AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024

ANALISIS SENTIMEN PUBLIK TERHADAP PELAKSANAAN PERSIAPAN PEMILU TAHUN 2024 MENGGUNAKAN METODE *MAXIMUM ENTROPY* (MAXENT)

Indra Ismawan[1], Mustikasari[2], Wahyuddin Saputra[3]

[1,2,3]Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Alauddin Makassar, Jl. H.M. Yasin Limpo No. 36, Gowa, Sulawesi Selatan, Indonesia, 92118.
Email: [1]60200117033@uin-alauddin.ac.id ,[2]mustikasari@uin-alauddin.ac.id,
[3]wahyuddin.saputra@uin-alauddin.ac.id

***Koresponden Author:** Indra Ismawan, 60200117033@uin-alauddin.ac.id
Accepted: 20 02, 2024 ; Revised: 26 02, 2024; Published: 28 02, 2024

Abstrak

Komisi Penyelenggara Pemilihan Umum (KPU) dan Badan Pengawas Pemilu (Bawaslu) bertugas untuk mengkoordinasikan dan mengawasi Pemilu di seluruh wilayah Indonesia. Meskipun demikian, pemilu tidak lepas dari berbagai sengketa, termasuk pelanggaran kode etik, dan administrasi. Pemilihan Umum adalah salah satu cara yang penting untuk masyarakat dalam menjalankan hak partisipasi politik mereka. Dalam era teknologi informasi, masyarakat semakin aktif dalam mengungkapkan opini dan sentimen mereka melalui media sosial seperti Twitter. dengan memanfaatkan data dari media sosial. Metode maximum entropy digunakan untuk mengklasifikasikan opini yang terdapat dalam tweet berdasarkan nilai entropy. Penelitian ini juga bertujuan untuk mengukur tingkat akurasi metode maximum entropy dalam melakukan klasifikasi sentimen. Dari hasil penelitian ini dengan menggunakan kurang lebih 1.400 data, maka didapatkan hasil akurasi sebesar 87,08% sudah cukup baik dan sangat memungkinkan untuk dikembangkan sehingga diperoleh hasil yang lebih baik.

Kata kunci: *analisis sentimen, klasifikasi, maximum entropy, pemilu 2024*

Abstract

[Public Sentiment Analysis Towards the Implementation of 2024 Election Preparations Using the Maximum Entropy Method] *The General Election Commission (KPU) and the Election Supervisory Agency (Bawaslu) are tasked with coordinating and overseeing elections throughout the territory of Indonesia. Nevertheless, elections are not without various disputes, including violations of ethical codes and administrative issues. General elections are an important way for the public to exercise their political participation rights. In the era of information technology, the public is increasingly active in expressing their opinions and sentiments through social media such as Twitter. Utilizing data from social media, the maximum entropy method is employed to classify opinions found in tweets based on entropy values. This research also aims to measure the accuracy level of the maximum entropy method in sentiment classification. From the results of this study, using approximately 1,400 data, an accuracy rate of 87.08% was obtained, which is quite good and highly feasible for further development to achieve better results.*

Keywords: *sentiment analysis, classification, maximum entropy, election 2024*

p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024

1. PENDAHULUAN

Indonesia menerapkan Demokrasi Pancasila sebagai fondasi sistem pemerintahannya, yang mengacu pada nilai-nilai Pancasila sebagai landasan Negara Republik Indonesia. Pemilihan Umum adalah salah satu bentuk implementasi demokrasi tersebut, yang memungkinkan masyarakat untuk berpartisipasi dalam pemilihan pemimpin sebagai bagian dari pemenuhan hak-hak asasi manusia.

Komisi Penyelenggara Pemilihan Umum (KPU) memiliki tanggung jawab untuk mengoordinasikan, menyelenggarakan, mengendalikan, dan memantau semua proses Pemilihan Umum, dengan dukungan dari Badan Pengawas Pemilu (Bawaslu) yang berperan sebagai lembaga pengawas dalam penyelenggaraan pemilu. Bawaslu bertindak sebagai lembaga independen yang mengawasi pelaksanaan Pemilihan Umum di seluruh wilayah Republik Indonesia.

Dalam kehidupan sosial, menyampaikan pendapat kepada orang lain telah menjadi rutinitas sehari-hari bagi banyak orang. Sentimen merujuk pada pendapat atau pandangan yang dipengaruhi oleh emosi yang kuat terhadap suatu hal. Seiring dengan kemajuan teknologi, saat ini masyarakat dapat menyampaikan pandangan mereka kepada publik melalui platform media sosial yang populer seperti *Facebook, Youtube, Twitter*, dan *Instagram*.

Data dari aplikasi *Twitter* dipilih sebagai sumber untuk analisis karena *platform* media sosial ini menjadi salah satu wadah utama bagi masyarakat untuk berbagi pendapat, pemikiran, dan sentimen mengenai berbagai isu yang relevan. *Twitter* memfasilitasi pengguna untuk dengan cepat dan mudah menyampaikan pandangan mereka melalui *tweet* yang singkat, sehingga menyediakan potensi besar untuk mengumpulkan data yang mencerminkan beragam opini dan pandangan dari berbagai kalangan masyarakat.

Informasi yang tersebar di *Twitter* sering kali mencerminkan tren, perubahan opini publik, dan bahkan potensi isu-isu penting yang mempengaruhi kebijakan dan tindakan pemerintah. Oleh karena itu, menggunakan data dari *Twitter* dalam analisis opini menjadi relevan dan bermanfaat dalam memahami dinamika masyarakat serta pembuatan kebijakan publik.

2. METODE

2.1. Data

Data yang diambil adalah *tweet* berbahasa Indonesia yang membahas tentang pelaksanaan persiapan pemilu tahun 2024. Pengumpulan data dilakukan dengan menyaring *tweet* dengan beberapa *keyword* yaitu pemilu 2024, persiapan KPU, dan pelaksanaan pemlilu 2024 di Indonesia. *Dataset* yang digunakan memiliki jumlah yang banyak yang nantinya menjadi data latih dan juga data uji dengan jumlah yang seimbang pada masing masing masing masing kelas positif dan negatif. Data yang diambil dari *twitter* kurang lebih 1.400 ribu data dari hasil *crawling* menggunakan bahasa *phyton*. Untuk pengujian yang dilakukan menggunakan data *training* dan data *testing*. Data *training* dan data *testing* sebelumnya sudah dilabelkan ke dalam kelas positif, netral dan negatif.

Untuk mengumpulkan data dan informasi yang akan digunakan dalam pembuatan aplikasi, maka dilakukan metode pengumpulan data dengan cara mengumpulkan jurnal, *paper* dan *browsing internet.*

2.2. Sentimen Analisis

Analisis sentimen adalah proses di mana data teks dipahami, diekstraksi, dan diolah secara otomatis untuk mendapatkan informasi tentang sentimen yang terkandung dalam suatu kalimat atau opini [1]. Analisis sentimen dilakukan untuk mengevaluasi pendapat atau kecenderungan opini seseorang terhadap suatu masalah atau

p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024

objek, apakah cenderung bersifat negatif, positif, atau netral.

2.3. Pemilu

Pemilihan Umum adalah sebuah proses demokratis di mana warga negara suatu negara secara langsung memilih wakil rakyat atau pejabat pemerintahan. Ini merupakan mekanisme kunci dalam sistem demokrasi modern yang memberikan kesempatan kepada rakyat untuk berpartisipasi dalam menentukan pemimpin dan kebijakan negara [2]. Tujuan utama dari pemilihan umum adalah memberikan kesempatan kepada warga negara untuk menyampaikan suara mereka dan memilih para pemimpin yang akan mewakili mereka di pemerintahan [3].

2.4. Klasifikasi

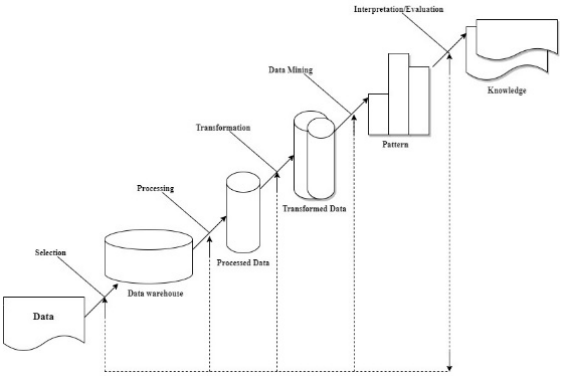
Klasifikasi adalah proses dalam menemukan sebuah model atau fungsi yang dapat menjelaskan dan membedakan data ke dalam kelas-kelas yang telah ditentukan sebelumnya. Ini melibatkan pemeriksaan karakteristik dari objek data dan menempatkannya ke dalam salah satu kelas yang telah didefinisikan sebelumnya [4].

Dalam klasifikasi, terdapat beberapa metode atau model yang digunakan untuk menyelesaikan kasus klasifikasi, di antaranya *naïve bayes, maximum entropy (MaxEnt), memory-based reasoning dan support vector machine (SVM).*

2.5. Data Mining

Data *mining* adalah proses penemuan pola atau informasi yang bermanfaat dari kumpulan data yang besar atau kompleks, menggunakan berbagai teknik dan algoritma komputasi. Tujuannya adalah untuk mengidentifikasi hubungan yang tidak terlihat secara langsung, tren, atau pola yang signifikan dalam data yang dapat membantu dalam pengambilan keputusan atau prediksi di masa depan [5].

Data *mining* adalah proses ekstraksi pengetahuan atau informasi yang bermanfaat dari sejumlah besar data yang tersedia. Ini melibatkan penggunaan teknik analisis statistik, matematika, kecerdasan buatan, dan pemodelan komputasional untuk mengidentifikasi pola, tren, atau hubungan yang mungkin tidak terlihat secara langsung dalam data mentah. Tujuan utamanya adalah untuk mendapatkan pemahaman yang lebih dalam tentang data dan menghasilkan pengetahuan yang dapat digunakan untuk pengambilan keputusan yang lebih baik atau prediksi di masa depan. Adapun proses dari tahapan data mining seperti gambar 2.1 berikut :



Gambar 2.1. Tahapan Data *Mining*

2.6. Preprocessing Data

Preprocessing dilakukan untuk membersihkan dan menyiapkan data agar sesuai untuk analisis lebih lanjut. Tujuan utamanya adalah mengatasi data yang tidak sempurna, menghilangkan gangguan atau noise, serta memperbaiki konsistensi data. Selain itu, pada tahap *preprocessing,* digit angka, huruf kapital, atau karakter lainnya seringkali dihapus atau diubah, dan teks dalam bahasa asing juga dihilangkan agar data menjadi lebih *homogen* dan relevan untuk analisis [6]. Hal ini dilakukan untuk menggali, mengolah, dan mengatur informasi, serta menganalisis hubungan tekstual dari data terstruktur dan tidak terstruktur. Adapun tahapan *preprocessing* adalah *case folding,*

p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024

text cleaning, stemming, stopword, dan tokenizing.

2.7. Term frequency-inverse document frequency

Metode *Term Frequency-Inverse Document Frequency* (*TF-IDF*) adalah sebuah metode yang digunakan untuk menghitung bobot setiap kata dalam sebuah dokumen atau koleksi dokumen. Pembobotan ini pertama kali diperkenalkan dalam sistem temu kembali informasi, yang merupakan salah satu bidang dalam pengolahan bahasa alami dan ilmu informasi yang berkaitan dengan pengambilan informasi dari koleksi teks yang besar [1].

Metode *Term Frequency-Inverse Document Frequency* (*TF-IDF*) terkenal karena efisiensinya, mudah diterapkan, dan hasil yang akurat dalam mengukur pentingnya sebuah kata dalam sebuah dokumen. *TF-IDF* adalah cara untuk memberikan bobot pada sebuah kata berdasarkan frekuensi kemunculan kata tersebut dalam dokumen dan kebalikannya dalam seluruh koleksi dokumen. Dengan kata lain, jika sebuah kata muncul lebih sering dalam suatu dokumen, tetapi jarang muncul dalam dokumen-dokumen lain dalam koleksi, maka kata tersebut dianggap memiliki bobot yang tinggi untuk dokumen tersebut.

Sebaliknya, jika kata tersebut sering muncul dalam banyak dokumen dalam koleksi, maka bobotnya akan lebih rendah karena kata tersebut dianggap kurang spesifik atau unik dalam mendeskripsikan dokumen tersebut [7]. Rumus pembobotan *TF-IDF* dapat dilihat pada persamaan 1:

$$W_{i,j} = tf_{i,j} \times log\left(\frac{N}{d_{fi}}\right) \quad (1)$$

Keterangan :

$W_{i,j}$= Bobot kata i pada dokumen j
$tf_{i,j}$=Jumlah kemunculan/frekuensi kata i dalam dokumen j
d_{fi}=Jumlah dokumen yang mengandung kata i
N = Total dokumen

2.8. Maximum Entropy

Entropy mengukur tingkat ketidakpastian atau keacakan dalam sebuah sistem. Dalam konteks teori informasi, *entropy* menggambarkan seberapa banyak informasi yang dibawa oleh suatu sumber informasi. Nilai *entropy* lebih tinggi menunjukkan tingkat ketidakpastian yang lebih besar atau kejadian yang lebih acak, sedangkan nilai *entropy* lebih rendah menunjukkan tingkat ketidakpastian yang lebih rendah atau informasi yang lebih teratur [8].

Entropy sering dikaitkan dengan nilai probabilitas. Ketika probabilitas suatu kejadian mendekati 0 atau 1, *entropy* cenderung rendah karena tingkat ketidakpastiannya rendah. Sebaliknya, ketika probabilitas suatu kejadian mendekati 0,5 (probabilitas yang sama untuk setiap hasil), *entropy* mencapai maksimum karena tingkat ketidakpastiannya maksimum. Dengan demikian, *entropy* dapat dianggap sebagai ukuran untuk seberapa baik kita dapat memprediksi kejadian berdasarkan pengetahuan yang kita miliki tentang distribusi probabilitasnya.

Maximum Entropy (*MaxEnt*) adalah model yang digunakan untuk menghasilkan distribusi probabilitas yang memiliki tingkat ketidakpastian maksimum atau *entropy maksimum*, sementara masih mempertimbangkan batasan-batasan yang diberikan. Dalam konteks analisis sentimen pada *tweet, maximum entropy* dapat digunakan untuk memprediksi probabilitas berbagai label sentimen (misalnya positif, negatif, atau netral) yang mungkin terkandung dalam suatu *tweet* berdasarkan pada fitur-fitur atau karakteristik tertentu dari teks *tweet* tersebut [9].

Dengan menggunakan *Maximum Entropy*, model akan mencoba untuk

AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE
p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024

menyesuaikan distribusi probabilitas terbaik yang mungkin untuk label-label sentimen berdasarkan pada informasi yang ada dalam data pelatihan. Kemudian, model tersebut dapat digunakan untuk memprediksi label sentimen dari *tweet* baru berdasarkan pada fitur-fitur yang diamati dalam teks *tweet* tersebut. Tujuannya adalah untuk menghasilkan prediksi yang seakurat mungkin berdasarkan pada informasi yang tersedia.

Fungsi umum dari *entropy* adalah sebagaimana pada persamaan 2:

$$Entropy\ (X) = -\sum_{i=1}^{n} P(X_i) x\ log_2\ P(X_i) \quad (2)$$

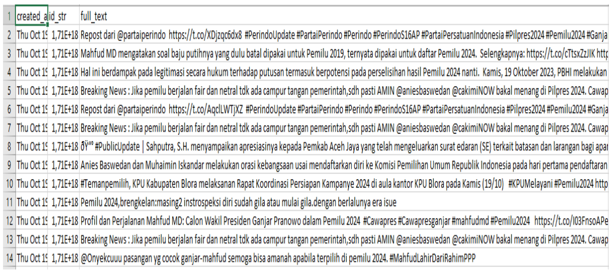
Keterangan :

Entropy (X) = Himpunan informasi dari suatu kejadian x

P(X) = Probabilitas dari kemunculan kejadian x

3. HASIL DAN PEMBAHASAN

Data hasil *crawling* (data mentah) menggunakan *twitter API* yang telah diberi label. Data disimpan dalam *microsoft office excel* menggunakan format *csv*. Jika sudah, maka datanya bisa dilakukan *preprocessing*. Berikut contoh data hasil *crawling* seperti gambar 3.1.



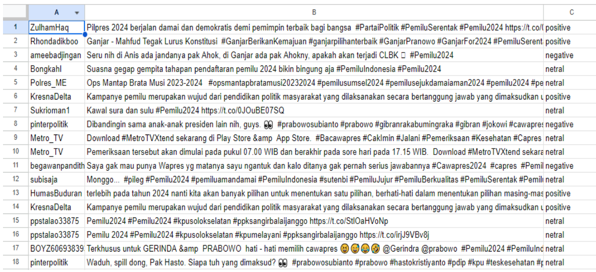
1	created_a	id_str	full_text
2	Thu Oct 1!	1,71E+18	Repost dari @partaiperindo https://t.co/XDjzqc6dx8 #PerindoUpdate #PartaiPerindo #Perindo #PerindoS16AP #PartaiPersatuanIndonesia #Pilpres2024 #Pemilu2024 #Ganja
3	Thu Oct 1!	1,71E+18	Mahfud MD mengatakan soal baju putihnya yang dulu batal dipakai untuk Pemilu 2019, ternyata dipakai untuk daftar Pemilu 2024. Selengkapnya: https://t.co/cTtsxZzJIK http
4	Thu Oct 1!	1,71E+18	Hal ini berdampak pada legitimasi secara hukum terhadap putusan termasuk berpotensi pada perselisihan hasil Pemilu 2024 nanti. Kamis, 19 Oktober 2023, PBHI melakukan
5	Thu Oct 1!	1,71E+18	Breaking News : Jika pemilu berjalan fair dan netral tdk ada campur tangan pemerintah,sdh pasti AMIN @aniesbaswedan @cakimiNOW bakal menang di Pilpres 2024. Cawap
6	Thu Oct 1!	1,71E+18	Repost dari @partaiperindo https://t.co/AqclLWTjXZ #PerindoUpdate #PartaiPerindo #Perindo #PerindoS16AP #PartaiPersatuanIndonesia #Pilpres2024 #Pemilu2024 #Ganja
7	Thu Oct 1!	1,71E+18	Breaking News : Jika pemilu berjalan fair dan netral tdk ada campur tangan pemerintah,sdh pasti AMIN @aniesbaswedan @cakimiNOW bakal menang di Pilpres 2024. Cawap
8	Thu Oct 1!	1,71E+18	ðŸ"° #PublicUpdate | Sahputra, S.H. menyampaikan apresiasinya kepada Pemkab Aceh Jaya yang telah mengeluarkan surat edaran (SE) terkait batasan dan larangan bagi apar
9	Thu Oct 1!	1,71E+18	Anies Baswedan dan Muhaimin Iskandar melakukan orasi kebangsaan usai mendaftarkan diri ke Komisi Pemilihan Umum Republik Indonesia pada hari pertama pendaftaran
10	Thu Oct 1!	1,71E+18	#Temanpemilih, KPU Kabupaten Blora melaksanan Rapat Koordinasi Persiapan Kampanye 2024 di aula kantor KPU Blora pada Kamis (19/10) #KPUMelayani #Pemilu2024 http
11	Thu Oct 1!	1,71E+18	Pemilu 2024,brengkelan:masing2 instrospeksi diri sudah gila atau mulai gila.dengan berlalunya era isue
12	Thu Oct 1!	1,71E+18	Profil dan Perjalanan Mahfud MD: Calon Wakil Presiden Ganjar Pranowo dalam Pemilu 2024 #Cawapres #Cawapresganjar #mahfudmd #Pemilu2024 https://t.co/I03FnsoAPe
13	Thu Oct 1!	1,71E+18	Breaking News : Jika pemilu berjalan fair dan netral tdk ada campur tangan pemerintah,sdh pasti AMIN @aniesbaswedan @cakimiNOW bakal menang di Pilpres 2024. Cawap
14	Thu Oct 1!	1,71E+18	@Onyekcuuu pasangan yg cocok ganjar-mahfud semoga bisa amanah apabila terpilih di pemilu 2024. #MahfudLahirDariRahimPPP

Gambar 3.1 *Crawling* Data

Setelah di *preprocessing* maka hasil seperti pada gambar 3.2

Gambar 3.2 Hasil *Preprocessing*

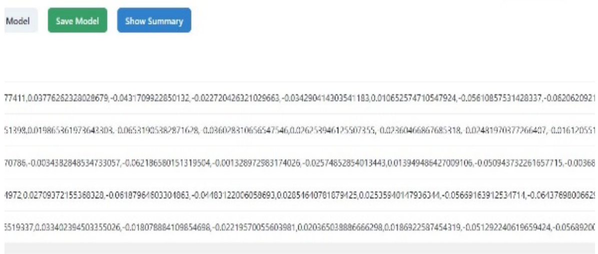
Data yang berhasil melalui tahap *preprocessing*. Tahap selanjutnya adalah melakukan *labelling* secara manual menggunakan *microsoft excel* yang seperti pada gambar 3.3. Hasil dari *labelling* datanya akan digunakan untuk klasifikasi *maximum entropy*.



	A	B	C
1	ZulhamHaq	Pilpres 2024 berjalan damai dan demokratis demi pemimpin terbaik bagi bangsa #PartaiPolitik #PemiluSerentak #Pemilu2024 https://t.co/(	positive
2	Rhondadikboo	Ganjar - Mahfud Tegak Lurus Konstitusi #GanjarBerikanKemajuan #ganjarpilihanterbaik #GanjarPranowo #GanjarFor2024 #PemiluSerent	positive
3	ameebadjingan	Seru nih di Anis ada jandanya pak Ahok, di Ganjar ada pak Ahokny, apakah akan terjadi CLBK ▯ #Pemilu2024	negative
4	Bongkahl	Suasna gegap gempita tahapan pendaftaran pemilu 2024 bikin bingung aja #PemiluIndonesia #Pemilu2024	netral
5	Polres_ME	Ops Mantap Brata Musi 2023-2024 #opsmantapbratamusi20232024 #pemilusumsel2024 #pemilusejukdamaiaman2024 #pemilu2024 #pe	netral
6	KresnaDelta	Kampanye pemilu merupakan wujud dari pendidikan politik masyarakat yang dilaksanakan secara bertanggung jawab yang dimaksudkan u	positive
7	Sukrioman1	Kawal sura dan sulu #Pemilu2024 https://t.co/0JOuBE07SQ	netral
8	pinterpolitik	Dibandingin sama anak-anak presiden lain nih, guys. 👀 #prabowosubianto #prabowo #gibranrakabumingraka #gibran #jokowi #cawapres	negative
9	Metro_TV	Download #MetroTVXtend sekarang di Play Store &amp App Store. #Bacawapres #CakImin #Jalani #Pemeriksaan #Kesehatan #Capres	netral
10	Metro_TV	Pemeriksaan tersebut akan dimulai pada pukul 07.00 WIB dan berakhir pada sore hari pada 17.15 WIB. Download #MetroTVXtend sekara	netral
11	begawanpandith	Saya gak mau punya Wapres yg matanya sayu ngantuk dan kalo ditanya gak pernah serius jawabannya #Cawapres2024 #capres #Pemil	negative
12	subisaja	Monggo... #pileg #Pemilu2024 #pemiluamandamai #PemiluIndonesia #sutenbi #PemiluJujur #PemiluBerkualitas #PemiluSerentak #Pemilu	netral
13	HumasBuduran	terlebih pada tahun 2024 nanti kita akan banyak pilihan untuk menentukan satu pilihan, berhati-hati dalam menentukan pilihan masing-mas	positive
14	KresnaDelta	Kampanye pemilu merupakan wujud dari pendidikan politik masyarakat yang dilaksanakan secara bertanggung jawab yang dimaksudkan u	positive
15	ppstalao33875	Pemilu2024 #Pemilu2024 #kpusolokselatan #ppksangirbalaijanggo https://t.co/StlOaHVoNp	netral
16	ppstalao33875	Pemilu 2024 #Pemilu2024 #kpusolokselatan #kpumelayani #ppksangirbalaijanggo https://t.co/irjJ9VBv8j	netral
17	BOYZ60693839	Terkhusus untuk GERINDA &amp PRABOWO hati - hati memilih cawapres 😆😅😂🤣 @Gerindra @prabowo #Pemilu2024 #PemiluInc	netral
18	pinterpolitik	Waduh, spill dong, Pak Hasto. Siapa tuh yang dimaksud? 👀 #prabowosubianto #prabowo #hastokristiyanto #pdip #kpu #teskesehatan #p	netral

Gambar 3.3 *Labelling Manual*

Setelah data dilakukan *vektorisasi* maka hasil seperti pada gambar 3.4.

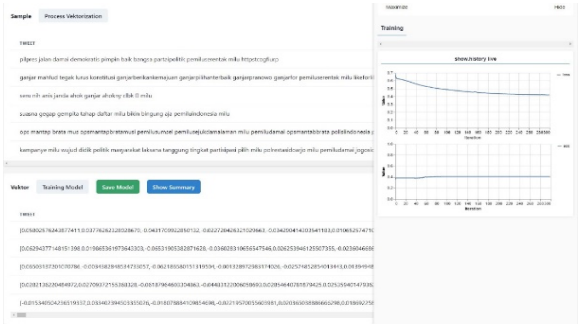


Gambar 3.4. *Vektorisasi*

Setelah melakukan training model maka akan tampil hasil seperti gambar 3.5 berikut.

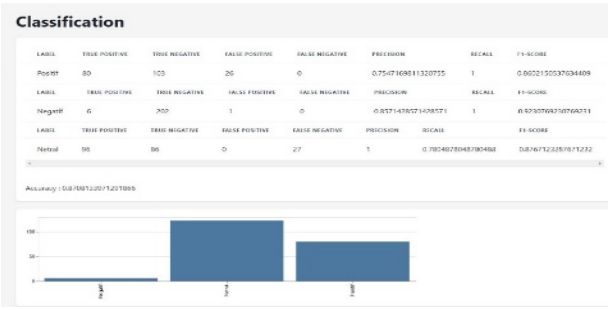
CC BY NC

p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024



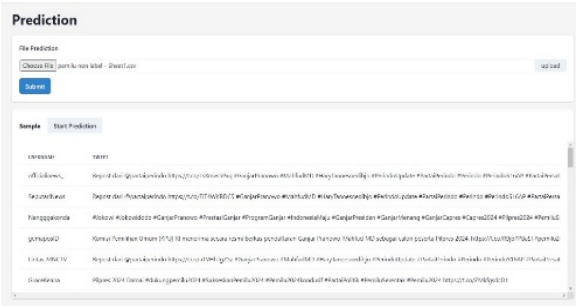
Gambar 3.5. *Training Model*

Hasil summary confusion matrix seperti gambar berikut 3.6.



Gambar 3.6 *Confusion Matrix*

Dataset *non* label yang akan di *processing* seperti pada gambar 3.7.



Gambar 3.7 *Data non labelling*

Result dan *dataset* yang telah di label otamatis sentimen positif, negatif dan netral. Seperti pada gambar berikut 3.8.

Gambar 3.8 *Data auto labelling*

Dan *result* juga akan menampilkan *wordcloud* positif, negatif dan netral seperti pada gambar 3.9.

Gambar 3.9 *Wordcloud*

Klasifikasi *Maximum Entropy perhitungan secara manual* sebagai berikut :

a) *Accuracy*

Akurasi merupakan nilai prediksi true (positif dan negatif) dari data yang digunakan, hasil *accuracy* adalah 87,08% didapatka n dari:

Accuracy=(80+103+6+202+96+86)/(80+26+0+103+6+1+0+202+96+27+0+86)*100% = 87,08%.

b) Precission

Precission merupakan nilai prediksi *true* positif dibagi dengan hasil keseluruhan nilai prediksi positif, berikut perhitungan *precession*:

Precission positif = 80/(80+26) = 0.75
Precission netral = 96/(96+0) = 1
Precission negatif = 6/(6+1) = 0.85

c) Recall

Recall merupakan nilai prediksi *true* positif dibagi dengan hasil keseluruhan nilai positif. Berikut adalah hasil perhitungan *recall:*

Recall positif = 80/(80+0) = 1
Recall netral = 96/(96+27) = 0.78
Recall negatif = 6/(6+0) = 1

p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024

d) F1-score

F1-score merupakan rata-rata dari *precision* dan *recall*. Berikut adalah hasil *f1-score*:

F1-score positif = 2*(0.75*1)/(0.75+1) = 0,86
F1-score netral = 2*(1*0.78)/(1+0.78) = 0,92
F1-score negatif = 2*(0.85*1)/(0.85+1) = 0,87

4. KESIMPULAN

Analisis sentimen menggunakan metode *Maximum Entropy* sangat bergantung pada dataset yang digunakan, representasi pada fitur yang digunakan, dan keakuratan model yang dilatih. Evaluasi yang baik terhadap metode *Maximum Entropy* juga penting untuk memastikan bahwa metode yang digunakan dapat mengklasifikasikan sentimen dengan akurasi yang akurat. Oleh karena itu, penting untuk menguji dan memvalidasi metode menggunakan set pengujian yang *independent* untuk mendapatkan kesimpulan yang dapat diandalakan.

Dari hasil pengujian didapatkan, data dibagi 1000 untuk data *training* dan 400 untuk data *testing*, akurasi yang didapatkan dari model yang dibagun yaitu 87,08%. Selain itu, untuk evaluasi *precision, recall, f1-score* didapatkan bahwa nilai evaluasi pada *class* positif 86%, untuk *class* netral 92% dan *class* negatif 87% yang dalam pengimplementasiannya dapat dianggap relatif sudah cukup baik namun masih sangat memungkinkan untuk dilakukan perbaikan, baik dari segi *preprocessing* data maupun pemilihan metode yang digunakan yang lebih sesuai dengan karakteristik data.

5. DAFTAR PUSTAKA

[1] M. F. Putra, A. Herdiani, and D. Puspandari, "Analisis Pengaruh Normalisasi, TF-IDF, Pemilihan Feature-set Terhadap Klasifikasi Sentimen Menggunakan Maximum Entropy (Studi Kasus : Grab dan Gojek)," *e-Proceeding of Engineering*. pp. 8520–8529, 2019.

[2] D. M. Liando, "Pemilu dan Partisipasi Politik Masyarakat (Studi Pada Pemilihan Anggota Legislatif Dan Pemilihan Presiden Dan Calon Wakil Presiden Di Kabupaten Minahasa Tahun 2014)," *J. LPPM Bid. EkoSosBudKum*, vol. 3, no. 2, pp. 14–28, 2016.

[3] N. M. Rahma, "Review Buku Sebuah Pembelajaran Untuk Pemilu yang Akan Datang," *J. Penelit. Polit.*, vol. 19, no. 1, pp. 67–72, 2022, [Online]. Available: https://ejournal.politik.lipi.go.id/

[4] M. Cindo and D. P. Rini, "Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) Literatur Review: Metode Klasifikasi Pada Sentimen Analisis," *Januari*, pp. 66–70, 2019, [Online]. Available: https://seminar-id.com/semnas-sainteks2019.html

[5] I. Pramudiono, "Pengantar Data Mining: Menambang Permata Pengetahuan di Gunung Data." 2016.

[6] D. B. Setyohadi, F. A. Kristiawan, and E. Ernawati, "Perbaikan Performansi Klasifikasi Dengan Preprocessing Iterative Partitioning Filter Algorithm," *Telematika*, vol. 14, no. 01, pp. 12–20, 2017, doi: 10.31315/telematika.v14i01.1960.

[7] V. Amrizal, "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)," *J. Tek. Inform.*, vol. 11, no. 2, pp. 149–164, 2018, doi: 10.15408/jti.v11i2.8623.

[8] B. Herwijayanti, D. E. Ratnawati, and L. Muflikhah, "Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity," *Jurnal Pengembangan*

p-ISSN: 2746-9204
e-ISSN: 2746-9190
Volume 4, No. 1, 2024

Teknologi Informasi dan Ilmu Komputer, vol. 2, no. 1. pp. 306–312, 2018. [Online]. Available: http://j-ptiik.ub.ac.id

[9] A. B. Alroy, P. P. Adikara, and S. Adinugroho, "Klasifikasi Hoaks Menggunakan Metode Maximum Entropy Dengan Seleksi Fitur Information Gain," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 9, pp. 9291–9298, 2019, [Online]. Available: http://j-ptiik.ub.ac.id

CC BY NC