yang didapatkan dari

perbandingan *training labelling* dengan perhitungan *labelling* hasil klasifikasi ditampilkan dalam grafik beserta jumlah masing-masing label.



Gambar 7 Grafik Label

Dari proses preprocessing dan pembobotan kata / ekstraksi fitur menghasilkan atribut sebanyak 2.382 atribut. Maka diambil 30 atribut yang jumlah kemunculannya paling banyak. Atribut tersebut masukkan dalam operator *wordlist* yang akan menyimpan data dalam tabel kata dan total jumlah kemunculan kata. Data pada *wordlist* kemudian ditampilkan menjadi 2 *wordcloud*, *wordcloud* pertama merupakan 5 kata teratas yang jumlah total kemunculan mencapai 200. *Wordcloud* kedua dimulai dari posisi 6 teratas hingga posisi 30.



Gambar 8 Wordcloud 5 Teratas



Gambar 9 Wordcloud Posisi 6 sampai 30 Teratas

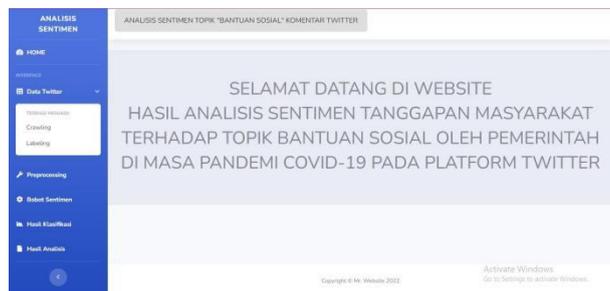
Berdasarkan hasil yang sudah didapatkan lalu dibuat *website* sebagai tampilan dari seluruh hasil klasifikasi. Pembuatan *website* menggunakan bahasa pemrograman PHP yang dihubungkan dengan *phpmyadmin* sebagai penyimpanan database untuk *website*.

```

1 <?php
2 /** ini koneksi ke database **/
3 $servername = "localhost";
4 $username = "root";
5 $password = "";
6 $dbname = "csv";
7
8 $connection = mysqli_connect($servername, $username, $password, $dbname);
9 if(!$connection)
10 {
11     die("Connexion Failed : ".mysqli_connect_error());
12 }
13 >
    
```

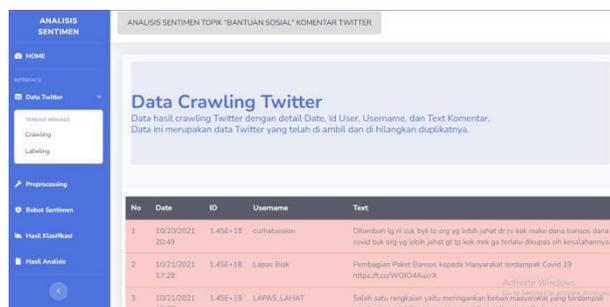
Gambar 10 Source Code PHP

Pada *website* tampilan jika di buka pada *homepage* terdapat judul *website* dan judul penelitian. Terdapat nav bar pada samping kiri yaitu menu yang dapat menampilkan hasil proses klasifikasi dari RapidMiner. Menu dalam *website* diberi nama sesuai dengan tahapan yang dilakukan dalam operator RapidMiner.



Gambar 11 Homepage Website

Pada menu Data Twitter, terdapat pembagian page untuk data hasil *crawling* dan data setelah *labelling*. Menu tersebut mengarah ke halaman yang menampilkan tabel data hasil *crawling* dan data hasil *labelling*.



Gambar 12 Halaman Data Crawling



Gambar 13 Halaman Data Labelling

Tampilan selanjutnya adalah tampilan dari menu Preprocessing yang menampilkan hasil pemecahan kalimat menjadi atribut yang bermacam.



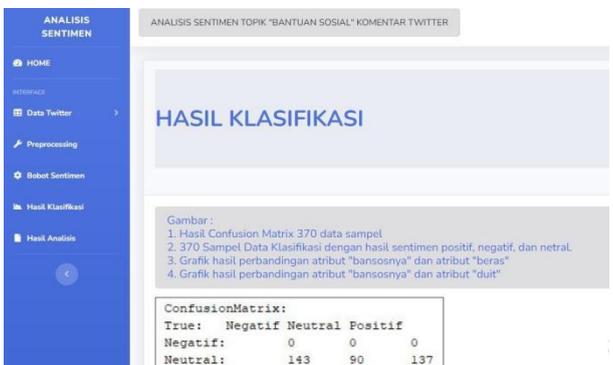
Gambar 14 Halaman Hasil Preprocessing

Hasil dari preprocessing kemudian diambil beserta nilai pembobotannya, untuk dibandingkan nilai atribut “beras” dan “duit” sebagai wakil dari keyword bantuan sosial barang dan bantuan sosial beras.



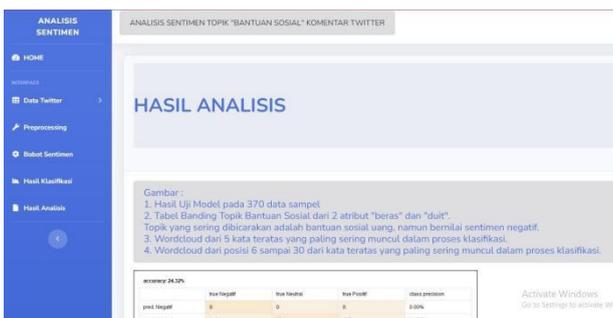
Gambar 15 Halaman Data Pembobotan

Hasil selanjutnya adalah hasil klasifikasi dari aplikasi RapidMiner berupa gambar-gambar grafik serta tabel.



Gambar 16 Halaman Hasil Klasifikasi

Dan menu terakhir menampilkan hasil analisis dari pembobotan 2 atribut simbol bantuan sosial dan menampilkan hasil uji model pada data sampel serta menampilkan hasil wordcloud.



Gambar 17 Halaman Hasil Analisis

4.5 Temuan Penelitian

Berdasarkan hasil analisis sentimen ini menghasilkan beberapa poin, yaitu :

1. Proses klasifikasi 747 data teks yang berlabel 255 kata Negatif, 196 kata Netral, 295 kata Positif menjadi 2.382 atribut kata. Banyak kata-kata yang tidak dapat dibaca sistem sehingga terseleksi pada proses *preprocessing* seperti kata yang disingkat-singkat.
2. Hasil performance vector dengan akurasi 24.32% data, dari label data manual maupun setelah dilakukan proses menggunakan klasifikasi Naive Bayes.
3. Topik bantuan sosial mempunyai banyak sekali bentuk bantuan, sehingga dalam penelitian hanya disimpulkan menjadi 2 bantuan sosial yaitu tunai atau barang dengan wakil atribut “duit” dan “beras”.
4. Frekuensi penggunaan kata yang ditemukan paling banyak adalah kata “bantuan“

6. Kesimpulan

Kesimpulan yang didapatkan berupa :

1. Algoritma Naive Bayes hanya menghasilkan tingkat akurasi rendah untuk 370 data sampel, yaitu 24.32%.
2. Nilai *Performance* yang dihasilkan dari klasifikasi data sampel adalah *accuration* 24.32%, *precision true neutral* 24.32%, *recall true neutral* 100%, yang berarti data kalimat pada sampel di dominasi oleh kalimat netral.
3. Terdapat 2 atribut yang diambil sebagai wakil dari 2 topik bantuan sosial, yaitu “beras” untuk bantuan sosial barang dan “duit” bantuan sosial uang. Kemunculan atribut duit lebih banyak dibandingkan dengan atribut beras, tetapi atribut duit ini memiliki nilai negatif yang tinggi. Kesimpulan, bantuan sosial uang/tunai lebih banyak menjadi pembicaraandari data Twitter, namun bernilai sentimen negatif.
4. Kelebihan dari penelitian, proses lebih efektif dengan satu aplikasi RapidMiner sebagai alat untuk proses *crawling* data, *preprocessing* data, pembobotan data, klasifikasi, hingga uji evaluasi.
5. Kekurangan dari penelitian, tidak dapat menyajikan data yang *real time* karena tidak bisa sekali *crawling* dengan jangka waktu yang diinginkan. RapidMiner memiliki keterbatasan waktu data, sehingga proses *crawling* data harus dilakukan berulang setiap jangka waktu tertentu.

7. Saran

Saran yang didapatkan berupa :

1. Membandingkan aplikasi pendukung klasifikasi agar menjadi pembanding kinerja proses yang lebih efektif dan efisien dalam penggunaan dan hasil akhir.

2. Klasifikasi dengan menggunakan algoritma lain agar dapat membandingkan tingkat keakuratan untuk hasil klasifikasi yang lebih baik.
3. Menggunakan data yang lebih rapi dalam segi kata-kata kalimat agar mudah dalam membaca data.
4. Membuat klasifikasi secara *real time* agar mendapatkan informasi analisis yang terbaru.

REFERENSI

- [1] Kemenkeu.go.id. (2020). Pemerintah Waspada Dampak Pandemi Covid-19 Terhadap Ekonomi Indonesia. SP – 27 /KLI/2020. Kementerian Keuangan. Jakarta.
- [2] Ting, S. L., Ip, W. H., & Tsang, A. H. (2011). Is Naive Bayes a good classifier for document classification. *International Journal of Software Engineering and Its Applications*, 5(3), 37-46.
- [3] Krisdiyanto, T. (2021). Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naive Bayes Classifiers. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu komputer dan Teknologi Informasi*, 7 (1), 32-37
- [4] Abdurrahman, M. S. (2018). Twitter dan Ruang Publik Pemerintahan Lokal yang Partisipatif (Telaah atas Komunikasi Politik Ridwan Kamil Melalui Twitter). *Jurnal PIKOM (Penelitian Komunikasi dan Pembangunan)*, 15(2), 152-168.
- [5] Republik Indonesia. (2019). Undang-undang Nomor 14 Tahun 2019 tentang Pekerja Sosial. Lembaran Negara RI Tahun 2019 No. 182, Tambahan Lembaran Negara RI Tahun 2019 No. 6397. Sekretariat Negara. Jakarta.
- [6] Menteri Keuangan RI. (2020). Peraturan Menteri Keuangan Republik Indonesia Nomor 43 Tahun 2020 tentang Mekanisme Pelaksanaan Anggaran Belanja atas Beban Anggaran Pendapatan dan Belanja Negara dalam Penanganan Pandemi Corona Virus Disease 2019. Kementerian Keuangan. Jakarta.
- [7] Tysara, L. (20 September 2021). Bansos yang Cair Bulan September sampai Desember 2021 dan Cara Mendapatkannya. *Liputan6.com*. <https://hot.liputan6.com/read/4662415/bansos-yang-cair-bulan-september-sampai-desember-2021-dan-cara-mendapatkannya>
- [8] Imron, A. (2019). Analisis sentimen terhadap tempat wisata di kabupaten Rembang menggunakan metode Naive Bayes Classifier. Skripsi. Yogyakarta: UII.
- [9] Suryani, PSM, Linawati, L., & Saputra, KO (2019). Penggunaan Metode Naive Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia. *Mayor Ilm. teknol. Elektro*, 18 (1), 145.Castleman, Kenneth R., 1998, "Digital Image Processing", Prentice Hall, New Jersey.
- [10] Attabi, A. W., Muflikhah, L., & Fauzi, M. A. (2018). Penerapan Analisis Sentimen untuk Menilai Suatu Produk pada Twitter Berbahasa Indonesia dengan Metode Naive Bayes Classifier dan Information Gain. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, 964X*.
- [11] Cahyono, Y., & Saprudin, S. (2019). Analisis Sentimen Tweets Berbahasa Sunda Menggunakan Naive Bayes Classifier dengan Seleksi Feature Chi Squared Statistic. *Jurnal Informatika Universitas Pamulang*, 4 (3), 87-94.
- [12] Aprilla, D., Baskoro, D. A., Ambarwati, L., & Wicaksana, I. W. S. (2013). *Belajar data mining dengan rapidminer*. Jakarta: Gramedia Pustaka Utama.
- [13] Jannah, M., & Sarwandi, C. C. (2019). *Mahir Bahasa Pemrograman PHP*. Elex Media Komputindo.
- [14] Melita, R. (2018). Penerapan Metode Term Frequency Inverse Document Frequency (Tf-Idf) Dan Cosine Similarity Pada Sistem Temu Kembali Informasi Untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Hadits Shahih Bukhari-Muslim) (Bachelor's thesis, Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta).
- [15] Muslehatin, W., Ibnu, M., & Mustakim, M. (2017). Penerapan Naive Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau. In *Seminar Nasional Teknologi Informasi Komunikasi dan Industri* (pp. 250-256).

Melani Asta Rosari, saat ini sebagai Mahasiswa program studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Tarumanagara

