

Perbandingan Algoritma *Naïve Bayes* Dan C4.5 pada Analisis Sentimen Presiden 3 Periode di Twitter

Faizal Albasithu^{1*}, Arief Wibowo²

^{1,2}Fakultas Teknologi Informasi, Teknik Informatika, Universitas Budi Luhur, Jakarta, Indonesia

Email: ^{1*}albashithu2@gmail.com, ²arief.wibowo@budiluhur.ac.id
(* : corresponding author)

Abstrak-Masyarakat Indonesia memperdebatkan kemungkinan adanya wacana presiden Indonesia 3 Periode, jadi menarik untuk diselidiki. Pembatasan dua periode masa jabatan presiden Indonesia oleh MPR merupakan upaya rakyat Indonesia di bawah MPR untuk mencegah presiden Indonesia tetap berkuasa lebih dari 10 tahun, agar menghindari penyalahgunaan dari satu pihak. Wacana ini menimbulkan pro dan kontra, salah satunya pada di platform media sosial Twitter. Perlu adanya analisis sentimen karena adanya Opini yang diberikan oleh masyarakat. Analisis sentimen adalah proses menganalisis atau klasifikasi pendapat, perasaan, sikap, perasaan, dan penilaian seseorang tentang pendapat masyarakat yang diperoleh dari berbagai platform media sosial maupun internet. Tujuannya adalah untuk mengetahui bagaimana sentimen di masyarakat terhadap Presiden 3 Periode di Indonesia melalui *tweet* di media sosial Twitter. Dalam penelitian ini dilakukan proses analisis opini masyarakat tentang Presiden 3 Periode dengan mengklasifikasikan opini ke dalam sentiment positif atau sentimen negatif. Proses klasifikasi dilakukan menggunakan metode algoritma *Naïve bayes* dan C4.5 dijadikan sebagai perbandingan. Identifikasi masalah pada penelitian ini yaitu bagaimana menganalisis sentimen berdasarkan pendapat masyarakat Indonesia di Twitter dan akurasi yang diperoleh algoritma *Naïve bayes* dan C4.5. Hasil klasifikasi pada data uji menunjukkan, algoritma *Naïve bayes* dan C4.5 lalu hasil klasifikasi dan algoritma C4.5 berhasil mendapatkan Akurasi sebesar 78%, sementara Algoritma *Naïve bayes* mendapatkan akurasi sebesar 85%.

Kata Kunci: twitter, analisis sentimen, algoritma c4.5, algoritma *naïve bayes*

Comparison of Naive Bayes and C4.5 Algorithm on Presidential Sentiment Analysis 3 Period on Twitter

Abstract-Indonesian people debate the possibility of a 3rd term Indonesian presidential discourse, so it is interesting to open it. The limitation of the two terms of office of the Indonesian president by the MPR is an attempt by the Indonesian people under the MPR to prevent the Indonesian president from remaining in power for more than 10 years, in order to avoid one party. This discourse raises pros and cons, one of which is on the Twitter social media platform. There is a need for sentiment analysis because of the opinions given by the community. Sentiment analysis is the process of analyzing or classifying opinions, feelings, attitudes, feelings, and judgments about public opinion obtained from various social media platforms and the internet. The aim is to find out how the sentiment in the community towards the 3rd term president in Indonesia is through tweets on Twitter social media. In this study, the process of analyzing public opinion about the President of the 3rd Period was carried out by classifying opinions into positive sentiments or negative sentiments. The classification process is carried out using the *Naïve Bayes* algorithm and C4.5 as a comparison. Identification of the problem in this study is how to analyze sentiment based on the opinion of the Indonesian people on Twitter and the accuracy obtained by yahoo bayes and C4.5. The classification results on the test data show that the *Naïve Bayes* algorithm and C4.5 then the classification and the C4.5 algorithm managed to get an accuracy of 78%, while the Nave Bayes algorithm got an accuracy of 85%.

Keywords: twitter, sentiment analysis, c4.5 algorithm, *naïve bayes* algorithm

1. PENDAHULUAN

Isu Amandemen Kelima terhadap Presiden 3 periode yang diusulkan saat ini menjadi bahan perdebatan publik. Pembatasan dua kali masa jabatan presiden Indonesia oleh MPR merupakan upaya rakyat di bawah MPR untuk mencegah presiden Indonesia tetap berkuasa lebih dari 10 tahun, agar menghindari penyalahgunaan kekuasaan dari satu pihak.

Sebelum perkembangan teknologi yang pesat, orang-orang menyampaikan pendapat dan kritiknya melalui media cetak, artikel-artikel yang dibuat oleh beberapa orang. Hal ini bekerja dengan sangat baik. Karena pada saat itu, media cetak sebenarnya merupakan media untuk mendapatkan informasi yang luas dan detail kepada masyarakat luas. Dengan pesatnya perkembangan teknologi, media cetak mulai sedikit demi sedikit mulai ditinggalkan. Berbagai media sosial seperti Twitter muncul. Masyarakat tidak lagi membutuhkan pers untuk menyampaikan pendapat dan kritiknya secara panjang lebar.

Pada penelitian ini dilakukan analisis pada masyarakat khususnya twitter mengenai Presiden 3 Periode dengan membagi opini masyarakat ke dalam sentimen positif dan negatif. Penelitian mengenai analisis sentimen dan text

mining banyak dilakukan oleh peneliti lainnya dengan berbagai metode. Penelitian ini berhubungan atau berkaitan dengan opini masyarakat dengan menggunakan klasifikasi algoritma *Naïve bayes* dan C4.5.

Analisis sentimen merupakan proses menganalisis atau mengklasifikasi pendapat, perasaan, sikap, dan penilaian seseorang tentang pendapat tertentu untuk menghasilkan nilai positif dan negatif tentang topik tersebut. Analisis sentimen merupakan area penelitian sedang berhubungan dalam text mining. Analisis sentimen juga dapat disamakan dengan penggalian opini karena berfokus pada opini, sikap, dan emosi yang mewakili pandangan individu terhadap suatu peristiwa atau isu tertentu [1]. Text Mining adalah sebuah cara mengekstraksi data dari teks yang tidak terstruktur. Proses Text Mining banyak mengadopsi dari penelitian Data Mining.

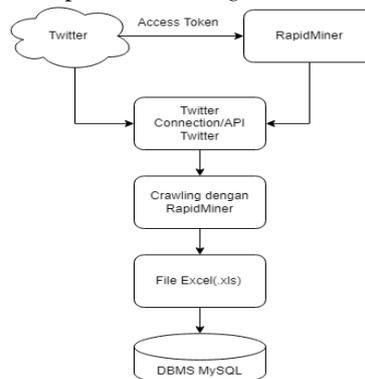
Identifikasi masalah pada penelitian ini yaitu bagaimana menganalisis sentimen berdasarkan pendapat masyarakat indonesia di twitter dan akurasi yang dihasilkan dengan algoritma *Naïve bayes* dan C4.5. Tujuan penelitian ini melakukan analisis sentimen mesyarakat di media sosial twitter terhadap *tweet* berbahasa Indonesia berupa pendapat masyarakat tentang isu presiden 3 periode di indonesia. Klasifikasi *Naïve bayes* dan C4.5.

2. METODE PENELITIAN

Tahap-tahap perancangan analisis sentimen pada penelitian ini yaitu *Crawling* (pengambilan data), pelabelan, pembagian data, *preprocessing* data, algoritma C4.5, algoritma *Naïve bayes* dan pengujian data.

2.1 Crawling

Crawling merupakan proses mengumpulkan data di suatu laman website bertujuan untuk dianalisis lebih lanjut. Dalam penelitian ini proses, *Crawling* dilakukan dengan menggunakan RapidMiner yang bertujuan dapat memperoleh data *tweet* pada twitter. Berikut proses *Crawling* di Gambar 1.



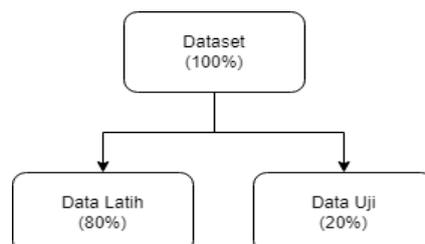
Gambar 1. Proses *Crawling*

2.2 Pelabelan

Pelabelan adalah proses pembagian kelas berdasarkan karakteristik di dalam suatu dokumen atau kalimat. Pada penelitian ini proses pembagian kelas dibagi menjadi sentimen positif dan sentimen negatif dengan cara melakukan *skoring*. Penilaian dokumen dapat masuk di kategori positif atau negatif, ditentukan dengan memanfaatkan sekelompok kata-kata positif dan sekelompok kata-kata negatif [2]. Pada penelitian ini proses pelabelan dilakukan dengan pelabelan manual. Pelabelan manual merupakan proses pemberian kelas. Proses pelabelan manual menentukan hasil akhir dari *tweet* termasuk ke dalam kelas positif atau negatif [2].

2.3 Pembagian Data

Pemodelan adalah proses menghasilkan pengetahuan berdasarkan data pelatihan tertentu. Data latih yang digunakan sebagai model dipilih dengan menggunakan teknik *quota sampling*. *Quota sampling* adalah proses pengambilan data untuk menentukan jumlah sampel data yang diambil dari suatu populasi yang memenuhi karakteristik atau kriteria tertentu sampai tercapai jumlah kuota data yang telah diinginkan [3]. Berikut proses pembagian di Gambar 2.



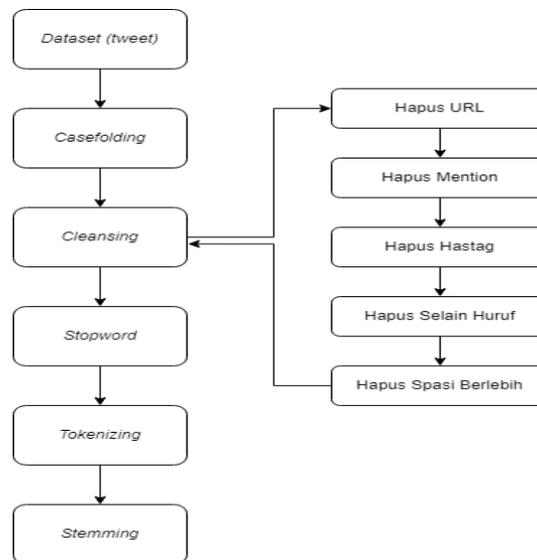
Gambar 2. Pembagian Data

2.4 Preprocessing

Preprocessing merupakan proses penghapusan *noise* dalam dokumen atau kalimat untuk bertujuan mempermudah dalam pengolahan data. Proses *preprocessing* juga untuk menghindari data yang tidak lengkap, data yang rusak, data yang saling bertentangan [3]. Data teks memerlukan proses konversi untuk mengubah menjadi data teks terstruktur, sehingga diperlukan *Preprocessing* data. Berikut adalah langkah-langkah untuk *Preprocessing* data.

- Case folding merupakan proses mengubah semua teks data *tweet* menjadi huruf kecil (*lowercase*).
- Cleansing* adalah proses penghapusan data *tweet* dari *noise* seperti *hashtag*, *username*, *url*, tanda baca, dan angka.
- Menghapus *Stopword*, merupakan tahap membuang kata yang dianggap tidak penting. Seperti kata imbuhan yang tidak memiliki kata seperti ‘kami’, ‘aku’, ‘kalau’, dan lain - lain.
- Tokenizing* adalah pemecahan kalimat asli menjadi kata-kata atau pemisahan urutan *string* menjadi beberapa bagian, seperti kata berdasarkan semua kata yang membentuk kata tersebut.
- Stemming* merupakan bertujuan untuk menghilangkan imbuhan sehingga merubah kata-kata dalam kalimat menjadi kata dasar.

Berikut proses *Preprocessing* di Gambar 3.



Gambar 3. Proses *Preprocessing*

2.5 Algoritma *Naïve bayes*

Naïve bayes adalah metode klasifikasi dari teorema dan probabilitas *Bayes* yang mengasumsikan setiap variabel *X* adalah independen. Yang artinya, *Naïve bayes* berasumsi keberadaan satu atribut (variabel) tidak ada hubungannya dengan keberadaan atribut lain (variabel). Dalam algoritma *Naïve bayes* setiap dokumen diwakili oleh sepasang atribut “*x*₁, *x*₂, *x*₃,...*x*_{*n*}” dimana *x*₁ adalah kata pertama, *x*₂ adalah kata kedua dan seterusnya. Sedangkan *V* adalah kumpulan kategori *Tweet* [4]. Selama klasifikasi, algoritma mencari probabilitas tertinggi di antara semuanya. Kategori Dokumen Teruji (VMAP) memiliki kesamaan adalah sebagai berikut:

$$P(V_j) = \frac{|\text{docs } j|}{|\text{kosakata}|} \quad (1)$$

Keterangan :

|docs *j*| = jumlah dokumen setiap kategori *j*

|contoh| = jumlah dokumen dari semua kategori

n_k = jumlah frekuensi kemunculan setiap kata

n = jumlah frekuensi kemunculan kata dari setiap kategori

|kosakata| = jumlah semua kata dari semua kategori

2.6 Algoritma C4.5

Algoritma C4.5 merupakan algoritma menyediakan sekumpulan data berlabel dan menghasilkan pohon keputusan sebagai *output* [4]. Algoritma C4.5 adalah algoritma klasifikasi dan merupakan hasil dari algoritma

induksi pohon keputusan ID3 (*Iterative Dichotomiser 3*). Diperkenalkan Quinlan (1996) sebagai versi perbaikan dari ID3 [5]. Langkah-langkah untuk membuat algoritma C4.5:

- Pilih atribut sebagai root.
- Buat cabang untuk setiap *record* atribut.
- Membagi kasus ke dalam cabang.
- Mengulangi proses untuk setiap cabang sampai semua kasus di dalam cabang adalah keputusan yang baik.

Dalam proses memilih atribut sebagai akar, perlu menghitung *entropy* dan *gain*. *Gain* dengan nilai tertinggi akan digunakan sebagai *root* atau akar awal [4]. Untuk menghitung *gain* dapat dilakukan dengan menggunakan rumus di berikut

$$Gain(A) = Entropy(S); \sum_{i=1}^n P_i * \log_2(P_i) \quad (2)$$

Di mana,

S = Himpun Kasus

A = Atribut

n = Jumlah Sampel

|S_i| = Jumlah Kasus pada Partisi ke-i

|S| = Jumlah Kasus dalam S

Selanjutnya untuk perhitungan nilai *entropy* dilakukan dengan rumus:

$$Entropy(S_1, S_2, \dots, S_n) = \sum_{i=1}^n P_i * \log_2 \quad (3)$$

Dengan,

S = Himpunan Kasus

n = Jumlah sampel

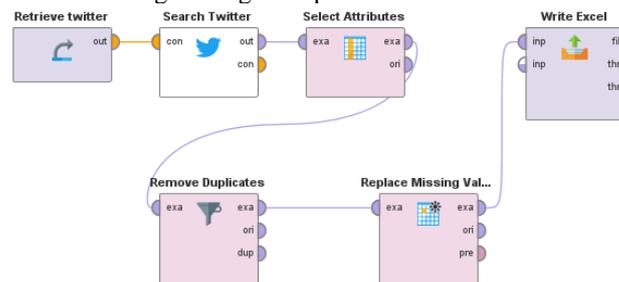
P_i = proporsi kelas

Setelah *entropy* dihitung untuk mencari *gain*, sampai ditemukan *gain* dengan nilai terbesar selesai, maka dapat dihasilkan pohon keputusan yang dapat diubah menjadi aturan yang dapat digunakan menggunakan aturan IF-THEN-ELSE.

3. HASIL DAN PEMBAHASAN

3.1 Crawling

Crawling pada penelitian ini menggunakan RapidMiner. Data penelitian bersumber pada media sosial Twitter. Data *tweet* yang berhasil dikumpulkan akan disimpan di sebuah file excel (.xlsx), lalu kemudian dimasukkan ke basis data (*database*) MySQL. Proses *crawling* diambil dengan keyword pencarian “Presiden 3 Periode”. Berikut proses *Crawling* dalam penelitian ini dengan dengan RapidMiner :



Gambar 4. Proses *Crawling* menggunakan RapidMiner

- Operator *Retrieve twitter* untuk menyambungkan RapidMiner dengan twitter.
- Operator *Search Twitter* untuk mencari data (*tweet*) dari kata kunci 'presiden 3 periode'.
- Operator *Select Attributes* untuk memilih atribut yang ingin diperlukan. Dalam penelitian ini diperlukan atribut *text* (*tweet*).
- Operator *Remove Duplicate* untuk menghapus *tweet* yang sama (*spam*).
- Operator *Replace Missing Value* untuk menghapus *tweet* yang tidak lengkap.
- Operator *Write Excel* untuk menulis (*export*) data ke *file excel* (.xlsx).

3.2 Pelabelan

Pada tahapan pelabelan dilakukan pemberian label (kelas) berdasarkan karakteristik pada sebuah dokumen atau kalimat. Dalam proses pelabelan dapat dilakukan dengan cara pelabelan manual. Tahap pelabelan manual ini

dilakukan oleh peneliti sebagai validasinya untuk memberikan label ke seluruh data yang berjumlah 210 *tweet*. Setelah tahap ini data akan mempunyai label negatif dan positif dengan label negatif berjumlah 136 *tweet* dan positif berjumlah 74 *tweet*. Hasil data yang sudah di label, kemudian diverifikasi oleh seorang ahli yaitu Dr. Bayquni, M.Pd, M.I.Kom. (Pakar Bidang Komunikasi dan Kebijakan Publik) dari Universitas Prof. Dr. Moestopo Beragama.

Tabel 3. Hasil Pengujian 1

No	Kalimat	C4.5	Naïve Bayes
1	Sayangnya presiden dundu justru minta 3 periode, saya dukung	positive	positive
2	Ratusan Mahasiswa Tolak 3 Periode Masa Jabatan Presiden divr.it/SMnhYy	negative	negative
3	Teriak gnti presiden juga itu urusan mereka dong ? Toh target mereka juga pemilu kala itu, Anda tidak adil, saat orang teriak gnti presiden anda sebut makar, tp org teriak presiden 3 periode anda bela dgn dalil amandemen,,	positive	negative
4	Dukungan terhadap Presiden Joko Widodo 3 periode menggema dalam acara Silaturahmi Nasional Asosiasi Pemerintah Desa Seluruh Indonesia (APDESI), di Istora Gelora Bung Karno, Jakarta, Selasa, 29 Maret 2022.	positive	negative
5	Btw naha sih si jokowi teh hayang 3 periode? Perasaan pas jaman sd diajarin nya kalo presiden cuma bisa 2 periode makanya dulu SBY gabisa jadi presiden lagi, lahh ieu si mahiwal ojol2, hayang terus we mereun maruk amat wkwkwkwk	negative	negative
6	adi presiden seumur hidup aja ... kalau 3 periode nanggung	negative	negative
7	hey 2 tahun doang. pak presiden aja betah jadi presiden 3 periode, masa kamu ga betah 3 periode?	negative	negative
8	Saya tidak setuju dgn kebijakan pemerintah Karena kami sebagai masyarakat atau rakyat indonesia semakin tertekan akibat pandemi dan disisi lain adanya wacana yg mengada2 presiden 3 periode ! Wow amat memilukan Jangan2 mau seumur hidup ! !!	negative	negative
9	Padahal rakyat nya banyak yg gizi buruk, wakil presiden nya malah blg gini. Gausah gimpi 3 periode deh	positive	negative
10	Karena si kafir ini gak bisa jadi presiden,direzim mukidi inilah si kafir ini berkuasa melebihi presiden,makanya minta 3 periode. SiKafir ini sadar,qlo dia gak bisa jadi presiden melalui pemilu,melalui rezim Mukidi ini dia berkuasa melebihi jabatan presiden.	negative	negative
11	Jokowi 3 periode??wkwkwkw.. Itulah usulan konyol para pencari muka dan penjilat pantat penguasa!Sudah jelas2 konstitusi dan UU mengatur presiden 2 periode,apakah ini langkah awal menuju penguasa tirani dan diktator baru?!!	negative	negative
12	Women support women katanya bullshit. Ya emang. Daripada support women gatau malu mending support presiden 3 periode. Penhouse aja sampe season 3, masa presiden kalah. Info buzzer mas mbak kebetulan jago berantem. Makasih	negative	negative
13	Jangan lupa njenengan sampaikan juga untuk tidak memperpanjang masa jabatan presiden atau merubah konstitusi untuk ambisi 3 periode. Walaupun dia sampaikan taat konstitusi tapi masyarakat sudah paham dan hafal bagaimana perilaku beliau selama ini, isuk dele sore tempe. #salamwaras	negative	negative
14	Masa g hafal kelakuan nya... Isuk dele sore tempe. Jokowi adalah salah satu contoh Manusia munafik yg ada di pemerintahan, jadi ketika dia berbicara 10 yg bener 1. Apalagi ini bicara perpanjangan dan 3 periode. Jadi presiden itu enak tinggal gunakan kepres bisnis jalan	positive	negative
15	udah gila apa 3 periode? buseeet di Indonesia 2 periode cukup pakdeee! udah deeh emg ga capek jadi presiden? ☐	positive	negative
16	Apa Pak Jokowi tidak takut Azab dari Allah SWT karena Melanggar Sumpah Presiden dibawah Al Qur'an saat dilantik menjadi Presiden, bila tidak tegas menolak Jabatan Presiden 3 Periode !!!!!. Semoga Allah SWT memberi Hidayah ke Jalan Yang Lurus.	negative	negative
17	@Angguneeee Sy sbg rakyat jg akan menyampaikan aspirasi. Silahkan pak Jokowi dan pendukung 3 periode makan TAI masing² 1 kg perhari. Ini jg aspirasi dari bawah. Jika mereka bisa mengabulkan aspirasi saya, maka akan sy dukung jd presiden seumur hidup. RUSAK RUSAK DAH twitter.com/Angguneeee/sta...	positive	negative
18	Naikin semua harga gitu, pantas ga masih dipercaya utk jd presiden? Pantas ga minta 3 periode?	positive	negative
19	Semoga Presiden @jokowi dan Wakil Presiden @Kiyai_MarufAmin segera pro-aktif menghentikan segala bentuk narasi dukungan 3 periode. Narasi dukungan 3 periode itu bentuk terorisme terhadap UUD.	positive	negative
20	Kami maafkan Pak @jokowi kita sama2 saling memaafkan.. Tapi jangan perpanjang masa jabatan Presiden dan jangan 3 periode ya Pak.. #2024Gantian	negative	negative
21	Tolak 3 periode!! Tolak penundaan pemilu 2024!! Tolak perpanjangan masa jabatan presiden!! Tolak melanggar konstitusi!!	negative	negative
22	pengen jadi presiden 3 periode	positive	negative
23	Maaf, Survei Menunjukkan Hanya 5% Warga yang Mau Presiden Jokowi 3 Periode	positive	negative
24	Dukungan terhadap Presiden Joko Widodo 3 periode menggema dalam acara Silaturahmi Nasional Asosiasi Pemerintah Desa Seluruh Indonesia (APDESI), di Istora Gelora Bung Karno, Jakarta, Selasa, 29 Maret 2022. Siapa tokoh di balik mobilisasi kades ini? #TempoThread	positive	negative
25	Gerakan tagar 2019 ganti presiden saja yang jelas-jelas sesuai momen saja ditangani dengan keras oleh aparat. Ini mestinya yang gaung2kan dan lontarkan ide Jokowi 3 periode bisa langsung di-dor saja. Jelas2 mau melanggar UUD.	negative	negative

3.3 Pembagian Data

Pada tahapan pembagian data, *tweet* yang telah berlabel akan dibagi menjadi data latih dan data uji. Pada penelitian ini proses pembagian data dibagi menjadi 20% data uji dan 20% menjadi data uji.

- Data uji adalah data yang tersedia dengan tujuan menguji tingkat keakuratan algoritma klasifikasi berdasarkan model latih.
- Data Latih adalah sebuah kumpulan data yang berasal dari data yang sudah melalui tahap *Preprocessing*.

3.4 Preprocessing

Tahap *Preprocessing* merupakan tahapan yang hanya dapat dilakukan setelah tersedianya satu atau lebih *Dataset* pada basis data (*database*) hasil dari tahapan pengumpulan data. Proses *Preprocessing* antara lain:

- Case folding dilakukan penyetaraan teks menjadi huruf kecil secara keseluruhan, misalnya: 'PRESIDEN' akan diubah menjadi 'presiden', 'Amandemen' akan diubah menjadi 'amandemen' dan seterusnya.
- Cleansing* dilakukan penyaringan dan pembuangan teks. Proses *cleansing* terdiri atas beberapa tahapan antara lain: menghapus *URL*, mention (@mention), *hashtag* (#hashtag), selain huruf.
- Menghapus *stopword* akan menghapus setiap kata yang kurang memiliki makna namun sering dijumpai dalam sebuah teks, misalnya kata: 'sama' 'yang', dan seterusnya. Proses penghapusan tersebut melibatkan kamus *stopword* yang terdapat dalam basis data (*database*).
- Tokenizing* adalah pemecahan kalimat asli menjadi kata-kata yang berasal dari tahap sebelumnya berdasarkan setiap kata penyusunnya. Berikut contoh *tokenizing* "saya menolak presiden tiga periode." Setelah proses *Tokenizing* pada kalimat tersebut, akan diperoleh lima kata yaitu: "saya", "menolak", "presiden", "tiga", "periode".
- Proses *stemming* merupakan proses mengubah kata yang berimbuhan menjadi kata dasar dengan melibatkan *Stemming* NaziefAdriani, misalnya kata: 'berdua' menjadi 'dua', 'pasangan' menjadi 'pasang', dan seterusnya.

3.5 Hasil Pengujian

Pada tahapan Pengujian data uji akan di uji melalui algoritma yang telah ditentukan yaitu algoritma *Naive Bayes* dan algoritma C4.5. Pada proses pengujian algoritma *Naive Bayes* dan C4.5 akan menghasilkan dua klasifikasi yaitu Positif dan Negatif. Hasil pengujian pada metode algoritma *Naive bayes* dan C4.5 dapat dilihat pada tabel 1 dan 2.

Tabel 2. Hasil Pengujian

26	@_Sridiana_3va Bingung saya, sejak dari mulai dilantik jadi Presiden sampai mau 3 periode menjabat kalau presiden ngomong artinya selalu terbalik...ya Allah tolonglah negara kami	negative	negative
27	Walaupun gw ga setuju masa jabatan Presiden 3 periode Atau Pemilu diundur... Tapi gw seneng sih kalo Pakdhe sampe 3 periode Seneng gw ngeliat kadrun GOBLOGGG makin kejang2 #AntiCapresYaman	negative	negative
28	Gua yakin, pak Presiden tak mau 3 periode. Pak Presiden taat konstitusi.	negative	negative
29	Terus kalo emang udah jadi 3 periode. Terus di tahun akhir masa jabatan ada rencana ganti konstitusi berapa masa jabatan lagi? Jabatan presiden kok kaya kontrak pemain bola sih:))	negative	negative
30	Jokowi juga hrs adil & bijaksana dgn tidak menafikan teriakan penolakan masyarakat thd dirinya utk menjabat 3 periode. Jadilah Presiden yg dikenang sbg negarawan.	negative	negative
31	Saya sangat tidak setuju jabatan Presiden 3 periode. Tp kalo perpanjangan waktu boleh lah, cuma 2x15 menit. Lanjut adu penalti	negative	negative
32	Pasal 29 ayat 4 huruf a dan b diduga telah dilakukan oleh Kades yang mendukung 3 Periode. Landasan Hukum menyampaikan Pendapat Dukung 3 Periode untuk jabatan Presiden sampai saat ini tidak ada di NKRI. Organisasi Masyarakat Kades diduga telah Ingkar Kepada UUD 1945 Pasal 7.	negative	negative
33	Aku berdoa biar aku bisa jadi presiden 3 periode	positive	negative
34	Sampe ngeluarin narasi "inflasi kan ada tiap presiden" buat justifikasi narasi 3 periode tuh kyk...gila lu sejauh apa coba dr realitas di sekitar lu. It's clear rakyat udah muak dan ini bahkan bkn ttg kenaikan harga doang. brooo touch some gras	positive	negative
35	Yg teriak siapa? Perasaan presiden sudah menegaskan menolak 3 periode. Kok digoreng terus?	positive	negative
36	Saya perwakilan De Heeren Zeventien menolak masa jabatan presiden menjadi 3 periode karena menyalahi konstitusi negara tercinta ini. TTD Lord Prince Ranga	negative	negative
37	Kpd Presiden Jokowi td siang sy sampaikan kekurang setujuan rakyat kpd Rencana unt MASA JABATAN PRESIDEN 3 PERIODE, Kata Presiden, SAYA TAAT KEPADA KONSTITUSI MASA JABATAN PRESIDEN TETAP 2 PERIODE DAN TIDAK AKAN PERPANJANG MASA JABATAN SAYA.. menunggu tweet Pak Menko seperti	negative	negative
38	Selama Jokowi presiden pendapatan masyarakat makin turun, banyak usaha bangkrut. Indo masuk negara dgn pendapatan rendah. Jurang antara miskin kaya makin lebar, pendapatan pajak terendah sepanjang sejarah. Utang bangak. Kok nekat mau bangun IKN dan ngotot 3 periode? Vanke !!!	negative	negative
39	Horeee...3 periode... Lumayan nunggu anak2... nya berpengalaman jadi gubernur.... 5 thn selanjutnya Baru salah satu anak nya... Jadi Presiden... Lanjut lagi 3 priode sampai anak cucu cicit nya besar....calon presiden... Wkwkwkwkwk...	positive	negative
40	Saya juga menolak presiden 3 periode apalagi cuma 2 periode, seharusnya presiden @jokowi menjabat seumur hidup, biar gerombolan2 @PKSejahtera bunuh diri masal.	negative	negative
41	Saya yakin tanpa mahasiswa rakyat mampu menolak perpanjangan jabatan presiden/tunda pemilu dan keinginan 3 periode ala rezim boneka	negative	negative
42	@lvOneNews @MataNajwa saya menolak presiden 3 periode atau tambah masa jabatan atau tunda pemilu.	negative	negative

3.6 Hasil Confusion Matrix

Pengujian adalah hal yang perlu dilakukan untuk mengevaluasi, menganalisis dan mengetahui tingkat akurasi atau kesamaan hasil yang diperoleh sistem yang di rancang. Dalam penelitian ini, algoritma *Naive bayes* dan

algoritma C4.5 digunakan untuk memprediksi label pada data uji dan pengujian yang telah dilakukan dalam memeriksa keakuratannya. Pengujian dilakukan dengan menggunakan Instrumen *Confusion matrix*:

Tabel 3. Hasil *Confusion Matrix*

Algoritma	Akurasi	<i>Precision</i>	<i>Recall</i>	F1 Score
<i>Naïve bayes</i>	85%	50%	16%	24%
C4.5	78%	53%	80%	63%

4. KESIMPULAN

Berdasarkan hasil pengujian dan evaluasi dari aplikasi yang dibuat menggunakan *Dataset* dan algoritma yang diusulkan, maka dapat disimpulkan berdasarkan 210 *tweet*, arah pandangan (sentimen) masyarakat Indonesia terhadap Presiden 3 Periode dari 5 kali pengujian nilai Algoritma C4.5 berhasil mendapatkan akurasi 78%, Sedangkan algoritma *Naïve bayes* mendapatkan akurasi 85%. Disimpulkan bahwa akurasi *Naïve bayes* lebih tinggi dari Algoritma C4.5. Tahapan utama yang terdapat dalam penelitian ini antara lain: *crawling*, preprocessing, pelabelan, pemodelan, pembagian data, dan klasifikasi Algoritma *Naïve bayes* dan C4.5. Tahap *Preprocessing* yang baik menjadi penentu terbentuknya hasil yang optimal untuk tahap selanjutnya.

DAFTAR PUSTAKA

- [1] A. P. Natasuwarna, “Analisis Sentimen Keputusan Pemindahan Ibukota Negara Menggunakan Klasifikasi Naive Bayes,” *Semin. Nas. Sist. Inf. dan Tek. Inform.*, pp. 47–53, 2019.
- [2] R. I. N. M. Wijaya, Tamora Nonia, “Analisis Sentimen Opini Publik Tentang Undang- Undang Cipta Kerja Pada Twitter,” *Jambura J. Electr. Electron. Eng.*, vol. 3, pp. 78–83, 2021.
- [3] F. V. Sari and A. Wibowo, “Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode *Naïve bayes* Classifier Berbasis Konversi Ikon Emosi,” *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [4] Z. E. Sholikha, E. Y. Puspaningrum, and W. S. JS, “Analisa Sentimen Pengguna E-Money Pada Twitter Menggunakan Algoritma C4.5 dan Naive Bayes,” *J. Inform. dan Sist. Inf.*, vol. 1, no. 3, pp. 1063–1071, 2020.
- [5] V. Purwayoga, T. Setiadi, and L. Zahrotun, “Pemanfaatan Text Mining Untuk Mengklasifikasi SMS Spam Porno Dengan Menggunakan Algoritma C4.5,” no. January, p. 5, 2020, doi: 10.32534/int.v12i01.1264.
- [6] E. Fitriani, “Perbandingan Algoritma *Naïve bayes* dan C4.5 Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan,” *Sistemasi*, vol. 9, no. 1, p. 103, 2020, doi: 10.32520/stmsi.v9i1.596.
- [7] F. Ratnawati, “Implementasi Algoritma *Naive Bayes* Terhadap Analisis Sentimen Opini Film Pada Twitter,” *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.
- [8] T. Krisdiyanto, “Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan *Naïve bayes Clasifiers*,” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 7, no. 1, pp. 32–37, 2021, [Online]. Available: <http://ejournal.uin-suska.ac.id/index.php/coreit/article/view/12945>.
- [9] W. A. Prabowo and C. Wiguna, “Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM,” *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 149, 2021, doi: 10.30865/mib.v5i1.2604.
- [10] Y. S. Mahardika and E. Zuliarso, “Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma *Naives Bayes*,” *Pros. SINTAK 2018*, no. 2015, pp. 409–413, 2018.